6th Conference on Advanced Space Technology

Book of Papers

10-12 NOV, 2015 Shanghai, China

Hosted by:

Chinese Society of Astronautics (CSA)

International Academy of Astronautics (IAA)

Contents

Researches on Multi-component Calibration-Free Trace Gas sensor in Manned Space Vehicle

.....Jiang Meng, Wang Xuefeng, Feng Qiaoling, Liang Hu, Member, IEEE 1

Application and Development of Enhanced Satellite-based Navigation Systems

..... Lianyin Deng and Shumin Gao 8

The Lightning Imager Instrument for Meteosat Third Generation

S. Lorenzini, R. Bardazzi, M. Di Giampietro, F. Boldrini, F. Feresin, D.M.A

..... Aminou, P. Willemsen 14

Space Detection System based on Ultrafast Optics

.....Liu Xun, Yang Bingxin, Ruan Ningjuan, Li Wei 28

Low-light-level Imaging Technology Based on Black Silicon CMOS Detector

..... Wei LI, Huaiyi WANG, Xun LIU, Ningjuan RUAN 32

SATTELITE IMAGERY FOR MARINE OBJECTS IDENTIFICATION UNDER HIGHLY PERTURBED SEA

.....V.P. Kulagin V.Ya. Tsvetkov M.P. Lapchinskaya 40

Information-analytical system of risk assessment and prevention of asteroid and comet hazard.....46

Kulagin V.P., Shustov B.M., Kuznetsov Yu.M., Kaperko A.F., Bober S.A., Obolyaeva N.M.,

.....Naroenkov S.A., Shuvalov V.V., Svettsov V.V., Popova O.P., Glazachev D.O 46

ADLAND – an Advanced Land Observation InfraRed Mission

..... Oertel, D., Hartmann M., Huang-Wolff, Y., Schierle, C., Scheiding, M., Ruecker, G. 55

The Technical Feature and Application of CBERS04-The International Cooperation Satellite

.....He Wei, Long Xiaoxiang, Zhang Minglei, Jia Xu 66

Multikernel Machine-based Hyperspectral Remote Imagery Processing System on Satellite Platform

.....XIE Xiaodan, JIA Jingcheng, JIA Yusheng 72

Space Optical Remote Sensing Application on Lands, Ocean and Atmosphere Observation

.....RUAN Ning-juan ZHANG Wen-yu 80

Some Thoughts on International Cooperation in Using GNSS Below the Sea Surface Level

.....Yean Joo CHONG 83

GNSS Based Time Transfer

..... Liu Yan Ying 86

S/X Dual-band Dual-polarized Microstrip Antenna Array for Satellite Applications

.....Binyun Yan, Zhen Hu, Weixing Sheng, Yubing Han, Yongpeng Kang, Yutao Yang 95

Spaceborne Ka/G Dual-frequency Cloud Profile Radar Discussion

..... Ping Wang, Biao Wang, Xia Ding, Yongjin Han, Haitao Wang 103

Study on RF MEMS Switch Matrix in Multifunctional Phased Array Antennas

.....Weiwei Zhou, Weixing Sheng, Zhen Hu, Xiaofeng Ma, Shihong Zhou 111

A High Performance Broadband Quasi-Yagi Antenna on LTCC Integration Technology

.....Yiyuan Zheng, Weixing Sheng, Renli Zhang, Zhen Hu, Yutao Yang, Yongpeng Kang 117

Research on Spacecraft in Orbit Early Warning Subsystem

..... YANG Haifeng1 Sun Yuan2 Sun Shoufu2 Cui Weiguang1 123

Study on high sensitive detection method of near-space hypersonic aircraft for space-based infrared detection system

.....Jia LI, He HUANG, Shaojuan LI, Shuang LIANG, Jingnan MA 136

A Novel Design for Satellite Communications Signal Spectrum Auto-Monitor Method

.....Zhang Fen ZuoYanqing Shen Ling 142

An Efficient Iterative Least-Squares Pattern Synthesis Method for Satellite Communication

Feng Danping, Ma Xiaofeng, Wu Yuqing, Sheng Weixing, Xiao Zhengming,

..... Shen Peng, Zhong Ruoqing 147

MINIMIZING LATENCY BY USING EXISTING GLOBAL GROUND STATION NETWORKS

.....Kenneth Olafsson 154

Research of Range Migration and Correction Algorithms in SAR Range-Doppler Imaging

..... Min Chen, Weixing Sheng, Zhen Hu, Xiaofeng Ma, Shihong Zhou 159

Experimental research on CPT atomic clock disciplined by GPS

..... Zhang Zhenwei, Yang Renfu, Zhang Xu, Wang Nuanrang, Zhao Huan, Chen Xing 168

Design of variable fractional delay filter using Discrete-cosine-transformation

..... Li Zeqing ,Li Bo,Liu Yanli,Yuan Jie 172

Research on Precise Remote Time Synchronization Based on Communication Satellite

WANG Haifeng, ZHANG Shengkang, YANG Renfu,WANG Xueyun,

.....WANG Hongbo, YI Hang, SHI Fan, SHANG Huaiying 178

Application of User Terminal in the Field of Space-based Internet

.....Li Haihao, Huang Yin , Chu Jianxiang, and Liu Xiaokuan 184

Application of ISLs(Inter-Satellite-Links) in the Field of Space-based Internet

.....Li Haihao, Chu Jianxiang, Huang Yin and Liu Xiaokuan 190

The Development of Forestry Terminal System Based on Beidou Multimode Navigation

..... Jinsheng Ni, Jinlin Yang 197

Research and application of intelligent processing technology of remote sensing data

.....Jinsheng Ni, Xue Yang, Ting Liu, Wenting Ma 203

Error Analysis and Experiment Validation of Height Retrieval for the Interferometric Imaging Radar Altimeter

.....Li Yin-wei, Lu Hu-lin, Wei Wei-wei, Wang Hai-tao 209

An Experimental Study on Airborne Polarimetric SAR Data Calibration

..... Peng Wang, Lei Yang, Zhi Tan 217

The Research on the Implementation and Transformation of GPS Technology Patents

.....Li Ping, Zhang Yue, Zhao Pengfei, Zhou Ziyan 226

Design of flywheel configurations for fast maneuver on single-axis of spacecraft

.....Ming Qi, Yi Yang, Dengyun Wu, Shaowei Zhang, Pengbo Zhang 237

An Improved Linearization Calibration Method of Space-borne Three-axial Magnetometer Based on Mathematical Platform

..... WANG Jiabo, CHEN Xi 246

High-precision Control of Mechanical Flywheel Based on RBF-PID Algorithm with Disturbance Compensator

.....Zhang Cong, Ma Jun, Han Shi-yu, Wei Hui, Tang Xu-dong, Zhong Qi-liang 253

SPACESTAR STAR TRACKER: Designed to be The Choice of Future Recurrent Platforms

.....F. Boldrini, G. Berrighi, D. Procopio 260

Sun Sensor on-a-chip: progress and final product capabilities

.....F. Boldrini, S. Brogi, D. Procopio 277

Studies on Fast Attitude Maneuver for Spacecraft

..... Li Yingbo, Sun Jingru, Zhang Zilong 289

QBITO DEVELOPMENT AND LESSONS LEARNED

..... Ignacio Barrios Tascón, Ana Laverón Simavilla , Efrén Moreno Benavides 306

KM5B – A LARGE SPACE ENVIRONMENT TEST FACILITY OF CHINA

..... Zhang Shiyi Chen Li Qi Xiaojun Peng Guangdong Kang Shuai Li Hongyu 318

Design and Manufacturing of Spacecraft Shell Structures Based on Automated Fiber Placement

.....Ke Wang, Zhihui Li, Li Chen, Tao Song 323

Design of Unification Dataflow on Spacecraft Based on CCSDS AOS

.....WANG Li-sheng, GAI Jian-ning, Chen Jian-yue，Tu Zhi-jun，HAN Xiao, Lai An-xue 329

Estimation of Star Sensor Alignments Using Landmarks

.....Yue Gu Qinghua Zhu Chensheng Cai WeiQing Wu 335

Research on Scheduling Method of Dynamic Resources with Satellites Parallel Test

.....LIU Yifan, LIN Shumin, WANG Huamao, YAN Jindong 344

An Algorithm for Detecting Projection Ellipse of Docking Ring Component Based on Arc Segment Combination

..... Liu Fucheng, TIAN Shaoxiong, LU Shan 352

New Concept Study on Electromagnetic Force Deorbit of Waste Satellite with Active Power-on System

..... Yue Sun, Shan Lu, Hailei Wu, Chaozhen Liu, Chenglong Jia 357

Binocular Vision Pose Measurement of Spacecraft in Short Range Based on Docking Ring

..... CAO Shu-qing, LU Shan, LIU Zong-ming, ZHANG Yu, YANG Guang 364

Development of an Equivalent Equipment for Satellite Power Supply and Distribution System based on LabVIEW

.....Xiaolong Liu, Yixin Geng, Shuhai Chen 376

Parameter Influence Analysis for Doppler Velocity-Measurement Target Source on Asymmetric Spatial Heterodyne Technology

.....Zhang Wei, Cheng Xin, Du Yang 384

Error Analysis of the Autonomous Celestial Navigation Based on the Spectrum Velocity Measurement

.....Huang Qinglong, Chen Xiao, You Wei 392

All-electric Propulsion Satellite Leading Revolution of Telecommunication Satellite Development

..... HU Zhao, ZHOU ZhiCheng, WANG Min 404

Orbit Control Manoeuvre Module for Thaichote Satellite

.....Manop Aorpimai, Pornthep Navakitkanok, and Veerachai Malayavej 410

Satellite Plate Structure Selection and Design Analysis

..... Wan Shuang, Rong Hua, Liu Yebao 419

Numerical Modeling for the Dynamic Tracking of Dynamic-hydrostatic Hybrid Gas Seals for liquid oxygen pump

.....ZHANG Shu-qiang, WANG Liang, CHEN Jie, ZHAO Wei-gang 429

Research on Dynamic Simulation of an Aerospace pyrotechnic device

.....YE Yao-Kun, LI Xiao-gang, Ding Feng1, Man Jian-feng 444

Application of GPS in TT&C Time Synchronization

..... ZHANG Yanping, SHANG Jianzhong, BAI Yungao, LI Haichao 453

Thermal Expansion Coefficients Calculate and Analysis for Unidirectional Fiber-reinforced Composite of Space Truss Structure

..... Jian Ma 459

Experimental Investigation for Single-phase Heat Transfer of Confined Jet Array Impingement on a Round Heated Surface in a Narrow Space

.....Zhongxuan Du, Xiang Zhang, Zaiteng Zhai, Han Chen, Shichen Jiang 467

ELECTRIC PROPULSION APPLICATION IN CHINESE TELECOMMUNICATION SATELLITES

.....Zhou Zhicheng, Wang Min，Zhong Xiaoqing 476

Research in the Conductor Brazing Technology for the Hall Thruster in High-Temperature Working Environment

.....Yang Jian, Chen Bo 483

Development of Non-Explosively Actuators

.....Liu Shiyi, Chen Jin 503

Integrating Design of Lightweight Flat Cable

WANG Zhibin, CHEN Mengjiong, HUANG sanbo, MA Jusha, CAO Jiaye,

.....LEI Gang，LU Jianfeng 512

Structural Analysis of the Interconnector used in Flexible Solar Array

WANG Zhibin, CHEN Mengjiong, HUANG sanbo, MA Jusha,

.....CAO Jiaye, CHI Weiying, LEI Gang 516

Study on Design Method for Distribution Systems Based on Maintainability

.....Du Zhanchao; Chen Dan 519

A Study on the Thermal Performance of the Motorized Thermal Shade System

..... Dong La, Han Chen, Zaiteng Zhai, JiangFeng Lu 526

Surface modification of RDX using conducting polymer: Preparation, properties and application

.....Zhiquan Shi, Yali Lu, Hongyong Huang, Rui Li 532

A TOOL FOR SIMULATION AND ANALYSIS OF SPACE PROPULSION SYSTEMES TRANSIENT PERFORMANCE

.....Chen Hongyu, Liu Hongjun, Hong Liu 542

Design and Fabrication of High Transparent and Conductive Optical Thin Films with Atomic Oxygen Resistant

.....FAN Xiang, CHI Weiying, LEI Gang, YANG Hongdong, SU Baofa 551

Thermal Analysis and Thermal Design Research of Space Power Control Unit Based on The Finite Element Method

.....Jiangyue, Tubo 554

A Bandwidth Efficient Phase Demodulator for On-Board Satellite Receivers

..... Bilal Hassan, Humayun Zahid, Zia ul Haq, Atif Noushad 562

Design of Annular Ring Patch Antenna for Circular Polarization Operation Using Dual Dielectric Substrate Layers

.....Zeenat Rajar, Saima Siddiqui 574

Design of Multi-Bus Communication Framework Using Multi-level 1553B Bus and RS422 Multi-Bus for Digital Systems in Satellites

.....Yixin Geng, Xiaogang Hu, Ye Rui 580

Fuel Cell Power System with High Reliability for Space Applications

.....Xiangan You and Chengan Wan 588

Research on Discharge Plasma Characteristics of An Ablative Pulsed Plasma Thruster via Emission Spectrometry 597

Lei Yang, Yuping Huang, Xiangyang Liu, Haibin Tang,

.....Chunming Wang, Zaiping Zheng 597

Analysis on the Out-of-Plane Mechanical Properties of Honeycomb Sandwich Structures and Experimental Validation

..... Xiaoyu Liu, Shuang Liu, Chaoqi Xu, Lei Wang 606

Geomagnetism Sensing & Calibration Systems of LING QIAO Satellite

..... CHEN Xi WANG Jiabo WANG Menglu 612

A Study of Intelligent Power Supply for CubeSat

.....Yu-Ying Yao, Wei Gong, Long-fei Li, Hui-sheng Liang, Yuan-Mo Liu 620

STEP M: Space to Effectively Prepare for Migration

.....Xiaofeng Zhao, Xia Wang, Shihai Li 627

Data Transmission Scheme for Micro-satellite Formation and Network

..... Li Peng, Zhong Lingling 641

The Opportunities and Challenges of China's Commercial Communication satellite in global landscape

..... Liu Yue 648

Trajectory Reconstruction Based on GPS Measure Data for Rocket Projectiles

..... BIAN Wei-wei, Ge Xiao-fei, XIN Zhen-fang, ZHAO Jin-xin 653

Digraph-model-based Dynamic Command Scheduling: An Emerging Paradigm for Smart Satellite Operation

.....Tian Zhi-xin, Cui Xiao-ting, Li Zhen, Qiao Yi-shi, Li Xiao-juan 661

Innovation Potential by Small Satellites for Spacecraft Design

..... Klaus Schilling 671

Small Satellite Formations for Earth Observation and Telecommunication

..... Klaus Schilling 678

Trajectory/General Parameters Optimization for Suborbital Launch Vehicle

..... Sun Shijie, Zhang Yechen, Min Changwan, Xu Chunling, Wang Guanyu 683

Innovational Design and In-orbit Test of Small SAR Satellite Electrical Power System

.....JIA Xiaodong, YAN Wanjuan, WANG Wentao, ZHANG Running 691

Study on Application of Wireless Ad Hoc Network for Moonlet

.....WANG Li-sheng, WANG Di, ZHU Bo, TU Zhi-jun, Lu Yi-bin, Lai An-xue 699

Extract Feature Frequency of Rocket Vibration Signal Based on EEMD and Correlate Coefficient Test

.....Zhang Jianhai, Han Yingchun , Liu Jun 706

Development and Testing of Nitrous Oxide/Propane Rocket Engine

.....Wang Dong1, Liang Guozhu2 714

Burnout Point Sensitivity Analysis for General Parameters of Nearspace Launch Vehicle

.....Zhang Yechen, Sun Shijie, Min Changwan, Chen Feng, Jiang Huan 722

Application of Model-based Systems Engineering in Small Satellite Conceptual Design – A SysML Approach

.....Muhammad Waseem, Muhammad Usman Sadiq, and Muhammad Frooq 729

Small Satellites and Non-State Actors: A New Challenge To State Sovereignty

.....George Anthony Long,. Managing Member, Legal Parallax 740

Hyperspectral satellite application in Asia-pacific regional space cooperation

..... Shendan, Gao Yufeng 747

Energy Management and Design of Vehicle Trajectory based on Optimized Section

..... Yang Ling-xiao, Min Chang-wan, Xiao Zhen, Zhang Ning-ning 753

Simulation Analysis and Experimental Study joint payload connection

.....CUI Shenshan ZHANG Tao CUI Zhandong WU Di XU Jian 764

Mass Optimization Design of Separation Devices for Pint-size Load Based on Multi-Island Genetic Algorithm

.....Zhao Shanshan, Zhang Wen, Zhong Jiehua, Li Jing 772

Tether Deployment Control Scheme Analysis and Design for Electrodynamic Tethered Upper-Stage Deorbit System

..... Zhang Feng, Wang Shu-ting, Wang Xiao-ding, Wu Sheng-bao, Shen Lin 773

Development and Discussion of the Air Launching Technology

.....Du Linfei, Wu Jihua 781

Experimental Research of Regression Rate of N_2O and HTPB Hybrid Rocket Motor

.....GUO Fang, LUO Zhan-fei, Zhao Hui-jie, Yang Fan, Zhang Ming-liang 786

STU-2: A CubeSat Mission for Earth Observation and Marine/Air Traffic Monitoring

Shufan Wu, Wen Chen, Caixia Chao, Chuanxin Zhang, Zhongcheng Mu,

..... Tao He, Zhiqiang Jia, Ziyi Wu, Ya Dai, Wenhe Liao, Xiang Zhang 797

Epsilon Launch Vehicle for Small Satellites

.....Yasuhiro MORITA 812

Researches on Multi-component Calibration-Free Trace Gas sensor in Manned Space Vehicle

Jiang Meng, Wang Xuefeng, Feng Qiaoling, Liang Hu, Member, IEEE

Abstract: for astronaut health care, trace ammonia and carbon dioxide concentration should be on-line detected. We presented a system with high sensitivity and stability based on quartz enhanced photoacoustic spectroscopy with hollow-core photonic bandgap fiber (HC-PBF) as reference cell for calibration free and laser wavelength locking. A 5m long HC-PBF splicing with single mode fibers and multimode fiber by optimizing fusion parameters is manufactured as the reference cell which filled with gas mixture. The total loss of the reference cell is less than 6 dB. The experimental results indicated that absorption spectra of reference cell maintained stable after 150 days at atmosphere. A normalized noise equivalent absorption coefficient (NNEA) of 1.18×10^{-8} $cm^{-1}W/\sqrt{Hz}$ was obtained with a DFB laser at 2003.5nm. This yielded a minimum detection limit of 0.36 ppm within a 1s lock in time constant.

Index Terms: gas sensor, hollow core photonic bandgap fiber (HC-PBF), ammonia, carbon dioxide, quartz enhanced photoacoustic spectroscopy.

I. INTRODUCTION

The exhalation trace gases measurement provides an efficient method for astronaut health care. Otherwise, this system also has more potential applications on early fire smolder gases monitoring and warning in manned spacecraft. At present, the ability to detect small amounts of trace gas plays an important role in many real world applications. Much scientific research is focused on environmental issues including air pollution, car exhausts, and noninvasive exhalation analysis in medicine. The medical significance of the presence of ammonia and carbon dioxide naturally occurring component in exhalations have been studied previously, demonstrating the usefulness of sensors to diagnose and monitor a variety of medical conditions, including chronic kidney disease (CKD), helicobacter pylori infection, and encephalopathy. Healthy individuals typically have a few hundred parts per billion by volume (ppbv) of ammonia in their exhalations, while patients diagnosed with CKD, for example, could have over one part per million by volume (ppmv) [1,4].

Photoacoustic spectroscopy is a well-established method used in trace gas sensing applications due to its high sensitivity. As one kind of photoacoustic spectroscopy technology, QEPAS gas sensors have advantages of high sensitivity, high selectivity and rapid responses. Tuning forks and micro resonance tubes are used instead of photoacoustic cells and microphones to realize quartz-enhanced photoacoustic spectroscopy detection. In particular, micro resonance tubes are used for signal enhancement. The total size of the system can be produced sufficiently small to be used for portable in-situ measurements.

In 2007, A.A Kosterev [2, 3] from Rice University, USA demonstrated a QEPAS system employing a 2μm DFB diode laser with power equal to 6.2mW. The minimum detection limits of carbon dioxide and ammonia are 18ppm and 3ppm, respectively. The influence of water vapor on the relaxation rate enhancement in human exhalation gas detection is discussed with a simple theoretical model that can be experimentally confirmed. In 2008, they also demonstrated a QEPAS

system for early fire detection with hydrogen cyanide and carbon monoxide in aerospace aircraft which supported by National Aeronautics and Space Administration via awards from the Jet Propulsion Laboratory, Pasadena, CA and Johnson Space Center, Houston, TX.

Previous works included the demonstration of gas detection with distributed feedback quantum cascade laser in middle infrared to improve the sensitivity, and performance improvement with different operating wavelength [4]. However the reference cell as an important component of the stable system also should be taken into account. As a gas sensor, the time taken for gas into the hollow core photonic bandgap fiber is too long for practice application. However as a reference cell, hollow core photonic bandgap fiber shows low transmission loss, long absorption light path compared with traditional self- focusing system.

In this letter, we report on a 5m HC-PBF used as a reference cell for wavelength locking to improve the precision and stability. The manufacture process includes fusion splicing and gas filling in normal pressure is described in details. Due to the wide self-broadening coefficient of ammonia, high pressure filling will lead to high modulation depth and blended spectra. By confirming stability of the reference cell, the transmission spectra are measured compared before and after 150 days. The QEPAS system with HC-PBF reference cell for ammonia detection results are presented here with minimum detection limit and normalized noise equivalent absorption coefficient. This work was supported by the National Natural Science Foundation of China “Research on calibration-free measurements of trace gas concentration in manned space vehicle” (Nos. 61203204) from 2012.

II. Manufacture of Reference Cell

The HC-PBF with cross section as shown in Fig.1 (a) was connected between a single mode fiber and a multimode fiber. The HC-PBF was prepared as the reference cell into the gas detection system. The diameter of the center hollow core is 15μm. A triangular array of air holes in a silica background is surrounded by an average spacing between the air holes of 4.8μm. The diameter of the cladding is 155μm. The splicing processing is operated using a conventional fusion splicer, and parameter settings for the PBF/SMF (including fusion current, fusion time and overlap) have been mentioned in reference [5].

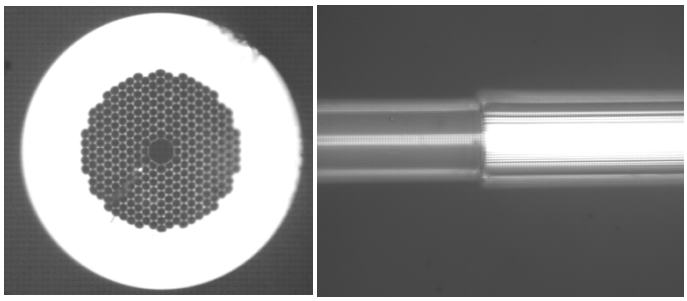


Fig. 1. Microscopy images of (a) Cross section of HC-PBF (b) Splicing region with SMF.

The optical microscope image of the fusion joint of the HC-2000-02/SMF is shown in Fig.1 (b). Although the core diameter of PBF was larger than SMF, but the MFD was 11μm, which was well matched to the SMF-28’s MFD 10.4μm. The end-coupling loss between the PBF and SMF in an optimal alignment can be as low as ~0.01dB [5]. The attenuation at 2μm is 20dB/km, so the attenuation of a 5-meter cable is 0.1dB. A 5m PBF was spliced to a conventional single mode fiber at the input, a multi-mode fiber at the output, and the total splice loss was experimentally measured to be 6dB, using optimized fusion parameters. Figure 2 demonstrated the transmission spectra

window of the hollow core photonic bandgap fiber, the transmission window is 1930nm to 2125nm. The spectra of HC-PBF filled with gas mixture are certified to be coincident with spectra simulated by the HITRAN data.

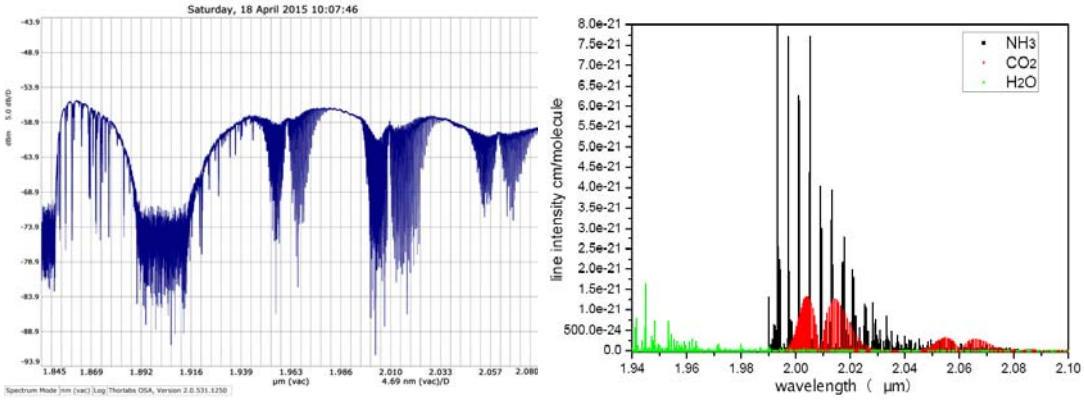


Fig.2. the spectra of the gas mixture filling in HC-PBF and simulated by HITRAN database.

The gas loading procedure required the fiber to be purged with nitrogen to remove residual gases in the PBG fiber, as well as to obtain a baseline. The PBG fiber, which uses splicing with a single mode fiber at one end, was fixed in gas housing. The housing was machined and tested to have a high quality vacuum. Each end has a switch to control airflow. One end of the housing was connected with a vacuum pump, and the other end with a pressure valve and gas cylinder. According to micro tube flow theory [6], when a hollow core is spliced with a single mode fiber, the volume of gas absorption cavity is cut down. Thus, the filling time of the gas can be reduced correspondingly. When the pressure of the housing is equal to the gas cylinder at 3×10^5Pa， and held for 1 hour, the gas in the PBG fiber can reach pressure equilibrium. Thus, the PGB fiber was removed, and spliced quickly with a multi-mode fiber in an ambient atmospheric environment. Any leakage from the un-spliced end can be ignored.

Figure 2 shows the spectrum of the gas absorption feature peaks which was filled in the reference cell. A super continuum light source is used as the input light. The transmitted spectrum is measured by optical spectra analyzer with a resolution of 0.02nm. Fig. 3 (a) shows the spectrum of independent gases, and indicates that the absorption line strength of ammonia is about 6 times stronger than the available carbon dioxide line at atmosphere pressure. Fig 3 (b) shows the spectrum for the mixture of gases in HC-PCF fiber. The gas mixture consisted of ammonia, carbon dioxide and nitrogen. After being filled with HC-PBF under high pressure, the line strength of carbon dioxide became much stronger to similar levels with ammonia, due to having a much faster vibrational relaxation rate under high pressure [2]. In this case, two overlapping absorption spectra of ammonia can be distinguished spectroscopically.

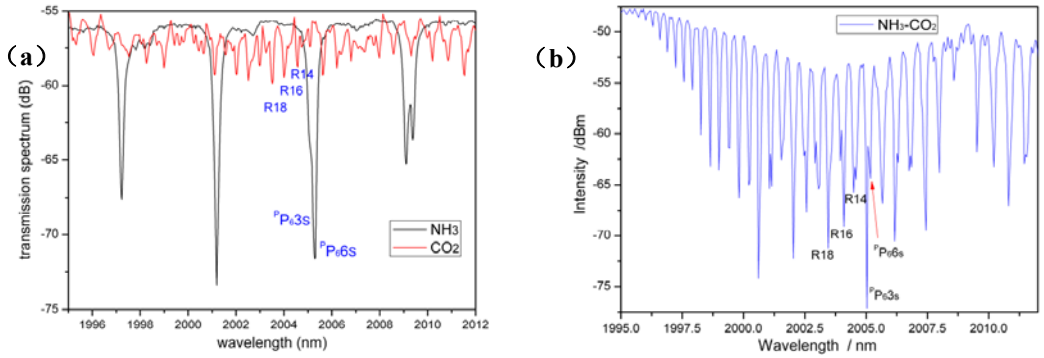


Fig. 3. The transmission spectra of carbon dioxide and ammonia gas filled in HC-PBF: (a) separated gases, (b) gas mixture.

Because of wide self-broadening coefficient of ammonia feature lines, the modulation depth will be larger if the filled gas under high pressure. Such spectra are found to be blended and overlapped by collisional broadening. As a reference cell, HC-PBF was filled with gases under normal pressure that had fast equilibrium attainment times, high resolution and stable calibration capability.

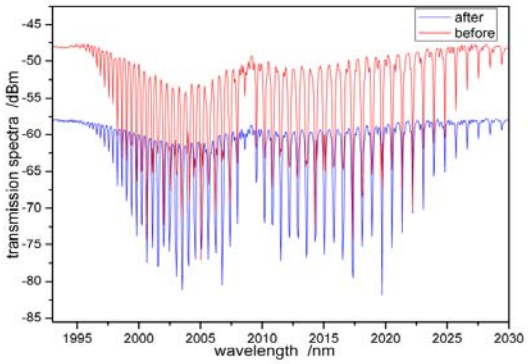


Fig.4. The transmission spectra of ammonia and carbon dioxide gas filled in HC-PBF before and after 150 days.

III. Experiment system

The gas detection experimental setup is shown in Fig. 5. A distributed-feedback (DFB) laser with a center wavelength of 2003.5 nm and narrow linewidth (<2 MHz) was used, and the output light is 3mW. The DFB laser is initially amplified by a Tm-doped amplifier module. The transitions ($^{3}F_{4}$->$^{3}H_{6}$) of trivalent thulium provide radiation near 2μm. Thulium doped fiber is a promising medium for building high power fiber laser sources and amplifiers [7]. The output power after amplification is 30mW, which can be tested with a power meter.

The output light was split into two using a 90/10 coupler. Light at the throughput port was used for absorption measurements of ammonia and carbon dioxide gases, while the other gas sample inside the HC-PBF was used for wavelength stabilization.

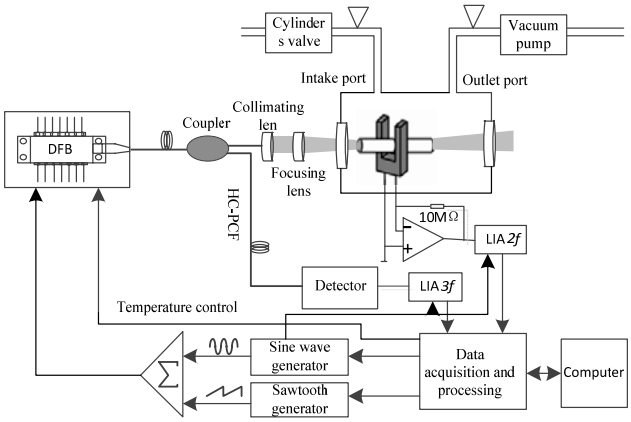


Fig. 5. Experiment setup for QEPAS detection

For wavelength stabilization, the DFB laser was modulated by applying a sinusoidal signal at 1/2 of the resonant frequency of the tuning fork through the laser driver. Light passed through a reference cell made of HC-PBF, and subsequently was detected using a photo detector. Signal from the PD was fed back to the laser driver after being processed by a lock-in amplifier. The first-harmonic signal detected by the lock-in amplifier was fed back as an error signal to the current source of the laser through the proportional-integrating (PI) controller circuit. The other path of light that passed through the gas absorption measurement was detected by a quartz-enhanced fork. The laser beam was collimated by a pair of fiber pigtailed collimator lenses. The AR coating was from 1.8μm to 2.4μm , and the focal length was 5.9mm. Subsequently the laser beam was focused between TF prongs using a BK7 lens. An acoustic micro-resonator can enhance the QEPAS signal, and hence increase the detection sensitivity. Therefore, stainless tubes with a length of 4.4 mm each and an inner diameter of 0.6 mm were mounted on both sides of the QTF to yield the highest signal-to noise ratio, which is reported to be higher than that of a bare QTF [8]. The signal was detected using a transimpedance amplifier to tune the piezoelectric current of the quartz fork to the voltage signal. The signal from the fork was passed through another lock in the amplifier and was delivered to a data acquisition card (Adlink PCI9826), and then connected to a personal computer.

IV. Results

The photoacoustic signal received by the detector [9] is：

$$S = KI\alpha Q\varepsilon \tag{1}$$

Here, K is the system constant correlated with the detection cell; I is the output power of the laser source; α is the absorption coefficient, which is proportional to the absorption line strength, absorption line shape and molecular density per unit volume; Q is the qualify factor of the quartz fork; and ε is the photo-acoustic transform efficiency.

To improve the detection sensitivity, two methods were used. The key factors included the absorption line intensity and the output power of the laser source. A Tm-doped amplifier module was used to improve the laser output power. The output power of DFB lasers can be improved from 3mW to 30mW without line-width extension. However, the normalized noise equivalent absorption coefficient (NNEA) at high light power is lower than that in low light power. The signal amplitude exhibits the opposite trend. Thus, the amplified times of the light power cannot be simply maximized, but instead should be optimized to a suitable value by tuning the pumped current of the amplifier module. Secondly, different models of quartz tuning forks were tested for higher quality

factors. Thirdly, the proper lines were chosen for carbon dioxide and ammonia detection with higher absorption line strength and less interference from water vapor. According to HITRAN database, within the tuning range of 2μm DFB lasers, the proper lines for CO_2 detection in this spectral region are R14, R16 and R18. In order to measure the CO_2 and NH_3 concentrations with one DFB laser with different modulation parameter simultaneously, the R14 line of carbon dioxide at 4988.654 cm^{-1} was selected (which has a line intensity of 1.323×10^{-21}cm/mole), as well as the transition line of ammonia at 4987.53cm^{-1} (which has a line intensity of 1.908×10^{-21}cm/mole). The laser radiation was on resonance with the CO_2 absorption line at a laser diode current of 76.2 mA when its temperature was set to 23.49℃, and NH^3 absorption line at the laser diode current of 48.66 mA when its temperature was set to 38.73℃. Thus, using a commutator that switches signals from the microcontroller and allows the laser parameters to cycle over two channels, improvements were observed in the output power I, the qualify factor Q, the absorption line intensity, and the sensitivity of on-line ammonia and carbon dioxide detection. Figure 6 is the second harmonic curve obtained from the spectraphone cell of the quartz tuning fork.

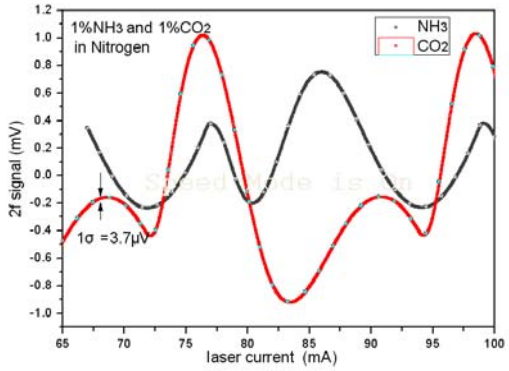


Fig.6. Harmonic curve of the ammonia and carbon dioxide

For the sensitivity evaluation, a photoacoustic 2f signal of 1% ammonia and carbon dioxide in nitrogen was acquired using 1s locking in time constant. A single point noise was calculated as the standard deviation of the measured data points
in the wing of the targeted absorption line, which yields a value of 3.7μV. This yields a minimum detection limit of 0.36 ppm for noise equivalent ammonia concentration at 2003.5nm. By normalization to the corresponding bandwidth of 0.75Hz and optical power of 30mW, a normalized noise equivalent absorption coefficient (NNEA= 1.18×10^{-8} cm^{-1} W/√Hz for the QEPAS system.

V. Conclusion

To improve the reliability and stability of carbon dioxide and ammonia detection, a reference cell is used to realize the lock in of the laser wavelength when the system is calibrated online. The reference cell manufactured from a hollow core photonic bandgap fiber with low loss connection and fast filling times was used to reduce the system volume. After the splicing section of the fiber that has been filled with a toxic gas and coated with a heat-resistant material, it will be a good candidate for portable and compact QEPAS systems.

We have demonstrated an optical fiber reference cell filled with a mixture of carbon dioxide and ammonia. A 5-m HC-PBG fiber splicing with single mode fiber with loss of 6dB at 2004nm was constructed. The small size and high sensitivity of the HCPBF-based QEPAS gas sensors would

find applications in miniaturization, robustness, accuracy and high sensitivity for biomedical, environmental, and aircraft applications.

REFERENCES

[1] M. E. Webber, R. Claps, F. V. Englich, and F. K. Tittel, "Measurements of NH_3 and CO_2 with distributed feedback diode laser near 2.0 μm in bioreactor vent gases". Applied Optics 40, 24 (2001).

[2] K. Owen, A. Farooq, "A calibration-free ammonia breath sensor using a quantum cascade laser with WMS 2f/1f", Appl. Phys. B 110, 5701 (2013).

[3] P. Patimisco, G. Scamarcio, F. K. Tittel, and V. Spagnolo, "Quartz-Enhanced Photoacoustic Spectroscopy: A Review", Sensors, 14, 6165 (2014).

[4] M. D. Wojcik, M. C. Phillips, B. D. Cannon "Gas-phase photoacoustic sensor at 8.41 μm using quartz tuning forks and amplitude-modulated quantum cascade lasers", Applied Physics B, 85, 307 (2006).

[5] S. M. Cristescu1, S. T. Persijn, S. T. Hekkert "Laser-based systems for trace gas detection in life sciences", Applied Physics B, 92, 343 (2008).

[6] S. Borri, P. Patimisco, A. Sampaolo, "Terahertz quartz enhanced photo-acoustic sensor", Appl. Phys. Lett., 103, 1 (2013).

[7] P. Patimisco, S. Borri, A. Sampaolo. "Quartz enhanced photo-acoustic gas sensor based on a custom tuning fork and a terahertz quantum cascade laser", Analyst, 139, 2079 (2014).

[8] G. Wysocki, A. A. Kosterev, and F. K. Tittel, "Influence of molecular relaxation dynamics on quartz-enhanced photoacoustic detection of CO2 at $\lambda = 2$ μm", Appl. Phys. B 85, 301 (2006).

[9] R. Levicki, G. Wysocki, A. A. Kosterev, and F. T. Tittel, "Carbon dioxide and ammonia detection using 2 μm diode laser based quartz-enhanced photoacoustic spectroscopy" Appl. Phys. B 87, 157 (2007).

Application and Development of Enhanced Satellite-based Navigation Systems

Lianyin Deng and Shumin Gao

Qian Xuesen Laboratory of Space Technology
Beijing Institute of Space Mechanics and Electricity

Abstract: Enhanced satellite-based navigation system,which has been recognized as one of the most important system in GNSS field ,is an efficient and reliable system to enhance the quality of service of satellite navigation system. Foreign enhanced satellite-based navigation system is introduced in this paper.The applications of enhanced satellite-based navigation system in aerospace, marine geophysical exploration and the railway industry are introduced.The necessity of building BeiDou enhanced satellite-based navigation system is stated.
Keywords: SBAS systems;Industry Application;Necessity

1 Introduction

Satellite-based augmentation system broadcasts ephemeris error, satellite clock error and ionospheric delay through enhanced satellite navigation signal repeaters of geostationary orbit (GEO) satellites.Developed countries compete for improving the positioning accuracy of existing satellite navigation systems .

A number of satellite-based augmentation systems has been established , such as the US Wide Area Augmentation System (WAAS), Russia's differential correction and monitoring system (SDCM), European Geostationary Navigation Overlay Service (EGNOS), Japan's multi-functional star based Augmentation System (MSAS) and India geostationary orbit enhanced GPS assisted navigation system (GAGAN). Significant economic and social benefits have been produced by these systems.

2 Foreign satellite-based augmentation systems

Depending on the data processing center processing method and use of observational data, GNSS augmentation system to enhance the system and into the wide area differential global differential enhancement system categories.

Considering the differences of observation data and data processing method, GNSS augmentation system is divided into Wide Area Augmentation System and the Global Differential Differential Augmentation System. Now American StarFire, OmniStar and British VERIPOS have been finished and produced significant economic and social benefits.US Wide Area Augmentation System (WAAS), Russia's differential correction and monitoring system (SDCM), European Geostationary Navigation Overlay Service (EGNOS), Japan's multi-functional star based Augmentation System (MSAS) and India geostationary orbit enhanced GPS assisted navigation system (GAGAN) are under construction.

2.1 Wide Area Augmentation System

（1）US Wide Area Augmentation System (WAAS)

Wide Area Augmentation System constructed by the US Federal Aviation Administration

(FAA) is mainly for aerospace applications.

July 10th, 2003, WAAS system has been widely used in aircraft, covering 95 percent of the continental United States and parts of Alaska. WAAS system service performance are as follows:

- ✧ Positioning accuracy ： 3m；
- ✧ Integrity missed alarm probability： 10^{-7}/h；
- ✧ Integrity alarm time：6.2s；
- ✧ Navigation continuity： 10^{-8}/h。

WAAS system components are shown in Figure 1. The coverage of the system is shown in Figure 2.

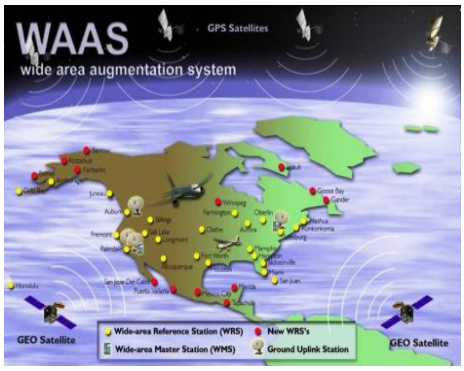


Fig. 1 WAAS components

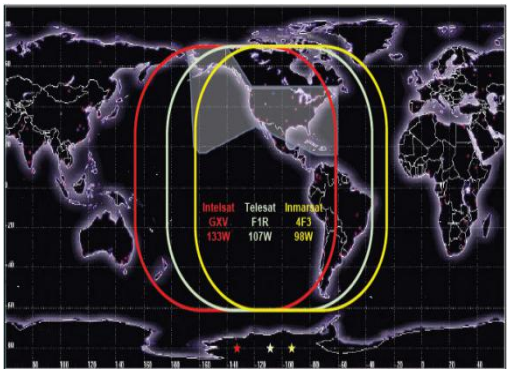


Fig. 2 WAAS coverage

(2) Russia The System for Differential Corrections and Monitoring

Since 2002, the Russian Federation has started to establish GLONASS satellite navigation augmentation systems that is the System differential correction and monitoring (SDCM).

SDCM system consists of three parts: the differential calibration and monitoring stations, central processing facilities and geostationary satellites for relaying differential correction information.

(3) European Geostationary Navigation Overlay Service（EGNOS）

EGNOS meet the high security needs of users by enhancing the positioning accuracy of GPS and GLONASS satellite navigation system.EGNOS is the The first step of European Global Navigation Satellite System plan and a prelude to the development of Europe's Galileo satellite navigation systems program.

EGNOS system service performance are as follows:

- ✧ Positioning accuracy ： 3m（95%）；
- ✧ Integrity missed alarm probability： 10^{-7}/h；
- ✧ Integrity alarm time：6.2s；
- ✧ Navigation continuity： 10^{-8}/h。

EGNOS consists of three parts which are the ground segment, the space segment and the user segment. The system structure is shown in Figure 3, the coverage is shown in Figure 4.

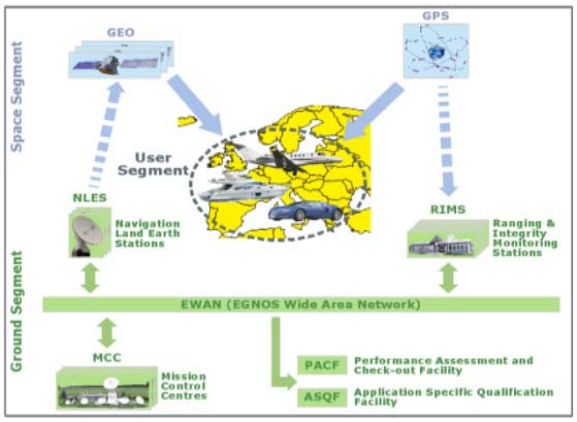


Fig. 3 EGNOS structure

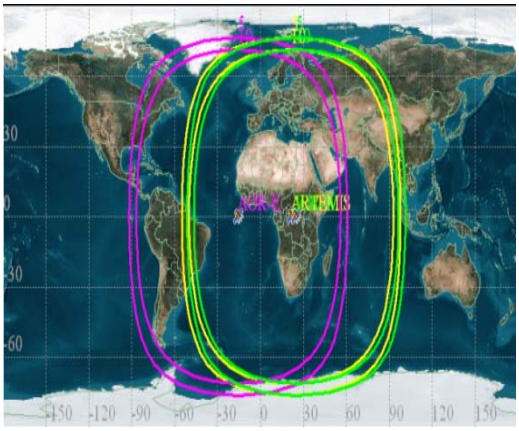


Fig. 4 WAAS coverage

2.2 Global Commercial Satellite Based Augmentation System

With the development of differential enhancement technology, worldwide commercial satellite-based augmentation system such as StarFire, OmniSTAR and VERIPOS has been established . Some systems have achieved decimeter even centimeter real-time positioning accuracy.

(1) StarFire

In 1999, a GPS-based StarFire network which is the world's first differential system with real-time decimeter accuracy launched. The system consists of L-band communications satellites and more than 50 global network of reference stations. In 2011, It's real-time positioning accuracy reached 5cm in global context..

Fig.7 Starfire reference station network and INMARSAT satellite distribution diagram

Starfire system has three differential positioning mode: WCT, RTG, RTG-RTK combination mode.

- ✧ WCT (Wide-Area Correction Transform) Services: Initial service mode of StarFire system, which requires reference stations and subscriber stations use the dual-frequency GPS receiver.It's positioning accuracy (1δ) is 35cm;
- ✧ RTG (Real Time Gipsy) services: global differential positioning service system,which used real-time correction algorithm and JPL GIPSY software, was based on WCT. Differential positioning accuracy of the system is up to 10cm (1δ);
- ✧ RTG-RTK combination services: RTG and RTK combined mode is a new differential positioning mode which launched recently.The mode overcome the disadvantage of long initialization time. Its positioning accuracy reached centimeter level.

（2）OmniSTAR

OmniSTAR includes more than 100 reference stations, nine satellite uplinks and two global network control center.It almost enables seamless coverage (Figure 9).It can provide 24 hours of high-reliability location-based services. OmniStar was acquired Tempo in 2011.

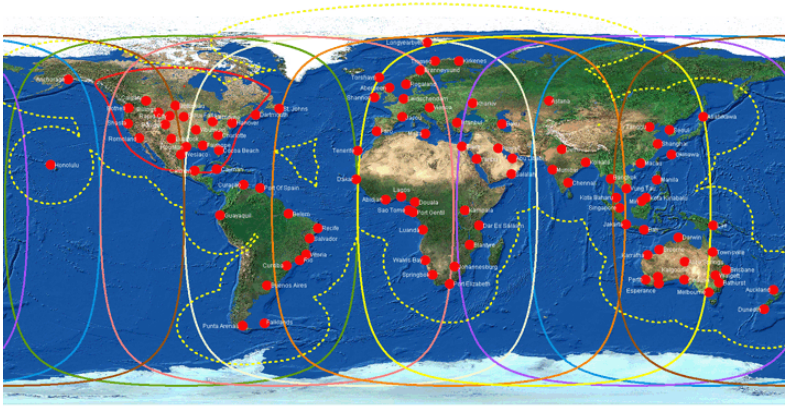


Fig. 9 Omnistar distribution station network

（3）VERIPOS

VERIPOS include 84 stations and seven GEO communications satellite(Fig.11).On 12th March 2014, VERIPOS announced that its global base station receiver would be upgraded to support multi-frequency GPS / GLONASS / Galileo / BeiDou four systems Septentrio receiver and are compatible Beidou navigation signals.

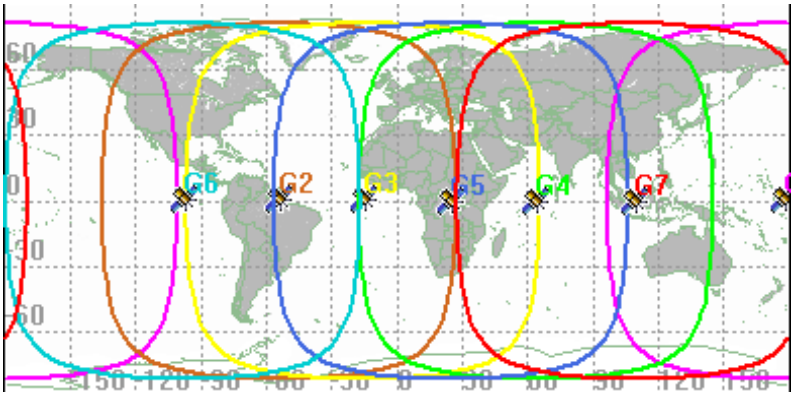


Fig. 11 VERIPOS satellite coverage

The main features of VERIPOS are as follows:

- ✧ data format: Improved RTCM SC104-44 type;
- ✧ transmission rate 1200bit / s;
- ✧ typical data update rate: 30s.

2.3 National Navigation Augmentation System

China carried out satellite navigation augmentation system from the end of the century. At the same time China started the construction of China's GPS augmentation system with a system accuracy of 5 to 10 meters. But it’s application is limited since 2000 because the United States announced the cancellation of SA policy.

China will continue to launch more satellites, improve the service performance of the system and expand coverage areas.By 2020, Beidou global navigation satellite system will consist of more than 30 satellites.It will provide high-precision and high reliable positioning navigation services. With the construction of the Beidou satellite navigation system, Beidou navigation satellite-based augmentation system will also be put on the agenda.

3 Industry requirements of satellite-based augmentation systems

Depending on the type of users , users can be divided into four categories, which are high integrity users, meter-level users, submeter-level users and centimeter-level users.

High integrity users are mainly the areas associated with human life such as civil aviation navigation, railways navigation monitoring and maritime transport services.Meter level users includes intelligent traffic, it has certain requirements in accuracy and integrity.Submeter users includes geological prospecting, exploration of marine resources, precision agriculture and environmental monitoring. Its high precision must be taken into account the integrity requirements.Centimeter-level users includes measuring, surveying, offshore drilling, seismic testing and other applications.It is more concerned about the accuracy requirements and requirements of integrity are lower.

Table1 User needs of Satellite navigation augmentation system

	Industry	Accuracy（95%，m）	Integrity		
			Alarm Limits（m）	Integrity Risk Probability	Alarm time（s）
High Integrity users	Air Navigation	16m(H) 4m (V)	40m(H) 10m(V)	$10E^{-7}$～$10E^{-9}$/h	6
	Railway navigation monitor	0.58	3	$10E^{-6}$～$10E^{-7}$/h	2
	Maritime Transport Services	0.6~10	15	$10E^{-7}$/h	6
Meter level users	Intelligent Traffic Management	1~2	3	$10E^{-5}$/h	10
Submeter-level users	Mineral exploration and investigation	0.1～1	1.5	$10E^{-5}$/h	10
	Geological disaster forecasting and monitoring of environmental monitoring	0.1~1	1.5	$10E^{-5}$/h	10
	Ocean mapping and geophysical services	0.1~1	1.5	$10E^{-5}$/h	10
Centimeter User	Precision mapping (urban line (road) measurement / cadastral / Precision Engineering Surveying/ Engineering actinomycetes (sample) / digital city data collection and map update	0.01~0.05	0.1	$10E^{-5}$/h	10
	Precision agriculture management	0.01~0.10	0.15	$10E^{-5}$/h	10
	Large construction machinery automation	0.01~0.10	0.15	$10E^{-5}$/h	10
	Large infrastructure projects deformation monitoring	0.01~0.10	0.15	$10E^{-5}$/h	10

4 Beidou satellite navigation system

4.1 Beidou satellite navigation system market space is limited

Satellite navigation technology has become an important tool for economic and social development of countries.Currently, the international satellite navigation industry has formed a relatively complete industrial system.

In 2012, the European satellite navigation Bureau issued a "global GNSS Market Report." The report noted that to 2012 global GNSS total revenue reached 100 billion US dollars.From 2010 to

2016, the global GNSS market will grow at a CAGR of 13%,

Industry experts predict that from 2010 to 2020 market size CAGR of GNSS is over 10%. In 2020 will reach $ 220 billion. Its high accuracy applications achieve 1/4 of the overall $ 50 billion market.

A breakthrough has been made in Chinese Beidou satellite system Construction. The Beidou satellite navigation system has played an important role in transportation, marine fisheries and other areas in China.In 2013, domestic market reached 10 billion yuan.The application of Beidou satellite navigation system will be greatly expanded in key areas of the national such as economy, industry and public services.

However, because the system construction is behind industrial development needs, the Beidou satellite navigation system application market space is severely squeezed. In 2013, the proportion of its output of the entire satellite navigation and location-based services in GDP is still less than 10%. Compared with the global share for GPS, Beidou satellite navigation system application market space is severely squeezed.

4.2 Beidou still unable to meet the needs of high-precision applications in the field of national strategic economic security

The position accuracy of China's satellite navigation rose from 10 meters to less than 8 meters and timing accuracy from 50 ns up to 30 ns.

However, in the aviation, rail, marine, mining exploration and other industries need to achieve real-time decimeter, centimeter-level satellite navigation and positioning accuracy.So Beidou still unable to meet the needs of high-precision applications in the field of national strategic economic security

5 Conclusion

While China is developing the Beidou satellite navigation system, the United States try their best to improve the global positioning system (GPS). Russia, European Union, Japan and India are also striving to develop satellite navigation systems.

According to the current international situation,It quite important to China to construct Beidou satellite-based augmentation system,which can greatly improve the competitiveness of the whole industry chain and promote the transformation of traditional industries.

References

[1] Contribution of the Compass satellite navigation system to global PNT users[J];Chinese Science Bulletin;201126

[2] Research and progress of Beidou satellite navigation system[J];Science China(Information Sciences);201212

[3] Gelb Arther Ed,"Applied Optimal Estimation",MIT Press, Cambridge, Mass,2014

[4] Brison, A.E.Jr.and I Io,Y,C,Applied Optimal Control-Optimization,Estimation and Control, Iemisphere Publishing Corporation, 2013.

[5] Stengel,R.F.,Stochastic Optimal Control-Theory and Application,Wiley-Interscience, 2014

The Lightning Imager Instrument for Meteosat Third Generation

S. Lorenzini[1], R. Bardazzi[1], M. Di Giampietro[1], F. Boldrini[1], F. Feresin[2], D.M.A Aminou[3], P. Willemsen[3]

1. Selex ES S.p.A. – A Finmeccanica Company, Italy
2. Thales Alenia Space, France
3. ESTEC European Space Agency, The Netherland

The Lightning Imager for Meteosat Third Generation is an optical payload with on-board data processing for the detection of lightning.

The instrument will provide a global monitoring of lightning events over the full Earth disk from geostationary orbit and will operate in day and night conditions.

The requirements of the large field of view together with the high detection efficiency with small and weak optical pulses superimposed to a much brighter and highly spatial and temporal variable background (full operation during day and night conditions, seasonal variations and different albedos between clouds oceans and lands) are driving the design of the optical instrument.

The main challenge is to distinguish a true lightning from false events generated by random noise (e.g. background shot noise) or sun glints diffusion or signal variations originated by micro-vibrations. This can be achieved thanks to a 'multi-dimensional' filtering, simultaneously working on the spectral, spatial and temporal domains.

The spectral filtering is achieved with a very narrowband filter centred on the bright lightning O2 triplet line (777.4 nm ± 0.17 nm). The spatial filtering is achieved with a ground sampling distance significantly smaller (between 4 and 5 km at sub satellite pointing) than the dimensions of a typical lightning pulse. The temporal filtering is achieved by sampling continuously the Earth disk within a period close to 1 ms.

This paper presents the status of the optical design addressing the trade-off between different configurations and detailing the design and the analyses of the current baseline.

Emphasis is given to the discussion of the design drivers and the solutions implemented in particular concerning the spectral filtering and the optimisation of the signal to noise ratio.

I. Introduction

Lightning Imager (LI) is one of the instruments of the Meteosat Third Generation (MTG) mission and in particular it will be installed on the MTG-I satellites with the Flexible Combined Imager (FCI).

The objective of MTG mission is to provide Europe, by extension, the International Community, with an operational satellite system able to support accurate prediction of meteorological phenomena and the monitoring of climate and air composition.

Thales Alenia Space as the MTG prime contractor is responsible for the procurement of the Lightning Imager instrument developed and manufactured by Selex ES. The primary objective of the Lightning Imager (LI) mission is to add complementary information relevant to the detection and location of cloud-to-ground and cloud-to-cloud lightning to those provided by existing/planned

ground based lightning detection systems. So, these continuous lightnings data for the whole hemisphere would represent a new set of data to be used in nowcasting, climatology and atmospheric research.

LI has no heritage in Europe. Two USA Low Earth Orbit missions (LIS and OTD) have already flown and one, the Global Lightning Mapper (GLM) of GOES-R, is currently under development.

The instrument requirements reported in Table I are challenging and they define, together with the concept adopted for the lightning detection, the main drivers for the LI design.

TABLE I – LI MAIN REQUIREMENTS

Parameter	Requirement
FOV	16° diameter shifted northward or 84% of visible Earth disk, including all Eumetsat member states
Spatial sampling	< 10 Km @ Latitude 45° and Sub-satellite Longitude
Dynamic range of Earth background (Lbkg)	0 ÷ 296.5 $W/m^2/\mu m/sr$ (night ÷ summer solstice at midday)
Optical pulse dynamic range (LLp)	6.7 ÷ 670 $mW/m^2/sr$
Optical pulse spectral range	777.4 ± 0.17 nm
Minimum optical pulse duration	0.6 msec
Optical pulse size	10 Km ÷ 100 Km circular pulse diameter
Maximum number of optical pulses in the FOV	25 in 1 millisecond 800 in 1 second
Instrument Average detection probability (IADP)	90% for latitude 45 deg 70% as average over the FOV 40% over EUMETSAT member states (goal)
LI Mass (total Optical Head and Electronic box)	93 Kg
LI Optical Head Envelope	718 x 1200 x 1456 mm^3

II. Lightning Detection Concept

The lightning detection is achieved implementing the following functions:

- Earth image acquisition for continuous monitoring of the lightning's presence in the FOV;
- calculation of pixel by pixel adaptive background to cope with non-uniformities and low terms variations of the image (oceans, clouds, area in night conditions and areas with daylight conditions) and to reject at the same time noise effects and spurious events;
- removal of the background level from the overall pixel signal to obtain the net lightning illumination level;
- use of adaptive threshold; lower thresholds can be used in low noise dark areas of the scene, using higher thresholds only in highly illuminated areas (with corresponding higher shot noise);
- pixels for which the difference between the pixel value and the estimated background signal exceeds the threshold are kept as Detected Transients (DT);
- collection of the DT video data and additional information for the ground processing with a dedicated processing electronics;
- in flight processing of DTs to reduce the number of False Transients (FT) to a level compatible with the platform downlink data rate constraints (30 Mbps).

In addition LI is capable to acquire, process and transmit to ground an Earth background image.

III. Instrument Overview

The Lightning Imager is composed of one Optical Head (LOH) and one electronic equipment, the LI Main Electronics (LME).

The LOH consists of four identical Optical Channels (OC), each one including (see Figure 2):

- a protective cover on the baffle aperture to prevent baffle and optics contamination during launch and pre-launch activities;
- a baffle for stray light suppression and thermal load minimization;
- a Solar Rejection Filter (SRF), to minimize both the background level and the thermal load inside the OC;
- a Narrow Band Filter (NBF) to reduce the bandwidth in the range of the lightning spectral pulse (Figure 1);
- an optical system with F# 1.73, 110 mm entrance pupil diameter (determined by radiometry required to achieve the IADP performance) and 190 mm effective focal length (determined by the targeted GSD of 4.5 Km at Sub Satellite Point - SSP - and the size of detector pixels);
- a CMOS detector with 1000 x 1170 pixels, 24 µm pitch, 1000 frame per second;
- a processing electronics implementing the detection functions.

Each OC images a different portion of the visible Earth surface with the four line of sights tilted 4.75° from the SSP toward North, South, West and East in order to achieve the required coverage (Figure 3).

The Main Electronics performs the overall payload functions, the interface to the platform, the configuration of the processing electronics, the data flow regulation and finally compacts and packetizes the scientific data.

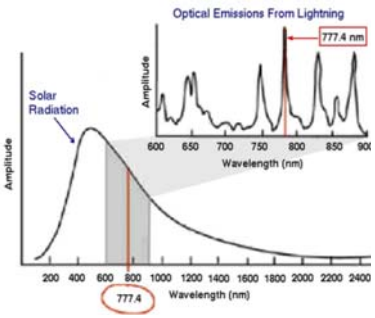


Figure 1 - Optical emission from lightning

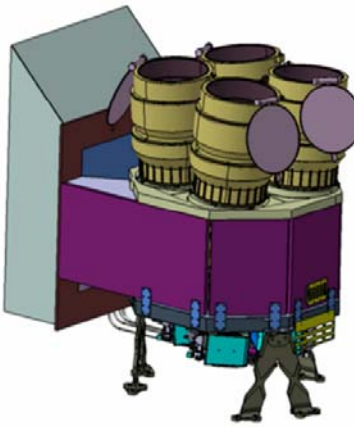


Figure 2 - LI Optical Head

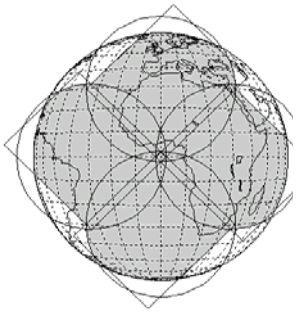


Figure 3 - LI coverage with 4 Optical Channel concept

IV. System Trade-offs

The main trade-offs for the selection of the LI configuration are:

- Single channel versus multi-channel architecture (this define the FOV of the optical channel).
- Narrow Band Filter (NBF) position within the optical path: close to entrance pupil for a parallel incidence beam working concept versus close to focal plane for a convergent incidence beam working concept.

- Optical system sizing: optimization of the entrance pupil diameter with respect to lightning Average Detection Probability (ADP), mass and volume constraints and definition of the Ground Sampling Distance (GSD) with respect to ADP and electronics processing load constraints and data rate bottlenecks.
- Optical system design: single primary optics with four relay systems; catadioptric approach; dioptric approach.

A. Single Channel Vs. Multichannel Architecture

The separation of the LI global FOV into multiple optical channel, taking into account the NBF positioned in the entrance pupil, allows mitigating the development risk of the critical items (NBF, detector and proximity electronics) and the optimization of the NBF performances.

Figure 4 shows the FOV layout for a single and double Optical Channel solutions fulfilling the coverage requirement.

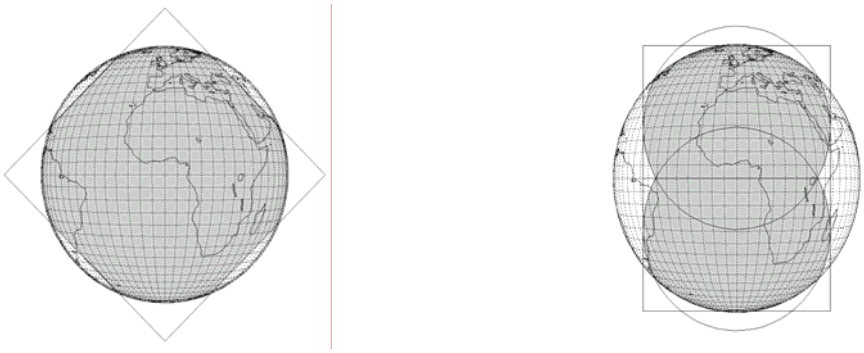


Figure 4 - LI coverage for Single and Double Optical Channels concept

The large FOV of the Single OC concept cannot be sustained by the NBF due to the blue spectral shift induced by the high angle of incidence which is not compatible with the bandwidth requirement. To limit such effect a Galileian telescope can be placed in front of the NBF (Figure 5) reducing the angle of incidence, but enlarging at the same time the filter diameter. This makes the achievement of the coating uniformity requirements more challenging. In addition this solution imposes the development of a very large detector array (about 5 Mpixels) and of a huge processing electronics (5 Gpixel per second to be processed in real-time) both considered unfeasible.

In case the global coverage is achieved by means of two or four Optical Channels, the detector size is reduced and the use of the Galileian telescope, requiring more mass ad envelope than available, is no more required.

The larger number of Optical Channels improves the NBF performance and the feasibility of detector and relevant processing electronics.

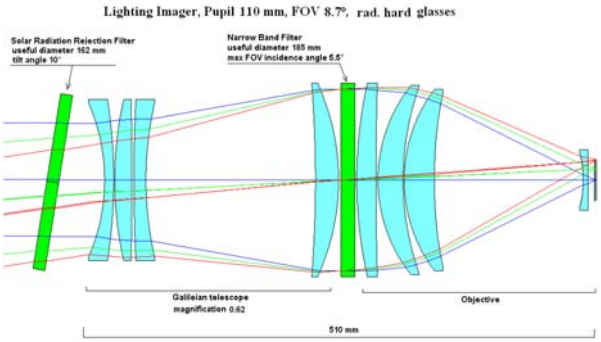


Figure 5 - Single optical channel layout with Galileian telescope

B. Definition of NBF Working Concept

The above mentioned NBF parallel beam concept has been compared with an alternative one, for which the filter is positioned close to focal plane (convergent beam optical configuration).

In this case to limit the spectral shift of the NBF bandwidth, the optical system must be telecentric and with an F# of 4.75, corresponding to a convergent beam of 6° maximum. The resulting pixel pitch of the detector is 65 µm producing a very large detector size and thus an increase of the effective focal length (~2.7X) compared with the NBF parallel beam concept.

However the trade-off between the NBF working concepts (parallel beam vs. convergent beam) is based on the Signal-to-Noise Ratio (SNR):

$$SNR = SLP / \sqrt{((SBKG + nro^2) * (1 + 1/NAVG))} \quad (1)$$

Where: SLP is the lightning pulse signal; SBKG is the background signal; nro is the read-out noise; NAVG is the background radiance averaging factor and corresponds to the number of background images averaged together.

Maximization of the SNR is achieved by proper specification of NBF bandwidth and in-band transmittance, taking into account, by means of a filter spectral response model (validated versus manufacturer test data), that minimizing the bandwidth leads to a decrease of the peak transmittance and vice versa.

Based on (1) and on the NBF spectral model, it can be demonstrated that SNR cannot be maximized for any illumination condition: in other words it is necessary to decide whether to increase SNR in night-time condition, maximizing the NBF transmittance, or in daylight condition, minimizing the filter bandwidth, the two solutions being in opposition.

The diurnal flash distribution provided by LIS data can be used to calculate a weighted SNR average, thus defining a Factor of Merit (FoM) that is not dependent on the illumination condition (see example in Figure 6), even if this statistics is not part of ESA/TAS requirement specification.

The FoM is calculated as a function of the NBF Full Width at Half Maximum (FWHM) for both the parallel and the convergent beam concepts. Results are reported in Figure 7 and Figure 8 respectively.

The parallel beam concept provides a better average FoM for any value of the FWHM.

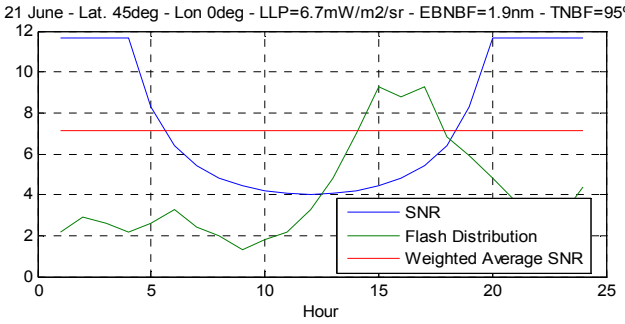


Figure 6 - SNR weighted average

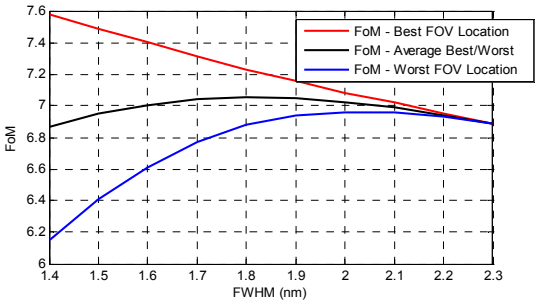


Figure 7 - FoM in parallel beam concept

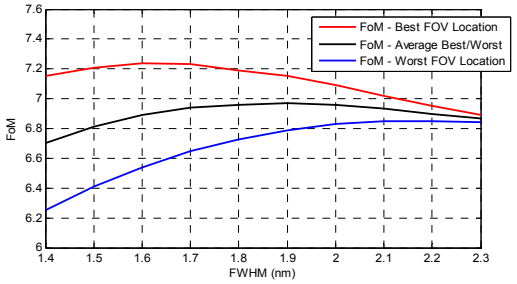


Figure 8 – FoM in convergent beam concept

V. Optical System Design

The results of the previous evaluations and trade-offs limit the study to optical configurations with four detectors and NBF in parallel beam.

A.Optical Configuration Trade-Off

A single primary optics with four relay and four detectors (in field splitting) has been evaluated.

With this configuration some Earth areas are not covered due to detectors borders. Therefore to get proper sampling of the image on the focal plane special arrangement is needed. For example in the following Figure 9 the image in the intermediate focal plane is split in four parts using a fore-optics, wedges, to avoid un-covered Earth areas, mirrors and relay optics.

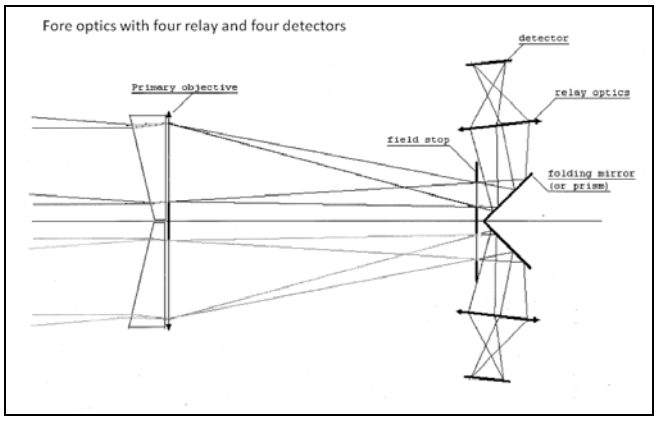


Figure 9 - single primary with four relay approach

The above first order concept might be realized by using a Schmidt-Cassegrain telescope as fore optics. Such an approach has the following advantages and drawbacks:

Advantages :

- The longitudinal envelope is shorter compared to others approaches
- Mirrors are free of secondary images and ghosts.
- The intermediate focal plane and the field stop avoid the light from outside the FOV to illuminate the relay reducing partially stray light contribution.

Drawback:

- Mirror shape and dimension drive the system out of weight and volume requirements. Moreover the relay optics dimension, schematized in Figure 11, are not negligible.

Although the present solution might appears attractive, due to its weight and volume this concept presents no advantages respect to a multi-channels approach that is described in the following.

The catadioptric approach consisting of a Schmidt telescope with relay optics is shown in Figure 10.

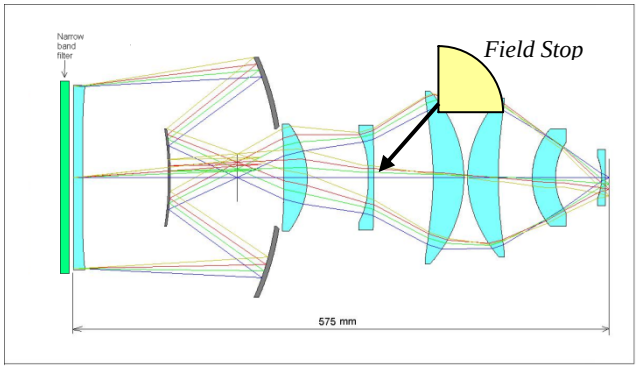


Figure 10: Catadioptric approach: Schmidt with relay

A field stop of suitable shape to image the required Earth portion, is placed in the intermediate focal plane.

Advantages :

- The longitudinal envelope is shorter than an all dioptric (with relay) approach
- The mirrors are free of secondary images and ghost.
- The intermediate focal plane and the field stop avoid the light from outside the FOV to illuminate the second part of the objective reducing partially stray light contribution.

Drawbacks :

- Due the central obscuration the filter diameter and the telescope transversal dimension are larger with respect to the equivalent pupil diameter.
- The obscuration ratio increases the radiation spread over the diffraction pattern outer rings deteriorating the nearby dimmer signals when strong in field sources (e.g. clouds) are focused on the focal plane.
- The FOV is vignetted as the entrance pupil is placed on the telescope Schmidt plate (located close to the narrow band filter) as shown in Figure 11.

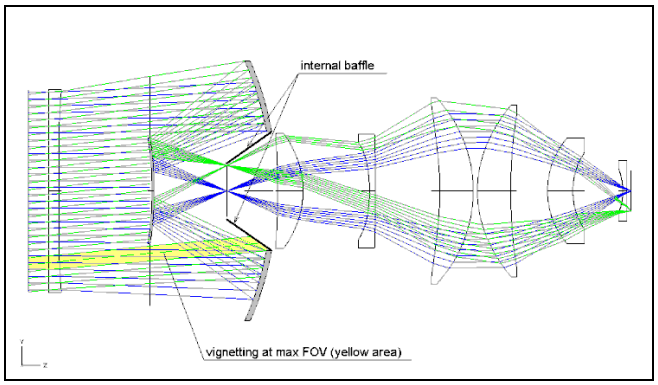


Figure 11: Schmidt telescope vignetting

The blue rays correspond to the axial FOV, the green rays are the ones of the maximum FOV. The yellow area, coming from the maximum FOV, is blocked from internal baffling. This loss of radiation increases with FOV and can be compensated by means of an extra pupil diameter to guarantee the required aperture at the maximum FOV. Nevertheless the lens equivalent aperture changes with respect to the FOV and it is maximum on axis and decreases with the FOV.

Some of the above drawbacks can be solved using a more complex telescope (three mirrors) increasing however weight, volume and complexity.

Considering the narrow spectral bandwidth, a dioptric design allows compact configurations with reduced complexity when compared to mirrors based solutions.

In Figure 14 a primary objective is coupled with a relay to realize an intermediate focal plane. In the intermediate focal plane a suitable shaped stop is placed while the interference filter is positioned in front of the fore optics.

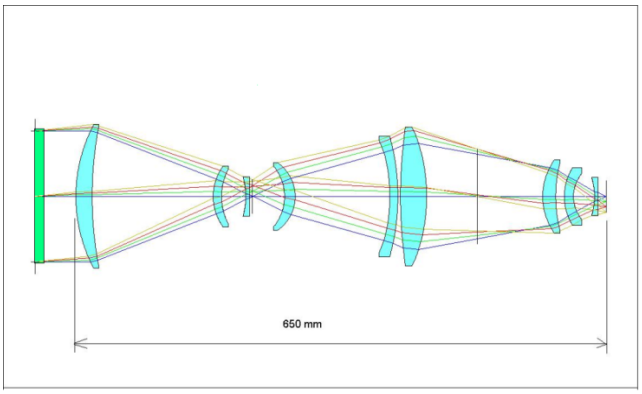


Figure 12: Dioptric with relay optical layout

Advantages :

- The narrow band filter is placed in front of the imager (green plate on Figure 14) and its size is minimized with respect to catadioptric solutions.
- The intermediate focal plane and the field stop prevent the light from outside the FOV to illuminate the second part of the objective reducing partially stray light contribution.

Drawback :

- Increase of longitudinal dimension and weight.

The advantages in stray light reduction do not balance the increase of mass and dimensions of the opto-mechanical system. Moreover, being in permanent view of the Earth, the first objective surface is the major contributor to stray light mainly due to contamination. In this case part of the scattered rays reaches the detector since they are close to incidence direction and cannot be baffled by the field stop.

The tight requirements on the longitudinal dimensions of the optical system and the need of a baffle to reduce fore optics illumination by off axis source are driving the design toward a single stage optical system.

B. Selected Configuration

The selected optical layout is composed by four optical channels, each one with an independent single stage lens, detector and baffle. Figure 13 shows the optical layout of the baseline solution.

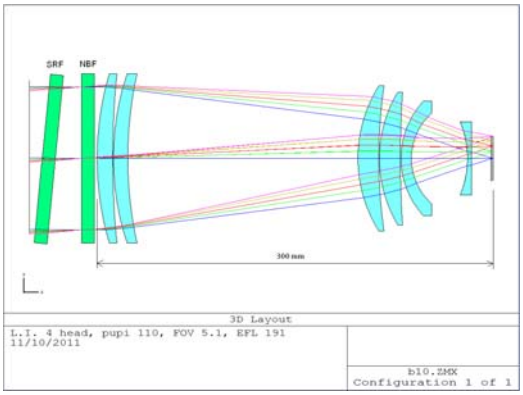


Figure 13 – Selected optical layout

The first two parallel plates are the Solar Rejection Filter and the Narrow Band Filter (both in green). The imager is composed of 6 lenses, all made of radiation resistant glass. All the lenses have spherical profile except one.

The lenses diameter are larger sized to insert some “light traps” and to limit the straylight caused from the lenses border and internal objective walls.

The Solar Rejection Filter, the first from left, is placed tilted to mitigate the ghosts images due to multiple reflections between filters.

Using a pixel projection model, the GSD corresponding to the above mentioned optical system parameters is represented in Figure 16 at SSP and at 45°N latitude and SSP longitude.

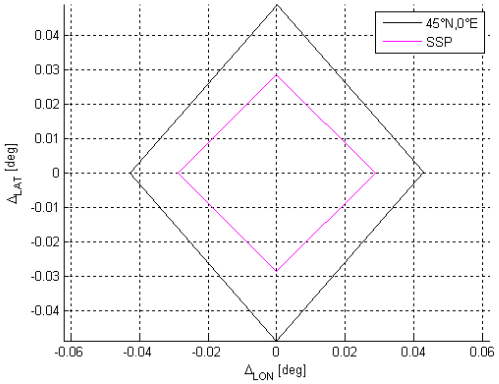


Figure 14: GSD model for reference LP event longitude and latitude

The nominal optical system performances at 20°C are shown from Figure 17 to Figure 21.

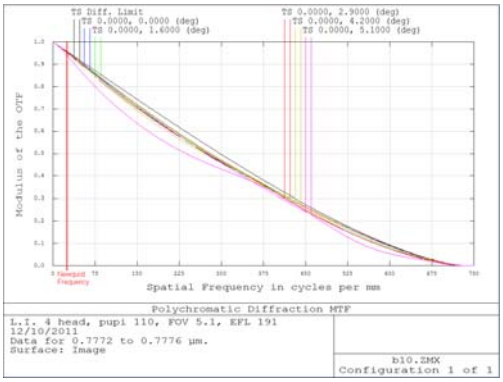


Figure 15: MTF

The RMS wavefront error, on axis (Figure 18), is 0.029 waves, corresponding to 22 nm.

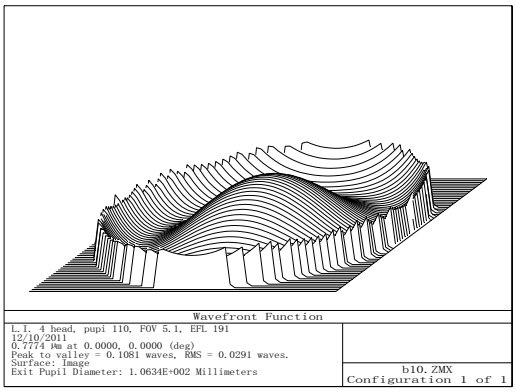


Figure 16: On axis RMS wavefront error function

The RMS wavefront error, at maximum FOV (Figure 19), is 0.072 waves, corresponding to 56 nm.

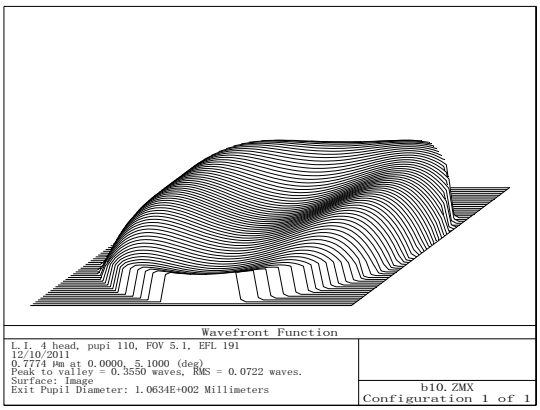


Figure 17: maximum FOV RMS wavefront error function

The instrument performance model indicates that the energy fraction in one pixel should be larger than 0.9 to ensure a proper ADP. Figure 20 shows that there is a margin with respect to the nominal design to account for image quality degradation due to manufacturing, alignment, thermal effects and other perturbations.

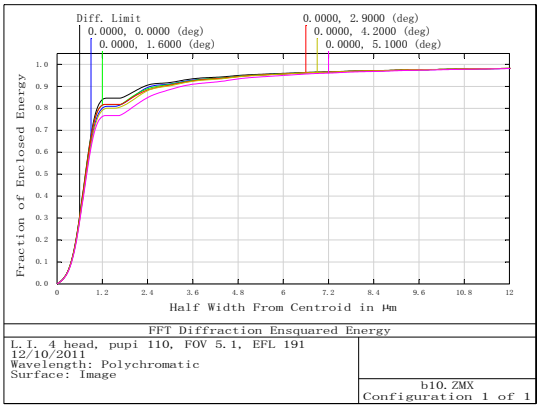


Figure 18: Ensquared energy function

The maximum distortion is less than 0.2% at maximum FOV as shown in Figure 19

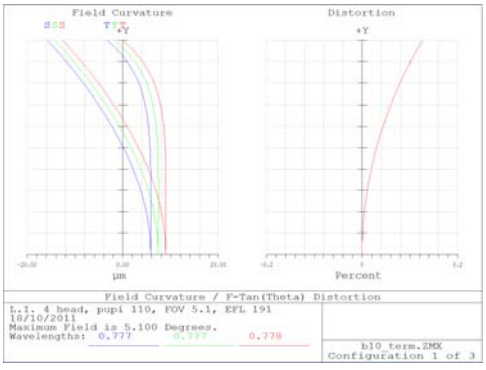


Figure 19: distortion and field curvature

The worst RMS wavefront error within the operative temperature range, on axis (Figure 24), is 0.039 waves, corresponding to 30 nm.

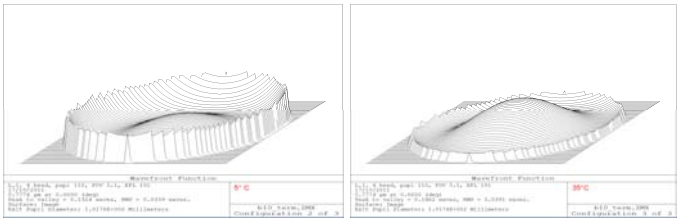


Figure 20: Wavefront function on axis at the operative temperature limits

The worst RMS wavefront error within the operative temperature range, for all the FOV positions (Figure 25), is 0.079 waves, corresponding to 61 nm at maximum FOV.

Figure 21: Wavefront function at maximum FOV at the operative temperature limits

The optical system performances at the limit of the operative temperature are almost unchanged with respect the 20°C ones.

The transmittance of the optical system at 777.4 nm has been maximized selecting an anti-reflective coating providing a transmittance of 99.78% for each of the 12 optical surfaces. Taking into account the transparency of the radiation resistant glass (99.7% at 700nm for 10mm thickness) and the overall lenses thickness (65mm), the overall transmittance of the 6 lenses is estimated as:

$$T(777.4nm) = 0.9978^{12} \cdot 0.997^{6.5} = 0.955$$

C. Baffle Design

The preliminary straylight analysis demonstrates that when the Sun directly illuminates the optical system components (i.e. when the Sun is within the baffle cut-off angle) the straylight becomes a critical issue for achieving the required absolute radiometric accuracy.

In order to limit to 2.5% the mission time when this criticality is present, a cut off of 16°, compatible with mass and envelope requirement, has been defined.

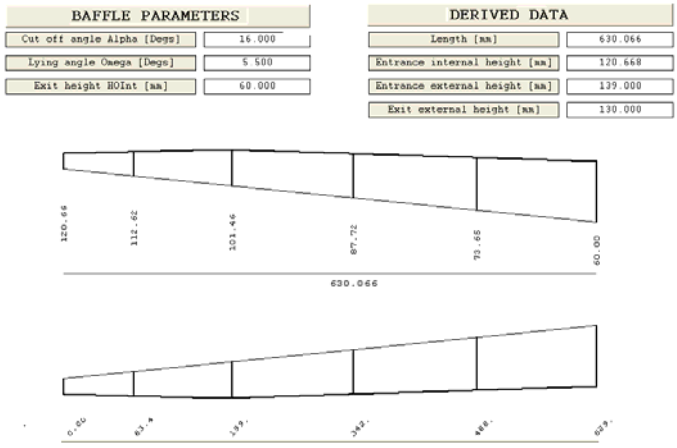


Figure 22: Baffle layout

D.LI Spectral Filters

The Solar rejection Filter (SRF) is devoted to reflect the solar radiation. It also works in synergy with the Narrow Band Filter (NBF) to obtain the required spectral filtering. It is directly exposed to the external space environment. Its mechanical mounting is thermally decoupled with respect the NBF and the optical system; its temperature range is much larger than optical system. The useful diameter is 118 mm.

In Table IV the performance requirements for SRF are provided in terms of transmittance at the scientific band and transmitted, reflected and absorbed energy in the 200 ÷ 2000 nm spectral range that includes almost the whole Sun radiance energy. The latter are estimated according the following formulas, where S(λ) is the Planck distribution for an equivalent black body at 5780K.

The calculation of SRF Total Energy Transmitted, Reflected and Absorbed are reported respectively in (1), (2) and (3).

$$TET_{SRF}\,\% = \frac{\int_{200}^{2000} S(\lambda) * T_{SRF}(\lambda) d\lambda}{\int_{200}^{2000} S(\lambda) d\lambda} * 100 \qquad (1)$$

$$TER_{SRF}\,\% = \frac{\int_{200}^{2000} S(\lambda) * R_{SRF}(\lambda) d\lambda}{\int_{200}^{2000} S(\lambda) d\lambda} * 100 \qquad (2)$$

$$TEA_{SRF}\,\% = \frac{\int_{200}^{2000} S(\lambda) * A_{SRF}(\lambda) d\lambda}{\int_{200}^{2000} S(\lambda) d\lambda} * 100 \qquad (3)$$

TABLE IV – SRF PERFORMANCE REQUIREMENTS

Parameter Description	Range (nm)	Value
SRF Transmittance	777.23 ÷ 777.57	> 95%
Total Energy Transmitted (TETSRF)	200 ÷2000	< 10 %
Total Energy Reflected (TERSRF)	200 ÷2000	> 84 %
Total Energy Absorbed (TEASRF)	200 ÷2000	< 6 %

The NBF, placed between the SRF and the first lens, performs the spectral discrimination of the lightning pulse from the Earth background radiance. It is supported by the lenses mounting so its operative thermal range is the same of the optical system. The useful diameter is 112 mm.

In Table V the performance requirements for NBF are provided in terms of:

- Transmittance in the scientific band. Its calculation is reported in (4).
- Equivalent bandwidth. Its calculation is reported in (5)

$$\overline{T_{NBF}} = \frac{1}{777.57 - 777.23} \cdot \int_{777.23}^{777.57} T_{NBF}(\lambda) d\lambda \qquad (4)$$

$$EB_{NBF} = \frac{\int_{772.4}^{782.4} T_{NBF}(\lambda)d\lambda}{\overline{T_{NBF}}} \quad (5)$$

The integral has been defined on a spectral range sufficiently wide to allow the manufacturer to optimize the coating design regardless of the proposed pass band shape.

Table V: NBF PERFORMANCE REQUIREMENTS

Parameter Description	Range (nm)	Value
NBF Transmittance ($\overline{T_{NBF}}$)	777.23 ÷777.57	> 90%
NBF Equivalent Bandwidth (EB_{NBF})	772.4 ÷ 782.4	< 1.9 nm

The ideal NBF bandwidth would be the scientific bandwidth (0.34 nm). Nevertheless, the actual bandwidth is specified in a figure of 1.9 nm to guarantee the required transmittance in the scientific bandwidth for all the working conditions and considering also the coating manufacturing errors.

These effects/errors are:

- Spectral shift due the AOI (0° to 5.1°)
- Deposition uniformity and wavelength centering error.
- Temperature effect.
- Radiation stability.
- Environmental stability (humidity, thermal cycling).

The out of band rejection requirements for the combined operations of the fully coated SRF and NBF (including substrate contributions) is specified in Table VI.

Requirements are given in the spectral range of the detector sensitivity (assumed from 200nm to 1100nm), with the exception of the 10 nm not covered by (5)

Requirements are provided in terms of:

- Mean out of band transmittance on the spectral range to minimize the background signal out of band, in the detector sensitivity range. The specified value provides a negligible contribution compared with the EBNBF figure.
- Maximum out of band transmittance to avoid any significant pulse spectral radiance contribution out of the required emission lines.

The Combined Out of Band Mean Transmittance calculation is reported in (6).

$$T_{OBM} = \frac{\int_{782.4}^{1100} T_{NBF} \cdot T_{SRF}(\lambda)d\lambda + \int_{200}^{772.4} T_{NBF} \cdot T_{SRF}(\lambda)d\lambda}{(1100 - 782.4) + (772.4 - 200)} \quad (6)$$

Table VI: SRF AND NBF COMBINED OUT OF BAND TRANSMITTANCE REQUIREMENTS

Parameter Description	Range	Value
Combined Out of Band Mean Transmittance (T_{OBM})	200 ÷ 772.4 & 782.4 ÷ 1100	< 0.01%
Combined Out of Band Maximum Transmittance (T_{OBMax})	200 ÷ 772.4 & 782.4 ÷ 1100	< 1%

VI. conclusion

The Lightning Imager will be the first instrument developed in Europe for the detection of lightning events from geostationary orbit. The challenging detection requirements in day and night conditions, combined with the large coverage and the tight mass and envelope make the Lightning Imager a complex instrument from an optical perspective.

Despite the complexity of the requirements an instrument based on four simple dioptric lenses has been identified from an exhaustive trade-off. The results of the optical analyses presented in the paper indicate that the selected configuration is very promising for a successful Lightning Imager mission.

Acknowledgement

THIS WORK was SUPPORTED/FUNDED BY THE EUROPEAN SPACE AGENCY (ESA) UNDER MTG CONTRACT PRIMED BY THALES ALENIA SPACE - FRANCE

References

[1] Steven J. Goodman, "Geostationary Lightning Mapper for GOES-R and Beyond", STAR Science Forum 2008, http://www.star.nesdis.noaa.gov/star/

[2] Ullrich Finke, "Characterising the Lightning Source for the MTG Lightning Imager Mission", http://www.eumetsat.int/groups/pps/documents/document/pdf_mtg_rep30.pdf

[3] Boccippio, D. J., W. J. Koshak, and R. J. Blakeslee (2002). "Performance assessment of the optical transient detector and lightning imaging sensor. Part I: Predicted diurnal variability. J. Atmos Oceanic Technol. 19, 1318–1332.

[4] Christian, H. J., R. J. Blakeslee, D. J. Boccippio, W. L. Boeck, D. E. Buechler, K. T. Driscoll, S. J., Goodman, J. M. Hall, W. J. Koshak, D. M. Mach, and M. F. Stewart (2003). "Global frequency and distribution of lightning as observed from space by the optical transient detector", J. Geophys Res. 108, 4005, doi:10.1029/2002JD002347.

Space Detection System based on Ultrafast Optics

Liu Xun, Yang Bingxin, Ruan Ningjuan, Li Wei

Beijing Institute of Space Mechanics & Electricity (CAST), Beijing 100094, China

Abstract: Ultrafast optics has been developed quickly, and many new technologies based on ultrafast optics have been proposed, which shows the applications in various fields. As the important method, the femtosecond laser provides the novel approaches for the development of space detection system. The femtosecond optical frequency combs can build the connection between the optical frequency and microwave frequency, and improve the stability and precision, which brings the breakthroughs in measurement field based on laser frequency. In this paper, the femtosecond optical frequency combs is introduced, and its applications in space detection are proposed.

Keywords: ultrafast optics, instantaneous optics, space detection, femtosecond, femtosecond optical frequency combs, high-precision

1. Introduction

With extreme space conditions and complex detection environment, it is vital to increase the detection accuracy of space physics quantity and build the detection standards, which will be important in future space activities. Ultrafast optics is the most advanced technology in the field of instantaneous optics and related to the micro time scale instantaneous phenomenon. The technologies based on ultra optics can be used to develop the measurement for time scale, absolute distance and spectrum, which could be applied in space detection, satellite constellation flight, relativity theory validation, extraterrestrial planet exploration, etc. In this paper, the ultra-fast optics and ultra-fast laser have been introduced firstly, and then according to the features, the new technologies based on the ultra-fast optics have been proposed. Finally, the space applications are summarized.

2. Ultrafast optics

Ultra-fast optics is related to the instantaneous phenomenon of the tiny time scale, which is the shorter than femtosecond. From middle 1980's, the ultra-fast optics has been developed quickly, which shows the application value in many fields. As the important method of ultra-fast optics, ultra-fast laser has two significant features. One is short pulse duration, and the other is high peak power. The short pulse duration is the level of femtosecond (10^{-15} s), and the high peak power can reach about ~10^{21} W/cm^2, which makes the ultra-fast optics show the application prospects, especially will bring more breakthroughs in space detection. The Schematic diagram of an amplifier system based on chirped-pulse amplification is shown in FIG.1.

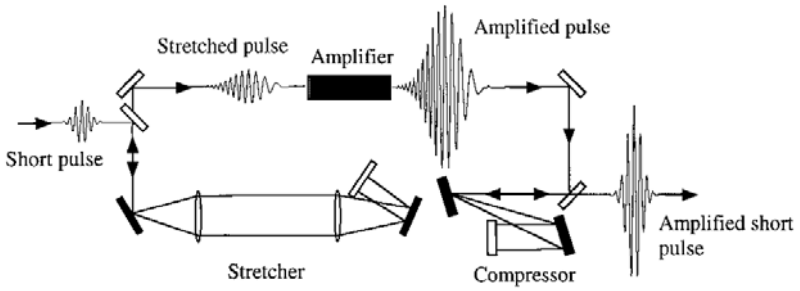


FIG. 1. Schematic diagram of an amplifier system
based on chirped-pulse amplification.

3. Femtosecond optical frequency comb

The frequency domain representation of a perfect frequency comb is a series of delta functions spaced according to

where is an integer, f_r is the comb tooth spacing (equal to the mode locked laser's repetition rate or, alternatively, the AM frequency)， f_0 is the carrier offset frequency, which is less than f_r. Combs spanning an octave in frequency can be used to directly measure f_0. Thus, octave spanning combs can be used to steer a piezoelectric mirror within a carrier envelopephase correcting feedback loop. Any mechanism by which the combs' two degrees of freedom (f_r and f_0) are stabilized generate a comb that is useful for mapping optical frequencies into the radio frequency for the direct measurement of optical frequency.

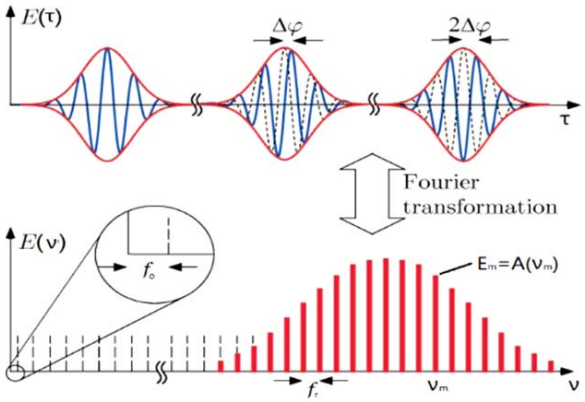


FIG.2 The features of optical frequency comb in time domain
(above image, ultra-short pulse) and frequency domain
(below image, narrow linewidth laser sequence)

Since the beginning of 2000, developments in femtosecond optical frequency combs have enabled the use of novel techniques in several applications covering the fields of science and engineering.The field of precision optical frequency measurements has undergone a major revolution during recent years due to the development of highly accurate frequency combs based on ultrashort pulse mode-locked lasers. Most of this work has taken advantage of the highly-developed Ti:Sapphire laser which either directly or in combination with nonlinear frequency broadening schemes can provide octave spanning continua. Locking both the offset frequency and repetition frequency of the resulting frequency comb to an atomic clock has permitted the development of absolute optical frequency references in the visible to near infrared wavelength regions. The

application of ultrashort pulse lasers operating at wavelengths longer than that of the Ti:Sapphire laser appears promising, for example, to establish time and frequency standards in the infrared for telecommunications by use of highly stable optical combs. Development of systems for frequency interval measurements based on a frequency comb is also of particular interest for frequency synthesis and metrology in the infrared. In such systems an octave spanning continuum is not mandatory and generally only a stabilization of the pulse repetition rate is required which is routinely attainable in most cases. If necessary, stabilization of the absolute frequency is achieved by locking to an optical reference.

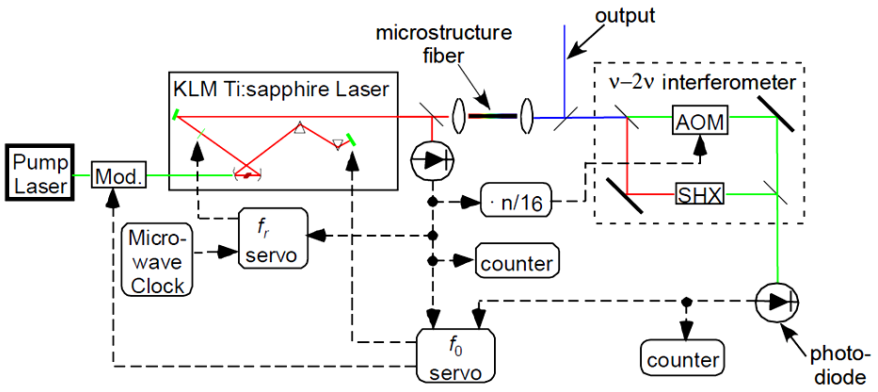


FIG.3 Schematic of a femtosecond comb generator.

4. Space application

Most notably the European Southern Observatory (ESO) is the precise calibration of astronomical spectrographs by frequency comb. Some scientists are also coordinating a number of projects with the German and the ESA to bring femtosecond optical frequency combs to space. The main applications are as follows.

4.1 Extraterrestrial planet exploration

The radial velocity measurement is the most productive technique for extraterrestrial planet hunting. The precision of radial velocity measurement has to be dramatically improved, with the evenly distributed calibration lines and a wide spectral range. The femtosecond optical frequency combs can be used to calibrate the high precision astronomical spectrometer, which can improve the precision more than two levels, reach the level of 1cm/s theoretically.

4.2 Astrophysical object composition detection

Based on opaqueness spectral detection of star radiation, the composition of astrophysical objects can be achieved. In the meantime, the detection of rock and metal on the planet's surface will also provide the planet’s information. The spectrometer based on ultra fast optics can provide the powerful ability to detect the astrophysical object composition. The laser induced breakdown spectroscopy based on ultra fast optics can reach the higher precision, and can realize the ultra-trace detection.

4.3 Satellite constellation flight

The satellites constellation flight demands for high precision absolute distance measurement. The rapid and precise measurement of long absolute distance is a critical technical support. Due to the locking of the repetition rate of the laser to a time standard, the distance between successive pulses is accurately known, providing direct traceability to the definition of the distance. Since the interpulse distance is typically of the order of 1 m, the range of non-ambiguity is large. Such accuracy is easily obtained with other methods like time-of-flight measurements.

5. Conclusion

The development of ultrafast optics provides the new methods for the space detection technologies and systems. The femtosecond optical frequency combs shows the applications in space detection, satellite constellation flight, relativity verification, extraterrestrial planet exploration, etc. However, there are some issues need to be researched. In the meantime, the miniaturization and utility are also need to be studied.

References

[1] Yu Zijiao, Han Hainian, Wei Zhiyi, Progress in dual-comb spectroscopy [J], Physics, 2014, 43(7):460-467.

[2] Wang Guochao, Yan Shuhua, Lin Cunbao, et al. Overview of large scale precision ranging by femtosecond optical frequency comb [J], Optical Technique, 2012, 38(6): 670-677.

[3] Han Hainian, Wei Zhiyi, Zhao Gang, Astro-frequency-combs and high precision measurement of the radial velocity of celestial bodies [J], Physics, 2012, 41(4): 249-255.

[4] Wu Xuejian, Li Yan, Wei Haoyun, et al. Femtosecond optical frequency combs for precision measurement applications [J], Laser & Optoetectronics Progress, 2012, 9, 030001.

[5] Jun Ye and Steven T. Cundiff, Femtosecond optical frequency comb: Principle, Operation, and Applications [B], Kluwer Academic Publishers /Springer, 2004.

[6] Giovana T. Nogueira and Flavio C. Cruz. Optical frequency comb for high-resolution and precision metrology [J], 2007 SBMO/IEEE MTT-S International Microwave & Optoelectronics Conference (IMOC 2007):644-647.

Low-light-level Imaging Technology Based on Black Silicon CMOS Detector

Wei LI, Huaiyi WANG, Xun LIU, Ningjuan RUAN

Beijing Institute of Space Mechanics & Electricity (CAST), Beijing 100094, China

ABSTRACT: The existing space optical earth observation system of China needs to improve the all-day detection ability, especially the ability to detect the dim target. The urgent need is to develop the high resolution low-light-level imaging payload technology and the new low-light-level detector technology. The hyperdoping black silicon detector has the advantages of high responsivity, extended infrared, mature technology and CMOS compatible, which can be developed to a new low-level-light imaging detector. This article first defines the conditions for low-light-level imaging, analyses the features and advantages for black silicon detector and summarizes the research situation in China and abroad. On this basis, it evaluates the performance level for the low-light-level imaging system based on black silicon detector. In addition, it gives suggestions to the development of the detector, which aims at improving the low-light-level imaging ability of China's space optical remote sensing system and laying a technical foundation for it.

KEYWORDS: low-light-level imaging; hyperdoping; black silicon; extended infrared; high responsivity

I. Introduction

The existing space optical earth observation system of China needs to improve the all-day detection ability, especially the ability to detect the dim target. It demands the high resolution low-light-level imaging payload, which will establish the base for improving the abilities of Chinese remote sensing [1, 2]. Some researches show that the low-light-imaging abilities mainly depend on the performance of low-light-imaging detector. Compared to the overseas detectors, the detectors manufactured in china needs to be developed quickly. The existing detectors cannot match the detect requirements. Therefore, the low-light-imaging detectors should be developed.

II. Low-light-imaging condition

Low-light-level is the moonlight, starlight, and airglow in the night. Low-light-imaging payload must need two imaging conditions. One is working in visible light and infrared band, the other is working in night. Under the condition of full moon, the moonlight is strong. The spectrum peak is 643nm, which prefer to infrared light, as compared with sun light. The reflective spectrum of moon is shown in Fig.1. With starlight and airglow but moonlight, the visible light is weak and the radiation intensity is less than 10^{-9}W/(cm^2·Sr um). The background noise is still dramatic though the high response detector is used. Under this condition, the intensity of near infrared, 1100-1200nm, is two-level higher than it of visible light. The existing detector works within 1050nm, which cannot response to the 1100-1200nm.

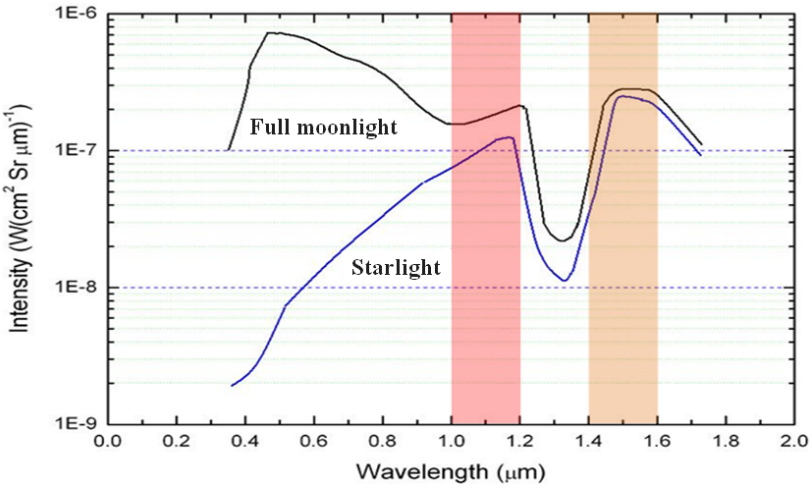


Fig.1 Radiation curve of full moon and starlight

The quantum efficiency of EMCCD-CCD-201 is shown in Fig.2. The maximum is 93%@575nm. The efficiency is low at the cut-off frequency, 20%@300nm and 10%@1000nm, respectively. Therefore, it needs to develop the advanced detected technology for improving the low-level-imaging abilities.

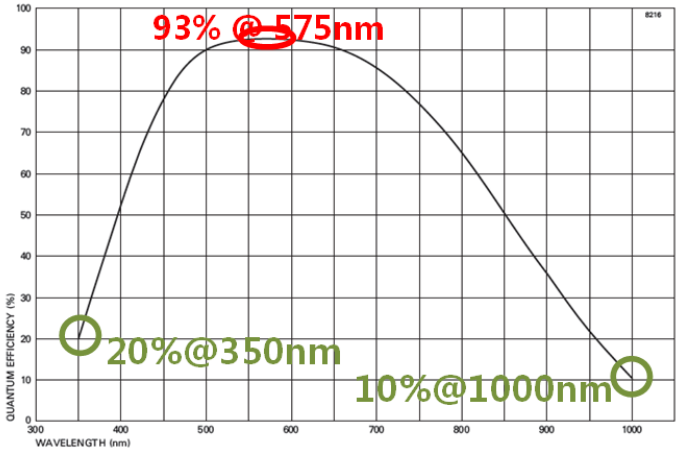


Fig.2 Quantum efficiency of EMCCD-CCD-201

III. Black silicon detector

Hyperdoping some special elements can make the silicon detector own the ultra- high responsibility and improve the detected ability in infrared band. After hyperdoping, the surface of detector appears black, so this type detector is called as “black silicon detectror” [3]. Two reasons contribute the improvements. One, highconcentration Si hyperdoping make the infrared absorptivity improved. The other, the crystalline cone appears on the detector’s surface. The crystalline cone is a micro-nano structure, which makes the detector own the antireflective effects [4]. Therefore, the black silicon shows the features as follows:

- Ultra-high responsibility
- Infrared extend
- Work at room temperature
- Low cost
- CMOS compatibility

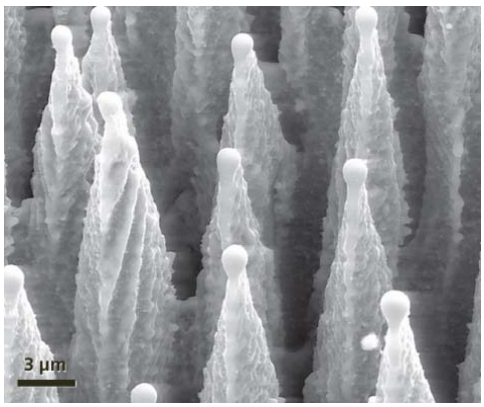


Fig.3 The crystalline cone on black silicon's surface

The change of spectral absorptivity in near infrared and mid and far infrared generated by hyperdoping is shown in Fig.4. The response curves under different bias voltage are shown in Fig.5 [5]. From the features above, the black silicon needs the requirements of low-level-imaging, which provides the original technology for payload development [6].

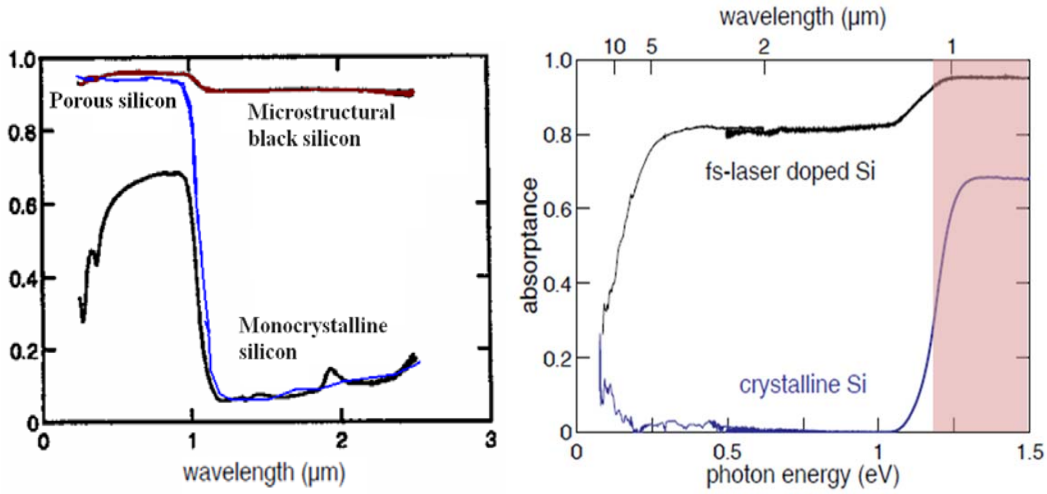


Fig.4 The change of spectral absorptivity in near infrared and min and far infrared generated by hyperdoping

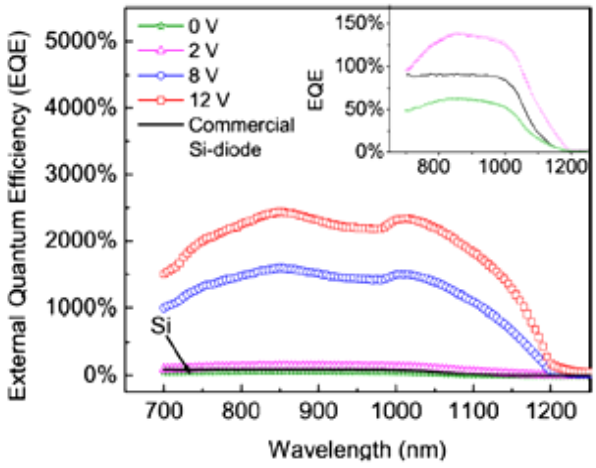


Fig.5 The response curves under different bias voltage

IV. Research status

4.1 Overseas

Harvard University developed the research of hyperdoping silicon detector firstly, and has achieved the best results around the world, as shown in Fig.6 [7]. The SiOnyx Company was established by Harvard University. Now, the CMOS imaging detector based on black silicon has been developed successfully, as shown in Fig. 7. The Hamamatsu Company in Japan has engaged in research for many years, but there is some difference compared with SiOnyx.

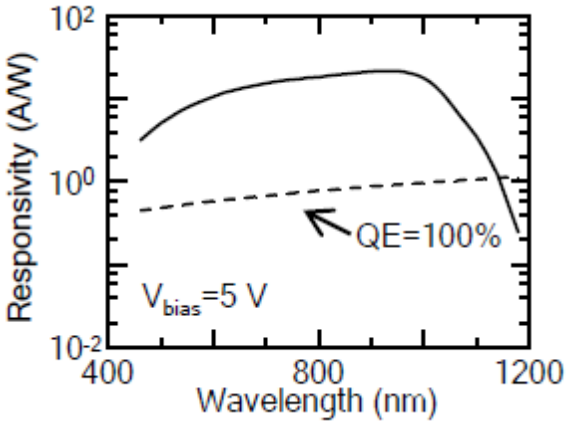
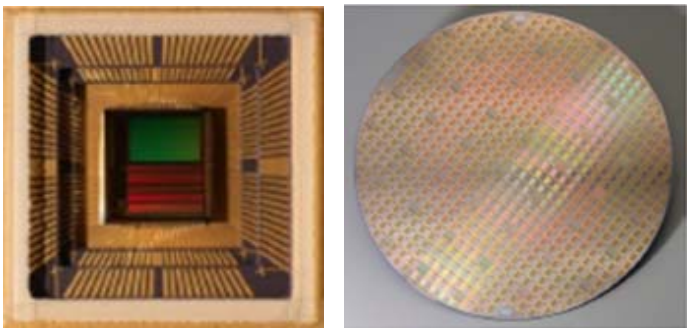


Fig. 6 Quantum efficiency by Harvard university



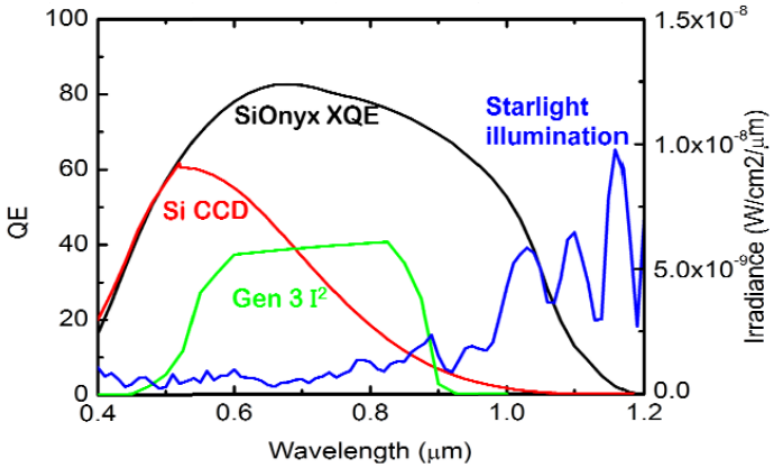


Fig. 7 Black silicon and its quantum efficiency developed by SiOnyx

SiOnyx Company has test the performance of black silicon detector. The test results show that the low-level-light imaging system owns the advanced abilities. The test results of low-level-light imaging system are shown in Fig.8. The parameters of SiOnyx XQE are as follows: resolution is 380×240, pixel is 11.2μm, integral time is 16ms. The only light source is a LED@940nm. Compared with standard CMOS detector, the SiOnyx XQE shows the perfect abilities. The test results of SiOnyx XQE-1310 without moonlight is shown in Fig.9. The parameters of SiOnyx XQE

are as follows: resolution is 1280×1024, pixel is 10.0μm, F number is 1.4, frame frequency is 30. The detector still shows the prominent low-level-light imaging ability [8, 9, 10, 11].

Fig. 8 The test results of SiOnyx XQE in the room

Fig. 9 The test results of SiOnyx XQE without moonlight

4.2 Inland

Many universities and institutes have joined in the research, including Institute of Semiconductors (CAS), Nankai university, Fudan university, Jilin university, etc. The Institute of Semiconductors (CAS) has established the equipments and study group. In the meantime, they have developed some excellent results. The response curves of black silicon detector developed by Institute of Semiconductors (CAS) are shown in Fig.10 [12]. The equipments of ion implantation and femtosecond laser are shown in Fig.11.

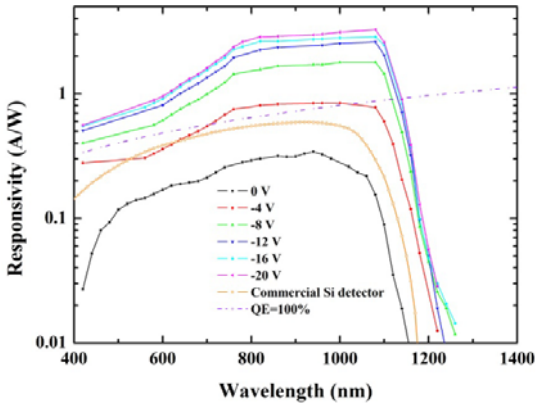


Fig. 10 The response curves of black silicon detector developed by Institute of Semiconductors (CAS)

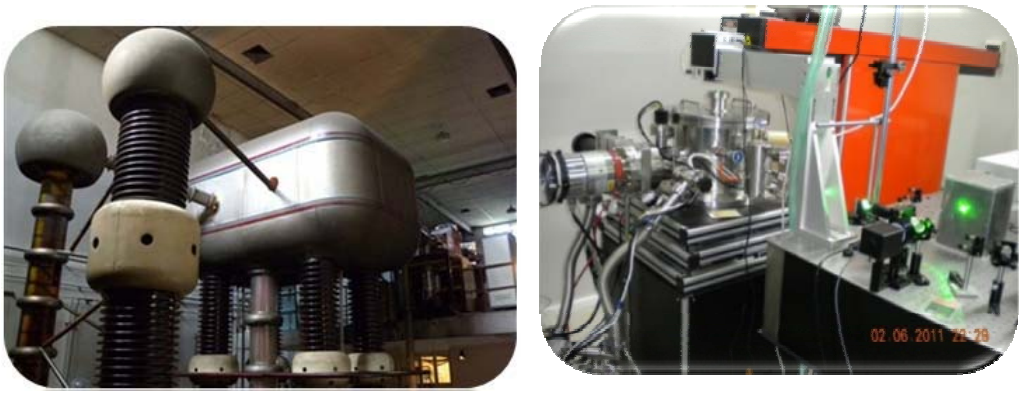


Fig. 11 The equipments of ion implantation and femtosecond laser

In theoretic research, there is little difference between foreign and interior institutes. In developed abilities, interior institutes have mastered the same methods. The foreign research results have been developed to the payload, while the technologies need to be developed in our institutes. According to the conditions above, Beijing Institute of Space Mechanics & Electricity (CAST) has established the group with Institute of Semiconductors (CAS) to develop the key technologies.

V. Low-level-light imaging based in black silicon CMOS detector

Low-level-light imaging is the high resolution optical remote sensing under the condition of low light. The development tendency is to improving the utilization efficiency of photon energy in the response spectrum. There are main methods: one, improving the quantum efficiency in the response spectrum; the other, extending the response spectrum more than 1000 nm.

Now, EMCCD is used in low-level-light imaging. According to the principle of EMCCD, the manufacturing process is complicated, and the cost is expensive. Therefore, the low-level-light imaging technology based on CMOS detector is the important development direction.

Beijing Institute of Space Mechanics & Electricity (CAST) has developed the research of Low-level-light imaging based in black silicon CMOS detector. The optical system is shown in Fig.12, which is composed of main optical system, revise lens group and black silicon CMOS detector. The parameters of optical system are listed in Tab.1.

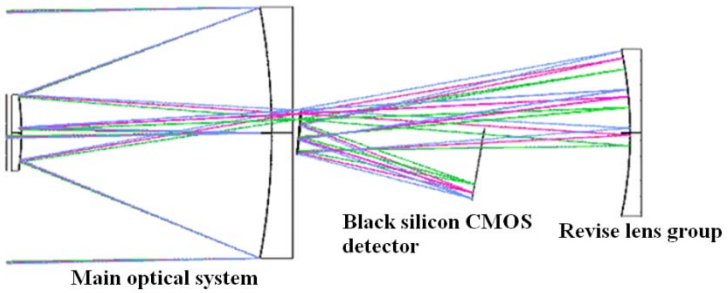


Fig. 12 The sketch map of space low-level-light imaging system

Tab. 1 Main parameters of optical system

View field	0.7°×0.7°
Clear aperture	1.8m
Focal length	15m
Band	0.4~1.2μm

The capability has been demonstrated and analyzed under different conditions of moonlight intensity. When the light intensity is more than 1/4 moonlight intensity, the imaging system based on black silicon detector has the same SNR with the system based on EMCCD, but the MTF can be improve about 10%, and the optical system can be simplified obviously. With the decrease of light intensity, the ratio of short wave infrared gradually increases, which makes the SNR of EMCCD detector reduces quickly and the payload cannot work. The SNR of black silicon detector is affected slightly, which shows the perfect performance. Expect the features listed above, the payload based on black silicon detector has some advantages.

(1) Integrated design of low-level-light payload

Because of the detector's restriction, the spectroscope is used to divide the optical system into two channels, low-level-light and high brightness, respectively. EMCCD is used in Low-level-light channel, and general CCD is used in high brightness channel. For assembling the different detectors and circuit board, the inverted telephoto optical system is used to guarantee the long enough back focal length, which makes the whole length be longer than the focal length.

The low-level-light imaging system based on black silicon detector owns the ability of weak light detection and non-weak light detection at the same time, which makes the payload be used in various fields. The design of single channel is used to reduce the design difficulty, and the dimension can be reduced by 30%.

(2) Reduce the thermal control difficulty

When the EMCCD is used, the detector needs to be cool so that the temperature can be kept at about -20℃. Under the room temperature, the black silicon detector works normally, which simplifies the thermal control system and reduce the thermal control difficulty. This feature also provides the benefits for the reduction of dimensions and weight.

With the development of black silicon, the latest research results show that the response spectrum can be extended to 1800nm [13], which can improve the performance of weak light detection payload.

VI. Conclusion

Hyperdoping black silicon detector shows the advantages in low-level-light imaging. We should develop the research as soon as possible, and pay more attention on the payload design. Foreign experiences can help us reduce the research risk and choice the technology route. We hope that the black silicon detector can be used in our space optical remote sensing in ten years.

References

[1] X. H. Hu, H. L. Zhang, et al. Design of Spaceborne Shimmer Tridimensional CCD Camera Imaging System [J]. Journal of Telemetry, Tracking and Command, 2014, 35(3) , pp. 43-48.

[2] M. L. Lian, S. T. Wang. Research on the Imaging Performance of Space Low Light Level Imaging System Based on ICCD [J]. Spacecraft Recovery & Remote Sensing, 2007, 28(3), pp. 6-10.

[3] Kay-Michael Guenther, Thomas Gimpel, et al. Excess carrier generation in femtosecond-laser processed sulfur doped silicon by means of sub-bandgap illumination[J]. Applied Physics Letters, 2014, 104, 042107.

[4] X. Li, J. E. Carey, et al. Silicon photodiodes with high photoconductive gain at room temperature[J]. Optics Express, 2012, 20(5): 5518.

[5] Aurore J. Said, Daniel Recht, et al. Extended infrared photoresponse and gain in chalcogensupersaturated silicon photodiodes[J]. Applied Physics Letters, 2011, 99, 073503.
[6] Zhihong Huang, James E. Carey, et al. Microstructured silicon photodetector[J]. Applied Physics Letters, 2006, 89, 033506.
[7] Jonathan P. Mailoa, Austin J. Akey, et al. Room-temperature sub-band gap optoelectronic response of hyperdoped silicon[J]. Nature Communications, 2014, 5:3011 | DOI: 10.1038.
[8] M. U. Pralle, J. E. Carey, et al. Infrared enhanced detection for laser imaging and biometrics[J]. Proc. of SPIE, 2013, 8734: 87340H.
[9] M. U. Pralle, J. E. Carey, et al. Black silicon enhanced photodetectors: a path to IR CMOS[J]. Proc. of SPIE, 2010, 7660: 76600N.
[10] M. U. Pralle, J. E. Carey, et al. IR CMOS: Infrared enhanced silicon imaging[J]. Proc. of SPIE, 2013, 8704: 870407.
[11] M. U. Pralle, J. E. Carey, et al. IR CMOS: ultrafast laser-enhanced silicon imaging[J]. Proc. of SPIE, 2012, 8353: 83533H.
[12] WANG Xi-yuan, HUANG Yong-Guang, et al. High Response in aTellurium-Supersaturated Silicon Photodiode[J]. Chinese Physics Letters, 2013, 30(3): 036101.
[13] J. P. Mailoa, A. J. Akey, et al. Room-temperature Sub-band Gap Optoelectronic Response of Hyperdoped Silicon[J]. Nature Communications, 2013, 5: 3011.

SATTELITE IMAGERY FOR MARINE OBJECTS IDENTIFICATION UNDER HIGHLY PERTURBED SEA

V.P. Kulagin[1] V.Ya. Tsvetkov[2] M.P. Lapchinskaya[3]

1. Doctor of technical science, professor in National research university "Higher school of economics", vkulagin@hse.ru;
2. Doctor of technical science, professor, rector advisor in Moscow institute of electronics and mathematics, cvj2@mail.ru;
3. PhD, associate professor in Moscow State University of Geodesy and Cartography, la.margo@miigaik.ru

Abstract.: The paper considers perturbed sea surface describing and analyzing method based on Earth's satellite images. Fuzzy sets method is the basis of the proposed method. The method allows obtaining models of perturbed sea surface that makes possible to identify sea objects from digital images. Surface model is superimposed to the image and then "subtracted" from the original one. On the rest, "clean" image, marine object identification is conducted.
Keywords: identification; modeling; perturbed sea surface; images; fuzzy sets

Introduction. In this paper, identification includes detection, parameters measurement and recognition of marine objects moving on perturbed sea surface is considered. It is widely used in vessels monitoring. Identification is one of the main tasks to solve by remote sensing and instruments mounted on the aviation and space platforms. There are different approaches to this problem solution. In [1] an approach based on construction of object environment model that is located on the sea surface is used. For the information processes analysis both at dynamic level (surface and air) of objects in the scene, and at objects' level that structures input information flows about occurring in the model process, mobile software agents are used. In [2,3,6] approaches of marine object identification together with its speed based on non-linear model built on radar images are described. In [4,5] algorithms of parameters estimation of dynamic model of sea objects that uses the extended Kalman filter are considered. Solution of marine objects recognition problems leads to necessity in new synthesis and probing signals processing methods design, at which for extended objects angle and amplitude of noise can serve as additional information sources that is used for classification and recognition. It should be mentioned that in [7, 8] necessity of space techniques' application to marine objects' management and identification is noted.

However, a common drawback of these approaches is that perturbed sea surface during the identification does not considered, and this is the subject of this paper.

Sea surface images models. Sea surface images are texture-homogeneous fields of the form $F(\Delta T,\Delta L_X,\Delta L_Y)$, where:

ΔT – variable characterizing non-derivative elements' optical density;

$(\Delta L_X,\Delta L_Y)$ – variables describing the spatial location of texture- homogeneous field for a particular underlying surface subclass of non-derivative elements'.

This set of variables describes geometric, statistical and physical properties of non-derivative texture elements and is an integral criterion of image's texture.

Experiments on texture homogeneous objects' recognition on the example of sea surface showed

that the main feature of texture images integral criterion is physical and linguistic uncertainty of direct visual deciphering features that appears as a variety of fuzzy properties. In particular, physical and linguistic uncertainties appear in the quantitative interpretation of lineal and area's size of non-derivative elements of possible situations on perturbed sea surface's image for subclasses such as "Ripple", "Ripple-on-the ripples", "Slick", "Foam". These images correspond to particular subclass of class "sea surface" with strongly marked structural features. Study of these images allowed to allocate fuzzy properties of deciphering criterion of image's texture "digital moiré" [10].

Consider informational situation [9] that will be called "Perturbation" and is characterized by a set of parameters such as the wavelength of the wave hill $\tilde{\lambda}$, the length of the wave hill's crest $\tilde{T}$, wave hill's square $\tilde{S}$, the wave hill's orientation $\tilde{\alpha}$.

These parameters in the same weather conditions do not have only one value and characterized by a certain spread. Parameters, that have uncertainties in the measurements, called fuzzy parameters. These parameters will be denoted with wavy line on the top: $\tilde{\lambda}; \tilde{T}; \tilde{S}; \tilde{\alpha}$.

Uncertainty leads to ambiguity in computer recognition of physical objects and phenomena on the underlying surface by using different texture features groups. When photographing, a real element on the sea surface will correspond to image's elements that will be called texture elements. These elements are divided into non-derivative, or basic elements, and derivative elements that are formed in terms of basic. Non-derivative elements play the information units role from which the model is formed. Fig. 1 shows a part of real situation on the sea surface, and the most important "perturbation" characteristics are marked.

The described method involves the use of computer-aided image recognition with characteristic features usage such as the size of non-derivative texture elements. Initial fuzzy wavelength and the length of the wave hill's crest parameters determine linear fuzzy size of non-derivative texture element p.

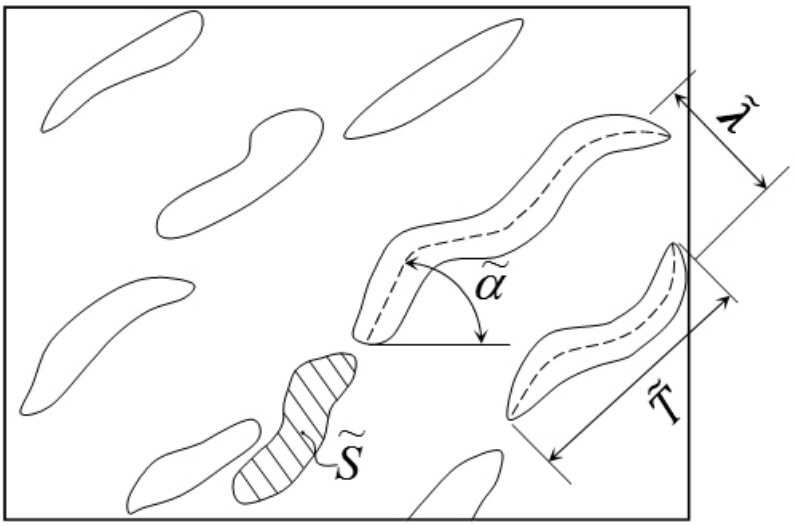


Fig.1. Fuzzy parameters of non-derivative texture image's element of the sea surface

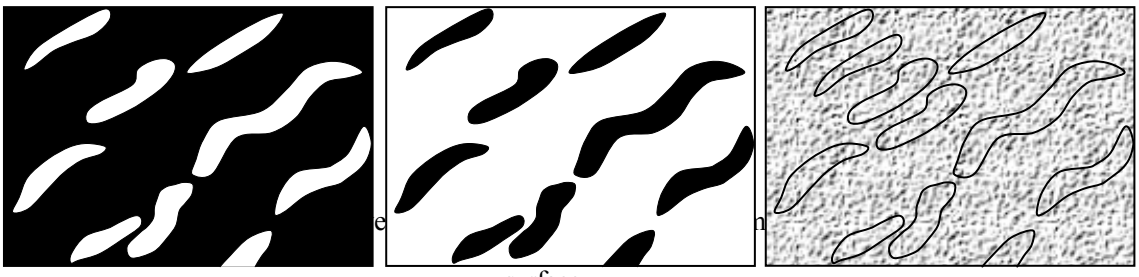
Approach to object identification using the moiré image. Perturbation on the sea surface, as it is fragmentary depicted on Fig. 1, refers to the moiré image. Moiré image is a picture obtained by two group's interaction of periodic structures in the image. In the formation of the moiré image of the sea surface, rotation angle α is uncertain. Information situation "perturbation" on the sea will correspond to information situation "digital moiré" on the image.

Various image contrasts also affects on a set of parameters formation for "digital moiré" situation calculated by the recognition program. In formation of the moiré images not only non-derivative texture

elements corresponding to physical objects involves but also some additional types of non-derivative texture elements arising from the specific imaging.

The first type of additional non-derivative elements is background corresponds to elementary physical objects. For instance, if an image part of the wave hills display in white (Fig.2), then background of non-derivative elements corresponds to negative wave hills display (Fig.2b).

The second type of additional non-derivative elements is image's graininess. For example, if non-derivative texture elements and, respectively, non-derivative background elements is displayed by a thin line, then derivative grain elements are superimposed on mentioned non-derivative elements and can be schematically shown in Fig.2c.



1) a) *3) b)* *5) c)*

Thus, three types of non-derivative elements are involved in moiré image formation that is the base for derivatives (dependent) elements construction:

the first type is non-derivative image's texture elements $\widetilde{p}_i$, corresponding to structural elements in the form of elementary physical objects of Earth's underlying surface class/ subclass (Fig.2a);

the second type is non-derivative background elements $\widetilde{p}_f$ from the non-derivative image's texture elements $\widetilde{p}_i$, in other words $\widetilde{p}_f = f(\widetilde{p}_i)$ (Fig.2b). Background elements occupy the main part of image and create significant interference in objects' identification that is comparable to the analyzed area size;

the third type is non-derivative graininess elements $\widetilde{p}_z$ (Fig.2c) that is the smallest in the image and creates interference in identification of small sea objects.

In accordance with the above mentioned for the informational situation "digital moiré", it is necessary to take into account all types of described non-derivative elements. According to the theory of fuzzy sets [11] these types of non-derivative elements ($\widetilde{p}_i$, $\widetilde{p}_f$, $\widetilde{p}_z$) could be written as a three functions of membership (Fig. 3). Where μ is a value of the membership function that is as probability lies in the ranges 0 to 1 and describes probabilities set for different values of fuzzy arguments. Highlighted in Fig. 3 areas, that are marked as "small", "medium" and "large", corresponds to ranges of non-derivative elements changes described above that are graininess, object (wave) and background, respectively.

Boundaries of "small", "medium" and "large" areas parameters L_1 and L_2 correspond to that the following physical meaning.

First, defining the boundary between graininess, image's background and object, these parameters define the most probable area of object identification. Besides this, parameters L_1 and L_2 define wave's size that makes possible to introduce the definition and criteria of "high perturbation". Under high perturbation we understand the following characteristic value ($\Delta = L_2 - L_1$), which is comparable in scale to the lateral dimensions of the identified object.

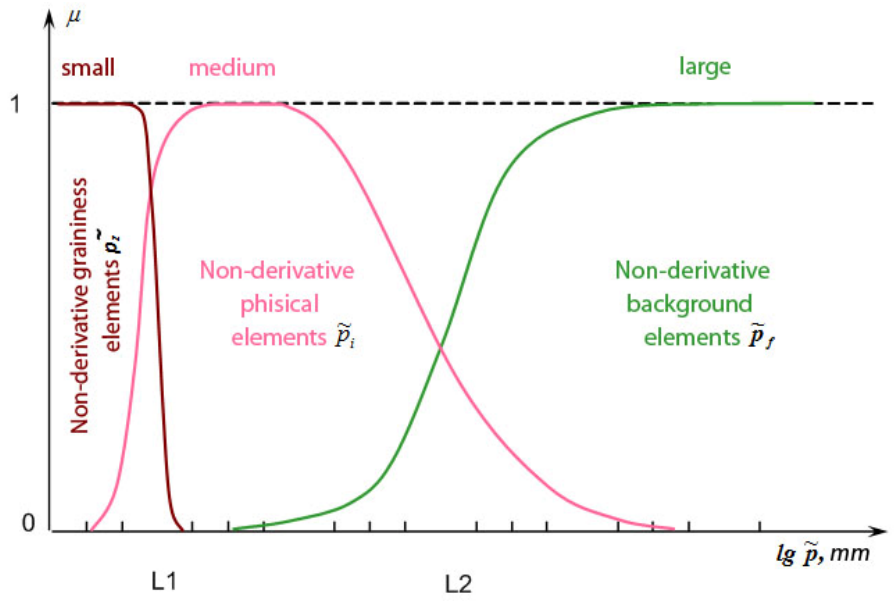


Fig. 3. Membership functions of non-derivative element p: $\tilde{p}_i$, $\tilde{p}_f$, $\tilde{p}_z$.

The automated calculation accuracy of "digital moiré" vector texture features parameters depends on quality of formed moiré image on a digital image fragment that depends on the area of non-derivative element $W = (\Delta L_X \times \Delta L_Y)$.

According to recognition algorithm of digital black-and-white images of the sea surface proposed in [12], the analyzed area may be equal to 256×256 pixels or more (Fig. 4a), and the digital image area that are currently being processed may be up to 64×64 pixels (fig.4b). This digital image processing method is similar to the well-known "moving average" method that means that at any given time, only one part of the image is processed, but the position of processing area varies according to a certain law. This approach significantly (an order) reduces the amount of processed information at a particular time, increases processing speed, and processing algorithm become independent of analyzed image's area.

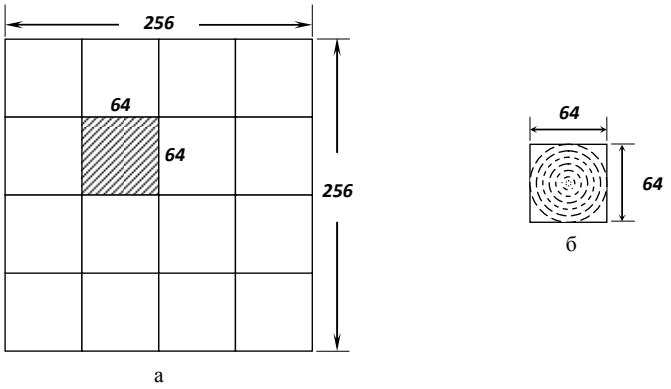


Fig. 4. Size of analyzed areas: digital image (a) and digital moiré image (b) in pixels

During the image's analysis, the image's scale determines the texture elements number and the texture element size given at area unit. Experiments have shown [10] (Table 1) that it is possible to form a textural object's features having only a digital image's fragment of non-derivative element's texture feature.

Thus, the automated analysis and image recognition quality depends on the number of non-derivative elements in the analyzed fragment of a digital image.

Taking into account specific features of fuzzy image's texture it become possible to apply statistical methods to generate a set of features for recognition. Based on the mentioned above, the proposed class "sea surface" is described by the three types of membership functions that characterize the graininess, background and object identification.

Table1. Experimental evaluation of recognition

Quantity of non-derivative elements in the area of digital image	Recognition quality
from 23 to 656	Moiré image is perfectly formed
from 1 to 23	Moiré image is poorly formed
from 0,48 (part) to1	Moiré image is badly formed
Less than 0, 48 (part)	Moiré image is not formed

Experiments on moiré images for various image subclasses of sea surface textures showed that among three types of non-derivative elements involved in the formation of moiré image, the highest type of visualization has a non-derivative image's texture elements due to the greatest contrast of its constituent elements. However, the contribution of background non-derivative elements $\widetilde{p}_f$ and graininess non-derivative elements $\widetilde{p}_z$ in formation of moiré image also exists, and it is significantly different for different images' subclasses of underlying surface that has been showed by the sea surface images.

Conclusions. The described approach, that benefit is in the ability of mathematical processing of textural features for which it is impossible to give a single quantitative assessment, allows passing from a qualitative analysis of perturbed sea surface to the quantitative and qualitative analysis. This approach is a part of the coordinate support of the Earth monitoring problem solution that further provides the basis for automatic recognition and objects identification on the perturbed sea surface. Based on the proposed approach it is possible to obtain the perturbed surface models that make it possible to identify the marine objects from the available images. This allows creating catalogues for perturbation of different sea state. In other words, from the available image the model of perturbed sea surface is "subtracted" that leads to the image "purification" and the more accurate identification of marine object.

It is expected that further development of this method will be the development of automated recognition methods of marine objects based on catalogs of the underlying surface.

REFERENCES

[1] Pavly`gin E`.D., Sosnin P.I. Mnogoagentnoe modelirovanie i vizualizatciia okruzhaiushchei` obstanovki morskogo sudna //Avtomatizatciia protcessov upravleniia. – 2010. – №. 2. – p.3-12.

[2] Skjetne R., Smogeli O. N., Fossen T. I. A Nonlinear Ship Manoeuvering Model: Identification and adaptive control with experiments for a model ship // Modeling, identification and control. – 2004. – V. 25. – №. 1. – p.3-27

[3] Gorobetc, V. N.; Gutneyk, V. G.; Zotov, S. M.; Kivva, F. V.; Shapiro, A. A. Matematicheskaia model` radiolokatcionnogo obraza korablia na morskom volnenii / V. N. Gorobetc [i dr.] // Radiofizika ta elektronika. - 2011. - Tom 2 (16), N 4. - S. 60-65.

[4] Fossen T. I., Sagatun S. I., Sørensen A. J. Identification of dynamically positioned ships //Control Engineering Practice. – 1996. – V. 4. , №. 3. – p.369-376.

[5] Chang S. J. Vessel identification and monitoring systems for maritime security //Security Technology, 2003. Proceedings. IEEE 37th Annual 2003 International Carnahan Conference on. – IEEE, 2003. –

p.66-70.
[6] Skjetne R., Smogeli O. N., Fossen T. I. A Nonlinear Ship Manoeuvering Model: Identification and adaptive control with experiments for a model ship //Modeling, identification and control. – 2004. – T. 25. – №. 1. – p.3-27.
[7] Buckley G. W. et al. Control system for a marine vessel : pat. 6273771 USA. – 2001.
[8] Clifford P. J., Hart N. R., Meulman C. B. Satellite system for vessel identification : pat. 7483672 USA. – 2009.
[9] Tsvetkov V. Ya. Information Situation and Information Position as a Management Tool // European Researcher, 2012, Vol.(36), № 12-1, p.2166- 2170.
[10] Lapchinskaya M. P. Algoritm classifikatcii odnorodny`kh tekstur zemnoi` poverkhnosti, osnovanny`i` na postroenii funktcii` prinadlezhnosti // Mezhdunarodny`i` nauchno-tekhnicheskii` i proizvodstvenny`i` zhurnal «NAUKI O ZEMLE». - №4-2012.- s.42-53
[11] Kofman A. Vvedenie v teoriiu nechetkikh mnozhestv. – M.: Radio i sviaz`, 1982, 432 s.
[12] Lapchinskaya M.P. Metod tcifrovogo muara. // Geodeziia i ae`rofotos``emka. - 1997. - №5. - s.86-113.

Information-analytical system of risk assessment and prevention of asteroid and comet hazard

Kulagin V.P.*, Shustov B.M.**, Kuznetsov Yu.M.*, Kaperko A.F.*, Bober S.A.*, Obolyaeva N.M.*, Naroenkov S.A.**, Shuvalov V.V.***, Svettsov V.V.***, Popova O.P.***, Glazachev D.O.***

*National research university "Higher school of economics",
**Institute of astronomy RAS,
***Institute of geosphere dynamics RAS

Abstract: The report discusses methods and software tools for the creation of information-analytical system (IAS) monitoring of dangerous celestial bodies and planning Asteroid-comet hazard (ACH). The article provides a description of the structure of the system and its functional components to enable rapid assessment of potential threats and forecast the effects of a collision dangerous space objects to the Earth. The result of the system is the integrated analytical information about the possible risks for the decision to reduce the possible damage and to identify effective measures for emergency management of cosmic origin.

I. Introduction

The role of international collaboration in scientific methods and technologies design for prevention of asteroid and comet hazard (ACH) and negative impact of space debris [1, 2] dramatically increases when population and territories protection from space origin emergency situation (ES) tasks are solved.

Such methods are based on data obtained from satellites and space monitoring stations. They are the basis for special tools creation for quick assessment of potential hazard and for effects prediction of dangerous space objects collision with the Earth. Their development allows a conceptual vision and complex integrated systems' architecture defining of decision-making informational and analytical support in the relation to expected emergency. This allows creating a preventive assessment of main directions of protection implementation and asteroid-comet hazard countering taking into account current regional infrastructure of area included in the risk zone.

Such a system design raises to the national security interests of many countries, especially those that have a large area where dangerous celestial body impact is most likely. Therefore, this problem solution requires international cooperation that allows providing more accurate information on emergencies of space origin risk's assessment.

Informational and analytical system (IAS) of hazardous celestial bodies monitoring for asteroid and comet hazard (hereafter - IAS ACH) preventing designed in Russia, uses technologies oriented to complex adjacent modeling support and statistics and monitoring data analysis that characterize various aspects and factors of ACH risks. It is an informational and analytical system of emergencies preventing and relieving the consequences integrated with already existing in Russia unified state system of emergencies prevention and liquidation.

II. Characteristics of asteroid and comet hazard

During IAS ACH design and development, specific features of asteroid and comet hazard should be considered, especially dangerous celestial body's size and trajectory's parameters. In addition, celestial body's basic parameters that in all cases affect to hazardous collision consequences are speed, angle of entry into the Earth's atmosphere, weight and type of the celestial body. Physical and chemical characteristics of the Earth's surface at the impact point also affect to hazardous consequences, for instance, collision with water basin could cause a tsunami, collision with hard surface could cause an earthquake, etc. It should be taken into account that even a collision with a small (on space standards) asteroid or comet part could lead to a huge loss.

Studies held within the framework of the United States program for asteroids search (http://impact.arc.nasa.gov) showed that only in respect of 4000 potentially dangerous bodies with sizes from 50 to 140 meters astronomers have sufficiently detailed information. Discovery of such celestial bodies are continuous, but there only a few percents of trajectories are known. Asteroid and comet hazard specific feature is that the frequency of immediate hazard detection will increase with a decrease in the average size of newly discovered potentially hazardous objects (asteroids and comets).

According to reports [3], asteroid with a size of 10 meters across that entered to the atmosphere at a speed of 20 km/s can cause the energy release of the order of 50 thousand tons of TNT. Therefore, increase in a number of cases when the Earth's trajectory intersects with large asteroids trajectories becomes a real danger. For this international cooperation becomes especially important to combine information about dangerous celestial bodies.

Due to the fact that asteroids and comets are the most chaotic Solar system's objects, one of the most difficult problem is their trajectories computation [4], collision with the Earth probability assessment and, as a consequence, potential hazard assessment. This problem get worse because of a lack of characteristics' information about near-Earth celestial bodies motion with a complex multi-parameter indicators of dynamic state and location in a given space region.

The main problem is to determine dangerous celestial body's trajectory, as a permanent change in their trajectories due to planets and Sun gravitational forces is typical for celestial bodies. In general, motion dynamics of such objects is described by random functions of time. However, there is a certain determinism in their motion's trajectories describing that allows developing of such objects behavior model and generating of predictive indicators for potential risks analysis and consequences assessment of their approach to the Earth.

III. Research Status

Evaluation of publication and patent activity in the field of hazardous celestial body monitoring and asteroid and comet hazard counteraction indicates that currently the most attention is given to researches related to the development of new methods and specialized software. In particular, it is primarily related to

- software for dangerous celestial bodies' trajectory calculation,
- software for celestial bodies' motion calculation and their behavior in the atmosphere,
- software for collision with the Earth's surface consequences calculation,
- software with geographic information systems usage.

Analysis of existing technical solutions indicates the following trends:

a) increase in calculations and visualizations accuracy,
b) earlier identification of hazardous space objects,

c) integration into multifunctional systems,
d) ability to work with large amount of data,
e) catalogues creation of hazardous space objects.

These trends, in turn, determine the following characteristics of the studies.

1. Focus on as earliest as possible dangerous space body determining (relatively to collision probability) is typical for ACH preventing methods.

2. Objects' trajectories correction is associated with numerical methods usage that allows considering physical laws, evaluate collision risk on predetermined time intervals, and perform calculations based on various celestial bodies' fluctuations. Software design in this class covers the following areas

- closest approach identification between controlled spacecraft and space debris to assess collision risk;
- satellites' motion models allowing trajectories design and calculations performance on this trajectories;
- trajectories' evolution study of real and virtual asteroids;
- trajectories' and the Earth with an objects trajectories intersection conditions calculation, and images creation on obtained data.

3. Calculation of celestial bodies' motion in the atmosphere is based on dynamic models usage and is associated with the large number of parameters when calculating actual trajectory and ballistic characteristics of asteroid (comets) with Earth collision. The key issue here is that all IAS ACH models are dynamic and allows interactive user influence on model's parameters and behavior at any stage of analysis and decision-making.

4. For GIS usage, it is important to provide graphical interface, and possibility to manage thematic layers for dynamical systems study, including systems of celestial mechanics. Similarly, imaging systems is characterized by increasing models complexity for which imaging is carried out, as well as increase in parameters that are available for user correction.

In this regard, the high priority task of international cooperation is to unite efforts in tools design allowing to provide automated collection of coordinate and non-coordinate information on near-Earth asteroids and dangerous celestial bodies cataloging, especially for those that trajectories lies on a collision trajectory with the Earth.

On the other hand, it is necessary to develop predictive methods for ACH effects assessment based on numerical mathematical modeling of processes accompanying celestial bodies with the Earth collision for land and for water, as well as the processes that take place when body decline in the atmosphere. Modeling is an effective and, in practice, the only research tool that allows to study the dynamics of relatively large celestial bodies with the Earth's atmosphere interaction.

IV. Simulation of asteroid

Observe the impact of large body is difficult due to the events rarity. Reproduce them in the laboratory is not possible due to event scale. Developments that is used for IAS ACH allow to depict demonstration models of preliminary damaging factors assessment for an accidents related to asteroid and comet hazard that based on calculations results and similarity principles.

Evaluation of celestial bodies interaction with the Earth's atmosphere and surface is usually carried out by impact processes modeling on the basis of calculations of all the main phases of the celestial body impact to the planet surface that includes phase of contact with shock waves

formation, shock wave distribution in the target, crater forming flow and shock emissions and their distribution in the atmosphere.

Figure 1 depicts sequential steps of 100 m in diameter asteroid's interaction with the atmosphere, descending on the Earth at a speed of 50 km/s. The figure shows

- temperature (left) and density (right) distribution on different altitudes;
- vertical axes represent the distance along trajectory;
- the black line on pictures is the boundary between vapor and air.

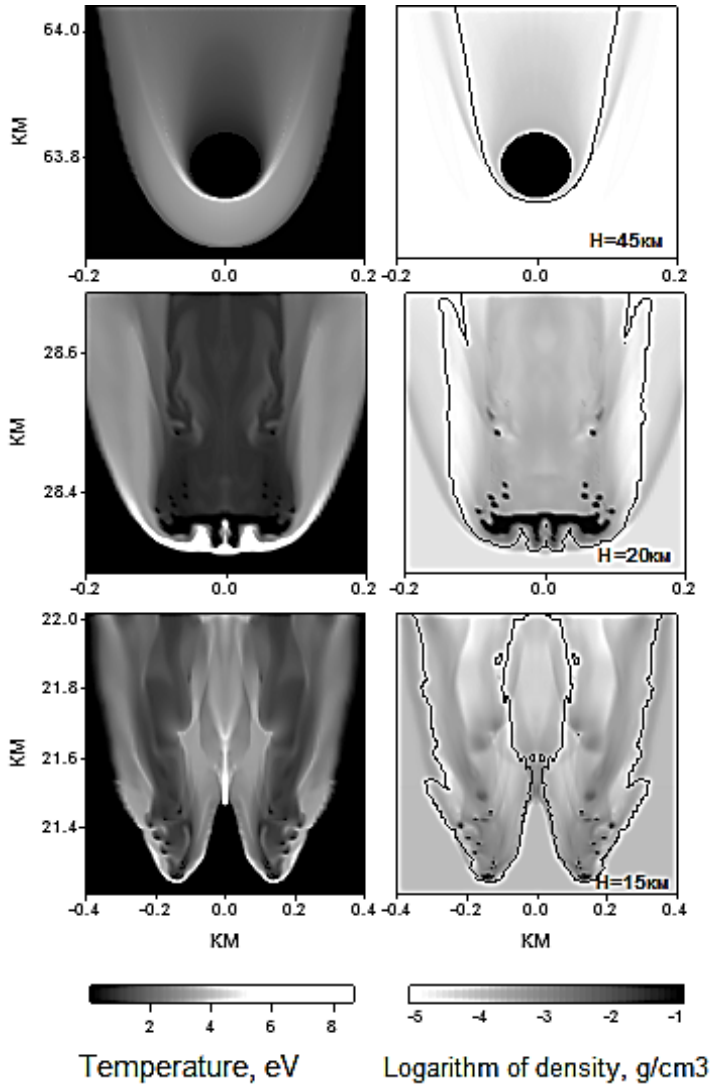


Figure 1 – Destruction of a comet-like meteoroid

A typical sequence of steps is as follows. At an altitude of about 30 km meteoroid begins to deform, on its surface wavy perturbations appears, the central part fills by vapor. Increase in aerodynamic load causes meteoroid flattening, at an altitude of 20 km it turns into a flattened structure.

At altitudes below 17 km, it is transformed into a jet consisting of a heated vapor at a shock airwave head and fragments of descending body. Thus, the destruction and fragmentation of the descending body occur before slowdown. Fragmentation leads to increase of the evaporation surface and, consequently, to increase ablation rate.

At an altitude of about 10 km meteoroid fragments completely vaporized, the jet turns into a pure gaseous. The heated air in the shock wave is under high pressure, so the jet expands. In the slowdown moment (at an altitude of about 2 km), the gas density in the jet is considerably lower than the ambient air density that leads to a phenomenon similar to the typical atmospheric explosions fireball.

At the time of meteor's full slowdown the bulk meteoroid material (in vapor phase, mixed with hot air) is concentrated at altitudes of 3-10 km. Next, all this material gets in the atmospheric plume and ejected into the upper atmosphere at an altitude of 100-1000 km [5].

For the preliminary calculation, it is convenient to use the principle of retrograde motion if celestial body and the Earth collision point is known. This allows restoring body's trajectory in interplanetary space if parameters of celestial body and the Earth collision are known (speed, trajectory's inclination, altitude, latitude and longitude of the entry point into the atmosphere).

The next step is to create a quick method of main adverse factors and possible damage assessment, based on descending body parameters, on which basis interpolation model can be built that allows quickly (within minutes) assess the full impact consequences.

V. Structure IAS ACH

Practical implementation of information and analytical system of dangerous celestial bodies monitoring is based on usage of advanced mathematical methods of information processing and high performance computing, and includes the following tasks:

- methods and models design of celestial body's motion to assess the possible interaction between asteroid and the Earth's atmosphere and surface;
- providing automated collection of coordinate and non-coordinate information on hazardous celestial bodies;
- software and tools design to calculate the trajectory of celestial bodies' motion, to estimate the probability of hazardous asteroids rendezvous and impact to the Earth, to determine impact point and behavior after descending to the atmosphere;
- data bank creation on consequences of celestial body impact on a certain area;
- visualization of computational and modeling-based information on the basis of geographic information systems that also provides creation of interactive electronic atlas of geographic information reference for emergencies prevention and localization tasks;
- models' integrated database design for all hazardous factors of such collisions with target celestial body.

Structural informational and analytical system of asteroid and comet hazard used in Russia includes the following components:

1) forecasting and analytical subsystem consisting of the following functional modules (software and algorithmic blocks):
   - software for study hazardous celestial bodies trajectories evolution that allows identifying dangerous trajectory of celestial bodies over long time intervals;
   - module of hazardous bodies trajectories determination;
   - module of body's descent trajectory determination that allows defining the impact point or air explosion point;

- module of probability assessment of hazardous bodies with the Earth collisions and rendezvous;
- software module of impact consequences assessment that allows estimating the level of a dangerous object on the terrain and infrastructure effect.

2)automated information collection subsystem on near-Earth asteroids, including
- module of coordinate information collection;
- module of non-coordinate information collection;
- module of input data processing;

3)data visualization subsystem, including
- visualization module of changes in dangerous bodies trajectories;
- module of reporting forms;
- module of digital maps preparation with territories that correspond to Russia territories classifier;
- module of calculation results visualization and data publication on the Internet.

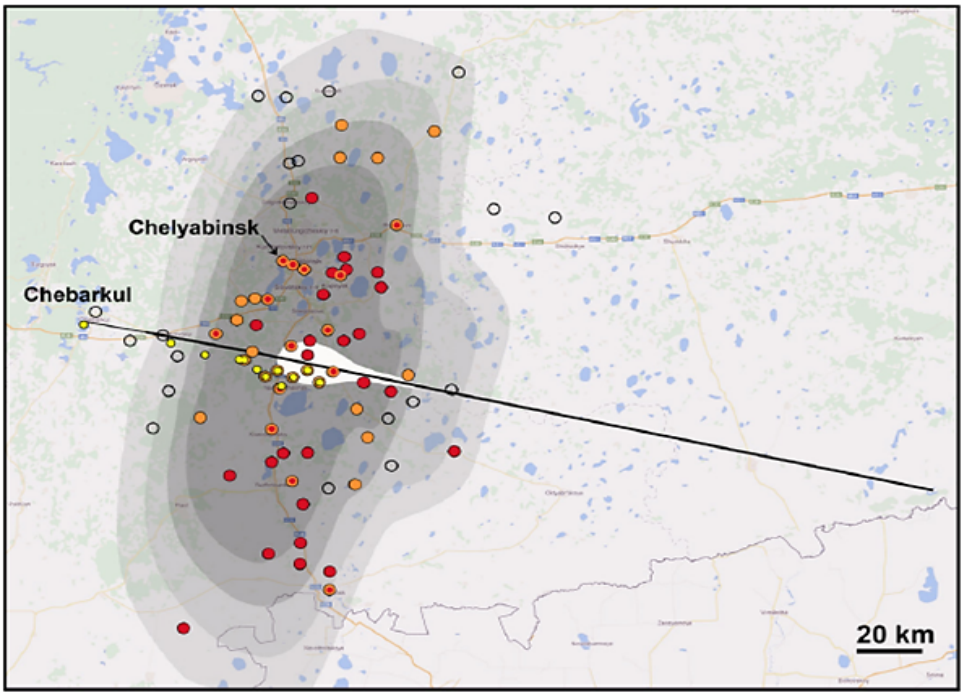
4)data stores, including
- database of potentially hazardous celestial bodies;
- data banks of possible celestial bodies impact consequences on certain territory;
- data banks of celestial bodies with the atmosphere or the Earth's surface interaction consequences;
- industry data storages reflecting infrastructure of territories security organization including ACH, and the types of hazarf;

5)administration subsystem that provides database maintenance operations performance, user accounts management, network infrastructure management, and data processing.

During its operation, IAS ACH performs continuous, regulated initial data collection on the current situation, using international observing sources on hazardous celestial objects. In information preparation, multi-criteria analysis and conditions assessment in certain indicators are performed. In automatic mode, operations on identification of near-Earth asteroids close rendezvous with the Earth are performed, probability of hazardous body collision with the Earth and place of a possible impact are defined. In the process of intelligent decision-making support by ACH IAS, consequences' assessment of space body impact to the Earth is held and possible options for impact consequences minimizing are formed.

VI. Research results

As an example of IAS ACH effectiveness, Figure 2 shows results of mathematical modeling of Chelyabinsk bolide motion superimposed on the actual consequences map.



Grey contours – over-pressure on the surface
Red – destructions (official data)
Orange – destructions, field data
Dark unshaded – no destructions
Yellow – meteorites

Figure 2 – Results of mathematical modeling of Chelyabinsk bolide motion

Figure 3 shows the results of virtual objects' trajectories calculation (asteroid's "copies") those details come from asteroid's data storage and processing.

(a) virtual object's calculated trajectory for Chelyabinsk asteroid in the coordinate system with a fixed Sun-Earth axis;

(b) virtual object's trajectory for Chelyabinsk meteorite on the Earth map.

Calculation time is 14 February 2013 23: 58: 52,815. Entry into the atmosphere time at an altitude of about 100 km is 15 February 2013 3:20:19.

Calculations were performed by designed software for data transformation, obtained from the storage and data processing on asteroids to the GMAT software script that allows to integrate asteroids motion equations taking into account all planets in the Solar system and the Moon, and to visualize the resulting trajectories.

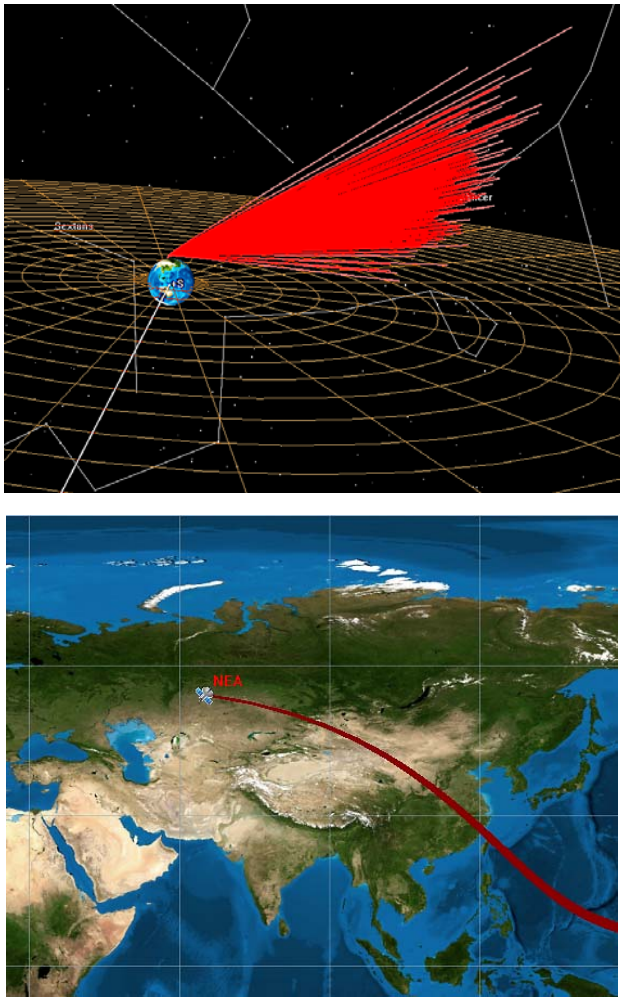


Figure 3 – Virtual objects' trajectories visualization (a) and their trajectories on the Earth's surface maps (b)

VII. Conclusion

For the IAS ACH development and international cooperation in this field there should be further software, effective models and custom applications development that involve working with large data flows and databases of astronomical observations obtained from satellites and ground-based monitoring stations. This is of particular interest regarding saturation of near-Earth space by debris [6] that has a more significant hazard.

In general, such a system allow to obtain new, more comprehensive information for quick assessment of cosmic origin emergencies' possible risks, consequences' prediction, effective measures identification to reduce potential damage.

References

[1] Barmin I.V., Kulagin V.P., V. P. Savinykh V.P., Tsvetkov V.Ya. Near-Earth Space as an Object of Global Monitoring // Solar System Research, 2014, Vol. 48, No. 7, pp. 531–535.

[2] Shustov B.M., Ryihlova L.V. i dr. Kontseptsiya sistemyi protivodeystviya kosmicheskim ugrozam: astronomicheskie aspektyi// Astronomicheskiy vestnik, 2013. - T.47, #4, s.327-340.

[3] Popova, O. P., P. Jenniskens, et al. (2013). Chelyabinsk Airburst, Damage Assessment, Meteorite Recovery, and Characterization. Science 342(6162): 1069-1073.

[4] Dunham D.W., Nazirov R.R., Farquhar R.W., Chumachenko E.N., Eismont N.A., Simonov A.V., Kosmicheskie missii i planetarnaya zaschita. – M.: FIZMATLIT.2013.276s. – ISBN 978-5-9221-1495-0.

[5] Shuvalov V.V., Artemieva N.A. Numerical modeling of Tunguska-like impacts // Planetary and Space Science. 2002. V. 50/2. P. 181-192.

[6] Barmin I.V., Dunham D.W., Kulagin V.P., V. P. Savinykh V.P., Tsvetkov V.Ya. Rings of Debris in Near-Earth Space // Solar System Research, 2014, Vol. 48, No. 7, pp. 592–599.

ADLAND – an Advanced Land Observation InfraRed Mission

Oertel, D., Hartmann M., Huang-Wolff, Y., Schierle, C., Scheiding, M.

Astro- und Feinwerktechnik Adlershof GmbH, Berlin Germany

Ruecker, G.

ZEBRIS Jürgen Brendel und Gernot Rücker GbR, Munich Germany

ABSTRACT: Current space-borne thermal infrared satellite systems aimed at land surface remote sensing retain some significant deficiencies, in particular in terms of spatial resolution, spectral coverage, number of imaging bands and temperature emissivity separation. The proposed Advanced Land Observation InfraRed Mission (ADLAND) addresses many of these limitations, providing multi-spectral imaging data with medium-to-high spatial resolution (60m Ground Sampling Distance /GSD/ from 600 km orbit altitude, or 80m GSD from 800km altitude) in the thermal infrared /TIR/ (up to 6 bands, between 8 and 11μm) and in the mid infrared /MIR/ (1 or 2 bands, at 3.5 - 4μm). These MIR/TIR bands will be co-registered with simultaneously acquired high spatial resolution (with less than 30 m GSD) visible and near infrared multi-spectral imaging data. To enhance the spatial resolution of the MIR/TIR multi-spectral imagery during daytime, data fusion methods will be applied, such as the Multi-sensor Multi-resolution Technique (MMT), already successfully tested over agricultural terrain. This new image processing technique will enable to deliver Land Surface Temperature (LST) Earth observation products with a spatial resolution of 30 x 30 m² for day time data records. ADLAND data shall permit to retrieve for High Temperature Events (HTE) such as vegetation- and peat-fires, coal seam fires, gas flares, and volcanic activity the Essential Climate Variable (ECV) T13 Fire Disturbance indicating the "Active fire location" and the "Fire Radiative Power" with less than 100 m spatial resolution. Together with the effective fire temperature and the spatial extend even for small fire events the innovative system characteristics of ADLAND go beyond existing and planned IR mission.

I. ADLAND HERITAGE

There is a longstanding German heritage in infrared (IR) land observation with push-broom sensor systems on board of micro-satellites.

The Bi-spectral InfraRed Detection (BIRD) satellite – developed by the German Aerospace Center (DLR) in Berlin Adlershof, and piggy back launched from India with a Polar Satellite Launch Vehicle (PSLV) in October 2001, demonstrated the detection and quantitative analysis of even small high-temperature events (HTE), such as wildfires, coal seam fires, and volcanic activities, more than ten years ago [1].

BIRD monitored coal seam fires in the Ningxia and Xinjiang areas in northern China in the years 2002 -2004

[2], contributing to joint Chinese-German project.

Figure 1 shows an example of coal seam fire detection in the day- and night-time BIRD images of the Rujigou coalfield that was verified by an on-ground inspection in September 2002. The daytime Mid InfraRed (MIR, at 3.8μm) image (Fig. 1, top, left) shows a few detected hot spots associated with stronger coal seam fires mapped on ground by in-situ inspections within a few days of this BIRD data take (Fig. 1, bottom, left). The night time BIRD MIR image data from January 2003, (Fig.1, top, right), allowed the recognition of most of the weaker coal fires in the scene. The detected at night time hot spots are overlaid on in-situ fire location maps, (Fig 1, bottom, right).

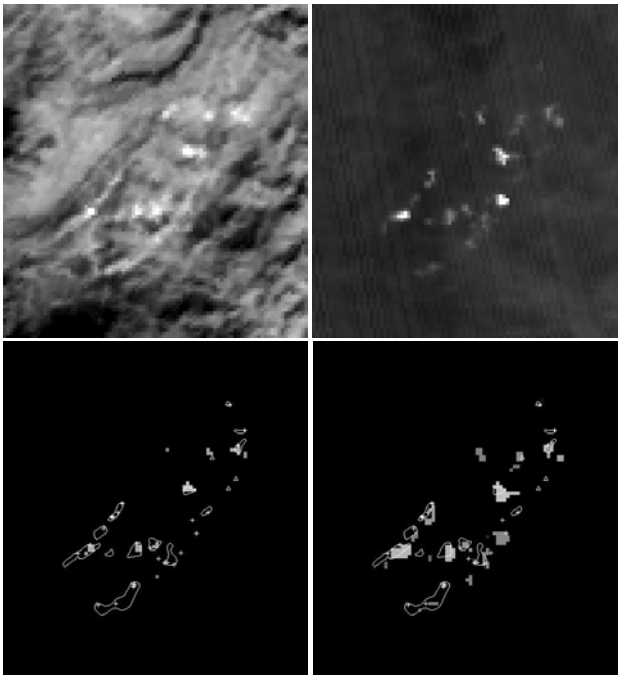


Figure 1. Coal seam fires detection in BIRD images and comparison with ground truth. Top, left: Day time MIR image from 21 of September 2002; Top, right: Night time MIR image obtained on 16 of January 2003; Bottom, left: Detected hot spots at day time overlaid on in-situ fire location maps; Bottom right: Detected hot spots at night time overlaid on in-situ fire location maps. White contours and crosses indicate location of coal seam fires verified by in-situ inspection in September 2002, while white triangles show the location of industrial chimneys [2].

BIRD was followed by the FireBIRD mission of the German Aerospace Centre (DLR), which is based on a BIRD type sensor. The first of two FireBIRD sensors was piggy back launched into a sun-synchronous orbit on 22 July 2012 on board of the German small satellite TET-1, and is regularly observing fires since autumn, 2013. TET-1 will be followed by a second satellite with similar specifications, BIROS, to complete the FireBIRD mini-constellation. FireBIRD provides FRP maps of actively burning fires derived at a spatial resolution of 250 m - together with the estimation of effective fire temperature and spatial extend even for small fire events many of which cannot be detected and/or characterized by existing and planned coarse resolution IR sensor systems [2].

The co-registered MIR, TIR and NIR imagery of BIRD, TET-1 and BIROS – all with a spatial resolution of 200-300 m – was, are and will be further used for the derivation of Fire Radiative Power (FRP), which correlates to combustion rate [3],[4].

Since BIRD ceased to take regular imagery (in 2004), several studies on future dedicated IR land observation studies have been conducted for ESA, for instance, a Study on Scientific Assessment of Space-borne High Temperature Event Observing Mission Concepts (ECOFIRE) [5], and the Dual-use European Security IR Experiment 2008 (DESIREX 2008) [6].

FireBIRD FRP is used to estimate combustion rate, and in conjunction with data from medium resolution sensors such as MODIS, to estimate greenhouse gas emissions from fires. Figure 2 shows a TET-1 FRP map of fires in a protected area in central Brazil.

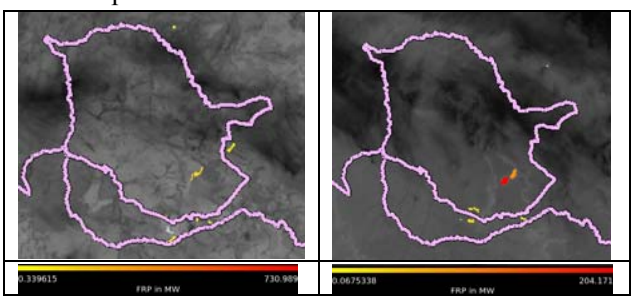


Figure 2. FireBIRD FRP map of fires in a protected area in central Brazil (purple polygon) obtained by TET-1 on 2014/10/17 at local noon (left) ad near midnight (right). The FRP value of the fire clusters - in MegaWatt (MW) – is colour coded – see scale below

.

II. LAND SURFACE TEMPERATURE

The FireBIRD sensors also provide IR imagery of the ambient Earth surface. However, the MIR- and TIR- imagery data from FireBIRD is only "bi-spectral", i.e. with one band centred at 3.7 and one at 9.2 µm, which is not optimal for land surface temperature estimation.

Medium-to-high spatial resolution multi-spectral space-borne images obtained by the Advanced Spaceborne Thermal Emission Reflection Radiometer (ASTER) on "Terra" satellite (launched in 1999) in the thermal infrared region of the electromagnetic spectrum is required to derive the Land Surface Temperature (LST). This allowed significantly to improve the knowledge on the carbon cycle, energy and matter fluxes of land surfaces, wildfire emissions and volcanic outbreakes as well as human activities.

LST is important for understanding energy fluxes between earth's surface and the atmosphere from the local to the global scale [7],[8],[9],[10]. With recent climate change the averaged radiation flux intensity has reached a value of 1.72 W·m-2 ("radiative forcing") and has thereby increased the upwelling infrared longwave radiation [11]. An accurate determination of the LST is necessary to derive information about space-time variations of the earth's radiation balance and for an improved understanding of the factors influencing climate change.

Acquiring more precise LST data is also crucial for environmental research [12]:

- Monitoring agricultural or forest areas and natural vegetation,
- Estimating evapotranspiration of plants,
- Monitoring urban heat islands,
- Detecting precursors of volcanic eruptions.

In addition, the data of the MIR/TIR bands can be applied to determining the mass balance of glaciers by monitoring the Ice Surface temperature (IST) or by deriving the Sea Surface Temperature (SST), to investigate the complex ocean currents to identify temperature changes and shifting flow patterns [13].

However, limitations of the spatial, temporal and spectral (especially for MIR/TIR bands) resolution of current satellite sensors remain a major challenge for obtaining sufficient, accurate physical data. Depending on retrieval methods, the estimated error range of LST products is between 0.5 Kelvin (K) and several K. If the dynamics of land surface emissivity (LSE) are not considered, the uncertainty increases significantly and can reach up to 6 K [14]. To include the dynamic LSE, NDVI-based emissivity methods or Temperature Emissivity Separation (TES) methods can be used [15], [16]. Especially for land surfaces, it is therefore necessary to advance the development of satellite-based observation techniques of LST, with improved spatial resolution, multi-band data registration in the TIR, and a corresponding repetition rate, to suitably reproduce space-time dynamics of land surface phenomena.

III. INFRARED OBSERVATION GAPS

Land applications were discussed at the ESA Sentinel Convoy for Land Applications workshop held at ESA ESTEC in November 2011 [17] with regard to mid- and thermal infrared observations for major land applications (see table 1).

Table 1: Science and operational gaps in mid- and thermal infrared observation for main land applications

Application	Earth observation variable(s)	IR observation gap
Carbon Cycle - fire sensing	Fire Radiative Power (FRP)	~250m resolution MIR/TIR - to sample smaller fires
Surface energy balance	Land Surface Temperature (LST)	< 100m resolution multi-spectral TIR, to account for spectral variations in emissivity
Water cycle	LST, Soil moisture	
Human population dynamics	Urban LST, urban emissivity	
Volcanoes	Thermal anomalies, pre-eruptive precursors & SO2 degassing plumes	< 60m resolution multi-spectral MIR/TIR
	Lava thermal characteristics and mapping	< 30m resolution multi-spectral SWIR/MIR/TIR

The Executive Summary and Final Report of the "1st International Earth Observation Convoy and Constellation Workshop" – conducted at ESA ESTEC in October 2013 - highlighted several specific convoy concepts. One of these highlighted convoy concepts is [18]:
"Higher resolution thermal infrared < 60 -250 m".

The following important aspects for this specific convoy mission concept are outlined in the Final Report of this workshop [18]:

- Measurements in the thermal infrared spectral range are needed for numerous Earth Science domains.

- Current in-orbit TIR instruments (e.g. Landsat-8-TIRS, MODIS, Terra-ASTER etc.) all have limited lifetimes and if replacements are not launched then there will be a strong possibility of a data gap later this decade.
- Multiple TIR channels are needed to separate emissivity and temperature.
- Mid infrared and thermal infrared are needed to measure fires and thermal hot spots (with additional visible channels for context).
- To measure thermal anomalies a high dynamic range is required to mitigate detector saturation.
- There is cross cutting science between land and atmospheric sounding communities regarding infrared imagery.
- Flying a thermal infrared imager with Sentinel-2 was highlighted.

IV. APPLICATION OF DATA FUSION

In [19] is described a good example of Land Surface Temperature (LST) retrieval based on Airborne Hyper-spectral Scanner (AHS) records, conducted over Barrax in Spain. Figure 3 shows the LST and the corresponding plant evapotranspiration (ET).

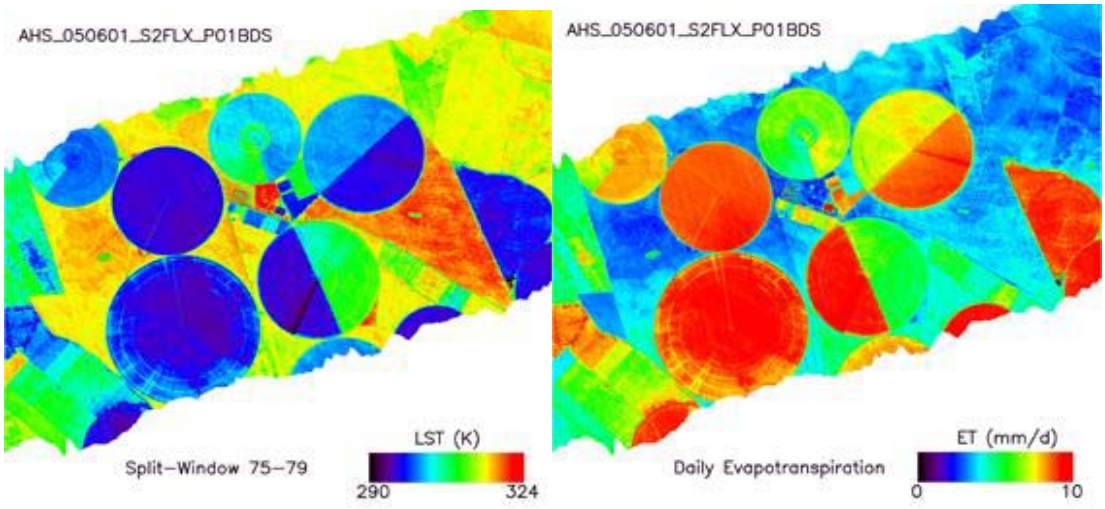


Figure 3. Top: Land SurfaceTemperature (LST), measured over Barrax (Spain) with the Airborne Hyperspectral Sensor (AHS), Bottom: Evapotranspiration (ET) of plants derived from LST [19].

Figure 3 shows false colour images as obtained from an aircraft – with a spatial resolution of better than 10 m.

It is a challenge for ADLAND, to deliver thermal images with a spatial resolution appropriate to the agriculture field scale (which is typically 30-50m).

For instrumental reasons, MIR and TIR imagers usually operate at coarser spatial resolution than do visible (VIS) and near infrared (NIR) band sensors on the same satellite platform, especially in case of a small satellites.

For space-borne LST observations, to be conducted by the multi-spectral thermal imager of ADLAND with 60-80m spatial resolution, data fusion methods shall be implemented to enhance the spatial resolution up to 30m. This kind of "thermal sharpening" may use, fortunately, an inverse relationship between LST and Vegetation Indices (VI's). One such possible data fusion method, the Multi-sensor Multi-resolution Technique (MMT), is well developed and tested over agricultural areas for multi- and hyper-spectral imaging data co-registered at day time with different spatial

resolutions [20]. Figure 4 illustrates the MMT Principal scheme, where the following elements are shown:

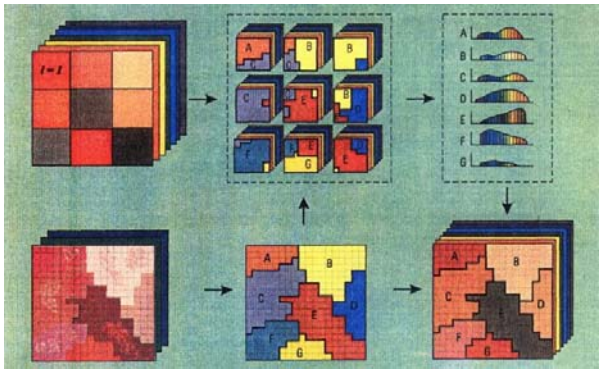


Figure 4. Principal processing scheme of the Multi-sensor Multi-resolution Technique (MMT) [20].

(1) a multi-spectral scene fragment from the spatial higher resolution VIS/NIR camera (the "classifying instrument") - in the lower row on left hand site,

(2) the classification map obtained from the multi-spectral VIS/NIR data - in the middle of the lower row, (3) a hyper-spectral or multi-spectral scene fragment from the spatial lower resolution sensor (the "measuring" instrument) - in the upper row on left hand site,

(4) an overlay of the classification result on the coarser resolution scene pixels of the measuring instrument - in the middle of the upper row,

(5) the classification result for each class (A-G) in the hyper-spectral domain – in the right hand site of the upper row, and

(5) the unmixed, or "fused", hyper-spectral (or multi-spectral) scene fragment with high spatial resolution borders between the classes – in the right hand site of the lower row.

V. ADLAND's GOALS AND ITS MAIN OBSERVATION FEATURES

The Advanced Land Observation Infrared Mission ADLAND addresses all gaps summarised in table 1. Further, ADLAND considers the above mentioned aspects for the highlighted in [18] convoy mission concept for "Higher resolution thermal infrared < 60 -250m". The ADLAND mission targets to provide:

a) Higher spatial resolution (with 60m ground sampling distance /GSD/ from 600 km altitude, or 80m GSD from 800km altitude) multi-spectral imaging TIR data (up to 6 bands between 8 and 12 µm) & imaging MIR data (1 or 2 bands at 3.5 - 4µm), and supplementary,

b) High spatial resolution (20 - 30 m) multi-spectral imaging data with two bands in the visible (VIS, at 0.5 and 0.6µm) and one band in the near infrared (NIR, at 0.86µm).

The main purpose of the VIS-NIR data is to provide a high resolution (~30m) classification map necessary for the application of the Multi-sensor Multi-resolution Technique (MMT) to the MIR/TIR data records of ADLAND acquired at day time. The application of the MMT to the ADLAND data sets requires the knowledge of the optical co-alignment of the instrument bands with an accuracy <0.2 of the GSD of the coarser resolution ("measuring") instrument [20].

VI. EXPECTED ADLAND DATA PRODUCTS

Data records from ADLAND will allow to retrieve spatially high to medium resolved Earth Observation (EO) imaging data products, including physical variables and Essential Climate Variables (ECV`s), such as:

A) Physical variables for normal temperature phenomena (NTP)

ADLAND data shall allow to derive the following physical variables for NTP:

Land Surface Temperature (LST), including urban surface temperature, Land Surface Emissivity, including urban emissivity, and Soil Moisture (SM).

LST and SM are strongly related to hydrology and land surface energy fluxes as well as to vegetation dynamics, including plant evapotranspiration, water stress and irrigation needs relevant for food production in water-scarce agricultural regions and sustainable management of water resources and, generally, the consumptive use of global freshwater.

B) Physical variables for high temperature events (HTE)

The Essential Climate Variable ECV T13 "Fire Disturbance" comprises: Burned Area, Active Fire (location) and Fire Radiative Power (FRP), which can be used for fire emission assessments.

ADLAND will be able to deliver data products on the location of active fires and their FRP for:

All vegetation and peat fires, including small ones with a flaming burning area from 10 m2 upward;

Gas flares of the oil industry, and

Most of the coal seam fires which burn partly underground and exhibit only part of its heat for remote sensing.

Additionally to these ECV T13 Fire Disturbance related data products the following additional fire attributes will be obtained from the ADLAND data: Effective fire temperature, effective fire area, fire line length, and fire line radiative strength, which all are key parameters for fire management and fire ecology [2], [3], [4], [5].

To predict and mitigate earthquake and volcanic hazards through detection of transient thermal phenomena, the following EO variables will be derived from ADLAND data: volcanic pre-eruptive thermal anomalies and thermal anomalies on fumarole fields, thermal characteristics of lava flows, such as: crust and lava temperature, and volcanic ash emissions.

VII. ADLAND's INNOVATIVE INSTRUMENT DESIGN FEATURES

ADLAND has a technically very innovative sensor and micro-satellite bus design concept, characterized by the following features:

Compact and flexible multi-spectral MIR/TIR- and VIS-NIR-sensor heads - both consisting of one entrance optics, one focal plane, with stripped spectral filters in front of one matrix detector array, each. Both sensors can be operated in push broom or in push-whisk broom modes.

Commercially available and space environment tested detector arrays are planned to be used. For the MIR/TIR sensor a Mega Pixel Integrated Detector Cooler Assembly /IDCA/ is considered.

Registration of multiple overlapping image fragments in all (up to 8) bands shall be implemented with at least two different integration times. Spectral band related digital binning of the data will be applied to improve the Signal to Noise Ratio (SNR). The use of minimum two different integration times is necessary to cover the large radiometric dynamic range in case of fire or active volcano observation, because these high temperature events may occupy up to 50% of a 60 x 60 m^2 pixel.

VIII. HIGH AGILITY OF ADLAND

There is an important recommendation of the Food and Agriculture Organization (FAO) and the Group of Earth Observations (GEO) [21] which is still not fulfilled, at least, for the spatial higher

resolution thermal infrared data:

"Within the next 5 to 10 years, the space agencies should develop and implement the next generation of operational moderate resolution sensing systems, working in concert to provide a truly integrated system, acquiring and providing global coverage of 60-10m cloud free imagery every 5-10 days."

To meet this recommendation in the optical and IR bands, it is necessary to have, either
- a constellation of several satellites (similar to Sentinel-2 A+B),

- a wide swath (> 600km) sensor system, or

- high agility of a single satellite & sensor system, with intelligent pointing of the sensors Line of Sight (LoS) in a wide Field of Regard (FoR) during one overpass - as illustrated in Figure 5.

In this example, the LoS of the push broom sensor is moving forward over the ground with a speed of ~20 km/s, which is approximately three times higher than the speed of the satellite over the ground (6.6 km/s).

The area covered by the sensor in Figure 5 is approximately two times larger than the area which the sensor would cover in a "classic" push broom mode, where the sensors LoS is moving forward with the speed of the satellite over ground (6.6 km/s).

This high agility will be possible for ADLAND due to the use of:

(a) A new multi-spectral MIR/TIR sensor design based on cooled Mercury Cadmium Telluride (MCT) matrix detector array in combination with stripped filters (allowing sampling with image fragment overlaps for supplementary data binning), and

(b) Innovative technologies in the Attitude Control System (ACS) of the ADLAND micro-satellite bus.

The ADLAND satellite bus development will be based on the experience obtained from the German micro-satellite missions BIRD, TET-1, and BIROS.

IX. ADLANDS's CONVOY AND CONSTELLATION MISSION OPTIONS

There are several interesting convoy and/or constellation mission options for ADLAND:

ADLAND is considered as an ideal candidate to be flown in convoy with Sentinel-2 A+B satellites to complement the Sentinel-2 Multi-Spectral Imager (MSI) high resolution VIS-NIR-SWIR data by mid infrared (at 3.5 -4.0μm) and multi-spectral thermal infrared (at 8 – 12μm) data.

The intelligent and agile pointing of the LoS of ADLANDS's sensors, will allow IR data acquisitions in the 580 km combined swath width of Sentinel-2A+B.
A Sentinel-2 IR-companion ADLAND would conduct data records nearly simultaneously to MSI acquisitions of Sentinel-2A+B at ~ 10:30 h local time. Further, the IR companion could be "switched on" at ~ 22:30 h local night time. The higher spatial resolution IR data products from the possible Sentinel-2A+B IR-companion could be used to retrieve Earth observation physical variables for normal temperature phenomena and for HTE applications as explained in detail in chapter 5 of this paper.

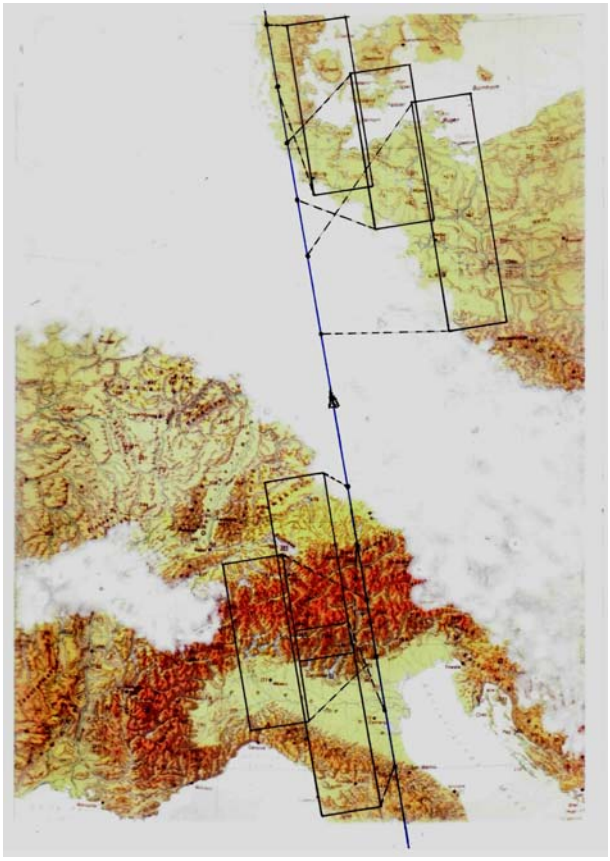


Figure 5. Illustration of intelligent and agile pointing of the Line of Sight (LoS) of a push broom sensor in a partly cloud-covered scene

ADLAND, flown in convoy with Sentinel-3, would complement the low spatial but high temporal resolution VIS-to-LWIR data of the Sentinel-3 Sea and Land Surface Temperature Radiometer (SLSTR), especially its land and fire data products [22]. In particular, ADLAND as an IR-companion of Sentinel-3 can be used for the validation of the FRP products from Sentinel-3.

It seems further reasonable to consider several ADLAND type IR-companions in constellation with Sentinel-2 and/or Sentinel-3, for instance: (a) one such IR micro-satellite in convoy with Sentinel-2 and/or Sentinel-3 (with Local Times of Descending Node /LTDN/ at 10:30h or 10:00h, respectively), and (b) another IR micro-satellite on an orbit with a LTDN at noon or at ~13:30 h, when both plant evapotranspiration and wildfire activity are at maximum.

X. SUMMARY

The ADLAND mission proposal targets on:

(i) higher resolution mid infrared and thermal infrared data - to close a major the land observation gap,

(ii) the ESA invitation to Sentinel convoy and constellation missions, and

(iii) a cost-effective micro-satellite based sensor system with high agility providing optical and infrared EO data also in case of partly cloud cover.

This may be an effective way to fulfil the recommendation of the Food and Agriculture Organization (FAO) and the Group of Earth

Observations (GEO) to secure a global coverage of 60-10m cloud free imagery every 5-10 days [21].

ACKNOWLEDGMENTS

The authors would like to thank the FireBIRD project team of the German Aerospace Center for providing FRP data and for maintaining a continuous dialogue on scientific exploration of the TET-1 data.

Our special gratitude is given to all scientists, companies, and organizations supporting the idea of a higher resolution thermal infrared Earth observation convoy mission to provide land surface data products which are still missing [23].

REFERENCES

[1] Briess, K., Jahn, H., Lorenz, E., Oertel, D., Skrbek, W. & Zhukov, B., (2003) Remote Sensing Potential of the Bi-spectral InfraRed Detection (BIRD) Satellite. Int. J. Remote Sensing, 24, pp 865-872.

[2] Ruecker, G., Lorenz, E., Hoffmann, A., Oertel, D., Tiedemann, and Halle, W., (2012), Upcominig and prospective fire monitoring missions based on the heritage of the BIRD (Bi-spectral InfraRed Detection) satellite, IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2012.

[3] Zhukov, B., Briess, K., Lorenz, E., Oertel, D., Skrbek, W., (2005), Detection and analysis of high-temperature events in the BIRD mission, Acta Astronautica 56, pp 65-71.

[4] Zhukov, B., Lorenz, E., Oertel, D., Wooster, M., Roberts, G., (2006), Spaceborne detection and characterization of fires during the bi-spectral infrared detection (BIRD) experimental small satellite mission, Remote Sens. Environment, 100, pp 29-51.

[5] Oertel, D., Zhukov, B., Wooster, M., Tank, V., Lorenz, E., Holzer-Popp, T., Goldammer, J., Martinez, S., Siegert, F., et al., (2005), ECOFIRE – Study on Scientific Assessment of Space-borne High Temperature Event Observing Mission Concepts, (ESA/ESTEC Contract 17690/30/NL/FF), Final Report, Berlin (Germany) German Aerospace Center (DLR).

[6] Sobrino, J. A., Bianchi, R., Paganini, M., Soria, G., Jimenez-Munzos, J., Oltra-Cario, R. et al. (2009), Dual-use European Security IR Experiment 2008 (DESIREX 2008), Final Report, Frascati (Italy), European Space Agency (ESA).

[7] Sobrino, J. A. et al. (2004), Land surface temperature retrieval from LANDSAT TM 5, Remote Sensing of Environment 90, No. 4, pp 434-440.

[8] Anderson, M. C. et al. (2008), A thermal-based remote sensing technique for routine mapping of land-surface carbon, water and energy fluxes from field to regional scales, Remote Sensing of Environment 112, No. 12, pp 4227-4241.

[9] Li, Z.-L. et al. (2013), Satellite-derived land surface temperature, Current status and perspectives, Remote Sensing of Environment, Vol. 131, pp 14-17.

[10] Merchant, C. J. et al. (2013), The surface temperature of the earth: steps towards integrated understanding of variability and change, Geoscientific Instrumentation, Methods and Data Systems. Discussion paper 3, pp 305-345.

[11] Solomon, S., Qin, D., Manning, M. et al. (2007), Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC), Summary for Policymakers.

[12] Kerr, Y. H. et al. (2000), The Second of ESA's Opportunity Missions: The Soil Moisture and Ocean Salinity Mission – SMOS, ESA Earth Observation Quarterly, Vol. 66, pp 18-25.

[13] Biastoch, A. and Böning, C. W. (2013), Anthropogenic impact on Agulhas leakage, Geophysical Research Letters 40, pp 1138-1143.

[14] Li, Z.-L. et al. (2013), Satellite-derived land surface temperature, Current status and perspectives, Remote Sensing of Environment, Vol. 131, pp 14-17.

[15] Gillespie, A. et al. (1996), Temperature/emissivity separation algorithm theoretical basis document, Vol. 2.4, pp 1-64, Maryland, USA (NASA/GSFC).

[16] Gillespie, A. et al. (1998), A temperature and emissivity separation algorithm for Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images. IEEE Transactions on Geoscience and Remote Sensing 36, pp 1113-1126.

[17] Remedios, J.J. Humpage, N. Ghent, D. (2012), Sentinel Convoy for Land Applications, Workshop Report & Executive Summary, University of Leicester, UK.

[18] Regan, A., (2014), Executive Summary and Detailed report on "1st International Earth Observation Convoy and Constellation Workshop", ESA/ESTEC, Noordwijk, NL, (ESA Ref. N°: EOP-SFT-2014-02-1789),

[19] Lagouarde et al.: The MISTIGRI thermal infrared project: scientific objectives and mission specifications, International Journal of Remote Sensing, 2012

[20] Zhukov, B., Oertel, D., Lanzl, F. and Reinhäckel, G., (1999), Unmixing-based multi-sensor multi-resolution image fusion, IEEE Transactions on Geoscience and Remote Sensing, Vol. 37, No. 3, pp 1212-1226.

[21] Justice, C.O. and Becker-Reshef, I. (Eds), (2007) Report from the Workshop on Developing a Strategy for Global Agricultural Monitoring in the framework of Group on Earth Observations (GEO), UN FAO, July 2007, Geography Dept. University of Maryland, 66, pp.

[22] Wooster, M., Xu, W. and Nightingale, T., (2012), Sentinel-3 SLSTR active fire detection and FRP product: Pre-launch algorithm development and performance evaluation using MODIS and ASTER datasets, Remote Sensing of Environment, 120, pp. 236-254.

[23] Ruecker, G., Menz, G., Heinemann, S., Hartmann, M., Oertel, D., (2015), VISIR-SAT – a Prospective Micro-satellite Based Multi-spectral Thermal Mission for Land Applications, Proceedings of the 36th International Symposium of Remote Sensing of Environment, Berlin, D, (paper ISRSE36-590-1).

The Technical Feature and Application of CBERS04-The International Cooperation Satellite

He Wei[1], Long Xiaoxiang[2], Zhang Minglei[1], Jia Xu[1]

（1 Institute of Spacecraft System Engineering,CAST,Beijing 100094, China）

（2 China Center for Resources Satellite Data and Application,Beijing 100073, China）

Abstract: CBERS（China-Brazil Earth Resource Satellite）is a series of earth observing remote sensing satellite jointly developed by China and Brazil. CBERS04 is launched at TSLC by CZ-4B in Dec7, 2014. This paper introduces design and technology feature of CBERS04 satellite. CAST(China Academy of Space Technology) and INPE(Instituto Nacional De Pesquisas Espaciais) each responsible for half of the funds and tasks, As an international technical cooperation, the two sides have established complete Product Assurance method. CBERS04 data have been successfully used in the geological research, agriculture, disaster, forestry and environment. CBERS embodies the important role in the international remote sensing data application.

I. Introduction

CBERS（China-Brazil Earth Resource Satellite）is a series of earth observing remote sensing satellite jointly developed by China and Brazil, which is called “exemplar of south-south cooperation” in high-tech field. CBERS04 is the CBERS01\02\02B follow-on satellites, intended to continue and improve the services of optical imaging and data, have a lot of technical progress. CBERS04 has 4 different cameras can provide 13 spectrums overlay visible light to infrared, and has as many as 6 kinds of resolutions, with the highest Swath width of more than 860km. CBERS04 data have been successfully used in the geological research, agriculture, disaster, forestry and environment.

II. Satellite Design

Satellite Design Overview

The CBERS04 consists of Service Model and Payload Model. CBERS04 is a 3-axis attitude stabilized spacecraft with a global coverage polar sun-synchronous orbit. The Platform of CBERS03 consists of STRU (Structure Subsystem), TCS (Thermal Control Subsystem), AOCS (Attitude and Orbit Control Subsystem), TT&C (Tracking, Telemetry and Command Subsystem) and OBDH (On-board Data Handling subsystem) etc. Appearance of CBERS04 is shown in Figure 23.

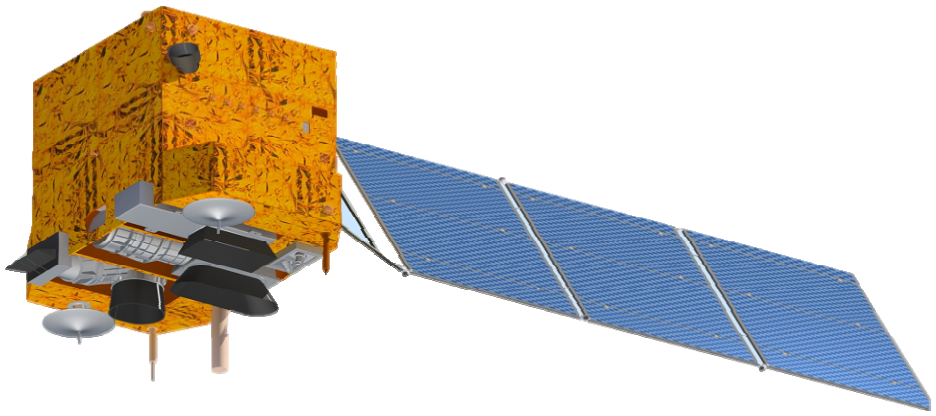


Figure 23 Appearance of CBERS04

The platform is designed to give support to the following main functions:

a）Carry a Panchromatic Multi-spectral High-Resolution (PAN) camera is one of most important remote sensor for surface observation, which operating in visible and near infrared bands for acquiring high resolution imageries of the Earth surface, it has the side-looking ability of 32° and can perform the in-orbit calibration and focal adjustment.

b）Carry a Multi-spectral (MUX) Camera operating in visible and near infrared bands for acquiring high resolution imageries of the Earth surface, it can perform the in-orbit calibration and focal adjustment.

c）Carry a Infrared Multi-Spectral Scanner (IRS) operating in visible, near, middle and thermal infrared bands for acquiring moderate resolution imageries of the Earth surface. This instrument has the radiometric calibration parts.

d）Carry a Wide Field Imager (WFI) for acquiring low resolution but wide swath imageries of the Earth surface.

e）Carry two set Data Transmission subsystem to transmit the image data from PAN camera, MUXCAM camera, IRSCAM camera and WFICAM camera at X-band to the ground receiving stations. It also transmits the beacon for ground image receiving stations tracking.

f） Carry a Digital Data Recorder (DDR) for record and playback image data from four cameras.

g）Carry a Data Collecting Subsystem (DCS) for relaying DCP data of Earth environment monitoring to the ground.

h）Carry a scientific Space Environmental Monitoring (SEM) for monitoring the particle radiation environment of the satellite and outer space.

Table 1 Main performance of CBERS04

Orbit	Sun-synchronous recurrent and frozen orbit
Height	778Km
Local Time at descending node	10:30AM
Repeat cycle	26days
Attitude Control	3-axis attitude stabilization, Earth pointing
3-axis Measurement accuracy	0.03°(3σ)
3-axis Pointing Accuracy	0.1°(3σ)

3-axis Stability	1×10^{-3} °/s(3σ)
Output Power of Solar Panel (EOL)	1500W
Ni/Cd Battery Capacity	50Ah×2
Mass of Spacecraft	2100kg
Spectral bands (μm)	
PAN camera	B1:0.51~0.85 B2:0.52~0.59 B3:0.63~0.69 B4:0.77~0.89
MUXCAM	B5:0.45~0.52 B6:0.52~0.59 B7:0.63~0.69 B8:0.77~0.89
IRSCAM	B9:0.50~0.90 B10:1.55~1.75 B11:2.08~2.35 B12:10.4~12.5
WFICAM	B13:0.45~0.52 B14:0.52~0.59 B15:0.63~0.69 B16:0.77~0.89
Pixel resolution(m)	5/10(PAN)
	20(MUXCAM)
	40/80 (IRSCAM)
	73(WFICAM)
Record/Playback Time	More than 15 min
Record/Playback ratio	1:1
Lifetime in Orbit	3 years

Operating Mode

CBERS04 Imaging and Transmitting Operating Mode as bellow:

1）Long time Mode：Only the Long-Time equipment work; DCS and SEM subsystem work all the time;

2）Sunlight Image：When the satellite fly over the coverage of ground image receiving stations, in daytime, the PAN camera, MUXCAM camera, IRSCAM camera and WFICAM camera will perform imaging and the image data will be transmitted to ground receiving station in real-time.

3）Shadow Image：During the night, IRSCAM will perform the imaging, the data will be transmitted to ground receiving station in real-time;

4）Recorde Mode: The 4 camera image in any area, DDR recorde image data;

5）Play back Mode: DDR and PIT&MWT transmitte the data when satellite fly over the coverage of ground image receiving stations.

III. Technology Feature

1）Payload of CBERS04 can provide 13 spectrums, and 6 kinds of resolutions.

CBERS04 has 4 different cameras can provide 13 spectrums overlay visible light to infrared, and has as many as 6 kinds of resolutions, with the highest Swath width of more than 860km.

2）CBERS04 can provide visible light and infrared image.

The IRS camera can provide infrared image,CBERS04 can obtain image from global day and night.

3）CBERS04 has rapid Revisit period and Overlay period

Using the 32° side-looking ability of PAN camera, CBERS04 has 3 days Revisit period, and using the Swath width of WFI camera, CBERS04 has 5 days Overlay period.

4）Design in accordance with the requirements of the public platform.

Satellite designs in accordance with the requirements of the public platform, and have a lot of extensible ability.

IV. Features of International Technical Cooperation

CAST(China Academy of Space Technology) and INPE(Instituto Nacional De Pesquisas Espaciais) each responsible for half of the funds and tasks. Starting from the first satellite, the two sides have established complete Product Assurance method.

1）Complete File System

CAST and INPE jointly develop the file list, include System Specification, Subsystem Specification, Design and Construction Specification, Environment Specification, EMC Specification, Interface Data Sheet, Interface Control Document(Spacecraft to Other Segment Interface Specification), etc. Both of us in the same file system governing the design.

2）Joint Review

In the important stage, in a fixed mode of joint review, such as PDR(Preliminary Design Review), CDR(Critical Design Review),FDR(Final Design Review), etc. During the AIT, TRRB(Test Readiness Review Board) and TRB(Test Review Board) must be conducted in the test before and after.

3）Deliver to JPO

In addition to finish acceptance work in their respective countries, products to the JPO(Joint Project Organization), delivery will be held by the expert review.

4）Engineering Change Request

We use the ECR(Engineering Change Request Form) to control product technical state changes. Each technical state changes should fill out the ECR, describes the content and influence of the changes, and confirmed by both parties.

5）OPEN- CLOSE

The problems need to be solved in the research, manufacture and AIT, we use closed loop management to implement. We fill the OPEN-POINTS sheet and verify the CLOSE termly.

V. Application on Orbit

CBERS04 remote sensing image is shown in Figure 24.

Figure 24 CBERS04 remote sensing image-Shenyang

In China, CBERS data have been successfully used in the geological research, agriculture, disaster, forestry and environment.

In Brazil, CBERS also have been used widely. Amazon deforestation in the Brazilian government to submit monthly report, the main data resources are CBERS. As present, the domestic users have exceeded 15,000, and circulated to Argentina, Chile, Bolivia, Venezuela, Uruguay, Paraguay, Peru, India, Italy, Spain, Canada, France, America, etc.

CBERS South Africa Station makes the 13 countries (with a total area of about $2,047\times107km^2$) in South Africa have a chance to use the CBERS data, greatly promote the local economic and social development.

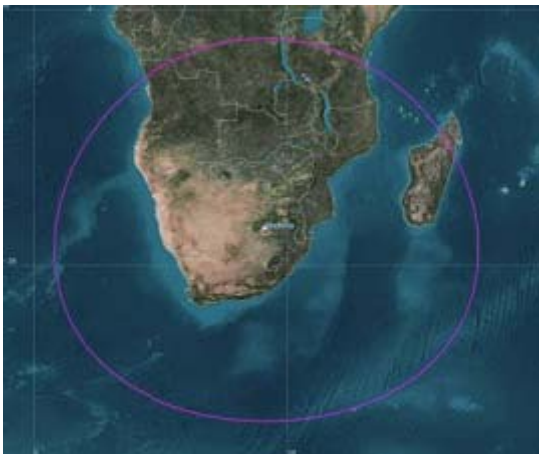


Figure 25 Schematic fig-CBERS Data Area of South Africa Station

As CBERS data sharing service platform, CHINA-ASEAN Remote Sensing Satellite Data sharing platform provide CBERS data to ASEAN countries, shared aerospace technique achievements, promote the development of CHINA-ASEAN aerospace science and technology cooperation.

CBERS embodies the important role in the international remote sensing data application.

References

[1] Zhang Qingjun, Ma Shijun. The technical feature and progress of CBERS[J]. Aerospace China, 2008,4: 19-24

[2] Zhang Qingjun, Ma Shijun. Achievements and Progress of China-Brazil Earth Resource Satellite [J]. Spacecraft Engineering,2009,18(4) : 1-8

[3] Wen Lan, Discuss of Product Assurance method of CBERS satellite[C]. //2nd China aerospace quality forum. 2008:182-185

[4] Xu Wen, Wang Zhigang, Wu Min. China international cooperation achievements in remote sensing satellite[J].Space International,2014,10:22-26

[5] Guo Jianning, Yu Jin, Yu Wenyong. New starting point of CBERS application, [J].Space International,2007,11:1-6

Multikernel Machine-based Hyperspectral Remote Imagery Processing System on Satellite Platform

XIE Xiaodan*, JIA Jingcheng, JIA Yusheng

Science and Technology on Optical Radiation Laboratory
Beijing 100039,P.R.China
Email: xiexiaodanbeijing@126.com
*Corresponding author

Abstract: Hyperspectral instruments with hundreds of contiguous spectral channels brings the developing of collecting remote imagery data . Hyperspectral imagery is the most popular remote sensing technology on satellite platform, with the prospective applications in military monitoring, energy exploration, geographic information, and so on. The preprocessing of hyperspectral sensing data is a feasible way through machine learning-based data analysis technology. The spectrum data in database is collected in advance, so it has inconsistency between the spectrum with the data collection. The inconsistency can be consider the nonlinear changing. The relationships of between spectral curves are the classical nonlinear relationship. So the classification is the nonlinear and complex classification problem. Kernel learning is a feasible and effective nonlinear learning method on hyperspectral sensing data. Optimizing the parameters does not effectively promote the kernel-based learning system. We present a framework of multiple quasiconformal mapping-based kernels learning for hyperspectral image classification. The learning system is improved adaptively the data structure in the kernel mapping to more precisely characterize the data for improving performance on solving complex visual learning tasks. The learning framework is applied to the hyperspectral image classification, and some experiments are implemented on two hyperspectral image databases to evaluate the performances of the system.
Keywords: Remote sensing, hypespectral imagery, kernel machine.

1. Introduction

Hyperspectral imagery is the most popular remote sensing technology on satellite platform, with the prospective applications in military monitoring, energy exploration, geographic information, and so on. Hyperspectral instruments with hundreds of contiguous spectral channels brings the developing of collecting remote imagery data. The increasing spectral and space resolution bring a large size of data, which brings two problems in the practical applications: 1) the bandwidth of the communication channel limits the transmission of the full hyperspectral image data for the further processing and analysis on the ground; 2) the demand of the real-time processing for some applications. Data compression is a solution to the transmission problem but no ability for the real-time analysis. So, machine learning-based data analysis technology is feasible and effective to produce one image from the full band of hyperspectral images through classifying the spectrum curve of each pixel according to the spectrum data of each object. The hyperspectral data machine learning system is implemented on the satellite platform. After the hyperspectral data collection, each pixel is classified and denoted to the different objects based on the spectrum database. The spectrum data in database is collected in advance, so it has inconsistency between the spectrum with

the data collection. The inconsistency can be consider the nonlinear changing. The relationships of between spectral curves is the classical nonlinear relationship. So the classification is the nonlinear and complex classification problem. Traditional classification methods are not effective to hyperspectral sensing data, among these machine learning methods kernel learning is a feasible and effective nonlinear classifier methods on hyperspectral sensing data. The data distribution is easy to classification owing to kernel mapping in the nonlinear feature space. The geometrical structure of the data is totally determined by the kernel function, has significant influences on kernel learning methods. The discriminative ability of the data could become worse with the inappropriate kernel. Researchers optimized the parameters of kernel function to improve KDA in the previous works [1], [2], [3]. Many kernel optimization methods for example, selecting the optimal kernel parameter from a set of discrete values. The geometry structure of data distribution in the kernel-based mapping space is not be changed only adjusting the parameters. Xiong proposed a data-depend kernel for kernel optimization-based machine learning [4], and Amari presents the support vector machine classifier through modifying the kernel function [5]. Moreover, multiple kernel learning methods are developed, for example, Sparse Multiple Kernel Learning [6], Large Scale Multiple Kernel Learning[7], lp-Norm Multiple Kernel Learning[8], and kernel machine learning is also applied in hyperspectral image classification[9].

2. Methods

For the hyperspectral sensing data classification, we present a multikernel machine framework for satellite platform hyperspectral remote imagery system, as shown in Figure 1. The satellite platform-based hyperspectral image system including three main parts, image collection, image processing online, image transmission. In our work, we pay attention to the second part. And in this part, we have the research on spectrum machine learning.

In the spectrum machine learning, motivated by the fact that kernel machine is effective to the nonlinear classification but the performance of kernel-based system is largely influenced by the function and parameter of kernel, we present a framework of quasiconformal mapping-based multiple kernels learning. Optimizing the parameters not effective to promote the kernel-based learning system owing to the unchanged data structure with the changing of the parameter of kernel function. No a universal single kernel is very effective way to detecting intrinsic information for the complicate sample data in the input data space. The learning system is improved adaptively the data structure in the kernel mapping from two facts: 1) quasiconformal single kernel structure changes the data structure in the kernel empirical space; 2) quasiconformal multiple kernels are combined to more precisely characterize the data for improving performance on solving complex visual learning tasks. The learning framework is applied to the hyperspectral image classification, and some experiments are implemented on two hyperspectral image databases. The procedure is shown in Figure 2. The procedure includes three procedures of multiple kernels optimization, training and testing for general kernel-based learning application.

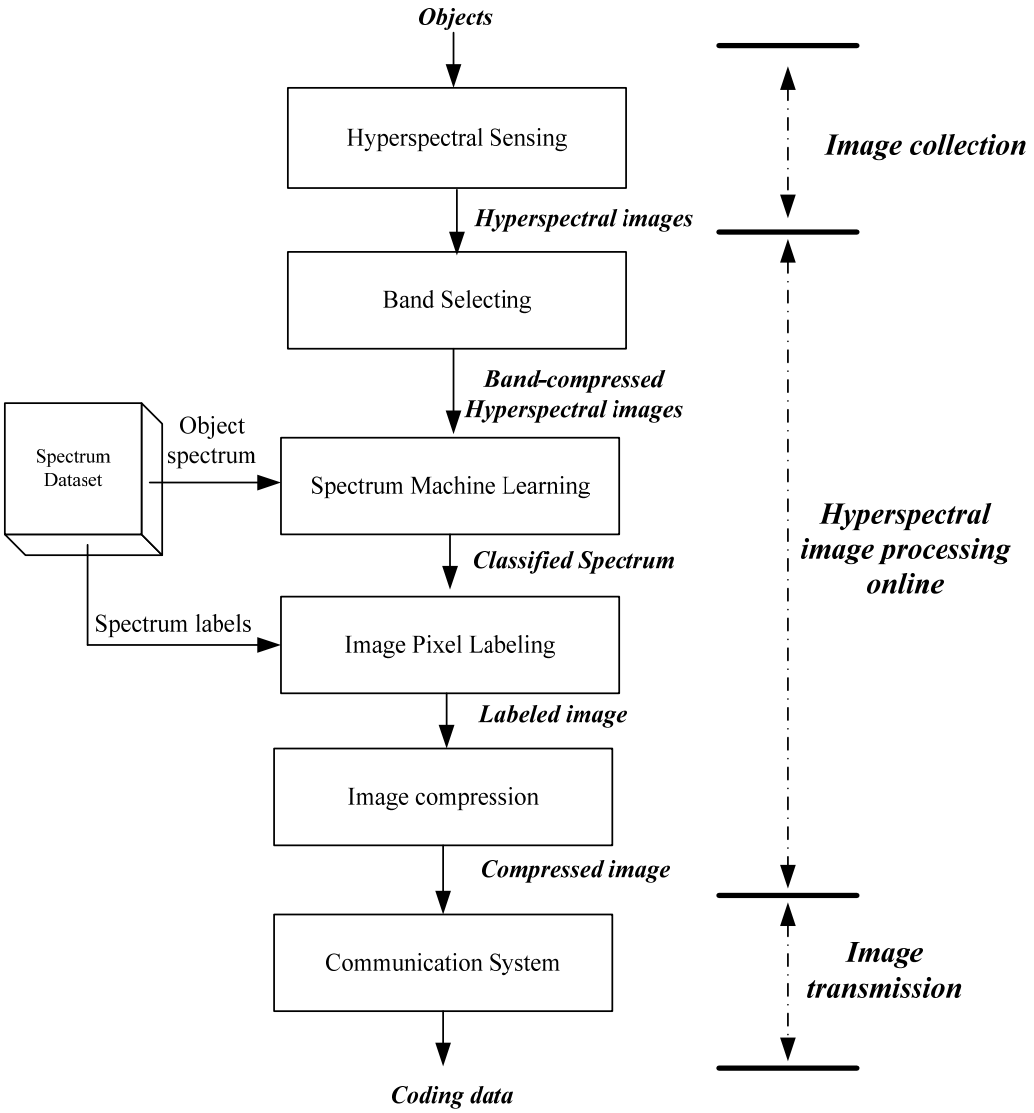


Figure 1. Multikernel machine-based hyperspectral remote imagery processing system on satellite platform

The procedure of solving the optimization equation costs more time than without optimizing kernel function. But the kernel optimization procedure can be implemented off-line. So the kernel optimization procedure will not cost much time on the online application. In the training steps, the optimal parameters is solved through iteration optimization, and the procedure will cost much time. While in the test stage, it needs the additional less time consuming. So, it has the little influence on the efficiency of learning system.

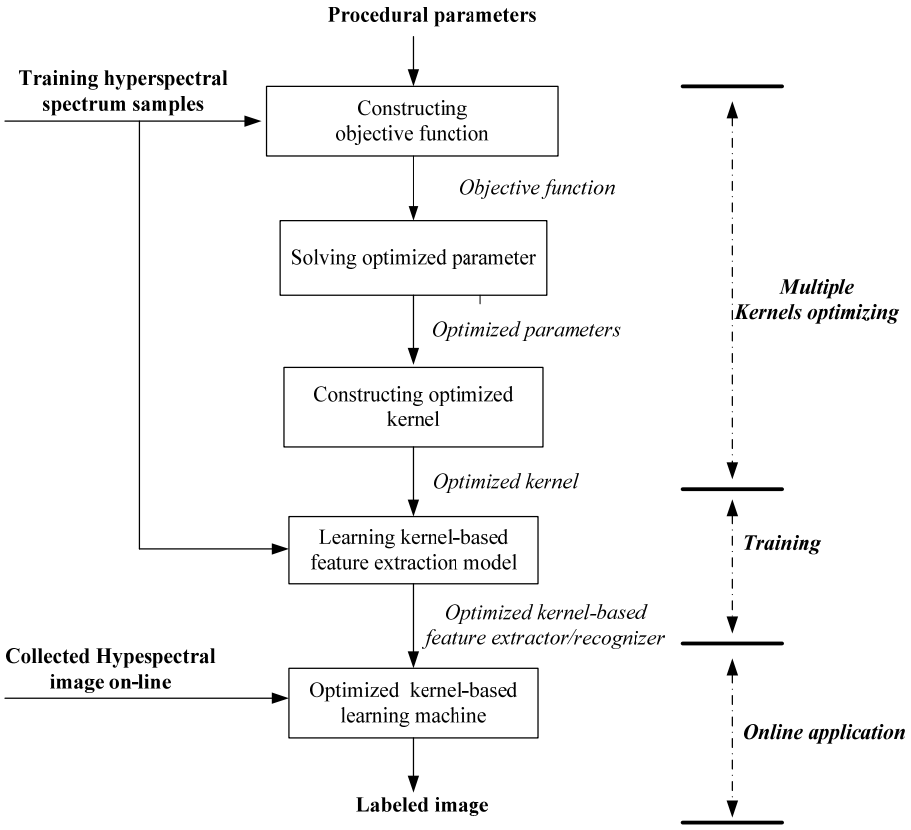
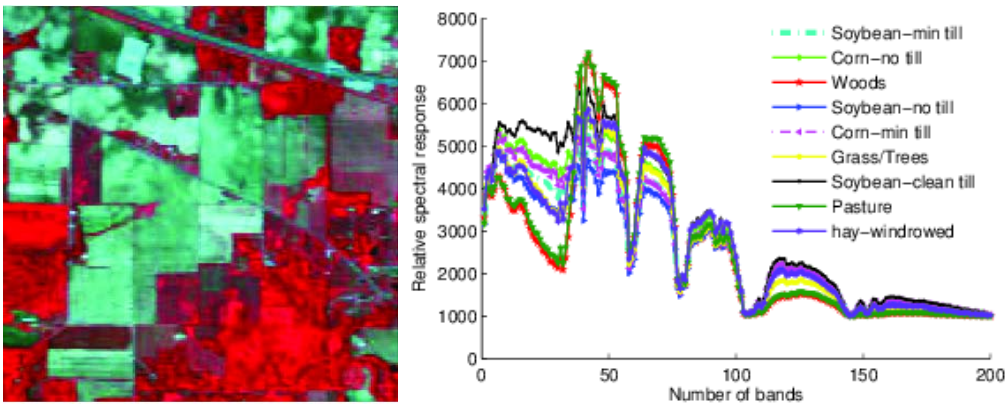


Figure 2. Procedure of multiple kernels based learning machine

3. Experiments and Result

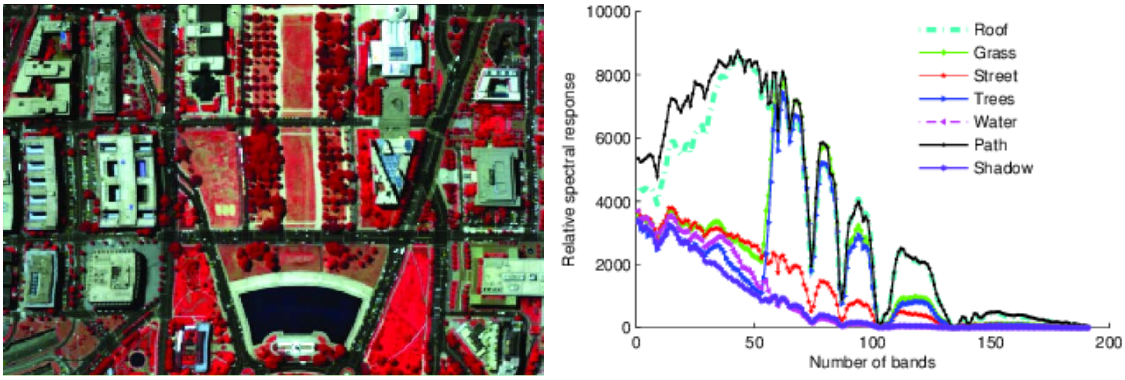
The experiments are implemented on the hyperspectral image database, Indian Pines dataset, which is collected under the various spectral and spatial resolutions, and the spetral curves denote the different remote sensing environments. The data cube have 224 bands of spectral resolution through 0.4-2.5 μm range , and it has the spatial resolution of 20m per pixel. In our experiments, we removed the noisy and water-vapor absorption bands and 200 bands of images are used in the experiments. The whole scene is consists of 145×145 pixels, and 16 classes of interested objects rang the size from 20 to 2468 pixels, but only 9 classes of objects are selected in the experiments. The experimental results show that the proposed system is effective and better performance are achieved compared with other technologies. Some examples are shown in Figure 3. D.C. Mall data was acquired under the airborne with hyperspectral digital imagery collection experiment (HYDICE) sensoron August 23, 1995. The image is size of 1280 307 pixels, and it has the spatial resolution of 1.5m, and 210 spectral bands is in the 0.4-2.4 region. Several bands influenced by the atmospheric absorption are ignored, and the rest 191 bands are implemented in the experiments. The image is resized to the size of 211 307 including 7 classes of land-covers namedly roof, grass, street, trees, water, path and shadow.Some examples are shown in Figure 4.

Different components of the image have different ability to present the objects. On the hyperspectral image databases, the averaged accuracy is as the final recognition accuracy. To reduce the computation stress, PCA is to reduce the dimensionality of the input hyperspectral image vector, and the 6 principal components are selected on the Indian Pines data. Some preprocessed images of 6 principal components are shown in Figure 5.



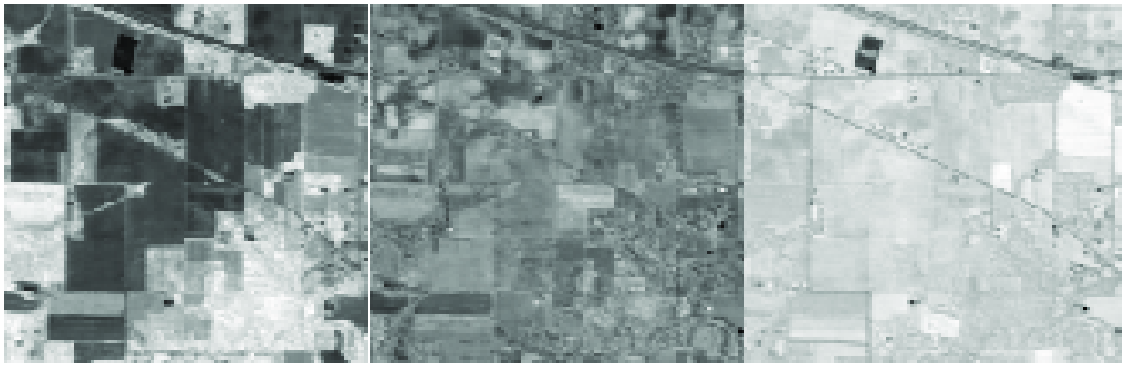
(a) Three band false color composite (b) Spectral signatures.

Figure 3. One example of Indian Pines data

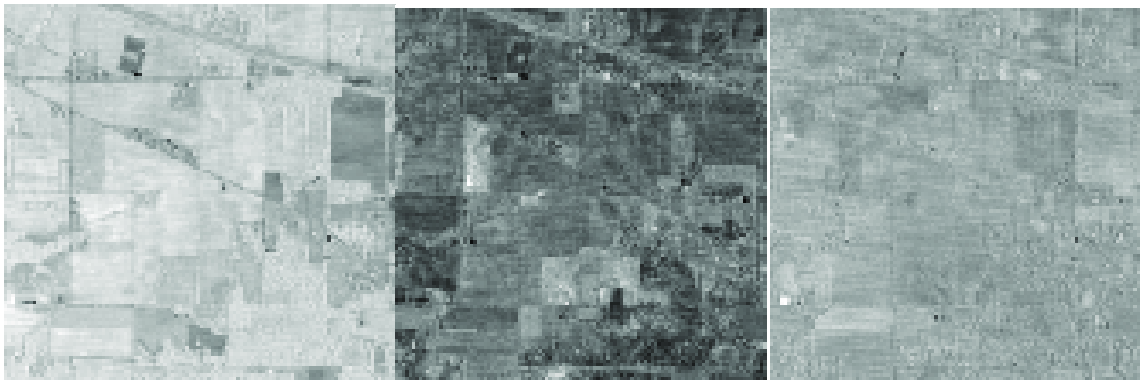


(a) Three band false color composite. (b) Spectral signatures.

Figure 4. One example from D.C. Mall data



1st principal Component 2nd principal Component 3rd principal Component

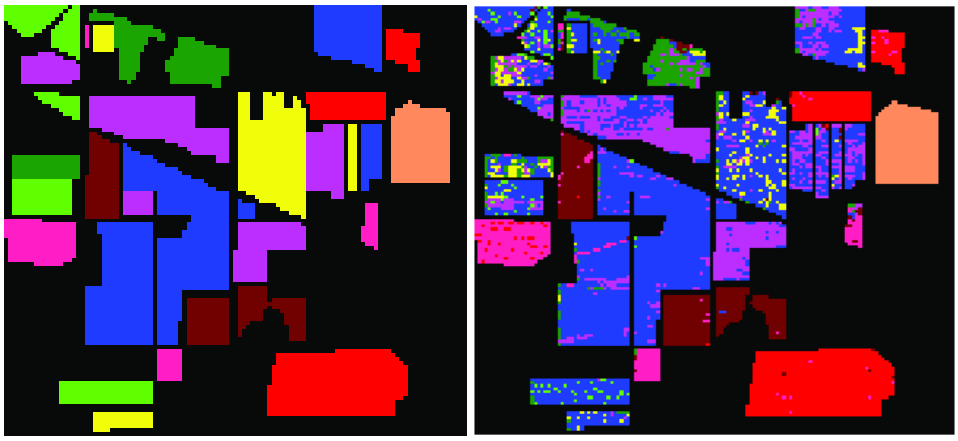


4th principal Component 5th principal Component 6th principal Component

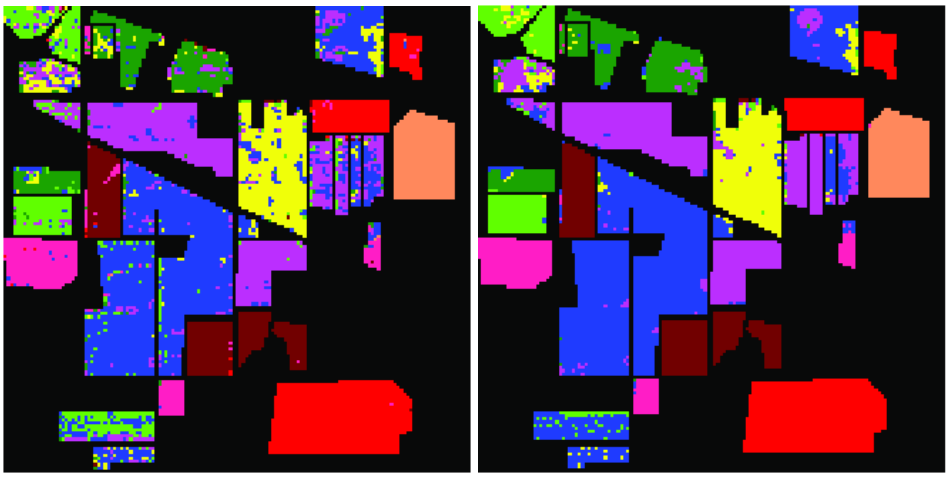
Figure 5. Six principal components of images on Indian Pines data.

Different components of the image have different ability to present the objects. On the hyperspectral image databases, the averaged accuracy is as the final recognition accuracy. To reduce the computation stress, PCA is to reduce the dimensionality of the input hyperspectral image vector, and the 6 principal components are selected on the Indian Pines data. Some preprocessed images of 6 principal components are shown in Figure 5.

Firstly, we implement Support Vector Classifier (SVC), Kernel Sparse Representation Classifier (KSRC) on data classification. We test the single kernel and quasiconformal multikernels for kernel classifiers on SVC and KSRC, that is, PK-SVC: Polynomial Kernel-SVC, GK-SVC:Gaussian Kernel-SVC, QMK-SVC: Quasiconformal Multi-kernels Based SVC, PK- KSRC: Polynomial Kernel- KSRC, GK- KSRC:Gaussian Kernel- KSRC, QMK- KSRC: Quasiconformal Multi-kernels Based KSRC. The experimental results are shown in Figure 6 and Figure 7. As experimental results on the performance of hyperspectral image databases, we can conclude that quasiconformal multiple kernels performance than other methods. The proposed framework is to solve the selection of function and parameter of kernel, which have the heavy influences on the performance of kernel-based learning system. Quasiconformal single kernel structure changes the data structure in the kernel empirical space. And then, quasiconformal multiple kernels are combined to more precisely characterize the data for improving performance on solving complex visual learning tasks, so the proposed framework outperforms others in the different datasets.



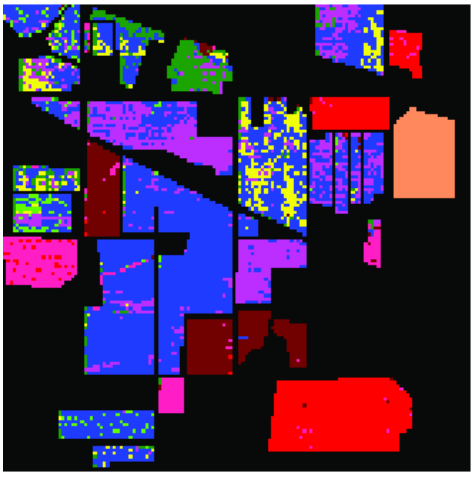
(a) Ground truth (b) PK-SVC



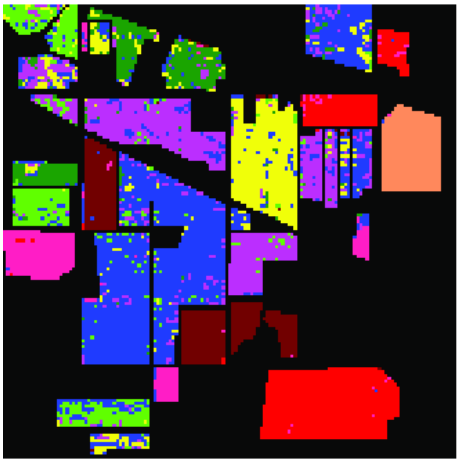
(c) GK-SVC (d) QMK-SVC

Figure 6. Classification maps of the SVC on Indian Pines data.

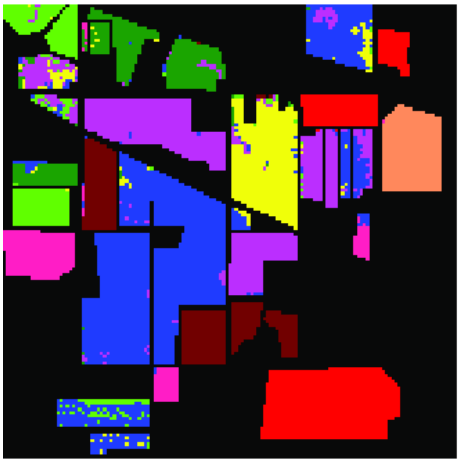
(a)Ground truth



(b) PK- KSRC



(c) GK- KSRC



(d) QMK- KSRC

Figure 7. Performance of KSRC on Indian Pines data.

4. Conclusion

This paper presents a novel framework of manifold multiple kernel learning, and applies multiple kernels model to increase the data description ability, and then introduce the quasiconformal kernel optimization to increase the ability of feature extraction. The proposed method is a promising dimensionality reduction method on data processing, especially on hyperspectral image processing, and it has its advantage of preserving the local and discriminative embedding, so the learning framework is feasible to the hyperspectral image classification, and moreover the framework can be used to other kernel-based systems in the practical applications. The computing efficiency of training the proposed framework is the problem in the practical system, because most applications are off-line training based model learning.

References

[1] Jian Huang, Pong C Yuen, Wen-Sheng Chen and J H Lai. “Kernel Subspace LDA with Optimized Kernel Parameters on Face Recognition”, Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004.

[2] J. S. Pan, J. B. Li, and Z. M. Lu, “Adaptive quasiconformal kernel discriminant analysis,” Neurocomputing, vol. 71, no. 13-15, pp. 2754-2760, 2008.
[3] Wen-Sheng Chen, Pong C. Yuen, Jian Huang, and Dao-Qing Dai, “Kernel Machine-Based One-Parameter Regularized Fisher Discriminant Method for Face Recognition”, IEEE Trans. Systems, Man and Cybernetics-Part B: Cybernetics, vol. 35, no. 4, pp. 658-669, August 2005.
[4] H. Xiong, M. N. Swamy, and M. O. Ahmad, “Optimizing the kernel in the empirical feature space,” IEEE Transactions on Neural Networks, vol. 16, no. 2, pp.460-474, 2005.
[5] S. Amari and S. Wu, “Improving support vector machine classifiers by modifying kernel functions,” Neural Networks, vol. 12, no. 6, pp. 783-789, 1999.
[6] N. Subrahmanya, Y.C. Shin, “Sparse Multiple Kernel Learning for Signal Processing Applications,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 5, pp.788-798, 2010.
[7] Sören Sonnenburg, Gunnar Rätsch, Christin Schäfer, Bernhard Schölkopf, “Large Scale Multiple Kernel Learning,” Journal of Machine Learning Research, vol.7. pp. 1531-1565, 2006.
[8] Marius Kloft, Ulf Brefeld, Sören Sonnenburg, Alexander Zien, “lp-Norm Multiple Kernel Learning,”Journal of Machine Learning Research, vol. 12, pp.953-997, 2011.
[9] Chen Chen, Wei Li, Hongjun Su, Kui Liu, “Spectral-Spatial Classification of Hyperspectral Image Based on Kernel Extreme Learning Machine,” Remote Sensing. Vol. 6, no. 6, pp.5795-5814, 2014.

Space Optical Remote Sensing Application on Lands, Ocean and Atmosphere Observation

RUAN Ning-juan ZHANG Wen-yu

(Beijing Institute of Space Mechanics & Electricity, Beijing 100094, China)

Abstract: Space remote sensing is the science of acquiring information of Earth surface by collecting emitted and reflected energy via optical payload and detector, which is applied on space vehicles (like satellite, spacecraft and space shuttle). This article will explain the fundamental theory, imaging chain and optical payload composition. And it will also introduce the essential target and element of remote sensing technology. At last, there will be an analysis of optical remote sensing payload on the current international development situation and commercial satellite application, which will provide suggestion for China's future remote sensing development.

1. Introduction

Optical payload is an instrument assembled on space vehicles, which focuses the collected electromagnetic radiation on a detector or receiver by using optical systems. According to the spectrum and technological system, the optical payload can be divided into ultraviolet payload, visible light payload, infrared payload, spectrometer and laser payload. They have been widely used in the observation of land, ocean, atmosphere and space detection, which improves human living quality and global economic development.

2. Overview of Optical Remote Sensing Technology

With the launch of the first man-made earth satellite on October 4, 1957 in former Soviet Union, the human society entered the space age. In the early 1960s, the satellite and spacecraft which carried the film camera and television camera system were sent to the space to collect the information for civil and military use, marking the beginning of the space remote sensing era. The space remote sensing has developed rapidly since 1970s. The payload resolution, observation spectrum and observation method were improved increasingly. Now it has formed two types: optical remote sensing payload and microwave remote sensing payload. The optical remote sensing payload is equipped on the space platform which can provide the orbit protection conditions, such as power supply, attitude stability and environment control. Cooperating with the ground station, it takes pictures of the earth or space object through optical band (UV, VIS, IR) and sends back information. The payload is an opto-mechanical instrument combining optics, mechanics, electricity, heat, control and information processing functions into it.

An early optical remote sensor used film as the information carrier which was replaced by photoelectric detector (like CCD, CMOS) later. Optical remote sensor usually consists of optical lens, detector & information processor, mechanical structure and thermal control device. All parts work together to acquire, process and transmit data.

Optical lens has the function of collecting light energy for objective signal and restraining stray light. The principal issue for the overall design of optical remote sensor is to change the operation requirement into technical requirement of optical lens, such as the working band, transfer

function and resolution of lens, and confirm the structural parameter of the lens, such as focal length, relative aperture and FOV. Therefore, the optical lens needs careful selection and design simulation. The detector and information processor are mainly responsible for transforming the optical signal into electrical signal which will be processed by amplification, filtering, quantization, coding and compression. After processing, the signal will be transmitted to ground station through satellite antenna. The mechanical structure fixes the optical lens, detector and information processor together. The thermal control device, which consists of thermal control coating, thermal insulation film, heat pipe, radiator, heater and temperature controller, controls the temperature of remote sensor. In order to make the remote sensor work under the spatial thermal environment, it needs active and passive thermal control measures to keep the required operation temperature range and gradient of the camera.

The data acquisition quality of the remote sensor is not only depends on its own performance, but also influenced by elements of the whole imaging link. The remote sensing imaging link includes target, energy source, transmission path and remote sensor. The electromagnetic wave emitted from energy source interacts with the target. The electromagnetic wave emitted or reflected from the target carries the object information. The remote sensor obtains the electromagnetic wave which carries the target information and processes it to get the remote sensing data. Through the inversion to the remote sensing data, we can get the target information.

3. Application Requirements for Remote Sensing Detection

AS an optical remote sensor can acquire global surface information in real-time and all-weather quickly and widely, it can detect the type, shape, spatial location, even the structure, kinds, quality and composition of the target via the features of electromagnetic wave emitted, reflected and scattered from different targets. Optical remote sensors can be widely used for land observation, marine observation, atmosphere observation and space detection.

Earth observation payloads, which can provide full-fledged data type, can be widely used in many realms such as resource survey, environment protection, planning and management and disaster reduction. The data requires high spatial resolution, high SNR, high revisiting capability, large dynamic range and wide bandwidth. Ocean observation mainly includes water color environment, dynamic environment, ocean management, and rights protection and marine law enforcement. Ocean elements observation is important for global climate and environment change. For ocean application, as there are various observed targets and elements of large amount, the data with different resolutions, large scale, high SNR, high sensitivity, small dynamic range and small band range are needed. At the same time, marine surveillance and management can be supported by and combined with land observation.Atmosphere application consists of global NWP, precipitation intensity monitoring, 3D stereo detection of clouds, 3D stereo detection of wind field, air quality monitoring, climate change, space climate monitoring, the middle and upper atmosphere detection, space radiation measurement, ionized layer anomaly monitoring and other aspects. Space environment monitoring includes space weather monitoring and debris monitoring.

4.The Development Status

VIS optical remote sensing camera is the earliest medium-resolution and wide-swath payload. Since 1972, USA has already launched many Landsat series satellites, and the latest Landsat-8 is a satellite with 15m/30m resolution and 180km swath. Since 1999, USA has started to

pay more attention on the research of high resolution payload. From the IKONOS (1999) to the worldview-3(2014), the resolution is improved from 0.82m to 0.31m. The short wave resolution of IR imaging payload can be 3.7m (worldview-3), middle wave resolution can be 5.5m (KOMPSAT-3A), and long wave resolution can be 100m (Landsat-8). The UV imaging payload is mainly applied for atmosphere observation, which is including ozone vertical observation, and the spectral resolution of UV spectrometer is 0.2 nm. The hyper imaging payload is mainly for surface feature composition analysis. For instance, the spectral range of EO-1/Hyperion is 0.4-2.5μm, and the spectral resolution is 10nm.It has 242 spectrum channels, and the ground resolution is 30m. The representative of Lidar payload is equipped on the ICESat, with single-band non-scanning mode, and the vertical resolution is 0.3m. ICESat has worked overtime until 2010.

With the development for several decades, the optical remote sensing payload of China has made a great achievement. The performance of visible light, IR, high spectral, UV and laser payloads continue to improve, and the spatial resolution of visible payload in orbit has improved from 3m of ZY-2 to 2.1m of ZY-3, then 0.8m of GF-2 satellite, meaning that we have entered into sub-meter era of high resolution. The scale of mapping payload has reached to 1： 50000（ZY-3），and the observation swath of HY-1 has reached to 3000km . The detection sensitivity of IR payload has improved from0.2K（FY-3A）to 0.1K（FY-3C）， and the spatial resolution of civil IR payload can go up to 40m(ZY-1-04) .Among high spectral payloads, the visible one can reach a maximum resolution of 2nm（HJ-1A）， and the IR spectral resolution can reach to 2.5cm^{-1}(short wave)/ 0.625cm^{-1} (long wave) of FY-3C satellite. UV payload of FY-3A has been successfully used in many fields such as Ozone detection. The distance measuring accuracy of the laser payload can go up to 0.3m（CE-3）.

The remote sensing satellite has been widely used in weather, agriculture, disaster monitoring, territorial planning, environment protection, forestry, ocean and mapping, and the organizations of GEOSS, EOS, GEO, GMES have been set up successively worldwide. The general trend of global earth observation system is that the spatial, spectral, time and polarization resolution will increase continuously, and the quality and accuracy of earth observation data will be improved accordingly. The combination of earth observation data with various resolutions will challenge remote sensing technology.

5. Conclusion

With the development of remote sensors and the application of remote sensing data, the remote sensing technology will better serve human beings, and the global remote sensing organizations will expand cooperation in this field, displaying their respective advantages to advance the development of the remote sensing technology together.

Some Thoughts on International Cooperation in Using GNSS Below the Sea Surface Level

Yean Joo CHONG

Commission 6, International Academy of Astronautics (2013-2017)
National University of Singapore (retired academic staff) y.j.chong@alumni.nus.edu.sg

Abstract: Global Navigation Satellite Systems (GNSS) are widely used. However because the electromagnetic signals used cannot penetrate through water, the coordinates of a point under water cannot be measured directly. For underwater scuba diving, simple methods can be devised to use GNSS for identifying underwater sites. Other uses of GNSS are also referred to.
The United Nations International Seabed Authority has granted over 27 contracts for seabed mineral exploration, including one contract involving the Keppel-NUS Corporate Laboratory of the National University of Singapore.
An article in Science, has drawn attention to the need to consider the impact that seabed exploration and mining may have on the Marine Environment. It is expected that GNSS would have wide uses for seabed exploration and mining and for studies of the their impacts on the marine environment.

1. Global Navigation Satellite Systems

The main Global Navigation Satellite Systems (GNSS) are:
the U.S. Global Positioning System;
the Russian GLONASS;
the Chinese Beidou (Compass);
and the European Galileo
In addition there are the Regional Systems of India and Japan.
At any instant of time, the GNSS is being used somewhere in the world.

However, because the signals used in GNSS cannot penetrate through water, GNSS cannot be used directly for the determination of the coordinates of a point below the surface of water.

2. Use of GNSS for Underwater Scuba Diving

In the sea waters around Singapore there are many sites below the surface of the sea water, where there are very beautiful coral reefs. For recreational scuba diving, some simple methods can be devised to use GNSS to locate from a boat on the surface of the water, the positions of interesting places below the water.

This is done by using a GNSS receiver to produce first a map which can subsequently be loaded onto a GNSS receiver.

This map is produced by a scuba diver physically going under the water to locate a site of interest.

The scuba diver then comes up vertically to the surface of the water, and a GNSS receiver on a boat would record the longitude and latitude of the point on the water surface directly above the place of the site below the water.

The depth of the point below the water could be determined by a water pressure gauge carried on the scuba diver. A sonar depth finder can also be used to determine the depth if a boat is fitted with such an instrument.

A map with the coordinates of interesting places can then be compiled and subsequently loaded onto a GNSS receiver.

This map could enable people on a boat using a GNSS receiver, to reach a spot just above an interesting underwater feature, to enable a scuba diver to go directly to the feature under the water.

It would be particularly useful for night diving, since the use of GNSS could guide a boat directly to be just above the point of interest below the sea surface .

3. Cooperation

This map could be uploaded onto the internet for others to download it into a GNSS receiver. Over a period of time, cooperation between the users could result in inputs from various interested parties. The map could be expanded and refined with more details.

This map could be supplemented with photographs and text commentaries as well as with videos and voice commentaries.

4. Other Uses of GNSS Below the Water Level

This example of Cooperation in underwater Scuba Diving could be used to discuss several issues concerning cooperation in the use of GNSS below sea level.

About 71% of the earth's surface is covered with water.

In addition to coral reefs, there would be interest in

sunken ships and treasures.

disasters involving objects under the water.

changes in the marine environment due to human activities, and to natural seismic events.

5. Issues on International Coperation

As different users are likely to use different Satellite Systems and different signal receivers, some important issues would include:

(a) Compatibility and interoperability of the various GNSS, their corresponding receivers and the software used by different people

(b) Standardisation in the types of measurements to be made in the Environmental monitoring of the impact of human activities on marine life

(c) International legal and political issues since different countries could be involved

(d) Use of GNSS in conjunction with other techniques for Deep Sea Mining

6. Deep Sea Mining

In recent years, the United Nations International Seabed Authority (ISA) has granted over 27 contracts for deep sea mining on over one million square meters of seabed in the Pacific, Indian and Atlantic Oceans.

For such deep sea mining operations, GNSS would be used in conjunction with various other techniques such as ultrasonics and robotics.

One of the recent contracts has been awarded to a Singapore Company. The following is a statement released by the Singapore Company:

Ocean Mineral Singapore (OMS), majority owned by Keppel Corporation, has entered into a 15-year exploration contract for polymetallic nodules.

Sponsored by the Singapore government and supported by the Keppel-NUS Corporate Laboratory, OMS will be conducting environmental studies and surveys for deposits of polymetallic nodules within the approved area of 58,000 square kilometres at a site within the Clarion-Clipperton Fracture Zone of the Pacific Ocean.

7. Impact of Deep Sea Mining on the Environment

On 9 July 2015, in the Journal, Science, an article was published, bringing to attention the need to take into account the impact of deep sea exploration and mining, on the marine environment.

GNSS would need to be widely used in studies of issues related to deep sea mining and its impact on the marine environment.

8. Proposal for an IAA Study Group

It would appear that an appropriate new Study Topic related to GNSS could be proposed for consideration by the International Academy of Astronautics, Commission on Space and Society, Culture and Education.

This Study could be carried out over a period of 2 years with interested participants from various disciplines in different countries.

The Study could be conducted through a combination of physical meetings, video conferencing and email exchanges.

9. Proposal for a Workshop at Technical Forum of UNISPACE IV

It is hoped that at the future Fourth United Nations International Conference on the Peaceful Exploration and Use of Outer Space (UNISPAVE IV),

there could be a Technical Forum just as there was a Technical Forum at UNISPACE III

It may be recalled that at the UNISPACE III Technical Forum, there was a Workshop whose details are shown below:

PROGRAMME OF THE TECHNICAL FORUM OF THE UNISPACE III CONFERENCE

Workshop on Small Satellites at the Service of Developing Countries

Monday 26 July 1999, 14:30 - 17:30

Organizer: Subcommittee on Small Satellites for Developing Nations of the International Academy of Astronautics (IAA)

Chair: Pierre Molette, Chairman of the Subcommittee

Rapporteur: Mr. Yean Joo Augustine CHONG, Singapore

GNSS Based Time Transfer

Liu Yan Ying

National Metrology Centre
Agency for Science, Technology and Research (A*STAR), Singapore
liu_yan_ying@nmc.a-star.edu.sg

Abstract: Any time scale should be harmonised with Coordinated Universal Time (UTC) and it is an important task of a time keeping laboratory. To do this, all the laboratories carry on the time transfer work daily to synchronize or steer its clocks and maintain the live traceability to UTC, including the custodian of GPS time and GLONASS time. In this paper, we will first introduce the time transfer general concept and the technical background of remote atomic clock comparison, and then present the time comparison result from multi-technique observation: GPS and GLONASS C/A (Coarse Acquisition) code, GPS P3 (ionosphere-free code by a linear combination of P code measurement) and GPS PPP (precise point positioning). Then, we will show some clock comparison results from above techniques using different time links between our laboratory and other laboratories in Asia and Europe with baseline from 3000 km to 10000 km. The data analysis also shows that using hydrogen maser can help improve the short term stability of time transfer.

I. Introduction

The international system of units (SI) for measuring time is the second. The interval of one second is defined by the duration of 9,192,631,770 cycles of the standard Cs-133 atomic transition between two energy levels, which is realized in a Cs atomic clock. Although the atomic clock based on modern technique will be off by one second only after millions of years, every clock is subject to aging and its performance changes during its life time. As such, the international time reference UTC (Coordinated Universal Time) is determined not by some individual atomic clocks, but by the readings of about 400 atomic clocks maintained by 69 time laboratories worldwide[1]. Similarly, each time laboratory defines its national time scale UTC(k) using a small number of atomic clocks. To align UTC(k) with UTC, the difference has to be known precisely in order to apply a gentle adjustment. The time laboratories undertake the time transfer work on a daily basis to get the difference of UTC-UTC(k) also called time comparison by means of GNSS (global navigation satellite system). Since the launch of GPS satellites in 1980's, time transfer has been done mainly through GPS. In more than 20 years, the remote time comparison technique has been upgraded rapidly from single-channel code only observation to multi-channel observation for both codes (including C/A and P codes) and carrier phase. During the same period, the data analysis method was improved from original common view to all in view, as well as precise point positioning (PPP). This enables the time transfer uncertainty decreasing significantly from previous 10 to 20 ns to 5 ns or less. All these result in the picosecond precision for clock comparison at two remote locations. Meanwhile, another method called Two Way Satellite Time and Frequency Transfer (TWSTFT) has been developed and deployed for TAI (International Atomic Time) calculation and UTC generation. Although the TWSTFT provides a best time transfer precision over GPS and GLONASS in short term interval, it is only implemented as the main time transfer tool by thirteen

percent laboratories because this system requires expensive user equipment deployment and high link fee of telecommunication satellite.

1.1. Principle of time transfer

For GNSS, each satellite continuously transmits data that indicates its location and its internal time referenced to a common time. All satellites synchronize operations so that these repeating signals are transmitted at the same instant. However, the satellite signal arrival time received by the ground device is different because the transmission path is different. If four satellites can be viewed by a receiver, its coordinates can be calculated. In the other words, if a receiver's position and satellite positions are known accurately, we can measure the signal arrival time with respect to the local laboratory time from each satellite. If all time laboratories implement the same practice, the time difference of any two remote sites can be obtained by subtracting the measurement data and cancelling the common satellite time.

1.2. GPS time transfer

GPS satellites transmit information on two radio signals designated as L1 and L2 which are modulated by two pseudo random noise codes, C/A code (on L1) and P code (on L1 and L2) respectively. The signal contains three different information — a pseudo random code, almanac data and ephemeris data.

In the early years, time transfer was carried out by means of a time receiver and it was only limited to C/A code signal observation with a standard format CGGTTS (Common GPS GLONASS Time Transfer Standard), as described in Allan & Thomas [2] [3]. For this type receiver, each measurement data is obtained from a 13 minutes observation corrected for hardware path delays. As C/A code is transmitted on L1 carrier frequency only, the signal path delay due to the ionosphere and troposphere could not be corrected perfectly and the measurement uncertainty was between 10 ns to 20 ns.

Later, when modern geodetic type receiver became available for time keepers, the time laboratories were able to conduct an ionosphere-free code measurement P3 observation using data from both C/A and P codes of the satellites. Therefore, it greatly improved time transfer accuracy by an order of two magnitudes [4] [5]. This is because the new generation geodetic receivers collect satellite data on C/A and P codes from the two L1 and L2 carriers, which allows us to remove the frequency-dependent ionosphere delay mostly by means of information originated from the two carrier frequencies. In the calculation of UTC, a more precise Klobuchar model is in use to compensate the ionosphere path delay instead of the receiver's auto-correction [6].

As the time transfer receiver measures the time offset between the satellite signal arrival time and local laboratory time reference, the code measurement accuracy is largely limited by its chip rate. To improve the resolution and sensitivity, carrier frequencies measurement was then introduced in the area because they are up to one thousand times higher than the code repetition rate. With the help of geodetic science in an effort to determine a position with sub-centimeter accuracy, a precise point positioning or PPP method was developed and implemented in analysing the observables from a combined observation of code and carrier frequencies. This GPS PPP technique has shown great advantage over the others by giving the accuracy at a hundred pico-second level. Today, about fifty present of time laboratories have implemented GPS PPP technique and the international atomic time (TAI) and UTC are calculated using the time transfer data from PPP technique.

1.3. GLONASS time transfer

The time comparison at two remote sites via GLONASS is still under study. This is mainly due to the satellite availability. The GLNASS was not in full operation until recent years. Another factor which makes it not so easy to use is the signal complexity. Unlike GPS, GLONASS satellites transmit the same code at different frequencies, a technique known as FDMA, for frequency division multiple access. This will cause different signal travel time delays when going through the receiver and the antenna, and then introduce frequency-dependent bias in the observational data. Some study has been done by Bureau International des Poid et Mesures (BIPM) [7] [8] and the solution is to make correction using precise satellite orbit and clock products from ESOC (European Sapce Operations Centre) of IGS (International GNSS Service) analysis centre for determining and removing the bias.

1.4. Common view and all in view

Before 2000, all time laboratories compared their local time references using GPS C/A code common view (CV) method. That means all laboratories receive only time signals transmitted from those commonly available satellites. This is done by configuring the time receiver to follow a tracking schedule. Therefore, the number of satellites commonly in view is only a few and the number of data collected during one day tracking is also quite limited. The precision by this technique is several nanoseconds for long baseline stations. It has been mentioned that compared to C/A code time receiver, the dual frequency geodetic type GNSS receiver provides more information as it receives not only C/A code, but also P code and carrier phase signals simultaneously. The more information we have the better precision we can achieve. As the deployment of geodetic receivers for time transfer, CV was phased out and all in view (AV) was developed. AV allows signal from all satellites in view to be collected and the increase of amount of data is very helpful to reduce the noise and improve the positioning and time accuracy.

TABLE 1 ASSOCIATED INFORMATION FOR THE TIME LABORATORIES

Lab TAI Acronym /Country	Receivers Type	Time Source & Reference Clock Type	Time Links Involved	Baseline between SG and Lab km)
SG Singapore	PolaRx2eTR	UTC(SG) H-maser	GPS PPP/P3	–
	AOS TTS-3		GPS C/A code GLONASS C/A code	
NICT Japan	PolaRx2 TR	UTC(NICT) H-maser	GPS PPP/P3	5310
TL Taiwan	Ashtech Z-XII3T	UTC(TL) H-maser	GPS PPP/P3	3060
PTB Germany	Ashtech Z-12T	UTC(PTB) laboratory Cs	GPS PPP/P3	10120
	AOS TTS-3		GPS C/A code GLONASS C/A code	

II. Observational data

In this section, we present time transfer comparison results between National Metrology Centre (hereafter SG) and other time laboratories, using GPS/GLONASS codes and carrier frequency time transfer techniques. A brief summary on all involved laboratories, their equipment [9] as well as the baseline distance is listed in Table 1. The laboratories selected maintain a local atomic

time scale realized by an ensemble of H-maser and Cs atomic clocks, which usually gives better frequency stability than time scale based on a single-running clock.

In order to conduct time transfer, all the receivers are connected to the same 1 PPS signal disseminated from local UTC(k). The required reference frequency is also from the same master clock. Here, we present the time link results obtained from GPS/GLONASS C/A code, GPS P3 and GPS PPP observations. All observational raw data used in this paper are from GPS all in view and GLONASS common view observation and have been processed with the IGS (International GPS Service) products to determine the geometric distances between satellite and the laboratories, as well as the satellite clock synchronization errors. Each point in GPS/GLONASS C/A and P3 data indicates the observation result from a standard 13 min tracking, while it represents a 300 s result in GPS PPP observation.

It is well known that hydrogen maser standard is much stable than a cesium atomic clock at time intervals shorter than 10,000s and thus a combined time scale including maser and cesium standards is always used by the time laboratory to derive a best performance of UTC(k) in short and long terms. This is also a commonly adopted method by those time laboratories that maintain a local atomic time scale [1]. We will verify this by showing the time link comparison charts below.

2.1 Time Transfer Through GPS and GLONASS C/A Code

The first time link comparison is the C/A code comparison between PTB and SG using data from a time receiver. Both laboratories use same type TTS-3 receiver for code time link. This time receiver is a dual frequency multi channels GPS and GLONASS code receiver. Using identical receiver produces less noisy time transfer result. Here, we used CV analysis method for GLONASS data because there is no information on its satellite clocks from IGS product and it is impossible to make precise correction for those data from non-common satellites. The comparison is made using data from two different periods as described in Figure 1 caption. The results are presented in the same figure. For time interval longer than half day, there is no obvious difference between GPS and GLONASS code time transfer. However, for time interval shorter than that, GPS C/A code time transfer shows better performance than GLONASS C/A Code. This is because the GLONASS data is more noise than GPS data, as clearly shown in Figure 2. The same finding applies to both cases that either a cesium clock or a hydrogen maser is used as the reference clock.

2.2 Time Transfer Through GPS PPP and P3

GPS carrier phase measurement is about hundred times more precise than the GPS code data and less sensitive to propagation multipath and allows a better estimate of the atmosphere effects. Receivers capable of doing this measurement are geodetic type receivers and are commonly used in many laboratories. This type receiver is able to measure GPS dual frequency carrier phase and P code signals, and hence can produce carrier phase and P3 observational data for timing application. Due to this reason, a number of time laboratories use this type receiver either as a backup for TWSTFT (Two Way Satellite Time and Frequency Transfer) link or main TAI (International Atomic Time) link by means of PPP (Precise Point Positioning) processing techniques. The laboratories involved in this paper are all equipped with such a geodetic receiver. In this section, we present the time transfer performance comparison result using data from PPP and P3. The Figures 3 to 5 are the comparison results from the link of TL-SG, NICT-SG and PTB-SG respectively.

TABLE 2 GPS/GLONASS DATA PERIODS USED FOR THE TIME TRANSFER COMPARISON BETWEEN DIFFERENT LINKS.

Time Link	Link Type	Data Periods
SG—PTB	GPS C/A code AV GLONASS C/A code CV	MJD = 56347 to 56380 (24 Feb. – 29 Mar. 2013) (Cs as master clock) MJD = 56832 to 56901 (24 Jun. – 1 Sep. 2014)
	GPS P3 & PPP AV	MJD = 57167 to 57201 MJD = 57167 to 57201 (25 May 2015 – 28 Jun. 2015)
SG—NICT	GPS P3 & PPP AV	MJD = 56347 to 56380 (24 Feb. – 29 Mar. 2013) (Cs as master clock) MJD = 57167 to 57201 (25 May 2015 – 28 Jun. 2015)
SG—TL	GPS P3 & PPP AV	MJD = 57167 to 57201 (25 May 2015 – 28 Jun. 2015)

III. Analysis and Discussion

In Figures 1, 3, 4 and 5, we show the direct time comparison results on frequency stability in terms of Allan deviation between SG and other laboratories based on TTS-3 time receiver and other type geodetic receiver respectively. To make comparison easily, time link results UTC-UTC(SG) calculated from BIPM Circular T are also presented in the same figures.

In Figure 1 and 2, we first display a PTB-SG link comparison between GPS code and GLONASS code time transfer. Obviously, observational data from GPS satellites is quieter than the one from GLONASS in short intervals and it produces a more stable time transfer result. Both systems show the same performance for long term intervals. This can be due to the fact that GPS data is obtained from all in view satellites and the noise is normally reduced through the averaging. But for GLONASS, only data from commonly seen satellites is used and the noise cannot be removed largely by the less sample.

The charts in Figure 3 to 5 show that GPS PPP time transfer technique is superior to GPS P3 at short term, especially at averaging time shorter than 40,000s, and the corresponding stability is at least a factor of 3 lower than that of GPS P3.

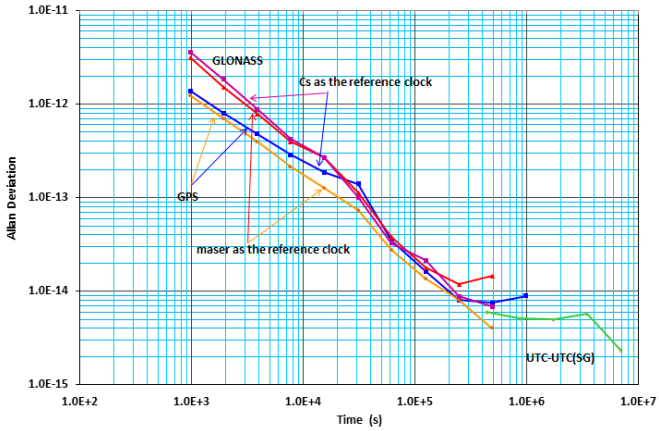


Figure 1. The time link stability comparison of C/A code between SG and PTB with identical receivers. Blue: GPS AV link from with a cesium as the reference clock; yellow: GPS AV link with a maser as the reference clock; purple: GLONASS CV link with a cesium as the reference clock; red: same as the purple one but with a maser as the reference. The green curve represents the stability of UTC-UTC(SG) [1]. Data interval for cesium as the reference clock: MJD=56347 to MJD=56380; data interval for maser as the reference clock: MJD=56832 to MJD=56901.

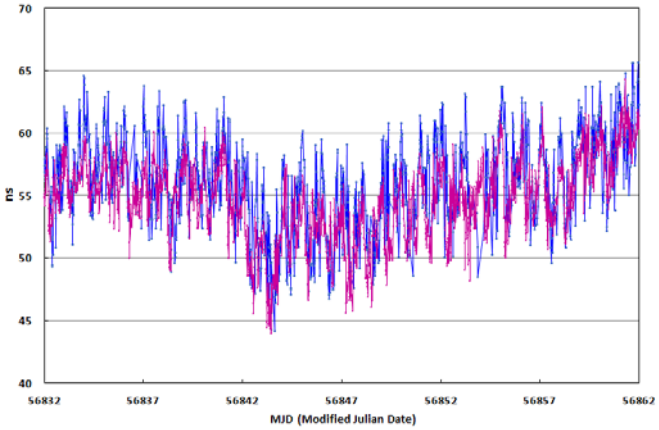


Figure 2. Time difference between UTC(PTB) and UTC(SG) calculated from C/A code links of GPS AV (purple) and GLONASS CV (blue) for one month period.

However, the measurement result from different links shows that the time transfer performance between GPS P3 and GPS PPP are quite similar at long term interval. This is also valid for SG and PTB link, which has the longest baseline distance, the time transfer stability is the same as the results from other two shorter baseline links. This can be explained by the fact that the GPS PPP method makes use of the carrier frequency observation which has much higher data acquisition rate (300s) and produces better stability as shown in Figure 6; however, long term accuracy and stability are dependent on the code measurement which is much nosier and varies by various factors like, multipath, hardware stability and atmosphere conditions etc.

It is also found that time transfer using maser as the reference is always better than that based on a cesium reference no matter which transfer technique in use.

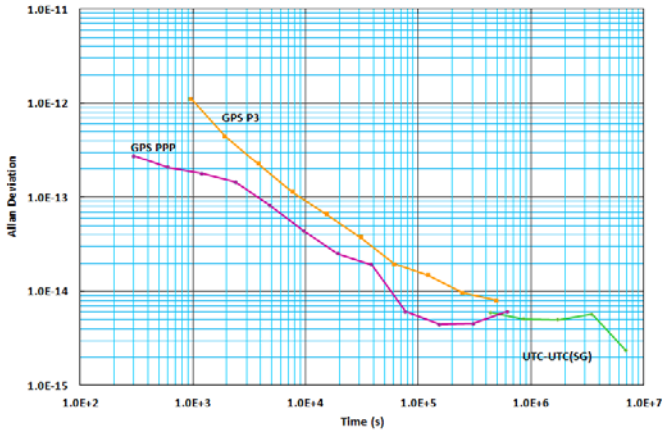


Figure 3. The time link stability comparison of GPS P3 code and PPP (carrier phase) time link between TL and SG using geodetic receivers. Results from GPS P3 and GPS PPP are indicated with different colours. Time scale stability of UTC-UTC(SG) is presented in green.

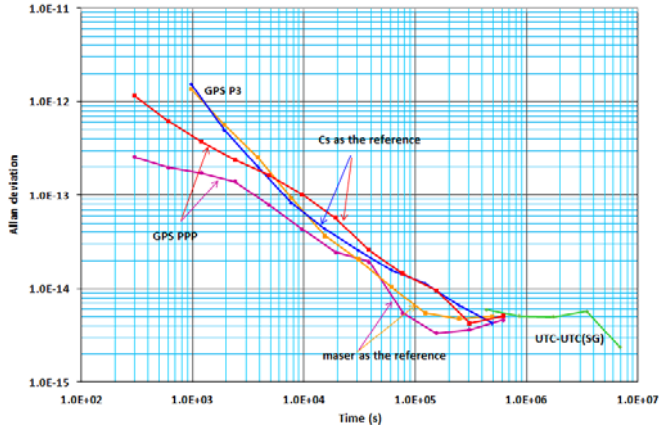


Figure 4. The time link stability comparison of GPS P3 code and PPP (carrier phase) between NICT and SG using geodetic receivers. Results from GPS P3 and GPS PPP are indicated with different colours. Yellow: GPS P3 link with a maser as the reference clock; red: GPS PPP link with a cesium as the reference clock from MJD=56347 to MJD=56380; purple: same as the red one but with a maser as the reference. The green curve represents the stability of UTC-UTC(SG).

IV. Summary

We have presented time link comparisons between SG and other timing laboratories through GPS C/A AV, GLONASS C/A CV, GPS P3 and GPS PPP techniques using at least one month observational data with a master reference of maser and cesium atomic clock. The measurement and analysis show that the performance of GPS C/A code time transfer is better than GLONASS method; both GPS P3 and PPP are superior to the C/A code observation, but only GPS PPP produces measurement result with lowest noise and gives the most precise solution than the others in regardless of the baseline distance of the two sites.

Actually, there is a lot potential to maximize the code and carrier frequency signal availability of GLONASS for time transfer application, especially with their full operation in future. The use of other alternative satellites in addition to GPS will results in an increased number of satellites in view and help us reduce signal acquisition time, improve positioning and time accuracy as well as mitigate the problems caused by obstruction etc. Therefore, BIPM has been doing comprehensive

study to include GLONASS observation in the TAI computation [10]. The experience will be vital to incorporate other upcoming GNSS such as Galileo and Compass for a combined measurement and analysis approach because increased number of measurement data provides more precise solution and maximum the benefit of TAI.

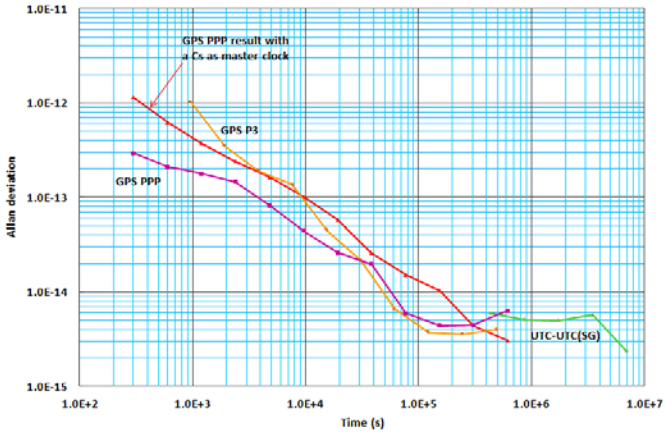


Figure 5. The time link stability comparison of GPS P3 code and PPP (carrier phase) time link between PTB and SG using geodetic receivers. The colors and indicators carry the same meaning as above.

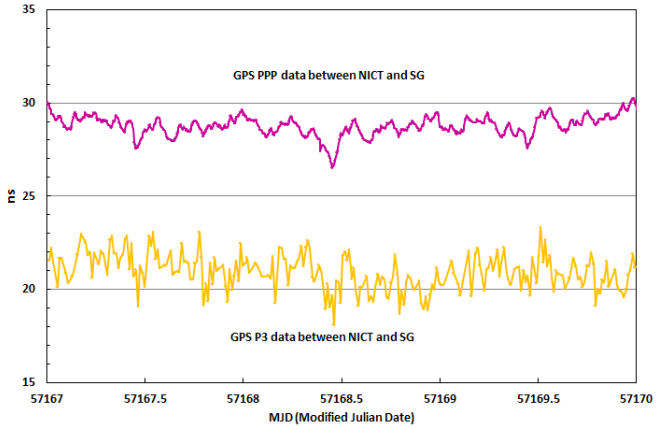


Figure 6. Time difference between NICT and SG calculated from GPS P3 and GPS PPP link between 25 and 28 May 2015. GPS PPP data is shifted 10 ns purposely for visual comparison.

References

[1] BIPM Circular T, http://www.bipm.org

[2] Allan, D W & Thomas C, Technical directives for standardization of GPS time receiver software, Metrologia 31, pp 69 1994

[3] Azoubib, J., Lewandowski, W., 1998, CGGTTS GPS/GLONASS data format Version 02, 7th CGGTTS meeting

[4] Defraigne P & & Petit G, Time Transfer to TAI using geodetic receivers, Metrologia, 40, pp 184, 2003

[5] J. Ray, K Senior, “Geodetic techniques for time and frequency comparisons using GPS phase and code measurements”, Metrologia 42, 2005, pp. 215-232.

[6] G. Petit, Z. Jiang, “Precise point positioning for TAI computation”, Internationa Journal of Navigation and Observation, Vol. 2008

[7] Jiang and W. Lewandowski, Accurate GLONASS time transfer for the generation of the Coordinated Universal Time, International Journal of Navigation and Observaion, Vol. 2012

[8] R. Dach, S. Shael etc, Combined multi{system gnss analysis for time and frequency transfer, 20th European Frequency and Time Forum (EFTF), 2006

[9] BIPM annual report on time activities, Vol. 2, 2008.

[10] Aurelie Harmegnies etc, Combined GPS and GLONASS in all-in-view for time transfer, Metrologia 50, pp 277, 2013

S/X Dual-band Dual-polarized Microstrip Antenna Array for Satellite Applications

Binyun Yan[1], Zhen Hu[2], Weixing Sheng[1], Yubing Han[1],
Yongpeng Kang[2], Yutao Yang[2]

(1 School of Electronic and Optical Engineering, Nanjing University of Science and Technology
2 Shanghai Institute of Satellite Engineering)

Abstract: A new design of S/X dual-band dual-polarized shared-aperture micro-strip antenna array is investigated for satellite applications. The operating frequencies are at the S band and X band respectively. For the whole antenna array, stacked-patch configurations were used to simplify the antenna and meet the bandwidth requirements of the two bands. For well aperture coupling, each S-band element is perforated to place four X-band elements within them. Transmission line feed was adapted in S band. Four symmetric orthogonal slots of u shape was used for slot coupling for X-band to broaden the band and minimize the cross-polarization. To achieve maximum gain in the main wave direction, all the X-band elements should accept feeding at the same phase and power, so that a new 8-way series-parallel mixed feed was generated for the 8x8 array of X elements. Test results are shown with good correlation between measured and predicted results, validating the design approach used.

Index Terms-Spacecraft payload, Array antenna, Dual-band Dual-polarized antenna, Shared-aperture antenna

I. INTRODUCTION

As an initiatively microwave imaging sensor, spaceborne SAR (Synthetic Aperture Radar) can gather information through all-day and all-weather imaging. With the development of SAR techniques, spaceborne SAR will realize higher resolution and wider swath. In the future, SAR system is heading to work at multi-mode and complex environment with lower cost and lighter spacecraft payload.

Antenna is playing an important role in the airborne SAR system. As reported in [1], to meet the increasing demands to enhance SAR system performance, relative techniques with low profile, multi-band, multi-polarization and structure simplicity need to be considered.

Dual-band dual-polarized (DBDP) shared-aperture microstrip antenna arrays have raised active research. In [2], a novel structure with half perforation is proposed to achieve a dual-band dual-polarized antenna element design, working at L and C bands. The element is aim to broaden the impedance bandwidth in the lower band with low profile. In [3], by adjusting the spacing between two slot antennas in series, an effective secondary resonance can be obtained. In the approximate circuit model, theory of mutual coupling is adopted and the results show similarity to the simulations. According to the antenna array in [4], active phased array antenna (APAA) can achieve smaller size and lighter weight. The array works at X band with microstrip line feeding and Ku band using the aperture coupled strip-line feeding. To minimize the size of the whole system, active channel blocks (ACBs) were directly combined with band pass filters (BPF) behind the radiators, which brought 25% off the size of the system. During the design of DBDP antennas, the compact

low-profile design investigated in [5] for SAR applications show good performance at C band but a little worse in L band. The new perforated stacked structure brought enlargement of bandwidth. Except that, appropriate arrangement of patches of different size is another validated method. A novel kind of DBDP common-aperture antenna is shown in [6], the part work at 60GHz radiated through the grid over the patches of 900MHz. In [7], sparse sub-arrays, operating at adjacent frequency ranges, are interleaved on a shared aperture, which increased the total bandwidth and reduce the mutual coupling.

All the designed antennas above considered, DBDP antennas can collect richer information in SAR system. Shared-aperture microstrip antenna arrays are capable of reliving spacecraft payload. In the most of the past research, big antenna arrays were usually investigated by theory and simulated by HFSS but still less fabricated. The work described in this paper focus on a DBDP shared-aperture antenna array working at S band and X band and the results are given.

The paper is organized into as followed. In Section II, the basic performance requirements for the array and the whole architecture are shown. Section III provides details about configuration of each part. In Section IV, measurement of the array is compared with theoretical model predictions and conclusions are given in part V.

II. Design Requirement

According to the working characteristics of SAR system, the requirements of the antenna array is provided in Table I.

TABLE I Design Parameters for Antenna Array

Parameters	S Band	X Band
Frequency	2.2GHz	8.5GHz
BandWidth	40MHz	20MHz
Polarization	Dual-linear	Dual-linear
Scan Range	>40°	>10°

To realize the shared-aperture work, each S band unit is rounded by 16 X band elements. As a consequence, the feeding and aperture-coupling needs to be solved. To obtain a design with miniaturization and low-profile, stacked symmetric configuration and slot-feeding method are utilized.

Stacked configuration is beneficial to broaden bandwidth and symmetric architecture is for minimizing the cross-polarization. Adjusting the space between the patches can reduce the coupling over the aperture.

The design in the paper uses ArlonAD270 substrate. The substrate dielectric constant is $\varepsilon_r = 2.7$ and the thickness of every layer is $d = 2.5\ mm$

All the works are based on three layers of substrates as shown in Fig. 1. The first layer are patches of two bands and under that is slot for feeding of X band. The third layer is covered with feed network for X-band and ground plane.

III. Antenna Configuration

IV. S/X Patches

Because the perfect radiation characteristics are under orthogonal polarization, patches of two bands are set to be square. After optimization based on calculation, the

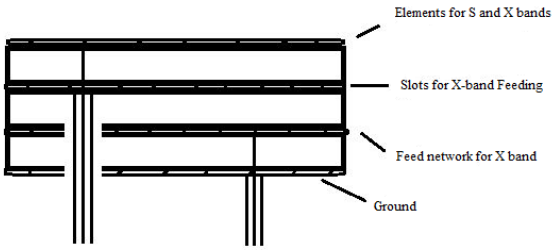


Figure 1. Geometry of stacked antenna array.

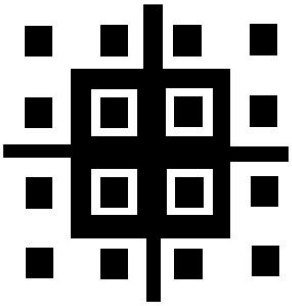


Figure 2. Geometry of the perforated S-band patches with 16 X-band patches.

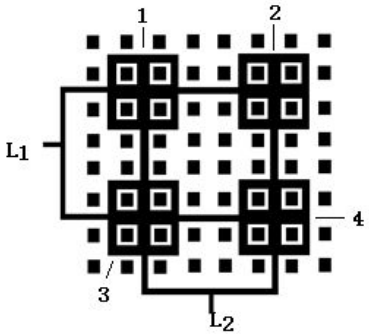
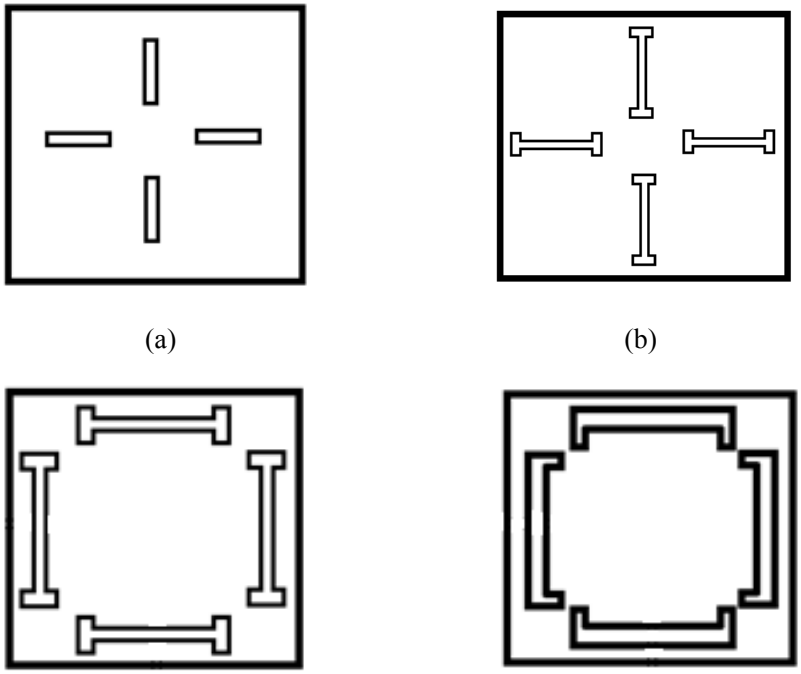


Figure 3. Geometry of feed network for S band with S/X-band patches.

length of S-band element is chosen as $l_s = 43.3$ mm and the length of X-band element is $l_x - 10.4$ mm .

V. FEED NETWORK

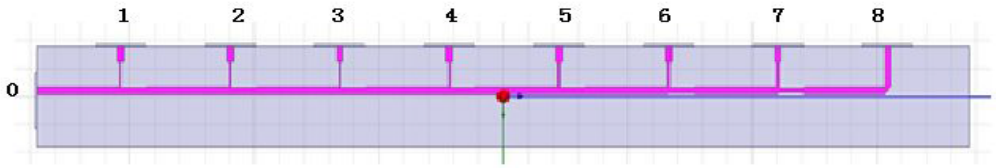
For the antenna array, feed network is more necessary than the elements as reported in [8]. Proper connection between elements to make the array work effectively is the function of feed network. The network is designed as simple as possible to function which means every element obtained appropriate excitation phase and amplitude, minimizing the loss by matching correctly.



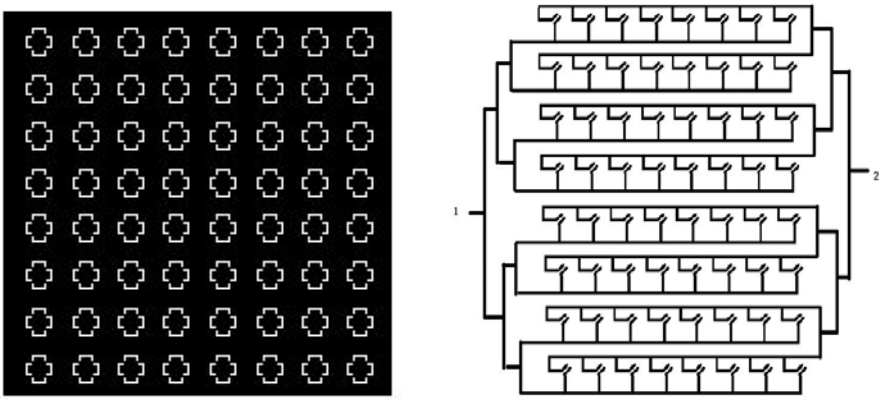
(a) (b)

(c) (d)

Figure 4. (a) Geometry of X-band patch with ordinary slots near the center. (b) Geometry of X-band patch with I-shape slots near the center.
(c) Geometry of X-band patch with I-shape slots near the side. (d) Geometry of X-band patch with U-shape slots near the side.



(a)



(b) (c)

Figure 5. (a) 8-way power divider model.
(b) Geometry of ground plane with slots. (c) Geometry of feed network for 8*8 X-band subarray

The whole antenna array in the paper is a 2*2 array for S band and 8*8 array for X band. The feed for S band is relatively simple.

The design requires complex feed network for X band to realize dual-linear polarization on the same plane. Orthogonal perforation are used to obtain slot-microstrip coupling and two symmetric slots are added to minimize the cross polarization. Four different perforations are built to check the influence on the performance of the element. According to the simulation results, I-shape and U-slot can both provides wider bandwidth. The size of X-band element is small for which resonant mode appears if slots are too close. As a result, the last one, the novel U-shape slot, shown in Fig. 4, is chosen as the final design.

For 8*8 array of X band, each element is supposed to be driven at the same phase and amplitude, which can't rely on ordinary feeding network. An 8-way power diver is investigated to combined to be the whole network. The 8-way power divider is shown in Fig. 5. The power is input from port0 and divided to port1-8. It's developed from 2-way and then 3-way dividers step by step.

VI. RESULTS

To validate the design, prediction of the antenna performance from HFSS 15.0 commercial software and test results of the fabricated array is compared as below. Considering the stacked configuration of the antenna, three layers are processed separately and bounded together by conductive and insulated adhesive.

Fig.6 shows the S11 parameter of the array in two bands are reaching similar results between simulation and measurement. The working frequencies are 2.2GHz at S band and 8.5GHz at X band. The bandwidth of S band reaches 40MHz in prediction but 60MHz in measurement, which is nearly 20MHz of X band. Actually, the performance of each element is at a high level, such as S11 parameter of the S-band element is about 140MHz but about 41MHz of the whole array.

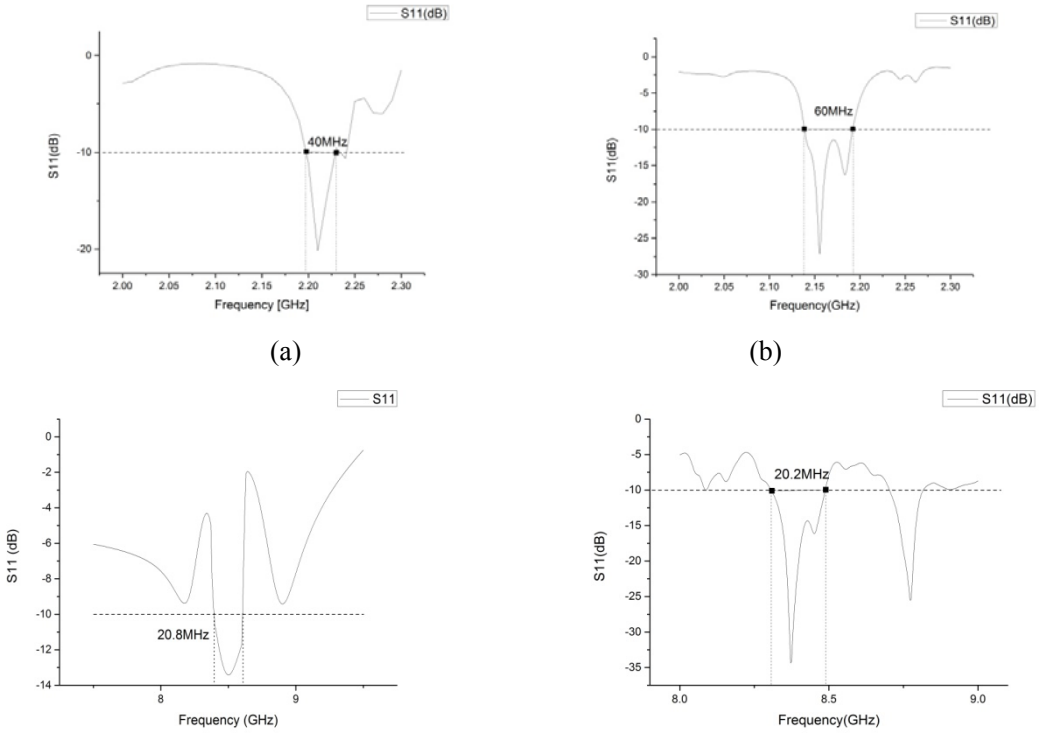


Figure 6. (a) Predicted return loss patterns of the S-band subarray. (b) Measured return loss patterns of the S-band subarray.(c) Predicted return loss patterns of the X-band subarray. (d) Measured return loss patterns of the X-band subarray.

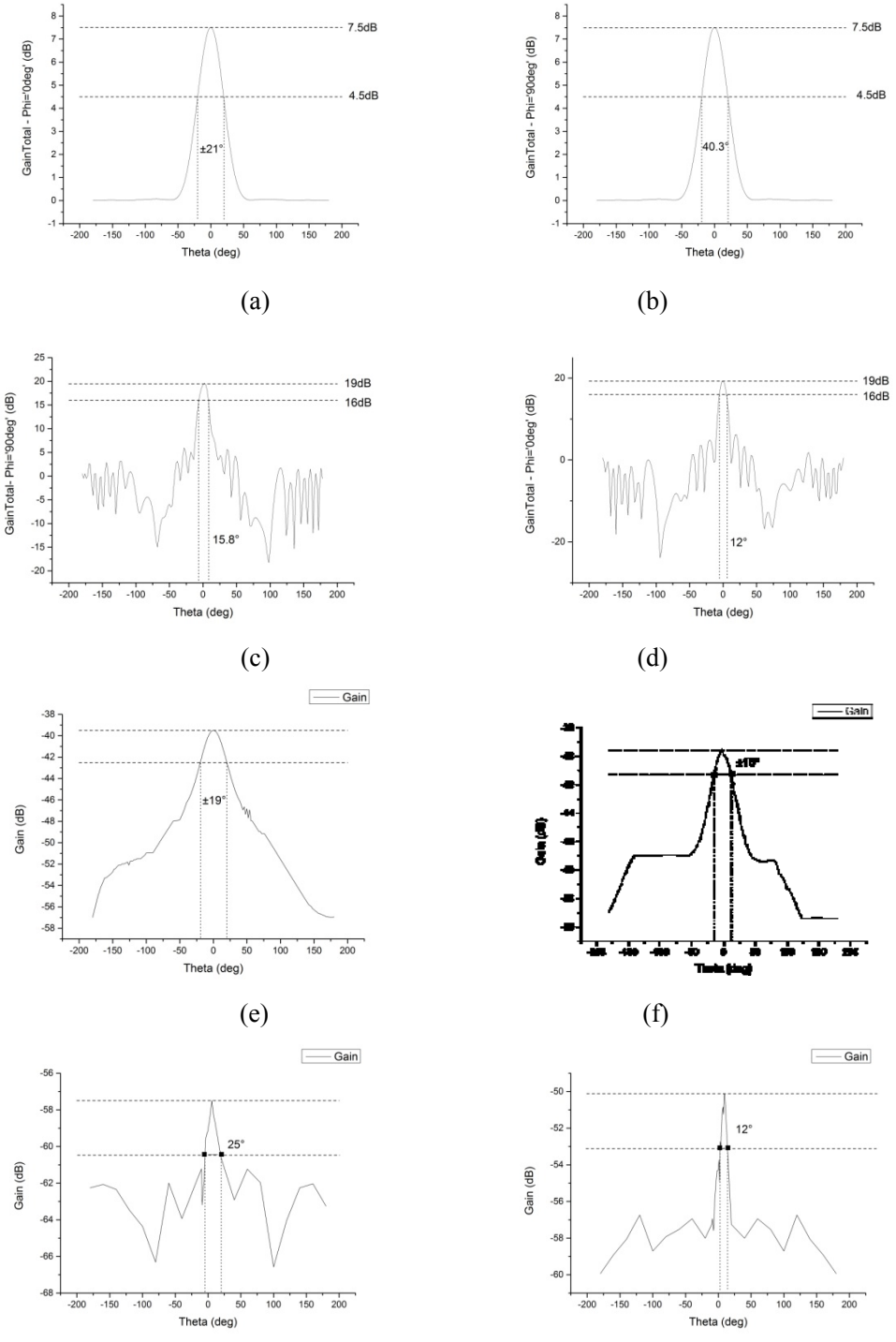


Figure 7. (a) Predicted radiation pattern on E-plane of the S-band subarray. (b) Predicted radiation pattern on H-plane of the S-band subarray.

(c) Predicted radiation pattern on E-plane of the X-band subarray. (d) Predicted radiation pattern on H-plane of the X-band subarray.

(e) Measured radiation pattern on E-plane of the S-band subarray. (f) Measured radiation pattern on H-plane of the S-band subarray.

(g) Measured radiation pattern on E-plane of the X-band subarray. (h) Measured radiation pattern on H-plane of the X-band subarray.

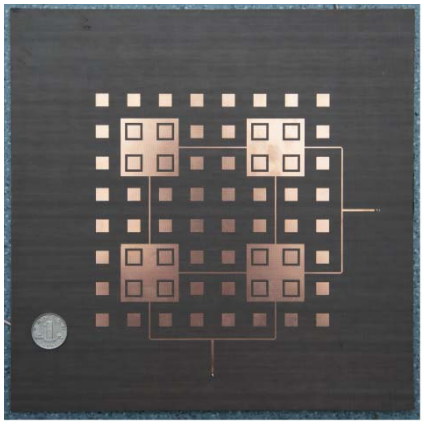


Figure 8. Fabricated antenna array

Obviously, the parameters of the whole array are poorer than the elements, because of coupling and other effects during the big array.

Reviewing the radiation patterns in Fig. 7, the results show good correlation with simulation. Scan angle on E-plane of the S-band subarray is about 40°and 15°at X-band subarray. The angle of the H-plane is smaller. Resonant modes appeared in X-band evidently because the X-band subarray contains 64 elements, arranged closely, but affecting the radiation characteristic at the working frequency tinily.

Fig. 8 shows the up face of the fabricated array. The whole size of this array is 280mm*280mm*1.5mm. The process of bounding is accomplished by hand. To avoid the influence on the geometry of the microstrip line, adhesive is only applied on the edge of each layer, which leads to interspace between layers. Meanwhile, adhesive can change the thickness of some area of the layers. Apparently, hand-mading causes variation on the whole structure, especially the location of the X-band feed network and the slots. According to the analysis above, errors exist absolutely, generating the difference between the estimation and measurement. These errors exit but are small. Generally, the results are in accordance with the predicted performance.

VII. CONCLUSION

A DBDP antenna array for satellite applications is developed and demonstrated. In the design, a novel feeding method for X band is investigated for the design of the S/X DBDP antenna array. U-shape slot and 8-way power divider combined perfectly to suit the stacked configuration. The predicted and test results of parameters of the array are compared. Test results were shown in good correlation between measured and predicted results, validating the design. Reasons of errors between them are still investigated by analyzing the process of the fabrication of the antenna array.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China under grant 61471196.

REFERENCES

[1] G. Krieger and A. Moreira, "Spaceborne bi- and multistatic SAR: potential and challenges," IEE Proceedings - Radar, Sonar and Navigation, vol. 153, pp. 184-198, 2006.

[2] C.-X. Meng, W.-H. Shen, Z. Sun, S.-S. Zhong, and K. P. Esselle, "A Novel L/C-band dual-polarized half-perforated antenna element for SAR arrays," ed.
[3] M. N. Mahmoud and R. Baktur, "A Dual Band Microstrip-Fed Slot Antenna," IEEE Transactions on Antennas & Propagation, vol. 59, pp. 1720-1724, 2011.
[4] Y. B. Jung, S. Y. Eom, and S. I. Jeon, "Experimental Design of Mobile Satellite Antenna System for Commercial Use," Consumer Electronics IEEE Transactions on, vol. 56, pp. 429 - 435, 2010.
[5] L. L. Shafai, W. A. Chamma, M. Barakat, P. C. Strickland, and G. Séguin, "Dual-band dual-polarized perforated microstrip antennas for SAR applications," Antennas & Propagation IEEE Transactions on, vol. 48, pp. 58-66, 2000.
[6] W. Menzel, M. Al-Tikriti, and M. B. Espadas Lopez, "Common aperture, dual frequency printed antenna (900 MHz and 60 GHz)," Electronics Letters, vol. 37, pp. 1059-1060, 2001.
[7] C. I. Coman, I. E. Lager, and L. P. Ligthart, "The design of shared aperture antennas consisting of differently sized elements," Antennas and Propagation, IEEE Transactions on, vol. 54, pp. 376-383, 2006.
[8] M. Yousefbeigi, A. Enayati, and M. Shahabadi, "Parallel–series feed network with improved G/T performance for high-gain microstrip antenna arrays," Electronics Letters, vol. 44, pp. 180 - 182, 2008.

Spaceborne Ka/G Dual-frequency Cloud Profile Radar Discussion

Ping Wang[1], Biao Wang[1,2], Xia Ding[1], Yongjin Han[1],Haitao Wang[1]

(1. Shanghai Radio Equipment Research Institute, Shanghai 200090, China
2. National Key Laboratory of Electromagnetical Environment Research, Shanghai 200438, China)

Abstract: Based on the current cloud remote sensing problem, we proposed a new remote sensing technique --- spaceborne terahertz cloud profile radar(CPR) in this paper, the design scheme is described, as well as its application prospect. Nonspherical scattering method and meteorological radar equation are given, which are the basis of backscattering cross section of ice cloud and radar. We have given the result of hexgonal column ice crystal and ice cloud composed by hexgonal column ice crystal in Gamma particle size distribution backscattering cross section in 94GHz and 215GHz. In addition, the ice-cloud retrieval algorithm based on a dual frequency radar is discussed, 215GHz is much more suitable for detecting ice cloud than 94GHz.

Keywords: ice crystal; hexagonal column; ice cloud; terahertz; cloud profile radar

1 Introduction

Clouds generally exist in the atmosphere, covering about 60%~70% of the earth's surface. Cloud significantly influence the Earth's short-wave and long-wave radiation budget through scattering, absorption and emission, and also constitute the major source of uncertainty in Earth's climate change and sensitivity. The relevant parameters of the cloud must be introduced in the global climate system. However, due to the complexity of the cloud composition, including water droplets and ice particles scale from microns to millimeters, microscopic and macroscopic physical structure of clouds determine the radiation properties. Due to the complexity of the convection and turbulence interactions, make it to be the major source of uncertainty in predicting climate sensitive. Nowadays, the difficulty of global climate model (GCM) and numerical weather prediction (NWP) depends not only on the experience and estimation but also on the introduction of cloud parameters. In order to improve the function of the cloud in GCM and NWP, The reliable observation value is introduced to satisfy the model requirement, so a new quantitative detection method of cloud microphysical parameters need to be developed. Here we only focus on the active cloud profiling radar observations.

The selection of radar wavelength requires a balanced consideration of the size of the cloud particles and the atmospheric attenuation. For radar with the same resolution and transmission energy, the shorter wavelength, the smaller particles can be detected, but the smaller the radar wavelength will lead to a lower atmospheric transmittance. Generally speaking, the detection of cloud particles with a diameter of 5μm-10μm requires a high frequency electromagnetic wave. Currently, the most commonly used cloud detection band include 35GHz, 94GHz and 24GHz, etc., the platform includes the foundation, shipboard, airborne and satellite borne. These Cloud Radar has many other functions, such as multi frequency, velocity scanning and polarization, with high spatial and temporal resolution, greatly enhance the inversion accuracy of cloud profile information. However, the detection of current cloud remote sensing in boundary layer clouds and snow ice

clouds, etc. is insufficient [7]. Based on the atmospheric attenuation, scattering characteristics and properties of ice particles, as well as the dual frequency inversion technology, we will introduce THz cloud profiling radar in this paper.

The paper is divided into five chapters. Introduction briefly introduces the importance and shortcomings of cloud detection, and suggestions for the development of new cloud remote sensing means. Chapter one introduces particle scattering theory and the weather radar equation; the second chapter present terahertz band atmospheric attenuation, hexagonal ice particle scattering, hexagonal ice crystals scattering characteristics and dual frequency ice inversion containing terahertz band, and illustrate that terahertz band is suitable for smaller ice particles detection; according to the previous introduction the third chapter propose Ka/G space-borne radar concept and relevant parameters are presented; the fourth chapter summarize the full text of this paper.

2 Basis of Cloud radar design

2.1 Atmosphere window

In the introduction, the selection of the weather radar frequency must consider the atmospheric attenuation and the scattering characteristics of the target. Therefore, the meteorological radar design should take into account the atmospheric window of the corresponding band. In general, the atmospheric attenuation of electromagnetic wave increases with the frequency. Slocum (2013) presented the study of attenuation characteristics of terahertz wave, as shown in figure 1. It can be seen that the atmospheric attenuation at the same humidity increases with increasing frequency. And 215GHz is located at atmospheric window. The atmospheric attenuation of 215GHz is less than 0.2dB/km; when the relative humidity increased to 90%, the atmospheric attenuation is quite serious and over 10dB/km. Therefore, the frequency range is not suitable for the meteorological observation of the higher humidity regions. However, the band can be used for the observation of the low water content targets in the middle and high atmosphere on the satellite platform.

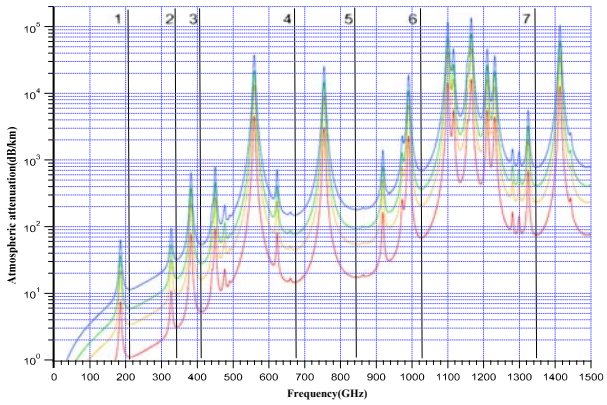


Figure. 1 The absorption coefficient of water vapor at different humidity of 296K

This scheme agree with that proposed by Illingworth(1999), who have studied the two-way attenuation characteristics of 35GHz, 79GHz, 94GHz, 140GHz and 215GHz wave at 4km, as shown in figure 2 and figure 3. The corresponding temperature and humidity profiles have also be illustrated. The discontinuousness of mid-latitude and tropic relative humidity profiles suggests that cirrus distribution can influence the relative humidity of the identical height.

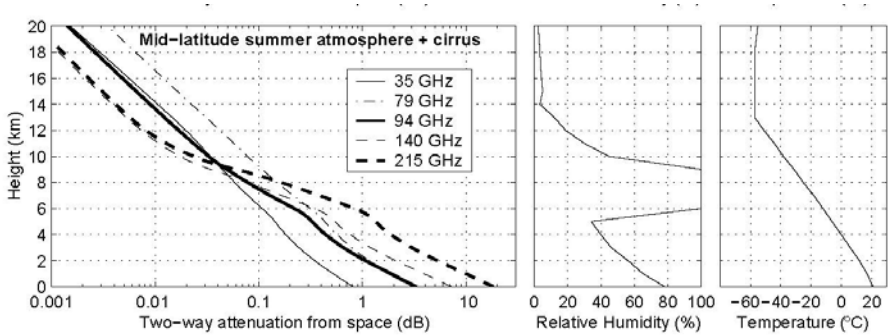


Figure 1 The two-way attenuation characteristics of various frequencies at typical mid-latitude summer situation (with cirrus)

The two-way attenuation characteristics of different frequencies at typical mid-latitude summer situation (with cirrus) and tropic situation (with cirrus) indicated that the maximum of 215GHz two-way attenuation from satellite platform to 4 km is about 2dB, approximately 9dB from satellite platform to 2 km and 22dB from satellite platform to ground. So 215GHz wave can be used for the cloud observation at the middle and high atmosphere on the satellite platform.

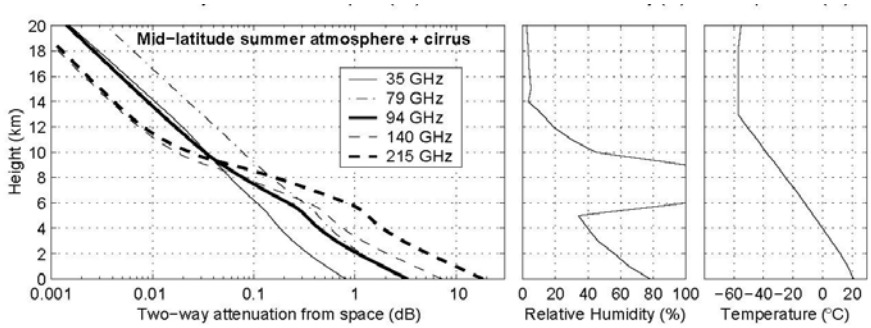


Figure 2 The two-way attenuation characteristics of various frequencies at typical tropic situation (with cirrus)

2.2 Effective radar reflectivity factor

The scattering characteristics of cloud particles can be described by range scattering matrix S while electromagnetic wave get in touch with cloud particles.

$$\begin{bmatrix} \boldsymbol{E}_{\mathrm{h}}^{\mathrm{s}} \\ \boldsymbol{E}_{\mathrm{v}}^{\mathrm{s}} \end{bmatrix} = -\frac{\mathrm{e}^{\mathrm{i}kr}}{\mathrm{i}\,kr} \begin{bmatrix} \boldsymbol{S}_{\mathrm{hh}} & \boldsymbol{S}_{\mathrm{hv}} \\ \boldsymbol{S}_{\mathrm{hv}} & \boldsymbol{S}_{\mathrm{vv}} \end{bmatrix} \begin{bmatrix} \boldsymbol{E}_{\mathrm{h}}^{\mathrm{i}} \\ \boldsymbol{E}_{\mathrm{v}}^{\mathrm{i}} \end{bmatrix} \tag{1}$$

Where $\boldsymbol{E}_{\mathrm{h}}^{\mathrm{i}}$ and $\boldsymbol{E}_{\mathrm{v}}^{\mathrm{i}}$ is the incoming electromagnetic field, $\boldsymbol{E}_{\mathrm{h}}^{\mathrm{s}}$ and $\boldsymbol{E}_{\mathrm{v}}^{\mathrm{s}}$ is the scattering electromagnetic field, the index h and v indicate horizontal and vertical polarization, respectively; k is beam, r is transmission distance of electromagnetic wave. S_{hh}, S_{vv} and S_{hv} are expressed as the scattering properties of matrix S, while S_{hh} and S_{vv} indicate the co-polarization properties, S_{hv} is the cross polarization properties. In view of the simple discussion of the cloud profile radar, we does not consider the issue of polarization. Taking into account the relation between backward scattering cross section and co-polarization properties of ice particle (as (2) shown) to study the ice particles and ice scattering characteristics.

$$\sigma_{\mathrm{hh}} = -\frac{4\pi}{k^2}\left|S_{\mathrm{hh}}\right|^2 \tag{2}$$

Assuming that the radar transmit horizontal polarized electromagnetic wave, receiving also adopts the horizontal polarization and ice clouds are composed of hexagonal ice crystals particles

(temporary not consider other morphology), using the discrete dipole approximation (DDA) method to study the single hexagonal particle scattering characteristics and backscattering characteristics. For ice clouds, it is composed of a large number of different scales of hexagonal particles, need to calculate the average backscatter section under certain distribution. Average backscatter section is given:

$$\overline{\sigma}_{\mathrm{hh}}=\frac{\int_{D_{\min}}^{D_{\max}}\sigma_{\mathrm{hh}}N(D)\mathrm{d}D}{\int_{D_{\min}}^{D_{\max}}N(D)\mathrm{d}D} \tag{3}$$

Where N (D) represents the size distribution of ice crystal particles; D_{max} and D_{min} are the largest and smallest size of D, respectively.

According to the relationship between the effective radar reflectivity factor (Z_e) and unit volume ice cloud radar scattering cross section, we study the effective radar reflectivity factor of hexagonal ice crystals.

$$Z_E=\frac{\lambda^4}{\pi^5\left|K_{\mathrm{w}}\right|^2}\int_{D_{\min}}^{D_{\max}}\sigma_{\mathrm{hh}}(D)N(D)\mathrm{d}D \tag{4}$$

Where λ is wavelength, D represent the particle size of ice crystals, $\left|K_{\mathrm{w}}\right|^2$ indicates the water dielectric factor.

2.3 Single base meteorological radar equation

The target of the weather radar is a 3D object, which needs calculate radar backscattering cross section. According to the ordinary radar equation (5), a single base meteorological radar equation can be transformed.

$$P_{\mathrm{r}}=\frac{P_{\mathrm{t}}G^2\lambda^2L_{\mathrm{r}}}{(4\pi)^3r^4}\sigma \tag{5}$$

Where P_t is the peak pulse power of radar, G is the antenna gain, λ is the wavelength, L_r is the two-way atmospheric attenuation, r denotes the distance between the target and radar, σ is the radar backscattering cross section of meteorological target.

Considering the relationship between the body target in a radar resolution unit and particle distribution (as shown in (6)).

$$\sigma=(c\tau/2)\theta\varphi r^2\cdot\int_{D_{\min}}^{D_{\max}}\sigma_{hh}(D)N(D)\mathrm{d}D \tag{6}$$

Where C is the electromagnetic wave propagation velocity, θ and φ are antenna beam widths.

Considering the relationship between radar reflectivity factor and the radar cross section of the target, and assume the radar beam is Gauss distribution, the single base meteorological radar equation can be written as:

$$\overline{P}_r=\frac{c\pi^3}{1024\ln(2)}10^{-18}\left(\frac{P_t\tau G^2\theta\varphi L_r}{\lambda^2}\right)\left(\frac{\left|K_w\right|^2Z_e}{r^2}\right) \tag{7}$$

Where $\overline{P}_{\mathrm{r}}$ is the average radar echo power, unit is W; C is the electromagnetic wave propagation velocity, unit is m/s; P_t is the peak power of radar transmitter, unit is W; τ is the time of pulse

emission, the unit is s; G is the gain of the antenna, unit is 1, θ and φ are widths of antenna beam, unit is rad; λ is radar wavelength, the unit is m; R is the distance between radar and cloud, unit is m, L_r illustrates the two-way atmospheric attenuation, unit is 1; $|K_w|^2$ is the meteorological target index, unit is 1; Z_e is the radar reflectivity, the unit is mm^6/m^3; 10^{-18} is the factor of international system of units; ln(2) is a factor due to the main beam integral of two-dimensional Gaussian distribution.

3 Non-spherical backscattering effects and retrieval methods

3.1 The backscattering properties of simple ice cloud model

3.1.1 The backscattering properties of hexagonal column ice crystal

The radar backscattering cross sections at 94GHz and 215GHz from DDA methods are computed. Hexagonal column ice crystal model is used in DDA methods, hexagonal column is divided into many cellular in Descartes coordinate showed in figure 4, where major axis of hexagonal column is same as the direction of electromagnetic propagation and is defined as Z axis, one minor axis is defined as Y axis (the arrow perpendicular to the paper in the right plane of figure 4), and the other is defined as X axis (the arrow parallel to the paper in the right plane of figure 4). It is not significant to distinguish Y axis and X axis, because hexagonal column is rotationally symmetric along its major axis.

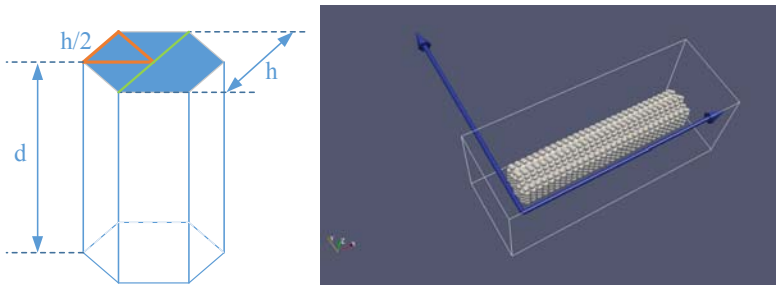


Figure 3 Hexagonal column geometrical model and cellular in DDA method

The dielectric factors for water are 3.359-1.929i at 94GHz and 2-0.8i at 215GHz. Horizontal polarization directs along major axis of hexagonal column and incidence direction of electromagnetic is 0°. After computed, the backscattering properties of single hexagonal column ice crystal at 94GHz and 215GHz is shown in figure 5.

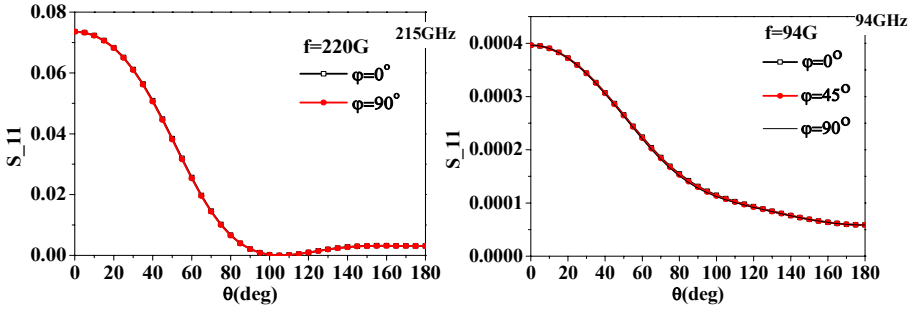


Figure 4 When electromagnetic incident direction is 0°, the scattering factor S_11 changes with scattering plane

S_11 is scattering factor when incident and scattering electromagnetic wave has the same polarization directing, θ is intersection angle between the direction of scattering electromagnetic wave and Z axis, φ is intersection angle between the direction of scattering electromagnetic wave and Y axis. From figure 5, you will find when θ is equal to 180° , the backscattering properties of

hexagonal column ice crystal at 215GHz is much bigger than at 94GHz. So 215GHz electromagnetic wave is more suitable for detecting small particle than 94GHz.

3.1.2 The backscattering properties of ice cloud

It is supposed that ice cloud is composed by a large number of hexagonal column ice crystals in gamma distribution which has a style as (8).

$$n(D_{\max}) = N_0 D_{\max}{}^{\mu} \exp(-\varepsilon D_{\max}) \tag{8}$$

Where, $n(D_{\max})$ is the number density of ice crystal particles with a D, N_0 is the intercept, $D_{\max}$ is crystal maximum size, μ is the dispersion usually ranging from 0 to 2, ε is the slope which is described by equation (9).

$$\varepsilon = \frac{b + \mu + 0.67}{D_{\text{median}}} \tag{9}$$

Where, $D_{\text{median}} = (D_{\max}/2)^3$. $b = 2.1$ is suitable for the tropical ice clouds and， $b = 2.3$ is suitable for the mid-latitude ice clouds. $n(D_{max})= 8.9336\times10^4/m^3$ and $N_0=3.33\times10^{-3}\mu m^{-1}$ has been chosen for gamma particle size distribution shown as figure 6.

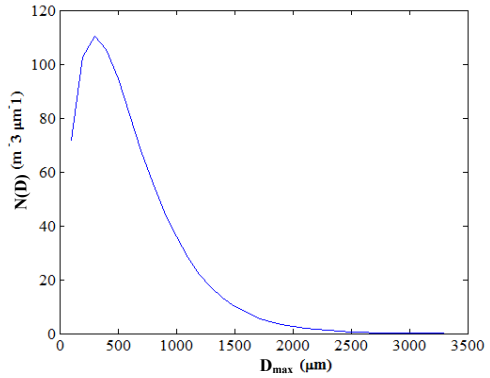


Figure 5 Gamma particle size distribution of ice cloud

Figure 7 has given the result of the backscattering properties of ice cloud constituted of gamma distribution hexagonal column at 94GHz and 215GHz. For each frequency, backscattering cross section is larger when D_{max} is larger. When D_{max} is fixed backscattering cross section at 215GHz is bigger than at 94GHz. So 215GHz is more suitable for detecting small ice particle.

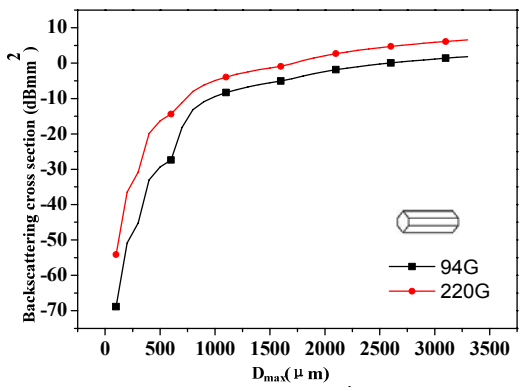


Figure 6 The backscattering properties of ice cloud as a function of maximum size

3.2 Dual-wavelength ratio method for ice cloud research

The use of dual-wavelength ratio (DWR) method for measuring mean crystal size in ice cloud was first proposed by Matrosov.The estimate of size can be more accurately than a single radar due to one radar in the Rayleigh regime while the other in Mie regime. This will lead to the sensitivity of one radar is different from the other, so this manner can be used to estimate mean crystal size. Figure 8 shown that DWR value is bigger for larger average particle size for the same pair and 35-215GHz combination has more potential for ice cloud research.

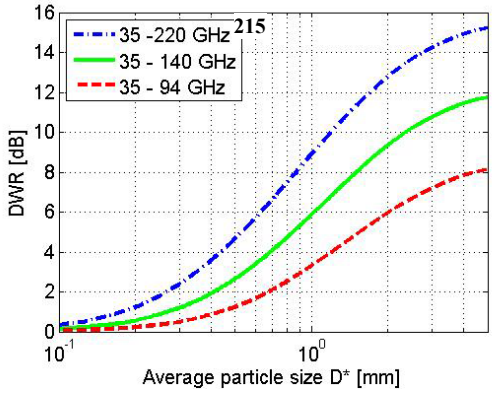


Figure 7 Dual-wavelength ratio as a function of average particle size for different pairs

4 The concept of spaceborne Ka/G dual-frequency cloud profiling radar

Dual-frequency radar included G-band has extremely potential applications of ice cloud research. Firstly, G-band radar has higher sensitivity than low frequency for the same condition. That is because reflectivity increases as the fourth power of frequency in the Rayleigh regime. Secondly, G-band is suitable for spacebonre platform due to its acceptable two-way attenuation. Lastly, the accuracy of cloud physical parameters retrieved including G-band is lower than any other low frequency. So we would like to propose Ka/G dual-frequency radar for observation ice cloud. As above, the baseline specifications of Ka/G dual-frequency radar are summarized in Table 1.

Table 1 Major design parameters of Ka/G dual-frequency cloud profile radar

Item	Specification	
Frequency	Ka-band(e.g.：35GHz)	G-band(e.g.：215GHz)
Altitude	550km	550km
Peak Power (@EIK)	3kW	1kW
Sensitivity	-10dBZ	-35dBZ
Vertical Resolution	250m	250m
Swath Width	25km	1km
Velocity accuracy	——	0.1m/s
Data window	0km ~25km	0km ~25km

5 Conclusions

The status and drawbacks of millimeter-wavelength cloud radar has been reviewed by many papers. Theoretically, for spaceborne single and dual-frequency radar, if ice clouds are primary observation targets then a G-band radar would be the preferred choice. In order to acquire mostly cloud profiling information, dual-frequency radar may be excellent scheme. At last, major design parameters of Ka/G dual-frequency cloud profile radar has been proposed.

References

[1] Solomon, S., Qin, D., Manning, et al. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [C]. Tech. rep., Cambridge University Press, Cambridge, UK and New York, NY, USA, 2007, 323: 326

Solomon, S., Qin, D., Manning, et al. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [C]. Tech. rep., Cambridge University Press, Cambridge, UK and New York, NY, USA, 2007, 323: 326

[2] Stevens, B. and Feingold, G.: Untangling aerosol effects on clouds and precipitation in a buffered system [J], Nature, 2009, 323(461): 607–613

[3] Stephens, G. L. Cloud feedbacks in the climate system: a critical review[J]. Climate, 2005, 323(18):237–273

[4] Lhermitte R., A 94 GHz Doppler radar for clouds observations [J]. Atmos. Oceanic Technol., 1987,4(1):36-48.

[5] Matrosov, S. Y. Possibilities of cirrus particle sizing from dual-frequency radar measurements [J]. Geophys. Res. ,1993,98(D11):343-344

Author: Ping Wang(1988-), Male, Master, Assistant Engineer, Engaged in terahertz active detection technology research.

E-mail: wangpingxixing@163.com

Study on RF MEMS Switch Matrix in Multifunctional Phased Array Antennas

Weiwei Zhou[1], Weixing Sheng[1], Zhen Hu[2], Xiaofeng Ma[1], Shihong Zhou[2]

(1 School of Electronic and Optical Engineering, Nanjing University of Science and Technology
2 Shanghai Institute of Satellite Engineering)

Abstract: According to the requirements of multifunction satellite applications such as low cost, low mass and small volume, this article conducted a study on the application of RF MEMS technique in multifunctional phased array antenna payloads. And the emphasis was put on the analysis of RF MEMS switch matrix configuration scheme for radio frequency aperture synthesis for communication payload and synthetic aperture radar (SAR) payload. A preliminary feasibility analysis of the proposed scheme was conducted, and at the same time the technical difficulties of applying RF MEMS switch matrix technique into multifunctional phased array antennas were pointed out together with other problems that should be solved before this technique comes into actual application.

Keywords: multifunctional payload; phased array antenna; RF MEMS; switch matrix

I. INTRODUCTION

The radio frequency microelectro-mechanical systems (RF MEMS) technique has drawn lots of attention since its born. Compared with conventional mechanical RF switch matrix, the switch matrix using RF MEMS technique has a massive reduce in both volume and mass but keeps a similar radio frequency performance at the same time[1]. Meanwhile, the RF performance of RF MEMS switch matrices is much better than present semiconductor switch matrices (such as p-i-n switch matrices)[2]. RF MEMS switches combine the advantageous properties of both mechanical and semiconductor switches. It offers the benefits of reducing system volume, mass and cost but promising a good RF performance. At the same time, RF MEMS switch matrices suits the applications in system integration and aperture synthesis. All the advantages above make RF MEMS switch matrix technique one of the best candidates in space borne and multifunctional phased array antenna applications[3].

Moreover, with the rapid development of satellite technology, the number of payloads in a single satellite has been increased dramatically. It's common that both the SAR payload and the communication payload are carried by a remote sensing or a telemetering satellite. However, the mass and cost will increase accordingly. RF MEMS switch matrices became one of the core techniques in realizing the RF domain integration of two or more payloads, which will bring a significant reduce of system mass, volume and cost[4]. This article conducts the study mainly on RF MEMS switch matrices in multifunctional phased array antennas, and the emphasis is the configuration scheme of RF MEMS switch matrices in the radio frequency integration of satellite SAR and communication payloads. A preliminary feasibility analysis is conducted to verify the proposed configuration scheme. The technical difficulties of applying RF MEMS switch matrix technique in multifunctional phased array antennas and the problems should be solved before this technique is available in practical applications are also pointed out in this article.

II. PAYLOAD ANALYSIS

Generally, a satellite consists of the payloads and the platform. For a specific kind of satellite, payload is a set of various space instruments and equipment which are needed to directly carries out satellite missions. Specifically, for the system architecture of a satellite communication payload, there are more than one receive antennas on the satellite, they receive various ground or inter-satellite signals respectively according to their own function. The shunt and transfer devices distribute received signals to receivers, and the RF signals are amplified in receivers, then usually they are down-converted in receivers. In transmit circuit, the signals are fed to transmit antennas by corresponding synthesize and transfer devices, and then they are radiated to user locations. A satellite payload is always divided into two subsystems: antenna subsystem and transponder subsystem. Antennas are designed to receive and transmit electromagnetic waves, and transponders amplify and process signals. For a SAR system, the space born image processing system is always placed on ground because the system is complex and the devices are cumbersome. As a result, the space segment devices of the satellite SAR system compose the payload, and it is the so called space borne SAR. Antenna subsystem is a pivotal part of space borne SAR payload. For plane array antennas which is mainly concerned in this article, they are composed of array plane and feed network. For distributed T/R array antennas, due to the fact that the power transmit module (T) and receive low noise pre-amplifier module (R) are distributed in the antenna feed network and become a part of the feed network, T/R module is also a part of antenna subsystem. This article will mainly discuss the RF MEMS switch matrices in RF domain of satellite communication and SAR payload, which is mainly the antenna subsystem, to realize a RF aperture synthesis. And planar phased array antennas are used in the two payloads mentioned above.

III. DESIGN SCHEME

In this chapter, this article will discuss the configuration scheme of RF MEMS switch matrix in multifunctional co-aperture phased array antenna of SAR and communication payloads. The aim to configure the switch matrix is to realize the multifunctional co-aperture phased array antenna of SAR and communication payloads, and then achieve RF integration of two payloads.

The RF integration of satellite SAR and communication payloads is a form of RF domain resource multiplexing. The multiplexing includes aperture multiplexing, RF components multiplexing, converter multiplexing and receiver/modulator multiplexing. However, an on-demand configuration scheme is needed to interconnect the multiplexed devices and their attached parts. So the application of switch matrices is vital here, which enables the connections between arbitrary inputs and outputs, and furthermore, the multiplexing is then realized.

The integration of RF system calls for an open system architecture, and the research of ISS indicates that the following open system architecture (figure 1) can be adapted[5].

From left to right, there are apertures, array interface unit (AIU), converters, receivers and digital IF processing modules. The RF apertures are divided according to their operating frequencies, and the aperture synthesis is realized in the same frequency domain. In the above architecture, microwave frequency apertures include array antennas operating between 2-18 GHz, and that's S, C, X and Ku bands. Moreover, for a SAR antenna, it always operates in L, S and C bands, in which C band almost suits for any application occasions[6]. At the same time, the communication payload antenna operates in Ku band, which is also within the range of 2-18 GHz, so this open system architecture is applicative for the multifunctional co-aperture design of this article.

In RF domain, RF MEMS switch matrices can be applied in the highlighted parts in figure 1. That's to say, if ISS open system architecture is applied, the switch matrices have the following

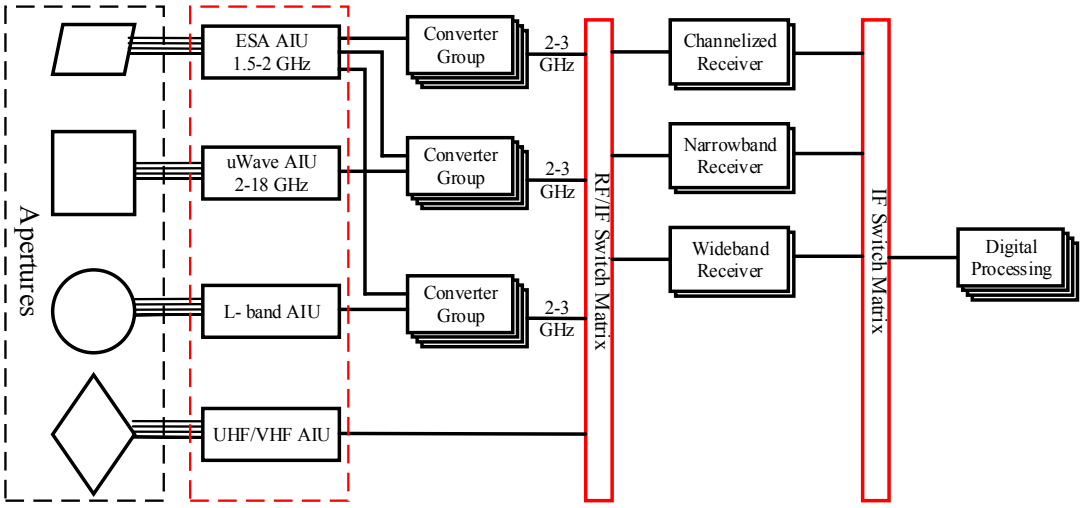


Figure 1. An ISS proposed open system architecture. The highlighted parts in red is possible parts that RF MEMS technology can be applied.

functions: construct the AIU, configure aperture, connect the corresponding converters and etc.; interconnect converters and receivers to realize multifunctional receiving; interconnect receivers and digital IF processors to process uniformly in order or conserve system resource.

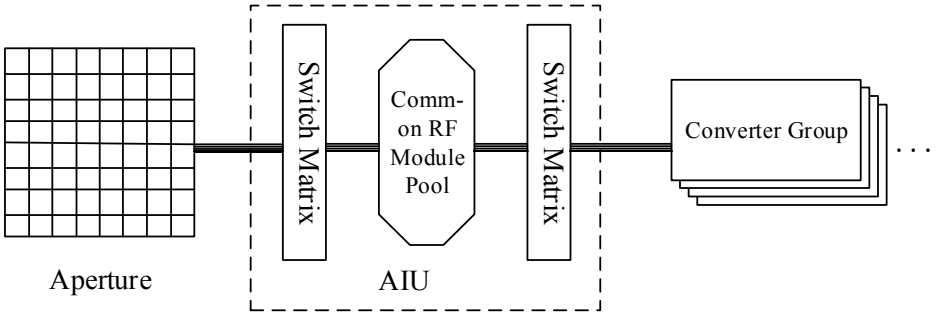
The AIU is shown as follows (figure 2), its function is to interconnect all apertures to functioning modules. As for multifunctional co-aperture phased array antenna of SAR and communication payloads in this article, since the aperture size is effected by its resolution, present SAR antennas in space always have the size of 20 m^2. But the size of communication payload antennas are smaller. As a result, part of the co-aperture array antenna element patches are designed using dual frequency bands (C and Ku) design. The output of dual frequency bands co-aperture antennas is connected with RF MEMS switch matrix, and the switch matrix is controlled by controlling algorithms to configure interconnections between corresponding antenna elements and RF units.

Before conducting RF integration design, it's necessary to sort the RF units of SAR communication payloads according to their functions and design levels. After sorting them, the common RF units form a common RF module pool. Sorting of these RF units is a complex and challenging task, and this task is still under way, so it will not be discussed in detail in this article. The common RF module pool connects to antenna apertures through RF MEMS switch matrix. The

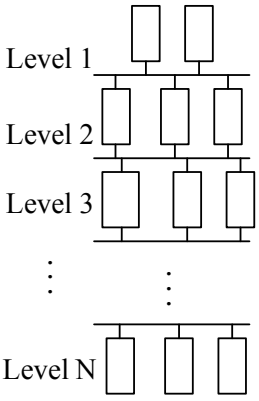
common RF units are configured level by level, as shown in figure 2b. The units on same level are not interconnected, while different levels are interconnected through RF MEMS switch matrices, this design reduce the complexity of switch matrices effectively. Or arbitrary RF units should be interconnected without level design mentioned above, which will dramatically increase the complexity of the matrices, or even impossible to implement. The output of common RF module pool is connected to converter group through RF MEMS switch matrix. The switch matrix in front of the converter group routes the input signal to corresponding frequency converters according to signal frequencies, and then the converters shift the frequencies to the same IF.

The RF MEMS switch matrix that interconnects converters and receivers route the uniform IF signal output by converter group to receivers, the receivers operate in the same frequency band but have different functions, and these receives make multifunctional receiving possible.

There are also RF MES switch matrices between receivers and digital processing end. Controlled by controlling programs and based on demands, the switch matrices feed the output signals to digital processing end to process the signal in order to obtain useful information.



(a)



(b)

Figure 2. (a) A more detailed construction of AIU, which consists of common RF module pool and RF MEMS switch matrices on both side to interconnect with other devices; (b) Common RF module pool. RF modules in the pool are arranged in levels to minimize the complexity of switch matrices

All RF MEMS switch matrices have their own corresponding controlling programs. Base on different system demands, the programs configure the switch matrices to realize different RF MEMS switch matrices functions mentioned above. Furthermore, the switch matrices in the

mentioned common RF modules pool can divide the RF units into different groups to get different functions and minimize mutual effects among different RF modules. And also the in-group controlling algorithms control the generation of different functions and resource allocation. Design of controlling programs and algorithms is one of the key points of the RF MEMS switch matrix design in multifunctional phased array antennas.

IV. FEASIBILITY ANALYSIS

Since that the concerning research of RF MEMS switch matrices in multifunctional phased array antennas is not mature yet, the article just proposes a possible application scheme and conducts a preliminary feasibility analysis. For the RF MEMS switch matrices applied in multifunctional phased array antennas of SAR and communication payloads, the technical problems is not all solved yet, and of course is not ready for practical application. However, this article tried to propose a RF MEMS switch matrices configuration scheme, which is a possible scheme for achieving aperture synthesis of two payloads and for realizing multifunctioning. The proposed scheme is based on the open system architecture proposed by ISS, which is already proved to be a feasible open system architecture. On basis of this architecture, RF MEMS switch matrix technique is used to achieve aperture synthesis, interconnection between converters and receivers and between receivers and digital IF processing end. So in the aspect of architecture, the scheme proposed by this article is feasible.

In the aspect of technique, present RF MEMS switch matrix technique is not adequate for applying the proposed scheme in practical satellite field though. For the devices applied in satellite field, it's necessary that a high reliability is promised, while at the time, reliability is one of the biggest drawbacks of RF MEMS switch matrix, because it's vulnerable to temperature, pressure and vibration[7, 8]. However, as mentioned, RF MEMS switch matrix technique can significantly reduce the size and mass of switch matrices, and at the same time keep RF performance excellent. This make RF MEMS technique an ideal technical scheme for satellite switch matrices. On the other hand, the power capacity is also a bottleneck of RF MEMS switch matrix. Generally, the power capacity of a single RF MEMS switch is several hundred milliwatts[9, 10], but it's highly possible that a higher power capacity is needed in satellite applications. Also, for large RF MEMS switch matrices, the architecture and topology design are great challenges[3], because as the order of switch matrices increase, the switching speed will decrease and the RF performance of the switch matrix will deteriorate[11]. However, though RF MEMS switch matrix technique remains the above technical issues, an increasing amount of efforts have been devoted into the research of it. A great amount of universities or research institutes have achieved many technical breakthroughs in RF MEMS switch matrix technique, which has been developing rapidly since then. So in the aspect of technique, the scheme propose by this article is feasible, and with the deep-in of concerning researches, this scheme will become practical available gradually.

V. CONCLUSION

RF MEMS switch matrix technique becomes an ideal technical scheme for satellite switch matrices since it has a small size, a low mass and an excellent RF performance. Though this technique itself remains some bottlenecks, it is approaching practical satellite application with unstopping technical breakthroughs achieved by researchers all over the world. In case of this background, this article conducted a research on RF MEMS switch matrices in multifunctional

phased array antennas and proposed a switch matrix configuration scheme for the application in aperture synthesis of SAR and communication payloads. Meanwhile, in architecture and technique two aspects, this article performed a preliminary feasibility analysis. The analysis indicates that the proposed scheme is feasible, and with the deep-in of the researches, this scheme will eventually come to practical application.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under grant 61471196.

References

[1] M. Daneshmand and R. R. Mansour, "RF MEMS Satellite Switch Matrices," Microwave Magazine, IEEE, vol. 12, pp. 92-109, 2011.

[2] G. M. Rebeiz and J. B. Muldavin, "RF MEMS switches and switch circuits," Microwave Magazine, IEEE, vol. 2, pp. 59-71, 2001.

[3] S. Di Nardo, P. Farinelli, T. Kim, R. Marcelli, B. Margesin, E. Paola, et al., "Design of RF MEMS based switch matrix for space applications," Advances in Radio Science, vol. 11, pp. 143-152, 2013.

[4] J. A. Lead, "JSF avionics architecture definition appendices," Arlington, USA: JAST Avionic Lead, 1994.

[5] R. Brousseau, A. Sanders, D. R. Huffman, H. Abercrombie, D. L. Hoffman, and S. Coleman, "An open system architecture for integrated RF systems," in Digital Avionics Systems Conference, 1997. 16th DASC., AIAA/IEEE, 1997, pp. 5.1-1-5 vol.1.

[6] W. G. Carrara, R. S. Goodman, and R. M. Majewski, "Spotlight synthetic aperture radar-Signal processing algorithms(Book)," Norwood, MA: Artech House, 1995., 1995.

[7] A. P. Lewis, M. P. Down, C. Chianrabutra, L. Jiang, S. M. Spearing, and J. W. McBride, "Lifetime Testing of a Developmental MEMS Switch Incorporating Au/MWCNT Composite Contacts," in ICEC 2014; The 27th International Conference on Electrical Contacts; Proceedings of, 2014, pp. 1-6.

[8] A. Lucibello, R. Marcelli, E. Proietti, G. Bartolucci, V. Mulloni, and B. Margesin, "Reliability of RF MEMS capacitive and ohmic switches for space redundancy configurations," Design, Test, Integration and Packaging of MEMS/MOEMS (DTIP), 2013 Symposium on, pp. 1 - 6, 2013.

[9] C. D. Patel and G. M. Rebeiz, "A high power (>5 W) temperature stable RF MEMS metal-contact switch with orthogonal anchors and force-enhancing stoppers," in Microwave Symposium Digest (MTT), 2011 IEEE MTT-S International, 2011, pp. 1-4.

[10] C. D. Patel and G. M. Rebeiz, "An RF-MEMS switch for high-power applications," in Microwave Symposium Digest (MTT), 2012 IEEE MTT-S International, 2012, pp. 1-3.

[11] A. A. Fomani and R. R. Mansour, "Monolithically Integrated Multiport RF MEMS Switch Matrices," Microwave Theory and Techniques, IEEE Transactions on, vol. 57, pp. 3434-3441, 2009.

A High Performance Broadband Quasi-Yagi Antenna on LTCC Integration Technology

Yiyuan Zheng[1], Weixing Sheng[1], Renli Zhang[1], Zhen Hu[2],
Yutao Yang[2], Yongpeng Kang[2]

(1 School of Electronic and Optical Engineering, Nanjing University of Science and Technology
2 Shanghai Institute of Satellite Engineering)

Abstract: In this paper, a high performance X-band quasi-Yagi antenna with broad bandwidth is presented based on low temperature co-fired ceramic (LTCC) integration technology. The proposed antenna utilizes a simple coplanar waveguide and balun structure to feed the power, ensuring the radiation efficiency at a high enough level. A truncated ground plane is inserted in the middle layer of LTCC substrate to act as the reflector of radiating elements. Three dimensional EM simulation results show that the antenna has 20% fractional bandwidth covering X-band with 15 dB return loss. The peak gain at the center frequency of 10.2 GHz is 7.5 dB. By virtue of the compact size and planar structure, this X-band quasi-Yagi antenna has great potential to implement integration and interconnect with receiver systems, and find a wide range of applications in various satellite communication systems.

Index Terms: Broadband antenna, planar structure, X-band, satellite communication systems, high gain.

I. INTRODUCTION

Yagi-Uda antenna, one of the most popular end-fire antennas, is widely used as a high-gain antenna in radar and communication systems [1]. Owing to the dipole elements made of metal rods, conventional Yagi-Uda antennas often suffer from heavy weight, relatively large size and inconvenience of fabrication, which restricts them to be applied into systems that have requirements for miniaturization and printed on substrate fabrication, such as wireless communication systems and satellite systems.

With the development of microwave theory and advancement of material, planar Yagi-Uda antenna combining the merits of both high-gain Yagi-Uda antenna and planar structure, also called as quasi-Yagi antenna, has received renewed interest recently due to its compact size, Light weight, high gain, high radiation efficiency, low cost and easy fabrication. Various designs of the planar quasi-Yagi antenna have been reported. In [2], Kaneda et al. firstly presented a microstrip-fed quasi-Yagi antenna, working at X-band with a gain of 3–5 dB and a cross-polarization level less than 15 dB. In [3], H. K. Kan proposed a coplanar waveguide-fed (CPW-fed) quasi-Yagi antenna with a 44% 10 dB impedance matching bandwidth at X-band. This design was realized on a high dielectric constant substrate and compatible with microstrip circuit and active devices. A modified planar antenna based on the classic Yagi-Uda dipole assembly was presented in [4], and experimentally validated radiation properties made this antenna a suitable candidate for phased arrays. A quasi- Yagi antenna based on LTCC substrate was discussed in [5], achieving high gain and good standing wave characteristics. Above all, quasi-Yagi antenna has shown great potential to

being used in many applications such as power combining, phased arrays, receiver systems, wireless communication systems as well as satellite communication systems.

In this paper, a broadband quasi-Yagi antenna based on LTCC integration technology is presented. The antenna utilizes a simple CPW feed and balun structure, creating 180° phase difference between the two coupled microstrip lines and offering a good impedance matching with the odd mode in the working frequency band. In addition, a truncated ground plane is inserted in the middle layer of LTCC substrate to act as the reflector of radiating elements. To accommodate the demand of specific receiver systems, the antenna is realized on a multilayer Ferro A6m LTCC substrate (ε_r =5.9, tanD=0.002), while the total thickness is 2.4mm. Three dimensional EM simulation results show that the proposed antenna has 20% fractional bandwidth covering X-band with return loss better than 15 dB. The peak gain at the center frequency of 10.2 GHz is 7.5 dB. By virtue of the compact size and planar structure, this X-band quasi-Yagi antenna will find a wide range of applications in various wireless and satellite communication systems.

II. ANTENNA DESIGN

The schematic of the proposed antenna is shown in Fig. 1. As can be seen from the figure, the proposed antenna consists of two director elements, a pair of driver elements, a microstrip to coplanar stripline (CPS) balun and a truncated ground plane inserted in the middle layer of LTCC substrate acting as the reflector. The use of the truncated ground plane results in a quite compact structure that can be easily integrated with any microstrip-based radio frequency systems, which has been confirmed with numer- ous experiments.

Since the driver elements and director elements are dipoles, a transition structure need to be added in order to prevent two ways signal counteracting, at the same time, keep antenna's radiation efficiency at a high level.

The microstrip to CPS balun, shown in Fig. 2, is an essential part of the transition. With it introduced, micro- strip feeding power can be equally divided into two ways continuous inverting excitation signals. An odd mode is able to be provided in the coupled microstrip line while the even mode is suppressed over broad bandwidth. In particular, balun's phase shifter creates 180° phase difference between the coupled microstrip lines at the center frequency around 10.2 GHz, equivalent to half wavelength in length difference. As is shown in Fig. 1, the length of 180° phase shifter's left and right arms are a and b respectively. They meet the following formula.

$$2\times(a-b)=\frac{\lambda_g}{2}$$

where λ_g denotes the guided wavelength at 10.2 GHz.

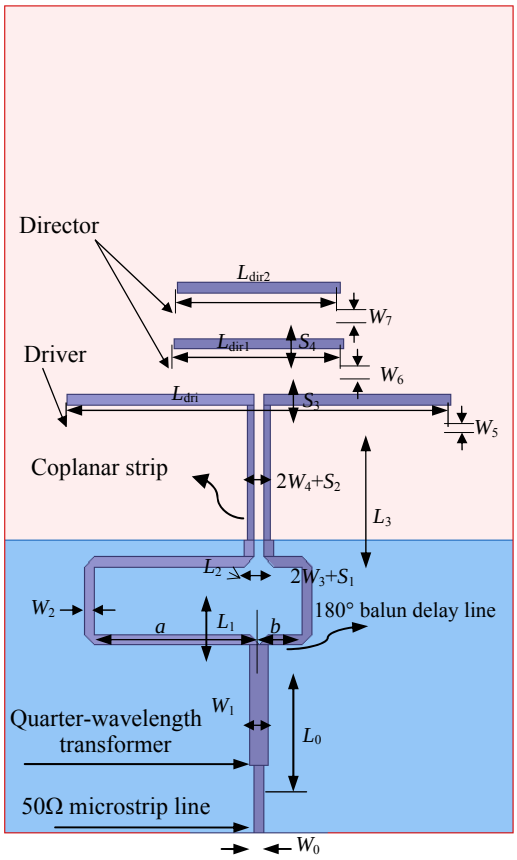


Fig. 1 The schematic of the proposed antenna

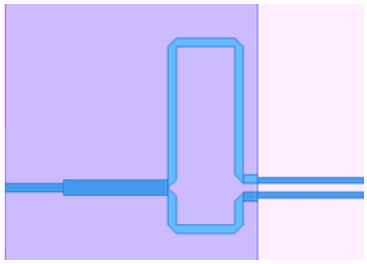


Fig. 2 The microstrip to CPS balun

A quarter wavelength transformer is also used for the good impedance matching between 50Ω microstrip line and balun structure. After two coupled microstrip lines connecting in parallel, the impedance becomes 50/2=25Ω. Therefore, the quarter wavelength transformer's impedance is set as $\sqrt{50 \times 25} \approx 35\Omega$.

Another important part of the transition is CPS which connects above balun and antenna dipoles. Bacause CPS does not support even mode, it will act as an open end for the even mode from the coupled microstrip lines and enables us to suppress the undesired modes excited from transmission lines in front of it.

Similar to the conventional Yagi-Uda antenna, the proposed quasi-Yagi antenna's driver dipoles and directors are parallel with each other and perpendicular to the radiation direction.

The length of the driver element is often approximately $0.45\lambda_g$, and the length of the director is usually $0.4\lambda_g$

It is noted that although the initial values of the dipoles' length are given above, the optimal length of the dipole elements have to be found by three-dimensional (3D) full-wave electromagnetic wave (EM) simulations.

In this paper, two shorter dipole elements are added as the director elements in attempt to improve the directivity of the antenna. These director dipoles not only increase the front-to-back ratio, but also play an important role in broadening the operation bandwidth of the antenna.

III. Simulated Antenna Properties

The overall 3D structure of the proposed antenna is shown in Fig. 3. This antenna is realized on a multilayer Ferro A6m LTCC substrate, and each layer has a post-fired thickness of 0.1 mm with a dielectric constant of 5.9 and loss tangent of 0.002. To accommodate the demand of specific receiver systems, the total substrate thickness is set as 2.4mm. Besides, a small copper block is attached beneath LTCC substrate for obtaining higher gain and better radiation effect.

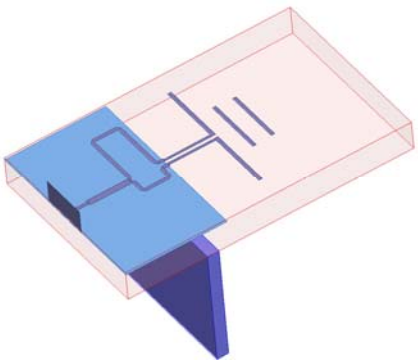


Fig. 3 Overall 3D structure of the proposed antenna

The final optimal parameters solved by 3D full-wave EM-simulator HFSS are:

W_0=W_2=W_3=W_6=W_7=0.3mm, W_1=0.52mm,

W_4=0.2mm, W_5=0.32mm, L_0=3.6mm, L_1=2mm,

L_2=0.5mm, L_3=4mm, L_{dri}=11.3mm, L_{dir1}=5mm,

L_{dir2}=4.8mm, S_1=S_2=0.3mm, S_3=S_4=1.35 mm,

a=4.86mm, b=1.26mm.

The total size of the substrate is 15mm×24.5mm, more compact than other structures which have been reported working at the same frequency band.

Fig. 4 shows the return loss properties of the proposed quasi-Yagi antenna. The antenna exhibits return loss of better than 15 dB over the working frequency band from 9.2 GHz to 11.2 GHz, owning good matching performance within a broad bandwidth.

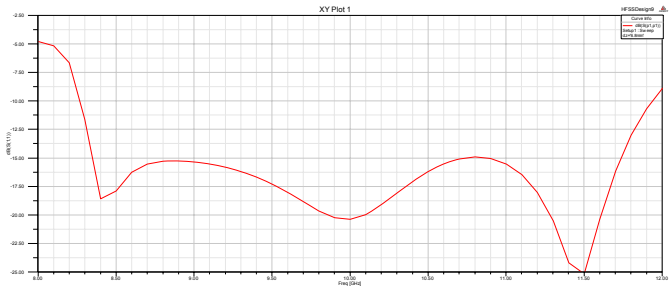


Fig. 4 Return loss properties of the proposed quasi-Yagi antenna

Fig. 5 shows the simulated antenna's gain. At the frequency band from 9.2 GHz to 11.2 GHz, the proposed antenna's gain varies from 7 dB to 8.3 dB. The peak gain at the center frequency of 10.2 GHz is about 7.5 dB.

E- and H- radiation patterns at 10.2 GHz are shown in Fig. 6. 3D radiation patterns at 10.2 GHz are shown in Fig. 7. As can be seen from the figure, the half-power beamwidth (HPBW) is 85° along the E-plane and 82° along the H-plane. The front-to-back ratio of the proposed antenna is 20 dB, greatly enhancing antenna's operating efficiency.

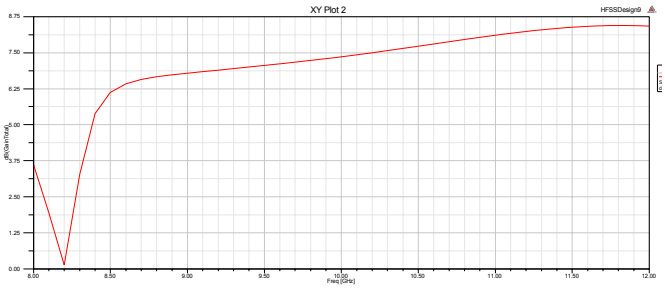


Fig. 5 Simulated antenna's gain

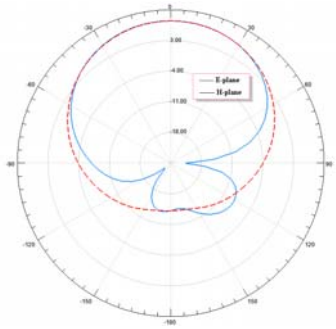


Fig. 6 E- and H- radiation patterns at 10.2 GHz

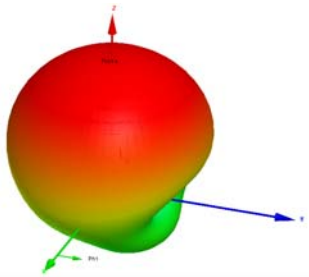


Fig. 7 3D radiation patterns

IV. CONCLUSION

A broadband quasi-Yagi antenna based on LTCC integration technology is presented in this paper. The antenna utilizes a simple CPW feed and balun structure, offering a good impedance

matching in the working frequency band and ensuring antenna's radiation efficiency at a high enough level. In addition, a truncated ground plane is inserted in the middle layer of LTCC substrate to act as the reflector of radiating elements. The antenna is realized on a multilayer Ferro A6m LTCC substrate with a copper block attached beneath it for obtaining higher gain. Three dimensional EM simulation results show that the proposed antenna has 20% fractional bandwidth covering X-band with return loss better than 15 dB. The peak gain at the center frequency of 10.2 GHz is about 7.5 dB. On account of the compact size, high performance and planar structure, this X-band quasi-Yagi antenna has great potential to implement integration and interconnect with receiver systems. Moreover, this X-band quasi-Yagi antenna will definitely find a wide range of applications in various wireless and satellite communication systems.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant No. 61471196.

REFERENCES

[1] Balanis, Constantine A, Modern Antenna Handbook. John Wiley & Sons, 2011.

[2] N. Kaneda, W. R. Deal, Y. Qian, R. Waterhouse, T. Itoh, "A Broad-Band Planar Quasi-Yagi Antenna," IEEE Transactions on Antennas and Propagation, vol. AP-50, no. 8, pp. 1158- 1160, August 2002.

[3] H. K. Kan, R. B. Waterhouse, A. M. Abbosh, and M. E. Bialkowski, "Simple broadband planar CPW-fed Quasi-Yagi antenna," IEEE Antennas Wireless Propag. Lett., vol. 6, pp. 18–20, 2007.

[4] A.P. Gorbachev, V. M. Egorov, "A modified planar Quasi- Yagi antenna for wireless communication applications," IEEE Antennas and Wireless Propag. Lett, vol.8, pp. 1091-1093, 2009.

[5] K. Zhang, J. Zhou, J. Fan, et al, "A Ka-band microstrip quasi-Yagi antenna based on LTCC technology," IEEE 2013 International Workshop on Microwave and Millimeter Wave Circuits and System Technology, pp. 157-160, 2013.

[6] C. A. Balanis, Antenna Theory Analysis and Design, 3rd ed. New York: Wiley, 2005, ch. 10.

Research on Spacecraft in Orbit Early Warning Subsystem

YANG Haifeng[1] Sun Yuan[2] Sun Shoufu[2] Cui Weiguang[1]

（1.Shandong Institute of Aerospace Electric Technology ,Yantai Shandong , 264003
2.Naval Aeronautical and Astronutical University,Yantai Shandong,264001）

Abstract:. When spacecraft is working in orbit, massive data that reflects status of spacecraft is produced. This information truly reflects the changing process of observe and control events and faults, which can be used on the early warning of the status development trend of spacecraft. Therefore, analyzing the feature of in orbit data and on which basis predicting its development trend, and making early warning of spacecraft's working status according to the prediction information make it possible to discover abnormal changes of spacecraft's features in early phase and solve problems in time so that to avoid major failure and reduce the risk of satellite's in orbit working. Using the next-generation communication satellite as an example, the early warning subsystem for communication satellite is designed. According to engineering application, using typical voltage and temperature data as an example, the early warning result is given, which shows the early warning effect is good.

KeyWords: spacecraft；early warning；on-board faults；satellite ； forecast

1. Introduction

When Spacecraft is working in orbit，after long-term working in the space environment, its performance and function will change with the factors of time, space environment, control events and faults. These changes are reflected by the telemetry data, so that we can analyze the law of the data changes of the spacecraft in orbit. On the basis of analyzing the telemetry data, we can make early warning according to the spacecraft performance and we can detect the abnormal changes of the spacecraft for timely and effective treatment to avoid major failure that may occur, and reduce the risk of spacecraft in orbit.

In this thesis, we design early warning system of telemetry data that can be used on satellite platform, by which, we can detect the abnormal changes of satellite performance, so that the satellite will make early warnings. In addition, it is of important significance to discover and conclude the law of the telemetry data changes according to the long-term forecast satellite historical telemetry data to make sure that the satellite will perform safe and stable in orbit and we can make health assessment and performance research of the satellite and improve the satellite designs.

2 Research on Satellite System

2.1 Unique of the satellite system

Satellite system has its particularity, which is mainly reflected in the following four aspects:

The system cannot be maintained. The system is exposed to complex electromagnetic environment, and it is easy to be disturbed；System capacity is low, usually there is only one system；Even if the fault occurred , the system will keep on working as far as possible.

These special causes lead fault to satellite system and make great risk to safe operation. The traditional satellite needs manual management on the ground station, because of the limited communication, it cannot make real-time control. With the demand of the real time and reliability of

satellite management, it is necessary to reduce the ground manual management and improve satellite autonomous control and autonomous fault management. The FDIR system designed to have the ability of fault detection, fault isolation and fault isolation, can make real time monitor and fault detection of the whole state of the satellite, locate the fault and take the measure of reconstructing to ensure the safe and reliable operation of the satellite.

2.2 Research on early warning system

The FDIR is used to monitor and deal with the satellite when the fault had happened, but some fault may have caused irreparable losses, so , before the fault has occurred, it is an effective method to take some measures to avoid the fault occurrence to manage a satellite in orbit.

It is an important direction for the satellite in orbit management to find the law that when satellite in orbit, its performance and functionality will change with the time and space environment, control events and failure factors according to the analysis of the in orbit satellite telemetry data. According to the analysis telemetry data in orbit and forecast the future development laws of the telemetry data, and applied to the forecasting of the data, it can make warning of the abnormal state the satellite may exist. It is a new direction of data analysis. Satellite telemetry data early warning system has achieved the function of telemetry data trend forecasting and early warning in ground support systems, and accumulates some experience and methods for satellite telemetry data processing, forecasting and early warning. Trend forecasting system run on the satellite avoids a lot of disturbance, it can get more real-time, accurate and extensive telemetry data to predict the trend of telemetry data, also it can early detect the abnormal changes of satellite performance, so that satellite early warning will work.

Through the analysis of satellite system, we put up the spacecraft in orbit early warning system, the system completes the forecasting and early warning tasks in all parameters of the satellite which has significant variation of telemetry parameter, mainly for the analog quantity, temperature and some other types.

1) Real time forecasting function：using of the past satellite telemetry data, we take accordance methods to make real time forecast for telemetry parameters.

2) Early warning function：Comparing the telemetry parameters predictive value, confidence degree and the theory threshold the user set, when the forecasting value deviates from the confidence level or beyond the theoretical threshold, the warning information is given in time.

Early warning system uses modular design, it consists by the parts as follows: the data receiving and distributing subsystem, telemetry parameter forecasting subsystem, early warning management subsystem, task allocation and the center control subsystem and in orbit telemetry database system. All the subsystems complete the on orbit telemetry data trend forecast and early warning function. The architecture is as Fig 1.

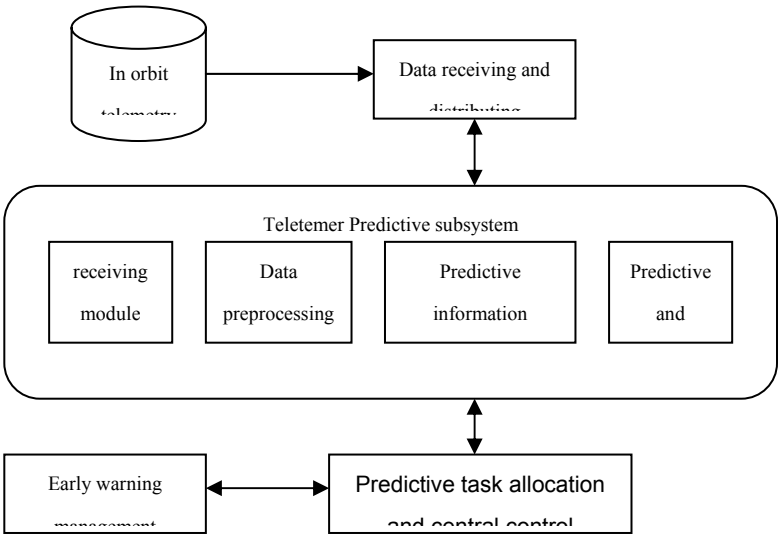


Fig1 general planning for the system

1) Data receiving and distributing subsystem: early warning system and exchange the information with telemetry database in orbit.

2) Telemetry parameter forecast subsystem： predicting, getting historical telemetry data, preprocessing of historical telemetry data，forecasting the task allocation, forecasting algorithm model configuration, forecasting analysis and information output.

3) Early warning management subsystem：completing the forecast data acquisition, comparison and analysis, when the forecasted values deviating from the confidence or beyond the theoretical threshold ,the real-time warning information is given.

4) Task allocation and the center control subsystem：Responsible for the control and scheduling of the whole system, the forecast results are given to the control terminal, which is used in orbit management, fault diagnosis and so on.

5) In orbit telemetry database system： Storage and management of the satellite in orbit telemetry data.

3 Research on implementation method of trend forecast

3.1 Data preprocessing method

In the process of collecting and distributing data of the in orbit telemetry data, the data is affected by the acquisition system, encoding, communication and some other facts, and there will be some outliers. The existences of outliers will affluence the results of data forecast. Therefore, it is necessary to preprocess the telemetry parameters data and eliminate the outliers. Analyzing the telemetry data to be forecast, we can use the following several outlier elimination method on the different characteristics of data preprocess, to eliminate the outliers.

1) Tucci median outlier elimination method：Processing data information changes stable.
2) Four point smoothing method for outliers：Processing data information changes smoothly.
3) Standard variance ratio method：Processing data information has strong periodicity.
4) Standard variance ratio method of plot：Processing data information has many uncertain factors

3.2 Classification and comparison of forecasting methods

Trend forecasting methods can be divided into the following three categories：trend forecasting

technology based on telemetry data, trend forecasting technology based on physical mode and trend forecasting technology based on knowledge，among which, the trend forecasting technology based on telemetry data can be divided into the trend forecasting of telemetry data based on mathematical model and trend forecasting technology of telemetry data based on intelligent algorithm。Table 1 compares and analyses the applicable objects, features and applications of 3 kinds of trend forecasting methods.

Tab 1 Comparison of all kinds of trend forecasting methods

Prediction type		Typical method	Applicable object	Characteristic
Based on telemetry data	Mathematical model	Least Square Method、ARMA model, Based on Fourier method and etc	The test data sequence changes with time and it has certain rules.	The method is based on the acquisition of the time series of teltmetry data, and it is easy to implement, and it is not necessary to describe the physical model of the device, and it does not require a priori knowledge of experts in various fields.
	Intelligent algorithm	Neural network、Support vector machine、SARIMA and etc		
Based on physical model		Kalman Wave filtering	System model is known, accurate system model can be established.	Although the proposed method has higher prediction accuracy, it is difficult to establish a precise mathematical model, so it has little application.
Based on knowledge		Expert system、Fuzzy logic	Be rich in expert knowledge, a complete knowledge base can be established .	The method can utilize the expert knowledge and experience in the field of object system, but it is limited to incomplete knowledge acquisition of expert knowledge， it is only suitable for qualitative reasoning and not for quantitative calculation, so its application in practice is Limited.

3 kinds of trend forecasting methods have their advantages and disadvantages; the trend forecasting technology based on physical mode needs a precise mathematical model，trend forecasting technology based on physical mode and trend forecasting technology based on knowledge is limited by expert knowledge acquisition. These two methods are seldom used in engineering practice, therefore the satellite trend forecast system uses the method of trend forecasting technology based on telemetry data .

In order to meet the needs of the satellite in orbit data analysis and trend forecasting, the forecasting system of the trend forecast system provides a variety of typical methods based on the forecast of the telemetry data. In the practical application of in orbit management, the appropriate forecasting method is chosen according to the data in orbit and parameters characteristic. Taking temperature and voltage of two typical telemetry parameters for example, based on engineering practice，the trend forecast method such as the least square method , the method based on Fourier and neural network will lead to good forecast effect.

3.3 Forecasting model and algorithm

In the forecasting process, the least square polynomial model, SumSin model、Fourier Series

model、Exponent model、Gauss model and Neural network model are used.

1) Least squares polynomial model

On the principle, the polynomial model can well fit the data with periodic law, and the low order polynomial model, especially the slow variation to the exponential form. Take the 4 order model as an example, the model expression is as follows:

$$y = \Phi(x,t) = a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4$$

In the model expression, $x = (a_1, a_2, a_3, a_4)$ is the model parameters to be solved, the number of model parameters $n = 4$, y is telemetry sampling data, t is corresponding sampling time sequence.

2) SumSin model

SumSin model is sin function and model, the model can well fit the data of various types of periodic law. Take the 6 order model as an example, the model expression is as follows:

$$\begin{aligned} y = \Phi(x,t) = {} & a_1 \sin(b_1 t + c_1) + a_2 \sin(b_2 t + c_2) + a_3 \sin(b_3 t + c_3) \\ & + a_4 \sin(b_4 t + c_4) + a_5 \sin(b_5 t + c_5) + a_6 \sin(b_6 t + c_6) \end{aligned}$$

In the model expression, $x = (a_1, b_1, c_1, a_2, b_2, c_2, a_3, b_3, c_3, a_4, b_4, c_4, a_5, b_5, c_5, a_6, b_6, c_6)$ is the model parameters to be solved, the number of model parameters $n = 18$, y is telemetry sampling data, t is corresponding sampling time sequence.

3) Fourier series model

The Fourier series model can fit well the data type of each periodic law, and is also a data model based on the harmonic frequency. Here is the model of the 4 order Fourier series, the model expression is as follows:

$$\begin{aligned} y = \Phi(x,t) = {} & a_0 + a_1 \cos(\omega t) + b_1 \sin(\omega t) + a_2 \cos(2\omega t) + b_2 \sin(2\omega t) + \\ & a_3 \cos(3\omega t) + b_3 \sin(3\omega t) + a_4 \cos(4\omega t) + b_4 \sin(4\omega t) \end{aligned}$$

In the model expression, $x = (a_0, a_1, b_1, a_2, b_2, a_3, b_3, a_4, b_4, \omega)$ is the model parameters to be solved, ω is Basic frequency, the number of model parameters $n = 10$, y is telemetry sampling data, t is corresponding sampling time sequence.

4) exponential model

Exponential model can better fit with the trend of changes in the index data, the expression is more concise than multi order polynomial model, but considering the index rose rapidly, we only consider the 1 ~ 2 order exponential model. Here is the model of the 2 order Exponent, the model expression is as follows:

$$y = \Phi(x,t) = a_1 e^{b_1 t} + a_2 e^{b_2 t}$$

In the model expression， $x=(a_1,b_1,a_2,b_2)$ is the model parameters to be solved， the number of model parameters $n=4$， y is telemetry sampling data， t is corresponding sampling time sequence.

5) Gauss model

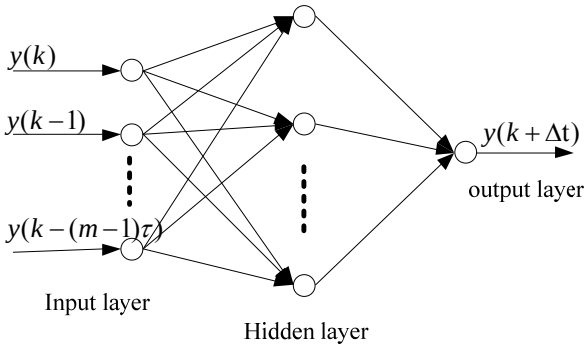
The Gauss model fits well with the data of the arch trend. Here is given the 3 order Gauss model, the model expression is as follows:

$$y=\Phi(x,t)=a_1\exp\left(-\left(\frac{t-b_1}{c_1}\right)^2\right)+a_2\exp\left(-\left(\frac{t-b_2}{c_2}\right)^2\right)+a_3\exp\left(-\left(\frac{t-b_3}{c_3}\right)^2\right)$$

In the model expression， $x=(a_1,b_1,c_1,a_2,b_2,c_2,a_3,b_3,c_3)$ is the model parameters to be solved，the number of model parameters $n=9$， y is telemetry sampling data， t is corresponding sampling time sequence.

6) BP Neural network model

BP Neural network model can better fit the data of non periodicity and square wave periodicity. It is known that there is a hidden layer of the three layer BP network, which can effectively approximate to any continuous function. The three layer network includes input layer, hidden layer and output layer. Considering the requirement of the network to predict the generalization performance of the network， Network design should adhere to the principle of reducing the complexity of the network as much as possible, using the three layer BP network model to predict the chaotic time series， The input layer node only transmits the input signal to the hidden layer， The hidden layer of the hyperbolic tangent function of S type of data processing to the output node， output nodes are simple linear functions.



Input data is： $Y(k)=[y(k),\ y(k-\tau),\cdots,\ y(k-(m-1)\tau)]$， output data is: $y(k+\Delta t)$， in the expression $k=1,2,\cdots,N$， m is input dimension， τ is cross step。

In the application process， we obtain the original parameters of SumSin model and Fourier series model respectively based on FFT and QR decomposition; The parameters of the Exponent model and the of Gauss model are studied by the characteristics of the index data and the Gauss data. Using genetic algorithm and Levenberg-Marquardt algorithm, we can obtain accurate parameters of the the SumSin model, Fourier series model, Exponent model, Gauss model and other models；

Using BP algorithm to obtain the accurate parameters of neural network model。

3.4 Forecast evaluation method

Prediction and evaluation of the early warning system to assess the use of the implementation of the results of the assessment, making the results more accurate, as small as possible to deviate from the real situation. In the practice, we calculate the degree of implementation by the application of statistical principle.

1) Confidence statistical evaluation methods

First, we do data preprocess and analysis of the historical telemetry data（we set the history data set for the telemetry parameters is $H = \{h1，\ h2,.....hi,.....hn\}$）， we get models and algorithms for data forecast. According to the model, the forecast value of the telemetry parameters is calculated（we set forecast data set for telemetry parameters is $P = \{p1，\ p2,.....pi,.....pn\}$），and compare it with the the corresponding telemetry parameters $H = \{h1，\ h2,.....hi,.....hn\}$，and then work out the error set, $E = \{e1，\ e2,.....ei,.....en\}$， according to the prior data analysis, a confidence degree is determined $\alpha \in (0,1]$. we define confidence interval value ∇ as $100*\alpha$ between $(H - \nabla, H + \nabla)$.The smaller the the α is, the value is closer to the truth. α =0,means hi , the forecast value of telemetry data equals to pi, the corresponding forecast value.， α =1 means All values are credible. $(P - \nabla, P + \nabla)$ can be used as a confidence interval for the true value of the subsequent forecast value, the confidence interval is α . By updating the historical telemetry parameter periodically, recalculating the predict value of the telemetry parameter, we can constant updating the confidence degree interval value ∇ ,whose confidence degree is α , so that we can be sure that the forecast value is within the latest and most accurate range.

2) Error statistics evaluation method

In the forecasting process, we can use the error statistical method in the statistical method to evaluate. We use the definite model and algorithm, according to historical data set of telemetry parameters （ we define the historical data set of telemetry parameters as $H = \{h1，\ h2,.....hi,.....hn\}$ ）, we can get the forecast value of the required telemetry parameters （we define forecast value set of the telemetry parameters $P = \{p1，\ p2,.....pi,.....pn\}$ ）, we use the data of the set H&P,， to calculate the data of the maximum absolute error, average error mean absolute error, we use the comparing benchmarks we set to analyze the mathematical statistics data, so we get the data to evaluate the system, which meet the user's request .

3.5 Satellite borne equipment

Through the summary and analysis of the computer on the application of the spacecraft in orbit, the software can be run in the following several environments.：

1) Software environment：Operating system Vxworks,Satellite borne flash file system OFFS.
2) Hardware environment：Computers equipped with 3 different CUP, BM3803、PC8245、

PC8548E.

Tab 2 list the CPU performance comparison among BM3803、PC8245、PC8548E

Tab2 performance comparion of CPU

Performance and index	BM3803（domestic）	PC8245	PC8548E
Core	SPARC V8	G2(Derivative version from PPC603e)	e500
Frequency	100Mhz	333Mhz	1.5Ghz
Performance	86MIPS 23MFLOPS	200～300MIPS	3065MIPS@1333Mhz
Cache	Data Cache：16KB Instruction Cache：16KB	Instruction cache：16KB Data cache：16KB	2level Cache： L1: Instruction：32KB, Data：32KB L2: Instruction：512KB, Data：512KB,
Error detection and correction function	EDAC	ECC	ECC
Memory interface	PROM,FLASH,SDRAM	PROM,FLASH,SDRAM	DDR,DDR2,FLASH,SDRAM
Peripheral interface bus	PCI2.2 interface	PCI	2 PCI or PCI-X +1 个 PCI-E
Power	1W@100MHz	2.3W Typical	5W@1000MHz Tj＝65℃ Typical
Device level	Space level	Up screening（Atmel orE2V）	M（E2V）
Support Vxworks	yes	yes	yes

Operating system Vxworks、satellite borne Flash file system OFFS, our institute has the research and development capabilities of the spacecraft related products, some related electronic products, developed on Vxworks、satellite borne Flash file system OFFS, have been tested on certain type of spacecraft, they work well and meet the design requirements

BM3803（domestic ）is space level product，it has the index of anti radiation. PC8245 cannot provide space level product present, when it needs to be used in spacecraft, screening should be done. Related electronic products have been tested in orbit, the products work well and meet the design requirements；PC8548E cannot provide space level products present, when it needs to be used in spacecraft, screening should be done, and project demonstration is ongoing on certain spacecraft.

3.6 Massive data storage research

The basis of the early warning system is historical data, if we make a long time forecast, a large amount of historical data is needed as the system forecast input.

For a long period of time, a large amount of historical data is needed to forecast the input. Therefore massive data storage device is needed to store the historical telemetry data. We can obtain the required information from the large capacity memory in the process of running the system. In the early of the system design, data volume and data storage analysis should be done, so that we can select memory, and the access method for telemetry data.

1) Data quantity evaluation：

At present, when the forecast evaluation algorithm of telemetry parameter early warning

system, it is more close to the trend of telemetry parameters, it takes the historical telemetry data of 3 months at most, to forecast the future telemetry data for less than 1 month. According to the statistical data of XX-2 satellite, the quantity of all the 1632 telemetry data of 3mongth is about 55.1G, if we use the original code storage, the amount of data can be reduced to 1/8 of the original. Comparing with the other data information, such as historical picture data and etc, telemetry information has the characteristic of low data transmitting rate, easy to be compressed and high compressed rate. So we can take compression algorithm in data storage to save storage space to increase the storage capacity for telemetry. We can use zlib algorithm, which has noting to do with the platform and easy to be transplanted, to store telemetry data. The algorithm uses the DEFLAE algorithm, LZ77 algorithm and Huffman encoding, it has mature technology, the lossless compression effect is good and the compression rate is at least 0.02。However, when using the compression algorithm, it may affect the efficiency and speed of data retrieval, which has a certain influence on the performance and real-time performance of the system.

2) Memory evaluation

Fast recording technology of massive data has been applied in the domestic and abroad, our institute has put the massive data storage unit of 80G on XX-2. At present, pre research on fast data storage prototype has been carried out, and the prototype has been completed (fig 2). Prototype technology index has reached the speed of reading and storing at the rate of 5Gbps, the Capacity is about 2.56Tb and it can be extended；the prototype machine support documented data management, by establishing an index table for storing files so that, the machine can locate the storage data quickly, the responding time is within 0.2 s, and it has function of file index, which greatly improve the performance of early warning system.

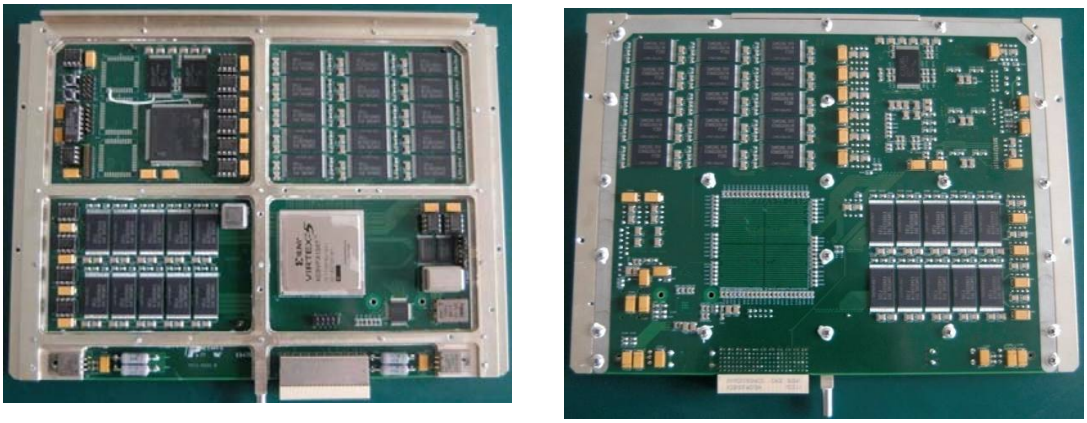


Fig 2 Principle prototype of high speed data memory

4 Early warning system

4.1 General planning

According to the requirement analysis, and basing on the technology and method of engineering practice, the early warning system can be divided into several requirement modules. They are configuration information management, telemetry parameter prediction, early warning, forecast task allocation, information display, external interface management and database interface. By main thread scheduling, they realize the user function of the forecast information management, telemetry parameters forecasting, early warning and interactive interface management. Structure diagram is as fig3 shows.

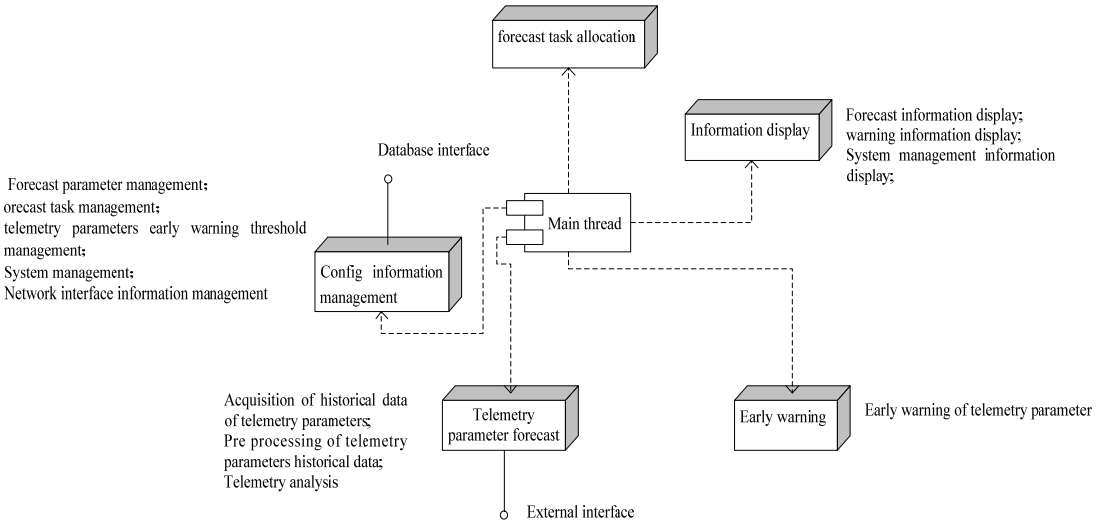


Fig 3 System structure diagram

Telemetry database manage system receives real time telemetry data information, and stores in database in chronological order, if the data to be stored is larger than the storage capacity assumed, the historical data that first stored should be erased. Forecasting of task allocation and center control subsystem launched an early warning tasks and the task information is sent to the data receiving subsystem. Data receiving and distributing subsystem according to the received early warning task and obtain the corresponding telemetry historical data information and distribute data information to the telemetry parameters prediction subsystem. The telemetry parameters forecast subsystem complete the information acquisition of historical telemetry data information，and then preprocess the historical telemetry data information, forecast the task allocation（input information of predictive algorithm model），process the historical telemetry data information, calculate the forecast telemetry data information, analyze the forecast telemetry data information, and then carry out the calculate the execution degree of the forecast information. The early warning management subsystem obtains the forecast data and the early warning threshold profile from the forecast task allocation and the central control subsystem, and completes comparing the theoretical threshold of the telemetry parameters, the confidence level and the user set up. When the forecast value deviates from the confidence level or beyond the theoretical threshold, the real-time warning information is given, and the corresponding information is transmitted to the forecasting task allocation and the central control subsystem. Forecast task allocation and central control subsystem transmit the predict results to the required control terminal through the network or bus, used in orbit management, fault diagnosis and other functions.

4.2 Early warning system

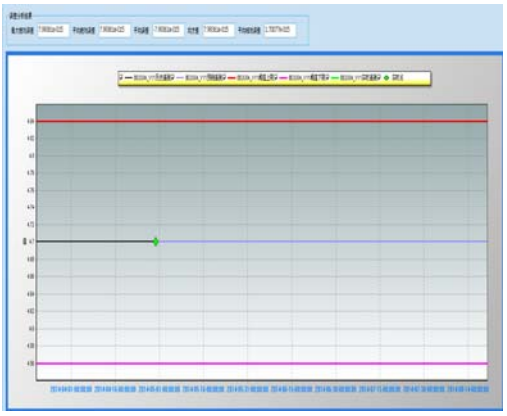
In early warning system, the least square polynomial method, LM+ Harmonic method, GA+Fourier method, BP neural network method and LM+ Index method are used. All the method can do forecast of telemetry data for spacecraft in orbit, we get 20000 data set. The manners and testing performance is as tab 3 shows. The time of getting data depends on the complex of storing the data, and the predict time depends on the complex of algorithm.

Tab 3 statistics form

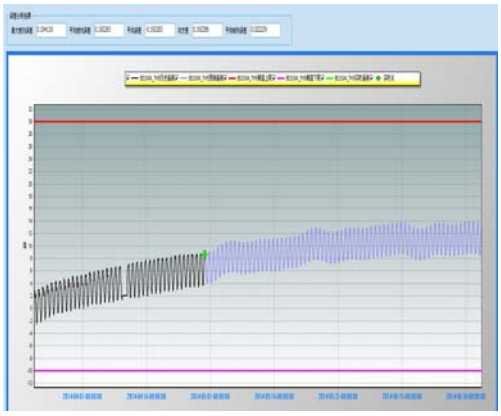
Algorithm name	Applicable data type	Data acquisition time（s）	Forecast time（s）
Least squares polynomial	With the simple monotonic change law is mainly applied to linear changes.	3.2	1.1
SumSin model	Has many kinds of superimposed periodic variation	5.9	16.8
GA+Fourier	Simple periodic variation	4.1	21.1
BPNeural network	With complex changes	4.0	29.3
Index method	Long monotone nonlinear variation law	8.0	7.9

Take four kinds of typical forecast values for the results show， as is shown in figure 4, the red line represents the upper and lower threshold of the telemetry parameters, black line represents the value of historical telemetry value, purple line represents the forecast telemetry parameter, green point represents the real time value of telemetry parameter.

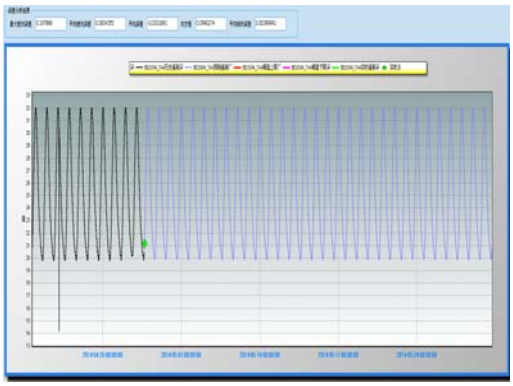
From figure 4, we can see that the trend is basically consistent with the actual data, and it can track parameter curve well.



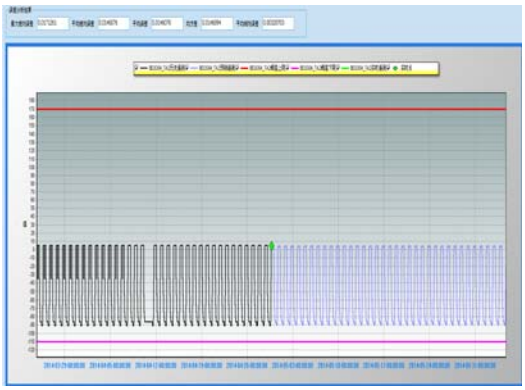
Least squares polynomial



SumSin model



GA+Fourier



BP Neural network

Tab 4 Forecast error analysis table

Algorithm name	Maximum absolute error	Mean absolute error	Mean error	Mean square deviation	Mean relative error
Least squares polynomial	7.9936E-15	7.9936E-15	-7.9936E-15	7.9936E-15	1.70077E-15
SumSin model	0.194119	0.192283	0.192283	0.192286	0.022222
GA+Fourier	0.197888	0.083485	0.033199	0.099027	0.003999
BP Neural network	0.017261	0.014507	0.014507	0.014689	0.003287

Forecasting and monitoring of early warning system is shown in figure 5,the blue point and dash line represents the prefabricated the user sets, red line represents the confidence interval, purple line represents the actual value，green line represents the predict value. When the forecast value exceeds the preset range of the user sets, system alarms and outputs the corresponding information to other systems.

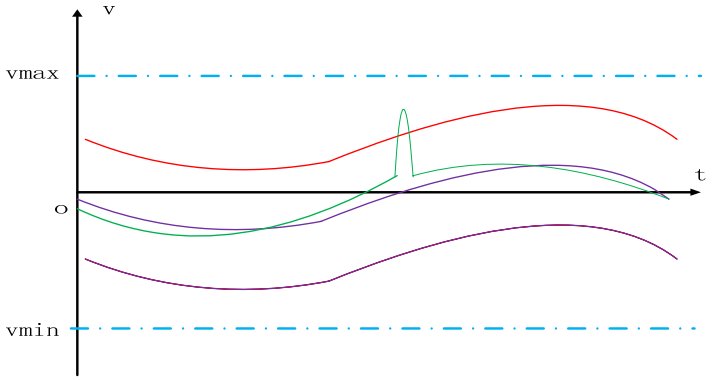


Fig 5 monitor of early warning system

5 Summary

The early warning system of satellite telemetry data has been realized in the ground support system, and it can carry out the function of trend forecasting and early warning of telemetry data. The system has already accumulated amount of experience and method to process, forecast and make early warning. If the system will be equipped in spacecraft, we should verify the availability of the forecast algorithm and the early warning system of satellite telemetry data in spacecraft environment by a prototype , which is made by BM3803（domestic）、PC8245、PC8548E CPU and massive memory.

References

[1] Zhang Bangxin，Suo Xunhua. Technique for diagnosing satellite failure[C]// 2001 Academic Conference for Simulation Technique. Beijing: China Computer Users Association, 2001: 213-215 (in Chinese)

[2] Tan Weizhi，Hu Jingang. Spacecraft systems engineering[M]. 2nd Edition. Beijing: China Science and Technology Press, 2009 （in Chinese）

[3] Caceres M. Satellites and launches trend down[J]. Aerospace America, 2004, 42(1): 18-20

[4] Zhang Yanduo，Jiang Xinwei，Huang Wenhu. Design of satellite TM system for fault simulator and alarm[J]. Journal of Vibration, Measurement & Diagnosis, 2002,22（3）：186-189 (in Chinese)

[5] Roemer M J, Byington C S, Kacprzynski G J, et. An overview of selected prognostic

technologies with reference to an integrated PHM architecture[C]. Proceedings of IEEE Aerospace. 2005:1~15

[6] Al-Smadi, A. The estimation of the order of an ARMA process using third-order statistics. International Journal of Systems Science vol. 36(15):975~980, Dec. 2005

[7] Richard C, Millar. A Systems Engineering Approach to PHM for Military Aircraft Propulsion Systems[M]. NAVAIR Propulsion & Power, 2007

[8] Rrook R S, David L A. A survey of serviceable spacecraft failures， AIAA-2001-4540[R]. Washington: AIAA, 2001

Study on high sensitive detection method of near-space hypersonic aircraft for space-based infrared detection system

Jia LI1,[2], He HUANG[1],Shaojuan LI[1],Shuang LIANG[1],Jingnan MA

1. Science Institute, Air Force Engineering University, Xi'an 710051, P.R.China;
2. School of Physics and Optoelectronic Engineering, Xidian University, Xi'an 710071, P.R.China

Abstract：Aiming at the problem of near-space hypersonic target detection for space-based infrared (IR) detection system, a temporal partial differential equations (PDE) based infrared dim small target detection method has been proposed. The highlight of the proposed method is that it utilizes the P-M model based PDE to deal with the temporal profile, which separate the small target from strong background clutter and obtain accurate detection result according to the difference of temporal profile among small target, stable background and the edge of cloud. In this paper, two real low signal-to-clutter ratio (SCR) infrared imagery sequences containing moving dim small target are applied to verify the efficiency of the proposed method. Compared with several classical methods, the experimental results show that the performance of proposed method is superior to several other methods both in aspects of subjective visual and objective evaluation index.
keywords：Infrared image, Target detection, Partial differential equations, Temporal profile

1 Introduction

Near space aircraft, which has rapidly developed recently, will become the potential threat to defense system because of its high velocity, high flight altitude and long voyage. Taking the advantage of the position, satellites are very suitable for the detection system as a platform to detect the target in near-space. However, due to the long distance from target to the space-based detection system, the target occupies a small area in the image and lacks the information of size, shape and texture; the signal-to-clutter ratio (SCR) of the image is low, which makes the target detection very difficult. Therefore, how to improve detection target accurately and stably from the IR image with low contrast and low SCR becomes a key technology.

In recent years, many researchers have made research on the problem of IR dim and small target detection, and several target detection methods have been proposed [1-3]. These methods can be classified into three main families [4]: spatial filter-based, transformation domain-based and temporal filter-based methods. Spatial filter-based methods utilize the grey level of target and background in the image to predict the background, and then detect the target. Typical spatial filter-based methods include max-mean/median filter [5], top-hat [6], two-dimensional minimum mean square error (TDLMS) [7], bilateral filter [8], bilateral two-dimensional minimum mean square error algorithm (BTDLMS) [9], etc. This kind of method has the advantages of simple principle and easy realization, but it is not effective for dim target detection in IR images with low contrast and heavy clutter background. Transformation domain-based methods decompose infrared image into a series of low-frequency and high-frequency sub-bands by transform image from spatial domain to other transform domains. The target and the background occupy the different sub-band, which can be used to suppress background and detect targets. Wavelet transform [10], shearlet transform [11] and Butterworth high-frequency filter [12] based methods are some classical transformation

domain-based methods. This kind of method perform well in target detection, however, there exist disadvantages of complex and time-consuming calculation. Temporal filter-based methods use the motion information of the target effectively. Traditional Temporal filter-based methods include triple temporal filter (TTF) [13], continuous wavelet transform (CWT) [14] and so on. These methods need to deal with three-dimensional data, so they have to suffer from the issues of large calculation. Temporal profile based method is an important method for detecting small target in temporal domain. It deals with one-dimensional data so that the amount of calculation will reduce greatly. The most representative method is connecting line of the stagnation points filtering method (CLSP) [15]. This kind of method is simple to complete and has a small amount of calculation, but some small fluctuations will lead to a degradation of detection performance.

Based on the drawbacks of the methods discussed above, a novel method is proposed to detect the moving target in clutter backgrounds through partial differential equation in this paper. Compared with several classical target detection methods, the experimental results show that the performance of proposed method is superior to that of other methods both in aspects of subjective visual and objective evaluation index.

2 Temporal profile model

2.1 Analysis of IR imagery sequence in temporal domain

IR imagery sequence containing dim small target can be modeled as follows:

$$I(x,y,k) = I_T(x,y,k) + I_B(x,y,k) + n(x,y,k) \quad (1)$$

where (x, y) represents the coordinate of one pixel; k denote the frames; $I_T(x,y,k)$ represents the dim small target image, $I_B(x,y,k)$ is the background image; $n(x,y,k)$ is the random noise.

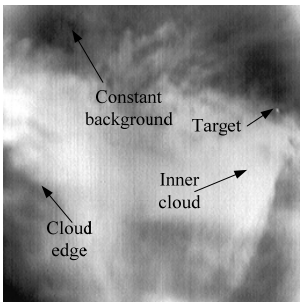
$I_T(x,y,k)$ has great changes between two adjacent images because of target movement. Meanwhile, the scene in imagery sequence collected by space-based infrared detection system is relatively stable. Namely, there are little changes between adjacent images within a certain period of time. Therefore, formula (1) can be rewritten as

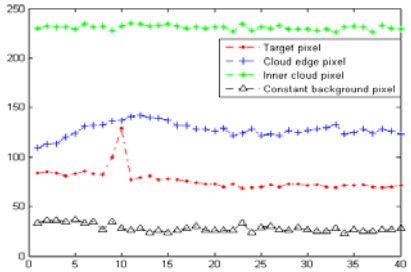
$$I(x,y,k) = I_T(x,y,k) + I_B(x,y) + n(x,y,k) \quad (2)$$

For certain fixed pixel (x,y) in the image, its gray value is a function of k. Obviously, the change of the gray value of a pixel in the imagery sequence is completely different in the case that a target moves across a pixel or not, which can be described using temporal profile.

2.2 Temporal profile

The temporal profile indicates variation of the pixel values over a period of IR sequences. For an infrared imagery sequence with moving small target, the temporal profile of the target pixels is a pulse-like curve. The width of the pulse is related to the speed of the target, and its amplitude is related to the brightness of the target. The temporal profile of the background pixels is a curve with stable values. The temporal profile of cloud edge pixels has an abrupt increasing or abrupt decreasing curve, but the variation of those pixels is smaller than target pixels. The temporal profile of various pixels is shown in Fig. 1.





（a） （b）

Fig.1: Temporal profiles of a target, clear sky, inner cloud, and cloud edge pixel. (a) One frame IR image of a sequence. (b) Temporal profiles

3 Temporal partial differential equation based target detection

3.1 Partial differential equation

The partial differential equation, which is represented by P-M model and proposed by Perona and Malik, has been widely used in various fields of image processing. This non-linear diffusion equation overcome the disadvantages of linear diffusion filter， P-M diffusion model can expressed as follows

$$\frac{\partial I_t(x,y)}{\partial t} = div[g(|\nabla I_t(x,y)|)\nabla I_t(x,y)] \qquad (3)$$

$$I_0(x,y) = I(x,y)$$

where, $I_t(x,y)$ represents the image at moment t, $\nabla I_t(x,y)$ denote the gradient at the pixel (x,y), $I_0(x,y)$ is the initial image, div () is the divergence operator. Function $g(\bullet)$ is diffusion function. In general, require that the function $g(\bullet) \geq 0$ and is a decreasing function. In addition , $g(0)=1$, $g(\infty) \to 0$. There are two classical form of function

$$g(s) = e^{-(s/k)^2} \qquad (4)$$

$$\text{or } g(s) = \frac{1}{1+(\frac{s}{k})^2} \qquad (5)$$

where , k is a threshold parameter.

3.2 Proposed method

From the analysis of section 2, there are obvious differences between temporal profile of target region and that of other regions, which can be used to suppress background precisely. In this paper, we use the partial PDE to smooth the temporal profile for obtaining the reference line of each pixel, and then the background can be predicted.

Assuming that $u(k,t)$ is the temporal profile of a pixel in an imagery sequence in a period of time. The diffusion function can expressed as follows

$$\partial u_t(k)/\partial t = \text{div}[c(\nabla u)]\nabla u\sqrt{b^2-4ac} \qquad (6)$$

where, k denotes the k-th frame, diffusion function $c(\nabla u)$ adopt the form as follows

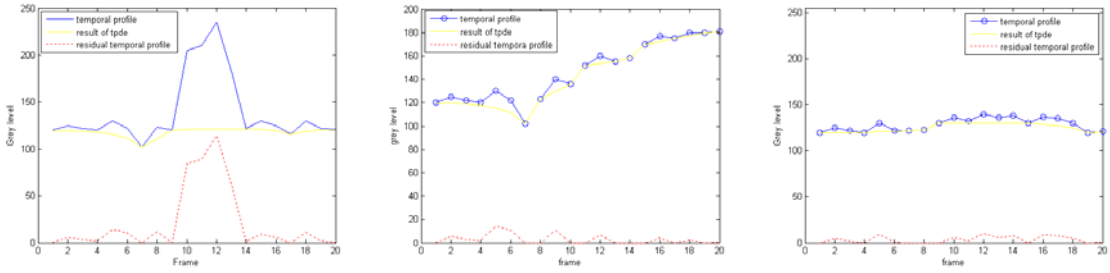
$$c(\nabla \mathrm{u}) = \exp[-(\frac{|\nabla u|}{k})^2] \qquad (7)$$

Since the temporal profile is discrete one-dimensional data. Then the formula (7) can be adapted as (8) when the partial differential equation is used to smooth the curve.

$$u^{n+1}(k) = u^n(k) + \alpha[c(\nabla L)\nabla L + c(\nabla R)\nabla R] \qquad (8)$$

where， $\nabla L = u(k) - u(k-1), \nabla R = u(k) - u(k+1)$ ， n is iterations.

By the method of partial differential equation, the reference line of the temporal profile can be obtained, and then, we can employ difference of the temporal profile and this reference line to remove background, reserve target impulse. As show in Fig. 4, the proposed method performs well in pixel of target, cloud edge and inner cloud.



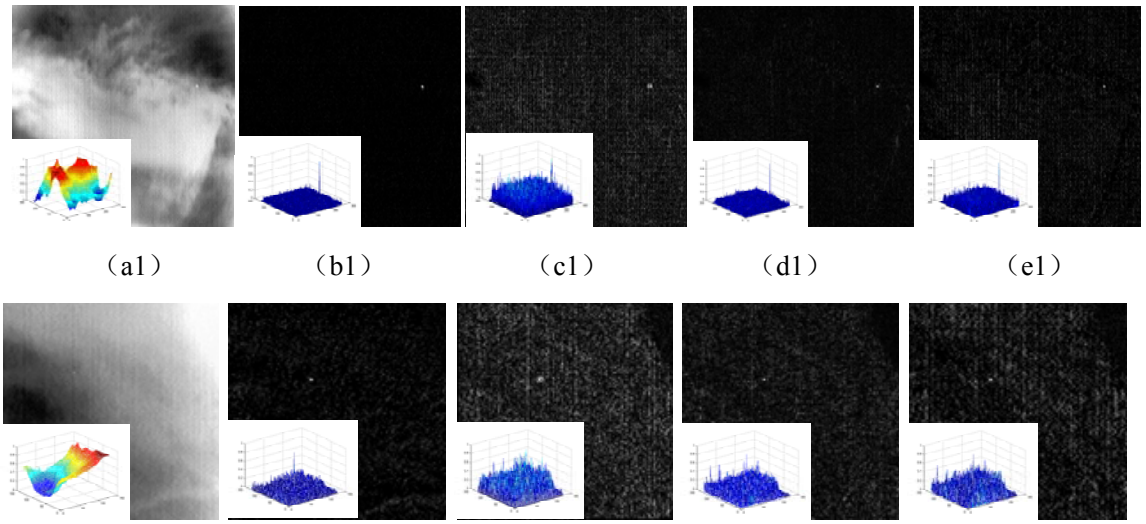
(a) (b) (c)

Fig.2. Result of PDF acted on temporal profile (a) target (b) cloud edge pixel (c) inner cloud

4. Experiments

To validate the effectiveness of the proposed method in this paper, the simulation experiments for two real low SCR infrared imagery sequences containing the moving dim small targets with sky background were done. The performance of different methods was appraised by metrics signal-to-clutter ratio gain (SCRG) and background suppress factor (BSF) [16]. In this section, three detection methods are tested with real IR imagery sequences.

Fig. 3 shows the detection results of these images using max-mean filter, top-hat and TDLMS respectively. Fig.3 (a1~a2) show the one frame image of original imagery sequence with moving target under clutter cloud backgrounds. The target hardly cannot be identified from the 3D map. From Fig. 3(c1~ c2), it shows that the max-mean based methods can enhance the target energy, but the images still contain a lot of clutter backgrounds. Top-hat and TDLMS based methods perform well at the first IR imagery sequence, but they have a poor ability to do well in low SCR image (see Fig. 3(d2) and (e2)). Target cannot be effectively separated from background clutter, which can make the false alarm rate increased. Compared with these three detection methods, the proposed method has a better effect on the clutter suppression and has the excellent results in low SCR imagery sequence.



（a1） （b1） （c1） （d1） （e1）

（a2） （b2） （c2） （d2） （e2）

Fig.3. Experiment results comparison between the classical methods and the proposed method.

（a1~a2）original image

results of (b1~b2) proposed method (c1~c2) Max-mean method (d1~d2) TDLMS method (e1~e2) top-hat method

Two objective quality metrics are shown in Tab.1 to assess the performance of different methods, in which the values in boldface indicate the higher quality. It is indicated that the proposed method has the larger values of SCRG and BSF. From the Tab. 1, it is clear that the proposed method not only has good effect in terms of improving the image contrast enhancement but in cluster suppression property, especially for low SCR images.

Tab.1 Performance comparison of different methods

Original image		Proposed method		Max-mean		TDLMS		Top-hat	
Index	SCR	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF
1	2.5456	4.5902	4.5681	3.2122	3.8955	1.7853	3.1563	3.1566	4. 1656
2	0.6786	7.4294	6.6268	4.0263	3.7366	1.8804	3.1947	2.0642	4.3606

5 conclusion

In this paper, aiming at the problem of detection of near-space hypersonic target for space-based infrared detection system, a temporal background suppression based infrared dim small target detection method has been proposed. The method takes full advantage of the motion information of the target, and uses the method of partial differential equations in temporal domain to predict the background and detect the target. Compared with several classical methods, the experimental results show that the performance of proposed method is superior to several other methods both in aspects of subjective visual and objective evaluation index.

References

[1] Revital H S, Ofer H, Rotman S R, et al. Compression of infrared imagery sequences containing a slow-moving point target, part II.[J]. Applied Optics, 2013, 52(8):1646-1654.

[2] Gu Y, Wang C, Liu B, et al. A Kernel-Based Nonparametric Regression Method for Clutter Removal in Infrared Small-Target Detection Applications [J]. IEEE Geoscience & Remote Sensing Letters, 2010, 7(3):469-473.
[3] Bai X, Zhou F, Xue B. Infrared dim small target enhancement using toggle contrast operator [J]. Infrared Physics & Technology, 2012, 55(2-3):177–182.
[4] Chengjun L, Ying W, and Zeling S, A small target detection algorithm based on multi-scale energy cross [C] //IEEE International Conference on Robotics, Intelligent Systems and Signal Processing, 2003: 1191-1196
[5] Deshpande S D, Meng H E, Venkateswarlu R, et al. Max-mean and max-median filters for detection of small targets[C]// SPIE's International Symposium on Optical Science, Engineering, and Instrumentation International Society for Optics and Photonics, 1999:74-83.
[6] Rivest J F, Fortin R. Detection of dim targets in digital infrared imagery by morphological image processing[J]. Optical Engineering, 1996, 35(7):1886-1893.
[7] Cao Y, Cao Y, Liu R M, Yang J. Small Target Detection Using Two-Dimensional Least Mean Square (TDLMS) Filter Based on Neighborhood Analysis [J]. International Journal of Infrared & Millimeter Waves, 2008, 29(2):188-200.
[8] Bae T W, Sohng K I. Small Target Detection Using Bilateral Filter Based on Edge Component [J]. Journal of Infrared Millimeter & Terahertz Waves, 2010, 31(6):735-743.
[9] Zhao Y, Pan H, Du C, et al. Bilateral two-dimensional least mean square filter for infrared small target detection[J]. Infrared Physics & Technology, 2014, 65(5):17–23.
[10] Wang A Z, Tian J, Liu J, et al. Small infrared target fusion detection based on support vector machines in the wavelet domain[J]. Optical Engineering, 2006, 45(7):076401.
[11] Qin H L, Li J, Zhou H X, et al. Infrared dim and small target background suppression using shearlet transform[J]. Journal of Infrared & Millimeter Waves, 2012, 30(2):162-166.
[12] Cunha A L D, Zhou J, Do M N. The Nonsubsampled Contourlet Transform: Theory, Design, and Applications[J]. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society, 2006, 15(10):3089-3101.
[13] Silverman J, Mooney J M, Caefer C E. Temporal filters for tracking weak slow point targets in evolving cloud clutter[J]. Infrared Physics & Technology, 1996, 37(6):695-710.
[14] Tzannes A P, Brooks D H. Temporal filters for point target detection in IR imagery[C]// AeroSense '97International Society for Optics and Photonics, 1997:508-520.
[15] D. Liu, J. Zhang, W. Dong, Small target detection algorithm based on average absolute difference maximum and background forecast, Interna-tional Journal of Infrared and Millimeter Waves 28 (5) (2007) 373–381.
[16] Shengxiang Qi, Jie Ma, Chao Tao, et al. A Robust Directional Saliency-Based Method for Infrared Small-Target Detection Under Various Complex Backgrounds [J]. IEEE Geoscience\s&\sremote Sensing Letters, 2013, 10(3):495-499.

A Novel Design for Satellite Communications Signal Spectrum Auto-Monitor Method

Zhang Fen ZuoYanqing Shen Ling

China Satellite Marine Track &Control Department,Jiangyin 214431,China

Abstract: A multi-channel spectrum on-line auto-monitor method for satellite communications signal is designed, which is different from the traditional single-channel method based on the spectrum analyzer. To realize the novel method, a suitable signal monitoring point is selected firstly, a system architecture of "monitoring point & directional coupler & integrated matrix network & spectrum test unit & monitoring terminal" is built secondly, and a comprehensive testing software is used for testing points selection and testing indicators control. As a result, real-time multi-channel spectrum signal analysis, threshold alarms and continuous sampling process saving is realized. Proven, real-time, and high data accuracy, improve work efficiency, the promotion and practical value.

Keywords: Spectrum monitoring, GPIB, integrated matrix network

1 Introduction

With the rapid development of varieties ofauto-monitoring technology,signal monitoring based on the computer-aided tools and the higher-degree auto-equipment becomes more effective, simple and flexible. In the satellite communications, the spectrum monitoring of the satellite communication station is an important way to check the work condition of bio-systems with RF. The spectrum analyzer plays an indispensable role.And the signal monitoring points are often used as the signal-test-ports of the frequency inverter, the division multiplexer and the modem under RF(RadioFrequency) or IF(Intermediate Frequency). In this paper, a novel method of the signal spectrum automatic monitoring is proposed, which is combined with both traditional spectrum monitoring methods and the spectrum analyzer GBIP ports.

2 Traditional spectrum monitoring methods

In the current field of the satellite communications, the main way of the real-time signal spectrum monitoring is that by linking the spectrum analyzer intothe communication-link to collect the RF or IF signal, the staff operates according to the image displayed on the spectrum analyzer, as shown in Fig. 1. This traditional method is commonly used. But there are several significant shortcomings. Firstly, the monitoring signal could not be switched automatically for the single input. Secondly, it depends on the operation of the staff, either the parameter setting or the data judgment. Thirdly, the monitoring system is lack of both the auto-feedbackand the alarm unit, which requires good care of the system. Fourthly, the monitoring information could not be auto-stored, and is inconvenient for the signal statisticsand analysis. At last, the information could not be processed automatically. Above all, the traditional method relies mostly on the manual work with low efficiency and difficult data-collection. This is not with the pace of the development of the modern information technology.

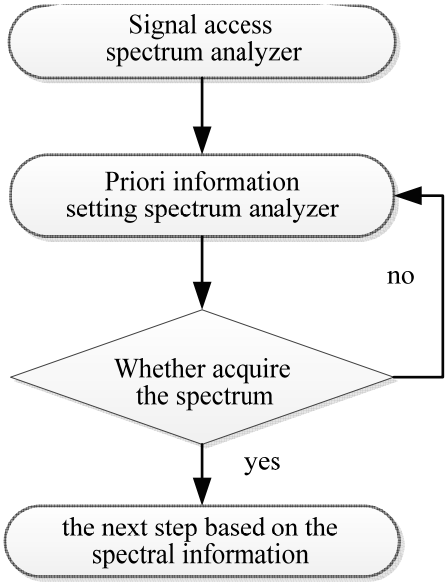


Fig. 1The workflow of the traditional method relied on the spectrum analyzer

3 Auto-monitoring method based on the GPIB port

GPIB(General Purpose Interface Bus) isa communication protocol for automated test systems between the various devices, to provide users with the remote automated monitoring capabilities. This auto-monitoring method is based on the spectrum analyzer GPIB port. By the input and output of GPIB, the automatic monitoring functions such as real-time acquisition, analysis, recording and alarm for spectrum monitoring data are realized. Thus,state awareness, analytical capacity and the monitoring efficiency of the satellite communications system is improved. Thesystem consists of hardware and software in two parts. The hardware of the system consists ofthe spectrum analyzer, GPIB-USB converter, and the auto-monitoring computer. The connection is shown in Figure 2. The software is the core of the auto-monitoring system. The platform is provided by the monitoring computer, and the application program is developed according to the requirements of users. The main functionsare as follows.

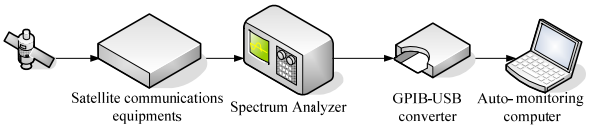


Fig. 2Hardware connection of the spectral auto-monitoring equipment based on the GPIB port

1）Data monitoring and early alarm

To realize the signal auto-monitoring, the relevant data is captured by GPIB of the spectrum analyzer. or the When the abnormal phenomena occur, such as the drastic fluctuations of signals and the loss of carrier wave, the monitoring system could offer an alarm and make a record.

2）Signal recording and analysis

The monitoring computer through GPIB port periodically sends commands to the spectrum analyzer, acquires the spectrum information continually and establishes a database. In this way, the real-time signal acquisition, recording and quantitative analysis is realized. On the edge of the satellite beam coverage area, the signal level changes very frequently. Thus, accurate and long-term signal level record is particularly necessary to acquire the satellite communication state in time.

3）Spectrum management

Establishment of the communication frequency resource library based on the current frequency data,is to analyze and compare the auto-acquired and real-time occupancy spectrum data. By this way, the spectrum conflict and communication interference could be prevented, especially for the problems as spectrum occupancy and spectrum sharing.

Software development is based on LabWindows and Visual C ++. The application programis composed as five functional modules: driver, data communication, real-time storage, real-time display, abnormal alarm. The whole function is realized by the integration and dispatch under the total controlled software.

4 A novel method of multi-channel spectrum on-line auto-monitoring

The novel auto-monitoring method is mainly realized by the introduction of the integrated switch matrix and the design of a multi-channel spectrum monitoring unit to achieve a comprehensive test for the communication signal. The main functions include the multi-channel signal monitoring and early alarmof satellite communication links. The comprehensive test computer(including software and hardware), the integrated matrix, the multi-channel spectrum monitoring equipment, spectrum analyzer and so on are needed. The system components are shown in Fig. 3. The comprehensive test computer is used for remote control of varieties of instruments and test deviceswhich connected to it via the control interface or network, to set the device parameters, to obtain real-time spectral data and to display the test and analysis results.The integrated network management system obtains various technical indexes and work status data of each satellite communication station device via the comprehensive test computer, and obtains parameters and state of each antenna via the antenna test subsystem.

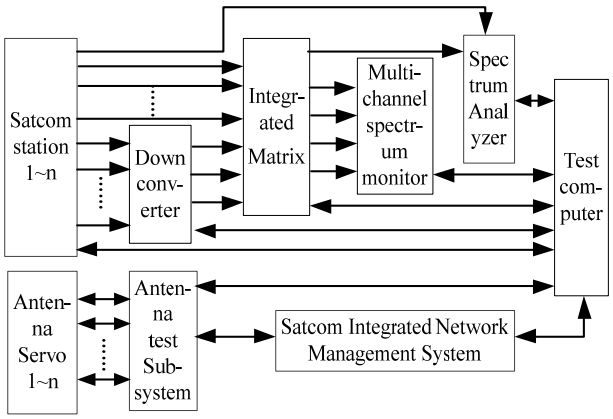


Fig. 3Block diagram of the integrated test system

4.1 Selection for the switch matrix

To solve the problem of different test points switched manually, the integrated matrix is added into the satellite communications link. Thus, boththe flexibility and the efficiency of the testisimproved. To switching the signal,it could be realized by the panel operation of the matrix unit, or by the smart software withthe data line and RS485. Both the 8 × 4 and 14 × 5 integrated matrix networksare provided for choice to realize the 8 inputs and 4 outputs with non-blocking and full exchange. And the function that the same inputs are chosen by different multi-output is included. As well,the port isolation of the switchmatrix, switching time, usage lifetime,frequency range, port SWR(Standing-Wave Ratio) must meet the application requirements.

4.2 Monitoring points

The input or output of the communication device is often chosen as the monitoring point of the integrated monitoring system. The up-link include the modem, the up-converter the output port of the power amplifier, while the down-link include the low noise amplifier, the down-converter output port.Theinput signal is leaded from the monitoring point to the matrix network, and the output signal is displayed the amplitude, frequency, spectral characteristics and other parameters by the terminal device. According to all the parameters, whether the output signal normal is judged, and an alarm is given if the parameters over the threshold. For the remote satellite communications station, LTE (Lightware Terminal Equipment) can be used to monitor the optical output spectrum.

The monitoring should be online test without affecting the normal communication. Therefore the signal from the monitoring point to the system requires a splitter or a directional coupler. The splitter isolation of the two output paths is around 20dB, and the channel loss is around 4dB. While the directional coupler (10dB) isolation of the two output paths is about 30dB, and channel loss is about 0.5dB. In contrast, the directional coupler has advantages such as high output isolation, low loss, less access to load and little effect on the communication system. So, the directional coupler is as the choice. Under the condition of normal communication, considering a monitoring point calculated in accordance with 0.5dB loss, the impact of the loss of communication can be completely eliminated through pre-designingthe channel gain.

4.3 Multi-channel spectrum monitoring unit

The output signal of the integrated matrix is accessed to the multi-channel spectrum monitoring equipment. According to the requirements of main program of the comprehensive test software monitoring the input signal could be selected and sent to the monitoring computer, to realize the sustained sample process.

4.4 Softwareworkflow

The basic test indexes for the signal spectrum monitoring are as follow: bandwidth, alarm forover high power, interfering signals, and so on. Under the condition of simultaneous monitoring for multiple points, the multi-object displaying mode is designed by software.Theworkflow of the online testing for the Satellite communication signal spectrumshown in Fig.4.Comprehensive testing software could display the real-time data online, and storethe data as the history record when needed, and load spectrum patternin an limited time to the cache for analysis and procession later. After the online monitoring started,the historical data maintains in the latest database without increasing, just one item, when the testing result is normal, while the data increases by one item when the testing result changessuch as the fault conditions change, or the fault state turn to the normal.

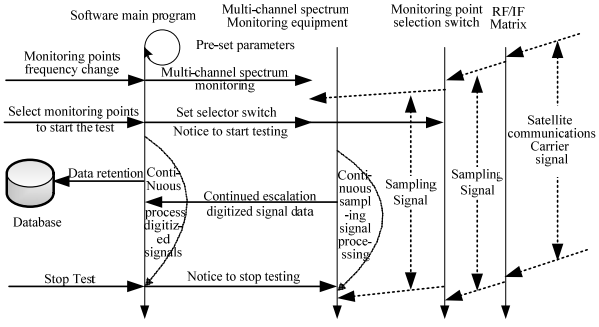


Fig. 4The workflow of multi-channel signal spectrum on-line auto-monitor

5 Conclusion

In the field of satellite communications, monitoring the signal spectrum is an important mean to monitor the work state of the RF or IF system. This novel signal spectrummonitoring of signal spectrum auto-monitoring method is improved greatly.In this method, there are definite functions with the integrated matrix unit, multi-channel spectrum monitoring unit and so on. Also, the abnormal data feedback alarm unit could make up the capture capability for the rapid interruption, and the analysis based on the database storage is up to the development of Big Data.

Reference

[1] Liu Mingbo, Chang Jishi. The design and implement of Satellite communications signal spectrum Auto-Monitor system.[J] Foreign Electronic Measurement Technology. 2012, 31(8)

[2] Liu Shunguo, He Wenxu. Virtual test system based on GPIB bus design and implementation.[J]. China Measurement Technology.2006(06):68-70.

[3] Zhang Wei. Design for Satellite Network IF Signal Frequency Spectrum Monitoring Equipment development. [D]. BUPT. 2011.

[4] Wang Qing. Full frequency electromagnetic spectrum monitoring equipment overall plan and design. [D].Xidian University. 2012.11.

[5] Zhou Haiying, Huang Jinyuan. Reliable Receiving of Meteorological Satellite Signal Based on Matrix.[J] Modern Electronic Technology. 2011,34:101-103.

[6] Zhang Pengxiang. The automatic test system of multi-port radio frequency network. [D].Soochow University. 2013.4.

[7] Du Weiwei, Wang Chenlin.The RF component of airborne inspection platform based on GPIB. [J]. Electronic Measurement Technology.2012.6:113-115.

[8] Yan Fengqing, Qian Cheng. RF Switch and its Applications in Communication System. [J].Chinese Journal of Electron Devices. 2005.3

An Efficient Iterative Least-Squares Pattern Synthesis Method for Satellite Communication

Feng Danping[1]，Ma Xiaofeng[1]，Wu Yuqing[2]，Sheng Weixing[1]，Xiao Zhengming[2]，Shen Peng[2]，Zhong Ruoqing[1]

（1. School of Electronic and Optical Engineering, Nanjing University of Sci. & Tech., Nanjing ,China)

（2.Shanghai Aerospace Electronics Co., Ltd., Shanghai, China)

Abstract：Low earth orbit（LEO）satellite communication phased array antenna or digital beam-forming antenna not only forms fixed multi transmit and receive beams covering the visible earth, but also has the ability to dynamically adjust the pattern according to requirements of communication task. In this paper, an efficient iterative least-squares pattern synthesis method is proposed with simpler iteration process and less adjustable parameters compared to the most popular iterative algorithms under the same synthesis accuracy and sidelobe level requirements. The proposed method can partially meet the online weight vector update requirement for satellite communication pattern dynamic synthesis.

Keywords: Pattern Synthesis; Iterative Least-squares; Efficient Method; Satellite Communication

I. Introduction

LEO satellite communication generally needs to form multi transmit and receive beams, covering the visible earth with isoflux and low sidelobe radiation patterns [1, 2]. Mainlobe service region isoflux is considered to realize equal sensitivity communication with the user in anywhere of the earth visible region, and low sidelobe is used to reduce the cochannel interferences of adjacent beams or other non-malicious sidelobe interferences [3].

There are many methods to synthesize a desired pattern, such as Woodward-Lawson method [4], Taylor and Chebyshev synthesis methods [4], stochastic optimization methods [5], alternating projection (AP) [6],least mean squares (LMS) and weighted LMS (WLMS) methods [7], weighted alternating reverse projection (WARP) [8], and so on.

With the application of active phased array and digital beamforming techniques, LEO satellite communication array antenna has the ability to dynamically adjust the pattern according to requirements of communication task. But the existing high performance pattern synthesis optimization algorithms proposed in [5-8] can't be directly used to space-borne online real-time computing, for the reasons of computation load, convergence speed and adjustable parameters' selection. In this paper, an efficient iterative least-squares pattern synthesis method is proposed with the demand of isoflux mainlobe and low sidelobe. Under the same synthesis accuracy and sidelobe level requirements，the proposed method has simpler iteration process and less adjustable parameters, which means less computation load and more robustness compared to the most popular iterative algorithms recently. The proposed method can partially meet the online weight vector update requirement for satellite communication pattern dynamic synthesis.

II. Problem Formulation

The antenna array is shown in Fig. 1. It is composed of N antenna elements and lies on lattices of equilateral triangle with 0.6 wavelength side in the xoy plane. All of the antenna elements are the same microstrip antenna with rectangle truncated corners and mutual coupling is considered through full-wave simulation.

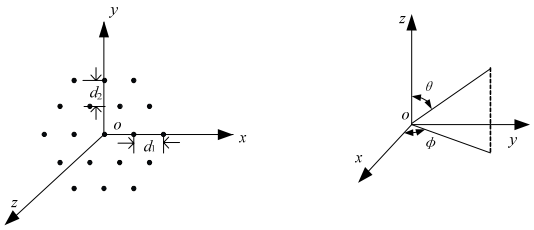


Fig. 1 Array geometry

The array antenna pattern can be denoted by

$$F(\theta,\varphi) = \boldsymbol{w}^{\mathrm{H}}\boldsymbol{a}(\theta,\varphi) \tag{1}$$

Where $w=(w_1, w_2, \ldots , w_i, \ldots, w_N)^{\mathrm{T}}$, i=1,2,…,N, is the weight vector; the superscript H denotes the complex conjugate transpose and the superscript T denotes the transpose; as shown in Fig. 1, θ is elevation angle and φ is azimuth angle. The elements $a_i(\theta, \varphi)$ of the geometry-dependent vector a(θ, φ) are

$$a_i(\theta,\varphi) = g_i(\theta,\varphi)e^{jk_0(x_i\sin\theta\cos\varphi + y_i\sin\theta\sin\varphi)} \tag{2}$$

Where (x_i, y_i) is the position of the ith antenna element and $g_i(\theta, \varphi)$ is the ith element radiation pattern considering the mutual coupling among elements. k_0 is the wave number $2\pi/\lambda$ and λ is the wavelength.

The array antenna pattern is divided into two parts: mainlobe region Q_S and sidelobe region Q_P. It is desirable that the resulting antenna pattern be close to the desired pattern in the mainlobe region and as low as possible in the sidelobe region. The problem is formulated as follows

$$\begin{aligned} J(\boldsymbol{w}) = &\int_{\Omega_S} \left|F(\theta,\varphi) - F_0(\theta,\varphi)\mathrm{e}^{\mathrm{j}\zeta(\theta,\varphi)}F_{\max}\right|^2 \mathrm{d}\theta\mathrm{d}\varphi \\ &+K\int_{\Omega_P}\left|F(\theta,\varphi)\right|^2\mathrm{d}\theta\mathrm{d}\varphi \end{aligned} \tag{3}$$

Where, the first item in Eq. (3) represents the deviation of the resulting pattern and the desired pattern in the mainlobe region. $F_0(\theta, \varphi)$ is the desired normalized amplitude pattern which satisfies isoflux coverage requirement; F(θ, φ) is the resulting pattern; ξ(θ, φ) and F_{max} are the phase of F(θ, φ) and the maximum value of |F(θ, φ)| respectively, updating in each iteration. Generally, the desired pattern is specified in a normalized form and its fixed magnitude and phase factor are disinterested. The second item in Eq. (3) represents the power of sidelobe region. Positive real number K is weight factor traded off between two items in Eq. (3).

Eq. (3) can be reformulated as

$$\begin{aligned} J(\boldsymbol{w}) = &\int_{\Omega_S} \left|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{a}(\theta,\varphi) - F_0(\theta,\varphi)\mathrm{e}^{\mathrm{j}\zeta(\theta,\varphi)}F_{max}\right|^2 \mathrm{d}\theta\mathrm{d}\varphi \\ &+K\int_{\Omega_P}\left|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{a}(\theta,\varphi)\right|^2\mathrm{d}\theta\mathrm{d}\varphi \\ = &\boldsymbol{w}^{\mathrm{H}}(\boldsymbol{R}_S + K\boldsymbol{R}_P)\boldsymbol{w} - \boldsymbol{w}^{\mathrm{H}}\boldsymbol{r}_S - \boldsymbol{r}_S^{\mathrm{H}}\boldsymbol{w} + \boldsymbol{R}_0 \end{aligned} \tag{4}$$

Where, each element of a(θ, φ) is defined in (2) and the substitutions are shown in (5)~(8).

$$\boldsymbol{R}_S = \int_{\Omega_S} \boldsymbol{a}(\theta,\varphi)\boldsymbol{a}^{\mathrm{H}}(\theta,\varphi)\mathrm{d}\theta\mathrm{d}\varphi \tag{5}$$

$$\boldsymbol{R}_P = \int_{\Omega_P} \boldsymbol{a}(\theta,\varphi)\boldsymbol{a}^{\mathrm{H}}(\theta,\varphi)\mathrm{d}\theta\mathrm{d}\varphi \tag{6}$$

$$\boldsymbol{r}_S = \int_{\Omega_S} \boldsymbol{a}(\theta,\varphi)\left(F_0(\theta,\varphi)\mathrm{e}^{\mathrm{j}\zeta(\theta,\varphi)}F_{\max}\right)^{\mathrm{H}}\mathrm{d}\theta\mathrm{d}\varphi \tag{7}$$

$$\boldsymbol{R}_0 = \int_{\Omega_S} F_0(\theta,\varphi)\mathrm{e}^{\mathrm{j}\zeta(\theta,\varphi)}F_{\max}\left(F_0(\theta,\varphi)\mathrm{e}^{\mathrm{j}\zeta(\theta,\varphi)}F_{\max}\right)^{\mathrm{H}}\mathrm{d}\theta\mathrm{d}\varphi \tag{8}$$

Then, calculate the gradient of J(w) with respect to w, and set it to zero.

$$\nabla J(\boldsymbol{w}) = \frac{\partial J(\boldsymbol{w})}{\partial \boldsymbol{w}} = (\boldsymbol{R}_S + K\boldsymbol{R}_P)\boldsymbol{w} - \boldsymbol{r}_s = 0 \tag{9}$$

Let $R_{SP}= R_S +KR_P$， which is determined when the mainlobe region and the sidelobe region are set. The optimal weight vector is

$$\boldsymbol{w}_{opt} = \boldsymbol{R}_{SP}^{-1}\boldsymbol{r}_s \tag{10}$$

It must be pointed out that Eq. (10) is optimal under the circumstance of certain fixed F_{max} and ξ(θ, φ). However under the normal circumstance, the maximal amplitude of the resulting pattern F_{max} is unknown and the phase vector of the resulting pattern ξ(θ, φ) is disinterested. Therefore, a new iterative method is proposed. The weight vector of previous iteration is used to calculate new F_{max} and ξ(θ, φ), substituting for the expected F_{max} and ξ(θ, φ) of the next iteration. The iterative formula for computing the weight vector is as follows

$$\boldsymbol{w}(k+1) = \boldsymbol{R}_{SP}^{-1}\boldsymbol{r}_s(k) \tag{11}$$

In summary, the process of the iterative optimization method is given as follows:

(1). Set the desired normalized amplitude pattern F_0(θ, φ) and initial w(0);（The w(0) can be selected as a(θ_0, φ_0), where (θ_0, φ_0) is the angle at maximum value of F_0(θ, φ)）

(2). Calculate desired F(θ, φ) in the mainlobe region. Calculate F_{max} and ξ (θ, φ) according to F(θ, φ);

(3). Calculate the kth r_s according to Eq. (7);

(4). Calculate the (k+1)th weight vector w(k+1) according to Eq. (11);

(5). If stopping criterion is satisfied, the algorithm would stop iterations, else go to step (2).

III. Synthesis Examples

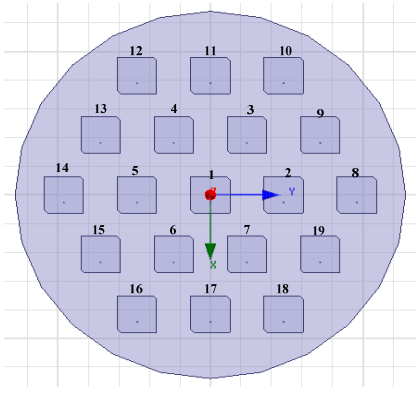


Fig. 2 The arrangement of the 19-elements array

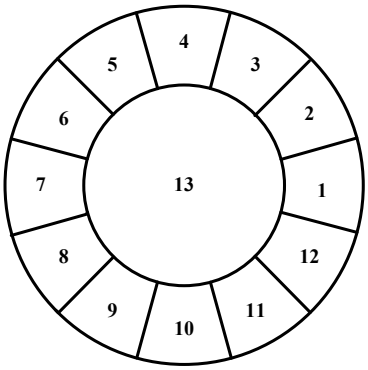
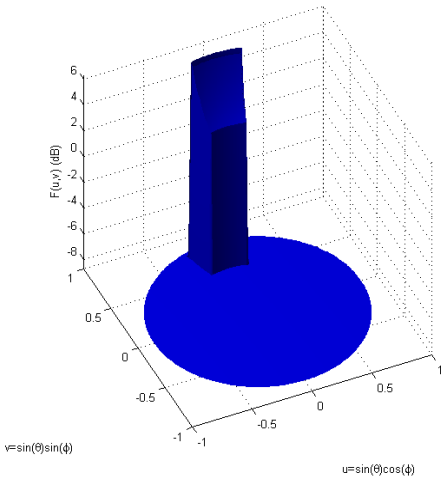
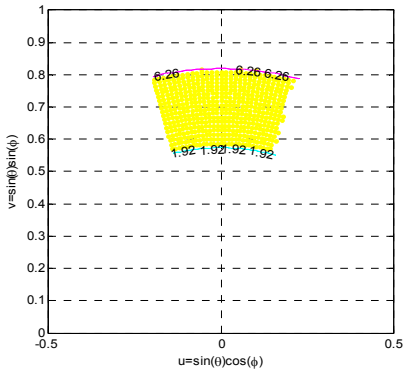


Fig. 3 Array antenna pattern coverage requirement

Fig. 2 shows the 19-element array of same microstrip element with rectangle truncated corners, and Fig. 3 gives the multi-beam array antenna pattern coverage requirement. We can see that the 13 patterns are classified into two types, central pattern and 12 edge patterns. We just select pattern 4 as example for page limitation. The shaped region of pattern 4 is a footprint with location of $\theta \in [35^0, 55^0]$ and $\varphi \in [75^0, 105^0]$. The desired synthesis pattern of the edge pattern is given in Fig. 4 which contains the normalized 3D patterns and the contour gain coverage plots. The pattern is plotted in UV plane, where u=sin(θ)cos(φ) and v=sin(θ)sin(φ).



(a) The normalized 3D pattern of the edge beam



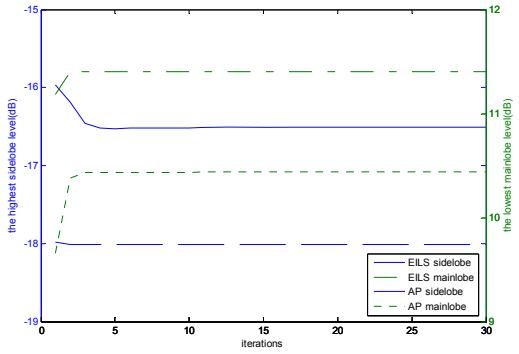
(b) The contour gain coverage of the edge beam

Fig. 4 The desired synthesis pattern

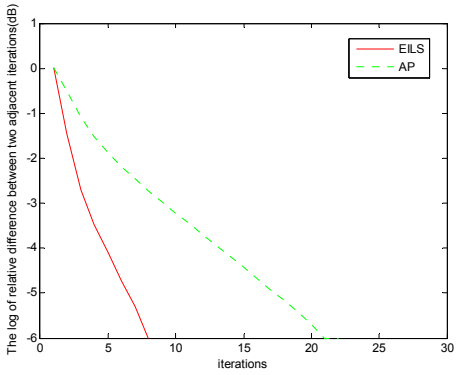
The mainlobe region of edge beam pattern 4 is covering $\theta \in [35^0, 55^0]$ and $\varphi \in [75^0, 105^0]$ with isoflux, the transient region is set to 10^0 and the rest is sidelobe region with low sidelobe

requirement up to -15dB. For comparison, the synthesis results of AP [6] are provided, with the same requirement as EILS and the mainlobe ripple is within ±0.5dB. To keep the balance of mainlobe level and sidelobe level, K in (3) should be chosen comprehensively in our method. The value of K should be larger when the mainlobe region is smaller and sidelobe level is lower. K=7 is set for this simulation.

The synthesized results for both EILS and AP are shown in Fig. 5 and Fig. 6. Fig. 5(a) depicts the tendency of the minimum gain in location of $\theta = 55^0$ at mainlobe region and the maximum sidelobe level over iterations. It is shown that mainlobe gain of EILS is up to 11.4 dB, while AP can just reach 10.4dB. Meanwhile, the sidelobe level of EILS is -18dB, 1.5 dB lower than AP. Fig. 5(b) presents the log of relative difference of the weight vector between two adjacent iterations. It is shown that EILS converges faster than AP. The synthesized patterns of EILS after 6 iterations and of AP after 21 iterations are shown in Fig. 6(a) and Fig. 6(b) respectively.

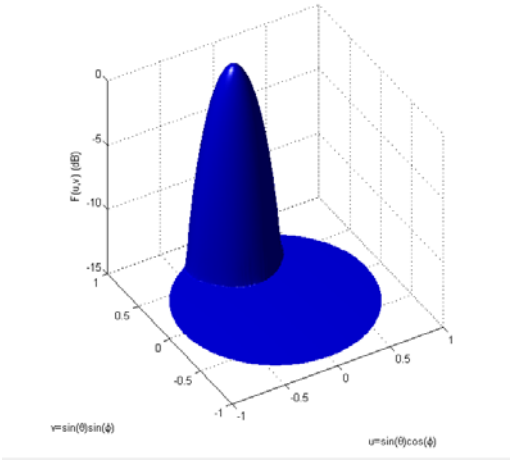


(a) Sidelobe level and the mainlobe level over iterations

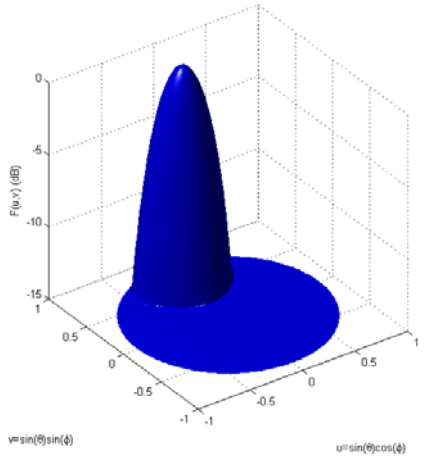


(b) Relative difference between two adjacent iterations

Fig. 5 The comparison of algorithm convergence



(a) EILS after 6 iterations



(b) AP after 21 iterations

Fig. 6 The synthesized pattern

The weight vector of EILS is optimal in each iteration under the previous iteration F_{max} and $\xi(\theta, \varphi)$. As a result, EILS has fast convergence. Furthermore, AP should take fine adjustment of every angle in sidelobe region to control sidelobe level, while EILS just need to adjust the value of K. So the computation of EILS in each iteration is much smaller.

It is assumed that the beam pattern sampling points in mainlobe region are P_s and in all shaped region are P_a. EILS takes $(2NP_s+5P_s+ N^2+C_{sqrt})$ complex multiplications and $(2(N-1)P_s+P_s+(N-1)N)$ complex additions in each iteration. AP yet needs $(2NP_a+4P_a+ N^2+C_{sqrt})$ complex multiplications and $(2(N-1) P_a+P_a+(N-1)N)$ complex additions. C_{sqrt} represents the computation load of square root. It is shown that the computation load of EILS and AP in each iteration is decided mainly on P_s and P_a respectively when the number of elements N is not so large. In fact, P_s is always smaller than P_a, so the computation load in each iteration is lighter than AP. The smaller the mainlobe region is, the lighter the computation load is. In this example, P_s=864 and P_a=32851, so the computation load of AP in each iteration is about 38 times larger than EILS. In general, the whole computation load of EILS is much smaller than AP for its faster convergence rate and lower computation load in each iteration.

Simulation results show that changing the maximum sidelobe level in AP cannot well improve the sidelobe level. In order to get the same performance as EILS, some basic parameters like range

of transient region and shaped mainlobe region have to be adjusted manually. Whereas, EILS just need to adjust the value of K, which is simpler and more efficient.

IV. Conclusion

In this paper, an efficient iterative least-squares pattern synthesis method is proposed. Simple synthesis example for certain LEO satellite communication coverage requirements is taken into account under an equally spaced triangular lattice 19 element array of microstrip antenna with rectangle truncated corners. The simulation and analysis results show that EILS is simpler, more efficient and robust compared to the most popular iterative algorithms such as AP, under the same synthesis accuracy and sidelobe level requirements. The proposed method can partially meet the online weight vector update requirement for satellite communication pattern dynamic synthesis.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under grant 61501240.

.

References

[1] Metzen, P.L. , Globalstar satellite phased array antennas, IEEE International Conference on Phased Array Systems and Technology ,2000,pp.207-210.

[2] Schuss, J.J. ,Upton, J. , Myers, B. , et al., The Iridium main mission antenna concept, IEEE Trans. on Antennas and Propagation, 1999,47(3),pp.416-424.

[3] Winston Li, Xinping Huang, Henry Leung, Performance evaluation of digital beamforming strategies for satellite communications, IEEE Trans. on Aerospace and Electronic Systems, 2004, 40(1):pp.12-26.

[4] C. A. Balanis, Antenna Theory: Analysis and Design, 3rd ed. New York, NY, USA: Wiley, 2005.

[5] W. B. Daniel and H. W. Douglas, Particle swarm optimization versus genetic algorithms for phased array synthesis, IEEE Trans. Antennas Propag., 2004, 52(3):pp. 771-779.

[6] Bucci, O.M., Elia, G.D., Romito, G.,Power synthesis of conformal arrays by a generalized projection method, Proc. Inst. Elect. Eng., Microw. Antennas Propag., 1995, 142(6): pp. 467–471.

[7] Vaskelainen L I. Iterative least-squares synthesis methods for conformal array antennas with optimized polarization and frequency properties. IEEE Trans. on Antennas and Propagation., 1997, 45(7):pp. 1179~1185.

[8] A. Haddadi, A. Ghorbani, and J. Rashed-Mohassel, Cosecant-squared pattern synthesis using a weighted alternating reverse projection method, IET Microw. Antennas Propag., 2011,5(15):pp.1789-1795.

MINIMIZING LATENCY BY USING EXISTING GLOBAL GROUND STATION NETWORKS

Kenneth Olafsson

[1]Kongsberg Satellite Services AS, Prestvannveien 38, P.O. Box 6180, N-9291 Tromsø, Norway,

Email: kennetho@ksat.no

ABSTRACT: It takes a Low Earth Orbiting satellite approximately 100 minutes to orbit the Earth. With a Ground Station near one pole it is possible to achieve a contact with this satellite on each orbit. However, for a number of applications 100 minutes is not an acceptable latency as disaster will have developed (disaster monitoring), clouds will have moved (Meteorology), threats may have developed further (Security) and polluters may have left the crime scene (oil pollution monitoring). In principle the problem is that satellites are too slow!

Adding satellites in a constellation is one solution; however it is much more expensive than adding ground stations. Another ground station near the opposite pole will reduce the latency to about 50 minutes and satisfy a number of applications like Meteorology. However, for some aspects this is not fast enough (security, oil spill monitoring, disaster monitoring etc.).

The world is not fairly distributed. About 75 percent of the land mass lie north of the equator, and 80 percent of the world's population, 85 percent of the capital and 90 percent of the conflicts are in these areas.

As optical satellites usually record their descending passes it makes sense to locate stations on the Southern tip of the land masses, receiving the data shortly after it has been recorded by the satellite. Commercial networks of stations near both poles and along the Southern fringe of the globes land mass exist. This results in a reduced latency for land applications. Broad communication lines (>155 MBit) are in place to many of these stations and data can be received anywhere in the world by adding a single Internet port.

KEYWORDS: latency, ground station, station networks

1. INTRODUCTION

Low Earth Orbiting Satellites circle the globe mostly in sun-synchronous orbits. This guarantees global coverage and assures that the on-board power system has enough energy available to operate the payload. It also means that on each orbit the satellite passes close to both poles, while the rotating Earth is covered in strips that enable a complete global coverage in a number of days, depending on the elevation of the satellite, the agility of the satellite, the payload characteristics and the exact orbit.

This makes it obvious that a location near a pole gives the best access to the satellite. As the Northern Hemisphere is much more accessible than Antarctica the first stations that were established in the late sixties were located in Alaska, Tromsø (Norway) and Kiruna (Sweden). The stations in Kiruna and Tromsø were the first to establish services on behalf of third parties (e.g. ESA). Later more satellite owners recognized the advantages of a Northern location and enabled the start of the commercial Ground Station business, where Northern Stations either host an antenna of

a satellite owner or where the satellite owner buys a time slot for the duration of the contact on an existing antenna that is owned by the Ground Station – this is called a satellite pass. The stations mentioned above are typically located at 65-70 degrees north and allow 10-11 contacts out of 14 possible each day.

This changed when KSAT established a first antenna in Svalbard in the late 1990s at 78 degrees north allowing 14 out of 14 contacts. The location became a great success when a redundant fiber connection allowed repatriating data with a latency of 90 minutes (each orbit) without the gaps for passes that cannot be received at other polar stations. Since then the KSAT Svalsat stations on Svalbard has grown to be the biggest commercial ground station in the world serving more than 22000 satellite passes per month and with 30+ full motion antenna systems installed. All frequencies (VHF, UHF, C-, L-, X-, S-, Ku- and Ka-band) relevant to operating Low Earth Orbiting Satellites are served from here.

As a logical next step KSAT decided to build a station in Antarctica at the Troll station. Troll is located at 72 degrees south, which means that 12 out of 14 passes can be seen. Today Troll has three full motion S/X-band 7,3 meter antennas and smaller 3,7 meter antenna. Data transfer has been dramatically improved recently as the data is sent via the geostationary Thor-7 telecom satellite launched in April 2015. It has a dedicated, KSAT owned transponder on-board that points a spot beam towards Troll. Through this upgrade KSAT is now capable of forwarding 1GB of data at a price that can compete with most last mile prices via fiber. For optical satellites Troll is a more logical download station than a Northern station as these satellites fill their memory with images on the descending orbit allowing them to dump the entire memory over Troll and to be ready for new tasks. With a Northern polar station only the latency will always be greater than 45 minutes as the satellite needs to cover the ascending orbit before reaching the Arctic area for download.

The availability of Troll allows satellite owners to half the redundancy to maximum 45 minutes which is satisfactory for many applications, but not for all. The Troll station sees 12 of 14 passes and is the key commercial Antarctic ground station with full operations round the year.

2. REDUCING LATENCY

The world is not fair. 75 percent of the landmass is located in the Northern Hemisphere, and so is 85 percent of the money and 95 percent of the conflicts. To reduce the latency further it is necessary to build stations on the Southern fringe of the Northern Hemisphere. KSAT has with its customers built a network of stations that are located in Hartebeesthoek (South Africa), Mauritius, Dubai and in Singapore. This KSAT owned network is complemented by further stations in California, Florida, Punta Arenas (Chile) and in West Australia owned by Universal Space Network. This global network of mid-latitude stations allows reducing the latency further to a maximum of 25 minutes for all orbits that cover the Northern Hemisphere. The exceptions are orbits where no mid-latitude stations exist and data is dumped at Troll.

Figure 1: Distribution between Northern and Southern Hemisphere with stations at Southern fringe

The KSAT global ground station network consist of a number of ground stations which forms a homogeneous and tailored network suitable for all LEO satellite owners worldwide.

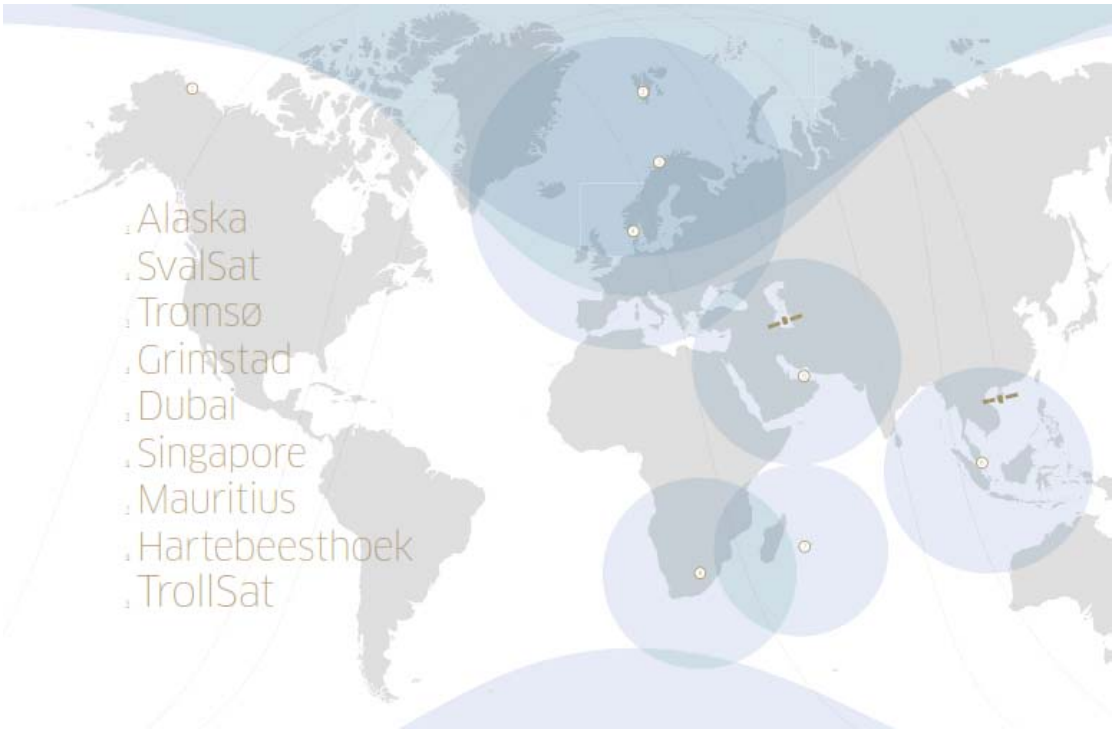


Figure 2: KSAT global ground station network

The logical next step to further reduce latency would be to install a global network of direct reception stations. This is, however, very costly and often such ground station locations would be expensive, unsafe or unavailable. Also for almost all applications latency between 20-30 minutes is satisfactory. Another possibility is to use geostationary satellites like TDRS, DRTS or EDRS. However, these require an additional costly payload on the satellite and they form a single point of failure. They also lack a guarantee for long term continuity. A relay satellite can be more expensive than the satellite(s) it serves. Building and operating the entire KSAT network costs only a fraction

of building and operating relay satellites and the onboard payload needed onboard the LEO satellites. Relay satellites can only become competitive with Ground station networks when the relay satellite and the LEO onboard payload are subsidized (not included in the service prize calculation). Additionally, relay satellites as EDRS can only serve one satellite at a time, whereas e.g. Svalsat can serve up to 30+ customers in parallel (one per antenna).

Another way to reduce latency is to build satellite constellations with multiple satellites that have an offset between orbits. Examples are Rapideye and the Cosmo Skymed missions that achieve 4-5 times lower latency than single missions. However, this is a very expensive way to reduce latency compared to using an existing commercial ground network as presented above. Of, course the combination of a constellation and a complete station network is ideal.

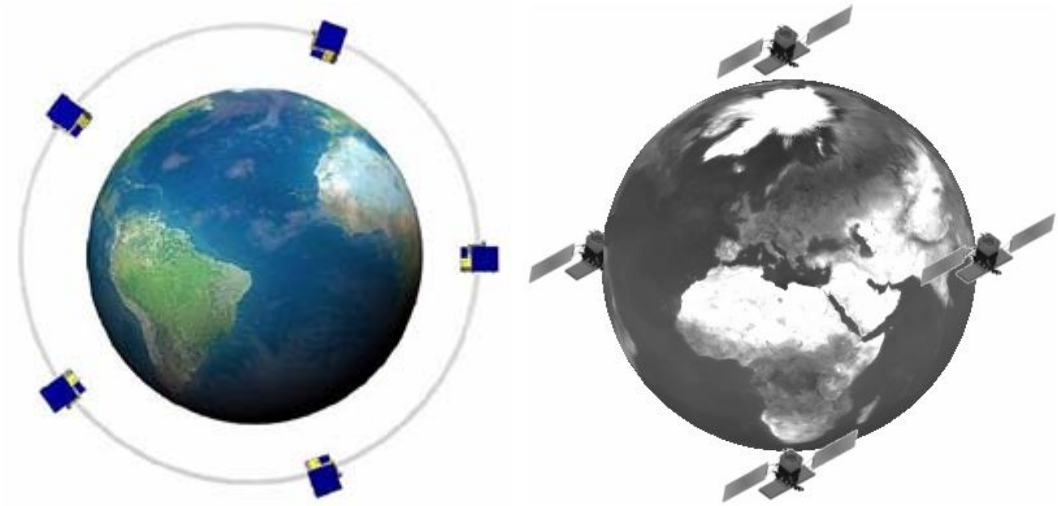


Figure 3: Concept pictures for Rapideye constellation (Image credit Rapideye) and Cosmo Skymed constellation (Image credit Alcatel Alenia Space)

The latency presented in this paper is the maximum for the time delay between image acquisitions of the satellite to reception by an existing ground station. This means that if an image is acquired over Myanmar and downloaded over Singapore the delay will be only a few minutes. However, for optical satellites the latency can be substantially higher as data is only acquired on descending orbits and due to the problem of cloud coverage. Optical images can only be taken during daytime with the right illumination. All these challenges are not applicable for radar satellites where data can be acquired at any time of day, through clouds and both on ascending and descending orbits. Radar satellites usually use dawn/dusk orbits, which are a special sun-synchronous orbit, where the local mean solar time of passage for equatorial longitudes is around sunrise or sunset, so that the satellite rides the terminator between day and night. Riding the terminator is useful for active radar satellites as the satellites' solar panels can always see the Sun, without being shadowed by the Earth.

What is not included in these latencies is the time delay between satellite acquisition order, the time it takes to uplink this order to the satellite and the time it takes until the satellite passes over the acquisition site. However, these latencies can be lowered by using agile satellites that can acquire data off nadir by turning into the direction of the acquisition target.

3. CONCLUSION

Through its unique ground station network KSAT has managed to help to exploit satellites in a more efficient manner for a cost that is several magnitudes lower than the cost of satellites. The service is scalable from a few passes to all satellite passes of a constellation. The Svalbard station, the Antarctic Troll station and finally the mid-latitude network have within the last 15 years managed to cut latency from several hours (during the gap when the satellite is not visible to "normal" polar stations) to only 20-30 minutes if Svalbard is combined with Troll and the mid-latitude network.

Research of Range Migration and Correction Algorithms in SAR Range-Doppler Imaging

Min Chen[1], Weixing Sheng[1], Zhen Hu[2], Xiaofeng Ma[1], Shihong Zhou[2]

(1 School of Electronic and Optical Engineering, Nanjing University of Science and Technology
2 Shanghai Institute of Satellite Engineering)

Abstract: The most widely used technology in radar imaging is synthetic aperture radar (SAR). It can achieve high range resolution by emitting large time bandwidth product signal and get high azimuth resolution using the relative motion between the radar antenna and the target. However, the point target in different azimuth can be compressed to different range positions, which results in the range migration and the coupling between range and azimuth, and decrease imaging quality. Based on the range curvature in SAR imaging, a linear phase multiply correction algorithm is proposed in this paper, and compared with the common interpolation correction algorithm. Finally, SAR imaging simulation of point targets has been done and the comparisons of the two correction algorithms have been provided.

Keywords: Range-Doppler algorithm, range curvature correction, a correction algorithm based on linear phase multiply, interpolation algorithm

1 Introduction

The most widely used in radar imaging technology is Synthetic Aperture Radar (SAR), it achieves high range resolution by emitting large time bandwidth product signal [1]. The relative motion between the radar antenna and the target, resulting in Doppler effect, on the one hand, makes the SAR get high azimuth resolution, on the other hand, can also cause the same point in different azimuths be compressed to the different range positions, that is the distance migration. In SAR imaging, range migration includes range bending caused by the relative movement between radar and the target and range walking generated by radar squint [2]. In this paper, only side-looking SAR is discussed, so there is no distance walking under any circumstances, but there always exists range bending. Range migration may result in the coupling between range and azimuth, and increase the difficulty of imaging. Therefore, imaging algorithm must include range migration correction [3].

Range-Doppler [2] is a classical algorithm of the imaging algorithms of side-looking stripe SAR. It meets the efficient modular processing requirements and has the simplicity of one-dimension operation [1] by operating in the range and azimuth frequency domain. Range-Doppler algorithm usually adopts interpolation methods to correct the range bending and then eliminates the coupling between range and azimuth. This algorithm decomposes the 2-D processing into two cascade of one-dimensional and reduces the difficulty of imaging. Its main disadvantage is that bring additional amount of calculation, and will appear new frequency modulation components in the range direction. So that the target response will expand again in the range direction, and the imaging quality of the range direction will fall [4]. In order to reduce the amount of calculation and the phase distortion, this paper also presents a simple linear phase multiply algorithm, without interpolation, only needs twice Fourier transform, and once complex multiply operation [3].

This paper begins with the principle of RD algorithm, analyzing the cause of range migration and deducing the range bending curve in frequency domain. On the basis of this, then elaborates the traditional interpolation method with the interpolation function of the nearest neighborhood approximation and the proposed rapid correction method. The selection of interpolation function is the nearest neighborhood approximation method. Finally, two methods are simulated and compared in point target model and the related conclusions are given.

2 Problem Formulation

2.1 Range bending equation in R-D domain

According to the Range-Doppler imaging principle [2, 5]. The base-band form of the point target's echo received by radar after orthogonal demodulation is

$$\begin{aligned} S_R(t,t_m) &= w_a(t_m - t_c) w_r(t - \frac{2R(t_m,r)}{c}) \\ &\times \exp\left[j\pi k_r (t - \frac{2R(t_m,r)}{c})^2 \right] \exp\left[-\frac{4\pi}{\lambda} R(t_m,r) \right] \end{aligned} \tag{1}$$

where t, t_m are fast time and slow time respectively, t_c is the moment of radar beam center irradiation to the target, $w_r(t - \frac{2R(t_m,r)}{c})$ is the echo amplitude of range direction, $w_a(t_m - t_c)$ is the echo amplitude of azimuth direction, k_r is the linear frequency modulation rate, λ is wavelength, c is speed of light, $R(t_m,r)$ is instantaneous slant distance between radar and the target in different slow time.

The distance between radar and the fixed target in the ground $R(t_m,r)$ which satisfies Fresnel approximation is

$$\begin{aligned} R(t_m,r) &= \sqrt{r^2 + (vt_m - x_T)^2} \approx r + \frac{(vt_m - x_T)^2}{2r} \\ t_m &\in \left[-\frac{T_{sar}}{2}, \frac{T_{sar}}{2} \right] \end{aligned} \tag{2}$$

where v is the platform's speed, r is the vertical distance between the target and flight direction of radar, x_T is the azimuth coordinate of the point target(usually x_T =0), T_{sar} is the time of synthetic aperture.

After range compression, scattered points in the same range cell have the same range migration track, but there is a translation in the azimuth time. While scattered points in different range cells have different range bending curvature, so it is difficult to carry out range bending correction in time domain [2]. In practical engineering, we usually correct the range bending in frequency domain. The Doppler frequency of the target is determined by the list equation

$$f_d = \frac{2v}{\lambda} = \frac{2}{\lambda}\frac{dR(t_m, r)}{dt_m} \tag{3}$$

According to (2) and (3), we can get the range bending equation in frequency domain

$$R(f_d, r) = r + \frac{\lambda^2 r}{8v^2} f_d{}^2 \tag{4}$$

The range migration in frequency domain is

$$\Delta R(f_d, r) = R(f_d, r) - r = \frac{\lambda^2 r}{8v^2} f_d^2 \tag{5}$$

The equation after discretization is

$$\begin{aligned} \Delta R(n,m) &= \frac{1}{8}\left(\frac{\lambda}{v}\right)^2 \left(r_B + \frac{mc}{2f_{sr}}\right)\left(\frac{n}{N} PRF\right)^2 \\ -\frac{N}{2} \le n \le \frac{N}{2}&, -\frac{M}{2} \le m \le \frac{M}{2} \end{aligned} \tag{6}$$

where m, n are the sampling ordinal number of range and azimuth direction respectively, M, N are the sampling numbers of range and azimuth direction, f_{sr} is sampling frequency of range direction, PRF is pulse repetition frequency, r_B is the distance between radar and the center of swath.

2.2 Traditional interpolation method

The idea of interpolation is: recover the data of true range bending curvature by the sampling data in range direction [6]. Assuming the signal before correction is $S(n,m)$, signal after correction is $S^{'}(n,m)$, thus

$$S^{'}(n,m) = S(n,m) + \frac{2\Delta R(n,m) f_{sr}}{c} \tag{7}$$

This is the range migration correction equation. Because $\frac{2\Delta R(n,m) f_{sr}}{c}$ is not always an integer and cannot achieve directly from the data after range compression, so we can estimate the $S^{'}(n,m)$ from original data using interpolation.

This paper adopts the truncated sinc interpolation. Assuming the number of the truncated sum terms are N, then the corresponding interpolation equation is

$$S'(n,m) = \sum_{i=N/2}^{(N/2)-1} S(n; m+m'+i) \times \frac{\sin(\pi(\Delta m + i))}{\pi(\Delta m + i)} \tag{8}$$

where $m^{'}$ is the integer part of $\frac{2\Delta R(n,m) f_{sr}}{c}$, Δm is the fractional part, $\Delta m = \frac{2\Delta R(n,m) f_{sr}}{c} - m^{'}$.

Interpolation will bring additional computation and appear new frequency modulation components in range direction, which leads to the target response expand again in range direction, and the fall of imaging quality of range direction [4].

3 Linear phase multiply algorithm

Based on the discrete time linear systems theory [7], the shift of time in range direction corresponds to the linear phase of frequency domain. In order to implement range migration correction in frequency domain, compensating the migration in frequency domain can be considered. The derivation of migration in Range-Doppler domain is as follows.

3.1 Fourier transformation in range direction

Fourier transform formula is

$$S_0(f_r,t_m)=\int_{-\infty}^{\infty}S_R(t,t_m)\exp(-j2\pi f_r t)dt \tag{9}$$

The phase in the integral equation is

$$\theta(t)=-\frac{4\pi f_0 R(t_m,r)}{c}+\pi k_r\left[t-\frac{2R(t_m,r)}{c}\right]^2-2\pi f_r t \tag{10}$$

In order to apply POSP (Principle of Stationary Phase) [1], we need to find the range time where derivative is zero.

$$t=\frac{f_r}{k_r}+\frac{2R(t_m,r)}{c} \tag{11}$$

The integral in (9) can be written as

$$\begin{aligned}S_0(f_r,t_m)&=w_a(t_m-t_c)w_r(f_r)\\&\times\exp(-j\frac{4\pi(f_0+f_r)R(t_m,r)}{c})\exp(-j\frac{\pi {f_r}^2}{k_r})\end{aligned} \tag{12}$$

where $w_r(f_r)$ is range spectrum envelope.

3.2 Fourier transformation in azimuth direction

Fourier transform formula is

$$S_{2df}(f_r,f_d)=\int_{-\infty}^{\infty}S_0(f_r,t_m)\exp(-j2\pi f_d t_m)dt_m \tag{13}$$

According to the phase in (12), we can achieve the phase of (13) as

$$\theta(t_m)=-\frac{4\pi(f_0+f_r)R(t_m,r)}{c}-\frac{\pi {f_d}^2}{k_r}-2\pi f_d t_m \tag{14}$$

By substituting the slope distance into (14), we will get the derivation for t_m, that is

$$\frac{d\theta(t_m)}{dt_m}=-\frac{4\pi(f_0+f_r)v^2 t_m}{c\sqrt{r^2+v^2t_m^2}}-2\pi f_d \tag{15}$$

The derivation is zero when

$$f_d = -\frac{2v^2(f_0 + f_r)t_m}{c\sqrt{r^2 + v^2 t_m^2}} \tag{16}$$

Finally, the result in (13) can be expressed as

$$S_{2df}(f_r, f_d) = w_r(f_r)w_a(f_d - f_{dc})\exp(j\theta_a(f_r, f_d)) \tag{17}$$

According to equation (14) and (16), we can obtain the phase after Fourier transformation in azimuth direction

$$\theta_a(f_r, f_d) = -\frac{4\pi r(f_0 + f_r)}{c}\sqrt{1 - \frac{c^2 f_d^2}{4v^2(f_0 + f_r)^2}} - \frac{\pi f_d^2}{k_r} \tag{18}$$

The radical sign in equation (18) is range migration factor. Thus, the correction function is

$$C_c = \exp(j\frac{4\pi f_r r)}{c}\sqrt{1 - \frac{c^2 f_d^2}{4v^2(f_0 + f_r)^2}}) \tag{19}$$

3.3 Range migration correction

Suppose RCM at least does not change with the distance in a limited area. At this point, range migration correction can be achieved by FFT, linear phase multiplication and IFFT.

According to the phase multiplier which has a given f_d and in the condition of broadside, the correction function can be approximated as [3]

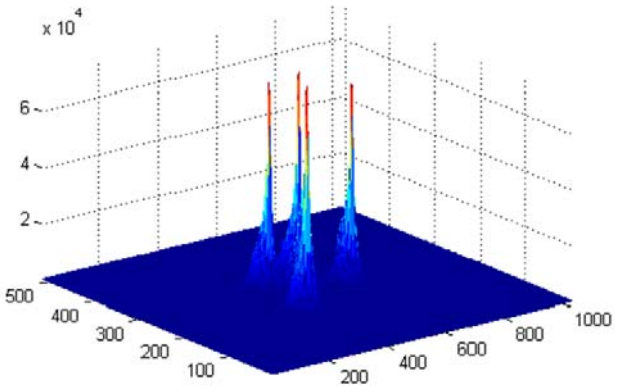
$$C_c \approx \exp(j\frac{4\pi f_r \Delta R(f_d, r)}{c}) \tag{20}$$

Using this range migration correction method, first of all must carry on the data block, each of the correction quantity should be set to a fixed value, and its drawback is that data block must overlap in the range direction, in terms of processing efficiency, this increase in complexity may do more harm than good.

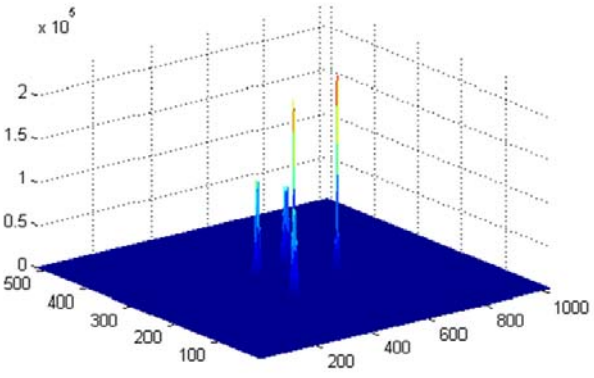
4 Simulation experiment study

Simulation parameters: three point targets, broadside, the distance between radar and the center of swath is R_0 =1200m, mapping range of range direction is [1200-150,1200+150]m, wavelength is λ =50cm, the bandwidth of chirp signal is B_r =150MHz, pulse width is T_r =1.33us, the sampling number of range direction N_x =512; the Doppler bandwidth is $B_d = 150$ MHz, length of synthetic aperture is L_{sar} =300m, sampling number of azimuth direction N_y =1024, mapping range of azimuth direction is [-200,200]m.

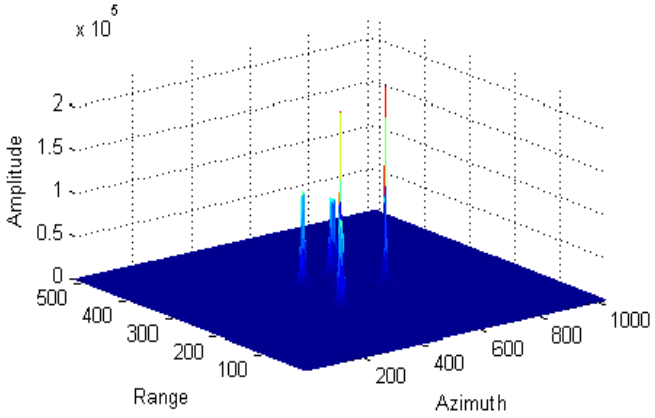
For qualitative quantitative analysis of two kinds of method, in this paper, three point targets are simulated. Figure 1 is the imaging result before range curvature correction as well as the results using truncated sinc interpolation and linear phase multiply method respectively. Figure 2 is peak-sidelobe ratio of each method. Table 1 is comparison results of the mainlobe width (ML) and the integral sidelobe ratio (ISLR).



(a)

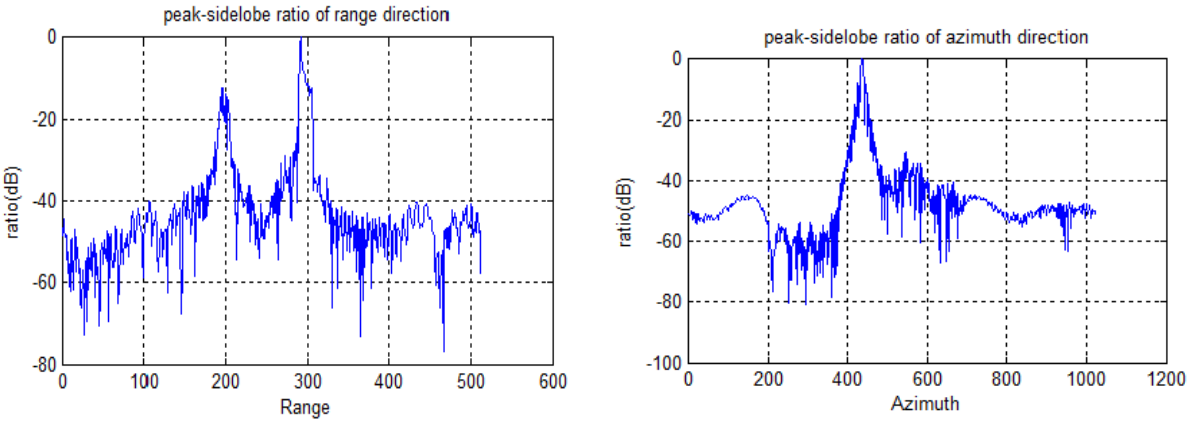


(b)

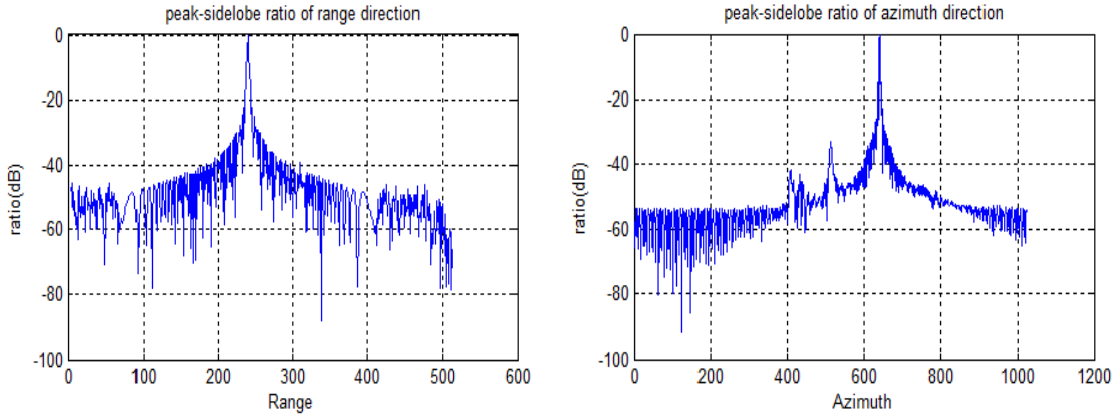


(c)

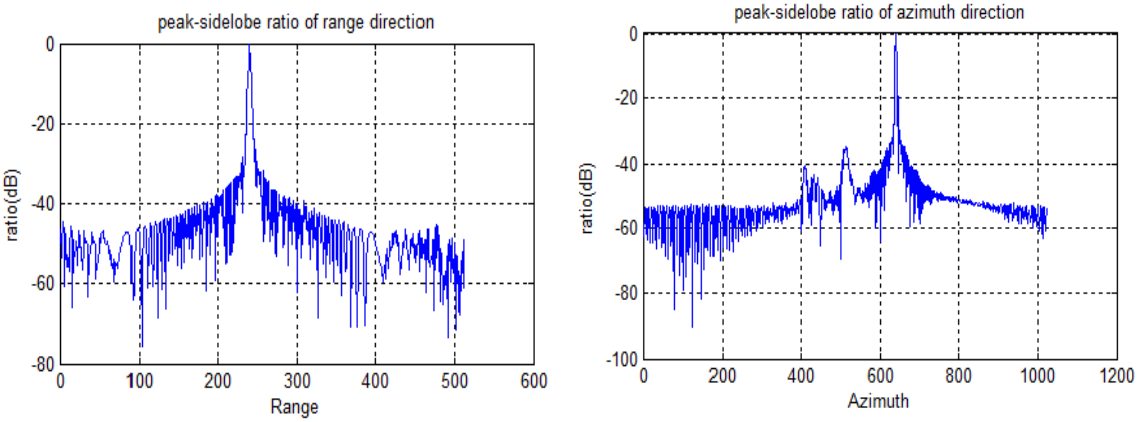
Figure 1 Imaging results comparison of uncorrected and two kinds of imaging algorithms
(a) uncorrected imaging result (b) truncated sinc interpolation (c) linear phase multiply



(a)



(b)



(c)

(a) uncorrected imaging result (b) truncated sinc interpolation (c) linear phase multiply

Figure 2 The peak-sidelobe ratio

Table1 Performance comparisons of different algorithms

Correction method	Azimuth direction		Range direction	
	ML/m	ISLR/dB	ML/m	ISLR/dB
uncorrected	2.0523	-11.8665	6.2294	-17.3620
sinc interpolation	1.4896	-16.1190	2.5507	-14.4002
linear phase multiply	1.6691	-16.6072	2.7575	-18.0896

From Fig.1, it can be seen that after correct the range migration, amplitudes become stronger and the sidelobe decrease. From three pulses , linear phase multiply method can get better image quality.

Fig.2 shows the peak-sidelobe ratio of each method and uncorrected situation. After correcting the range migration, the side lobe decrease and the ratio obviously improved.

From Table 1, we also can see that linear phase multiplication has better imaging result. After the range curvature correction, the peak-sidelobe ratio of each direction has improvement compared with uncorrected result. At the same time, the mainlobe width in range and azimuth direction are both narrow down compared with uncorrected image. Thus, the requested of high resolution can be realized.

5 Conclusion

In the condition of broadside, this paper completes three tasks. First, deducing the range curvature equation in R-D domain. Second, doing research on the correction algorithms. Finally, comparing and analyzing the imaging results of point targets in terms of mainlobe width、 peak-sidelobe ratio and integral sidelobe ratio. The simulation results show that the imaging results using linear phase multiply method are better compared with the nearest neighborhood approximation method, and can clearly see that there are four peak values similar to impulse. Interpolation kernel function changes with distance, so the computation is large and processing efficiency is not high, as well as can cause phase precision loss and decrease the image quality.

Acknowledgements

This work is supported by the National Natural Science Foundation of China under grant 61471196.

References

[1] Ian G. Cumming, Frank H. Wong: ‘Digital Processing of Synthetic Aperture Radar Data Algorithms and Implementation’ (Artech House, Norwood, MA, 2005)

[2] Tao Zeng, Wenfu Yang, Zegang Ding, Luosi Liu: ‘ Advanced range migration algorithm for ultra-highresolution spaceborne synthetic aperture radar’, IET Radar, Sonar and Navigtion., 2013,7,(7), pp.764–772

[3] Tao Zeng, Wenfu Yang, Zegang Ding, Luosi Liu: Huang Yang, Li Chunsheng, Li Jingwen, Zhou Yingqi: ‘An Airborne Squint SAR Processing Algorithm’, IEE Radar, Sonar and Navigtion., 2011,11,(5), pp.235–239

[4] R. Bamler, ‘A comparison of Rang-Doppler and wave-number domain SAR focusing algorithm’, IEEE Trans. Geosci. Remote Sensing, 1992, (30), pp.134-139
[5] Wang Linyu, Zhang Yanxia: ‘An improved Algorithm of Range-Doppler for Air-borne Synthetic Aperture Radar’, International Conference on Transportation, Mechanical, and Electrical Engineering, 2011, pp.1713-1716
[6] Eineder, M., Adam, N., Bamler, R., Yague-Martinze, N., Breit, H.: ‘Sapceborne spotlight SAR interferometry with TerraSAR-X’, IEEE Trans. Geosci. Remote Sens., 2009,47,(5),pp.1524-1535
[7] Oppenheim AV, Schafer R W. Digital Signal Processing [M]. Englewood Cliffs：Prentice-Hall，1975.

Experimental research on CPT atomic clock disciplined by GPS

Zhang Zhenwei, Yang Renfu, Zhang Xu, Wang Nuanrang, Zhao Huan, Chen Xing

(1. Science and Technology on Metrology and Calibration Laboratory, Beijing 100854;
2. Beijing Institute of Radio Metrology & Measurement, Beijing 100854)

Abstract: To realize high precision clock，CPT atomic clock disciplined by GPS is developed. We develop CPT atomic clock with GPS disciplining function that can improve frequency accuracy from 1.5E-8 to 1E-12. The experimental scheme has compact configuration and high measurement resolution. Kalman filter is applied to eliminate GPS 1PPS noise, so that the signal-to-noise ratio improves 12dB. In state of GPS disciplining, long-term frequency stability is improved, reaching 10^{-13} level in τ>10000s. 1PPS accuracy stays in the offset ±30ns all the time. Frequency stability is consistent with which at free running in τ<100s. However in 100s<τ<1000s, due to the infection of the GPS 1PPS noise, frequency stability is deteriorated about 2E-11.
Keywords: CPT atomic clock, GPS, 1PPS

1. Introduction

CPT atomic clock is based on coherent population trapping (CPT) principle, with advantages of small size, light weight, low power consumption[1-3]. As secondary frequency standard, frequency accuracy of CPT atomic clock has long-term frequency drift. To improve the frequency accuracy and long-term frequency stability, high-level frequency standard is wanted to calibrate CPT atomic clock. GPS atomic frequency standard has excellent frequency accuracy, long-term frequency stability and global-service capability. So it is very suitably used to calibrate CPT atomic clock. CPT atomic clocks disciplined by GPS become an inexpensive high-performance atomic clock. Combined with inertial navigation equipment to be Micro-PNT, applications domain of CPT atomic clock is abroad.

2. Experiment

CPT atomic clock utilizes Rb87 as the operation atomic. In constant static magnetic field, the ground state level of rubidium atoms appears Zeeman split, $5S_{1/2}F=2,m_F=0$ and $5S_{1/2}F=1,m_F=0$. The frequency difference of the two hyperfine states is quadratically caused by the magnetic field intensity[4]. Adjusting the coils current, the magnetic field intensity can be varied, so that CPT spectrum shift occurs.

According to the relationship of coils current and CPT spectrum shift[9], Coils current can cause CPT spectrum shift about 50Hz, so the corresponding output frequency of CPT atomic clock varies 0.15Hz. Therefore CPT atomic clock frequency accuracy must be better than 1.5E-8.

The experimental scheme of CPT atomic clock disciplined by GPS is shown in Fig.1. The receiver receives GPS signals, and exports 1PPS signal as a reference time in time comparison. TDC measures the time distance between CPT atomic clock 1PPS and reference 1PPS. FPGA is applied to filter and feedback. The disciplining process consists of two stages: in the calibration stage, CPT

atomic clock frequency is calibrated fast according to the time difference; in disciplining stage, CPT atomic clock frequency is adjusted imperceptibly in real time.

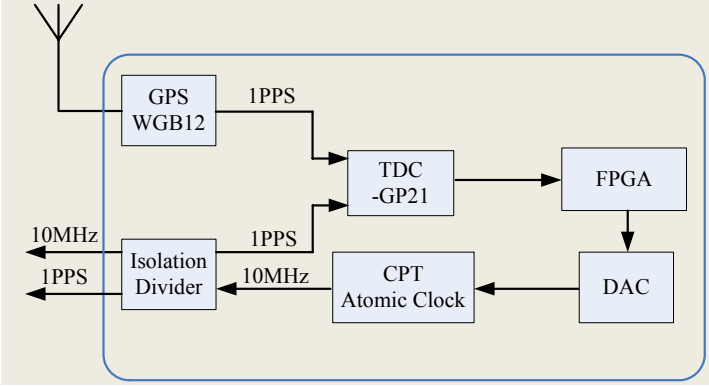


Fig.1 the experiment scheme of GPS disciplining CPT atomic clock

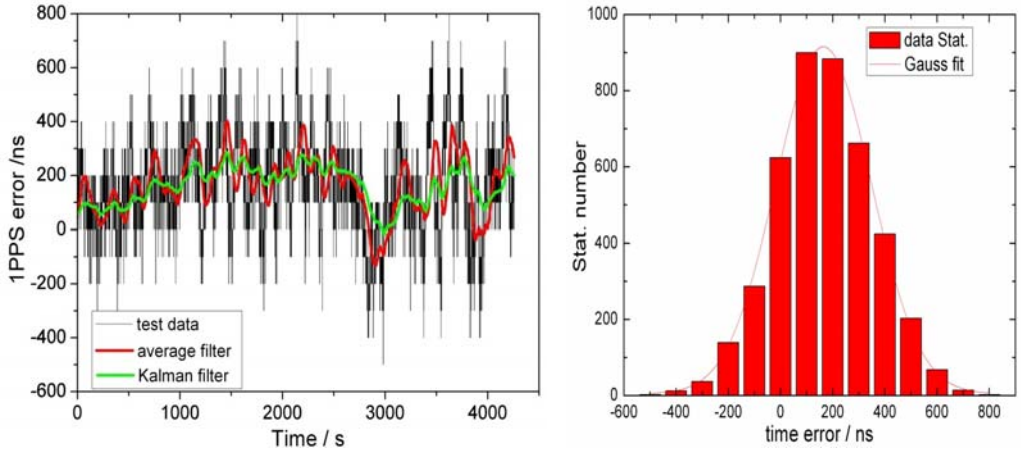
Since the ionosphere and troposphere delay error, multipath and receiver noise interfere, the GPS 1PPS also includes error and noise, and according with Gaussian distribution[5]. If the GPS 1PPS is directly used to calibrate CPT atomic clock, the frequency stability of CPT atomic clock will be deteriorated. Therefore, the GPS 1PPS must be filtered before calibrating CPT atomic clocks.

Kalman filter algorithm is derived by the minimum mean square error principle, used to solve optimal linear filter and estimation problems. ,According to the current state observational data and previous state evaluation kalman algorithm calculates current state evaluation. The Kalman filter algorithm parameters are defined in the papers[6,7]. Because of not storing observation data in the calculation process, kalman algorithm expands very little storage space, so that is widely used in the field of GPS disciplining crystals and atomic clocks[8-9].

3. Results and discussion

In the experiment, H-maser frequency standard is used as frequency reference to test frequency stability.

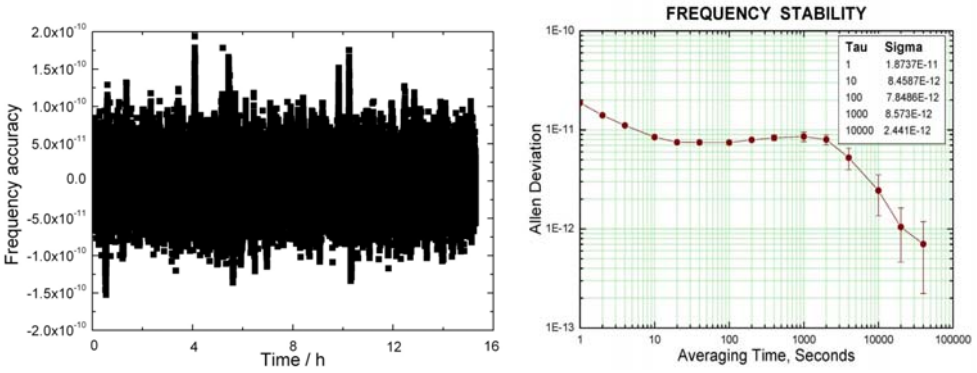
Kalman algorithm is used to filter GPS 1PPS noise. The results is shown in Fig.2-a. Black line is test data, green line is the filter data. Before filter, the noise is relatively obvious and similar with Gauss distribution shown in Fig.2-b. After filter, the time jitter is obviously lower, the signal-to-noise ratio improves 12dB.



（a）before VS after Kalman filter applied （b）statistical chart of time difference data

Fig.2 the experiment result of kalman filter

CPT atomic clock frequency drift is almost neglectable in state of disciplining, shown in Fig.3-a. 1 hour average frequency accuracy is better than 1E-12. Frequency stability result is shown in Fig.3-b. Setting reasonable PI algorithm feedback parameters, we can ensure frequency stability is not affected in $\tau<100$s, consistent with which at free running. However in $100s<\tau<1000$s, due to the infection of the GPS 1PPS noise, frequency stability is deteriorated about 2E-11. After $\tau>1000$s, because of excellent GPS timing precision, frequency stability begins to improve obviously.



（a）result of frequency accuracy （b）result of requency stability

Fig.3 test result of CPT atomic clock on state of GPS disciplining

In the autonomous navigation system application, 1PPS timing precision is a key parameter. We test the comparison between CPT atomic clock 1PPS and UTC, shown in Fig.4. In the GPS tame state, 1PPS accuracy of CPT atomic clocks keep gradually improving, and staying in the offset ±30ns. According to navigation principle, position accuracy is equal to the product of the light speed and timing precision. So timing precision is higher and position is more accurate. Without GPS disciplining function, timing precision of CPT atomic clock gets worse, and time accumulate error bigger and bigger, but still keeps 50ns in 3 hours.

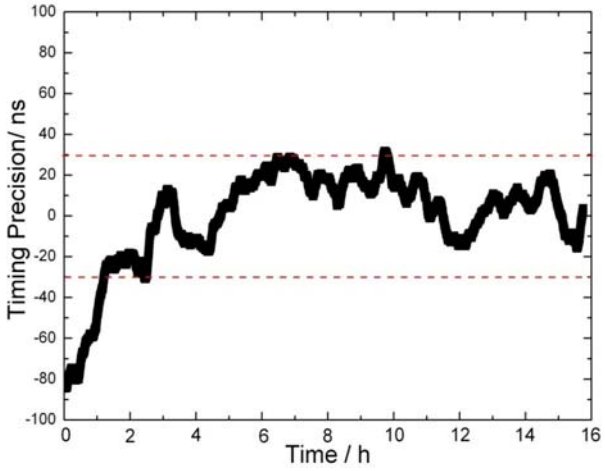


Fig.4 time precision on GPS disciplining state

4. Conclusions

We developed a CPT atomic clock with GPS disciplining function that can improve frequency accuracy from 1.5E-8 to 1E-12. The experimental scheme has compact configuration and high measure resolution. Kalman filter is applied to eliminate GPS 1PPS noise, so that the signal-to-noise ratio(SNR) improves 12dB. In state of GPS disciplining, long-term frequency stability is improved, reach 10^{-13} level in $\tau>10000$s. 1PPS time accuracy of CPT atomic clocks improves, and stays in the

offset ±30ns all the time. Because of these advantages, CPT atomic clock disciplined by GPS can be applied in high-performance time & frequency source, high-precise time synchronization, or other fields.

Reference

[1]. N. Cyr, M. Breton. All-Optical Microwave Frequency Standard: A Proposal[J]. IEEE Transactions on Instrumentation and Measurement, IM-42,1993.42(2):640~649

[2]. R. Wynands, The atomic wrist-watch, Nature, 2004, Vol429, p509-510.

[3]. V. Shah, S. Knappe, J. Kitching, Subpicotesla atomic magnetometry with a microfabricated vapour cell, Nature, 2007, Vol 1. p649-652.

[4]. J.Vanier. Atomic clocks based on coherent population trapping: a review[J]. Applied Physics B. 2005.81(4):422~423

[5]. Cui B J, Wang Y Z. Study on data processing method of GPS disciplined rubidium frequency standard[J]. Journal of Electronic Measurement and Instrument.2010.24(9):808~812 (in Chinese)

[6]. R.E.Kalman. A New Approach to Linear Filtering and Prediction[D].1960

[7]. GUO H R, YANG Y X. Determination of Covariance Matrix of Kalman Filter Used for Time Prediction of Atomic Clocks of Navigation Satellites[J]. Acta Geodaeticaet Cartographica Sinica.2010.39(2):146-150 (in Chinese)

[8]. C. Green. Advantageous GPS disciplining of the OSA BVA Oscillator [J].IEEE. 2012: 263~266

[9]. Yang X H, Zhai H S. Study on GPS Disciplined Rb Clock Based on New Frequency Accuracy Measurement Algorithm [J].Chinese Journal of Scientific Instrument. 2005.26(1): 41-44 (in Chinese)

Design of variable fractional delay filter using Discrete-cosine-transformation

Li Zeqing[1] ,Li Bo[2],Liu Yanli[1],Yuan Jie[1]

(1: Shanghai Spaceflight Institute of Electronic Technology, ShangHai, China, 201109
lize_qing@126.com
2: North University of China, Taiyuan, ShanXi, China, 030051, libo@nuc.edu.cn)

Abstract: In this paper, design of variable fractional delay (VFD) FIR filter is investigated. First, the interpolation formula of discrete-time sequence is derived by using discrete cosine transformation (DCT). Second, the DCT Interpolation method is utilized, with suitable index mapping, to design a VFD FIR filter. Then, examples are demonstrated to show the effectiveness of the designed filter with the help of magnitude, group delay and NMRS. Results of the designed VFD FIR Filter have an acceptable limitation over a wide range of frequency. Finally, the realization of VFD FIR filter is proposed.

Keywords: variable fractional delays (VFD), discrete cosine transform (DCT)， Finite Impulse Response(FIR)

I INTRODUCTION

Application and design of digital fractional delay filter have been widely studied in the open literature by Lakso [1].FIR filter with fraction of unit delay is required in many areas of Digital Signal Processing, including speech coding [2], speech assisted video processing [3], time delay estimation [4], and communication. Fig.1 shows a signal processing model for a vibrating string. The pitch of the sound is controlled by the length of the delay line in the feedback loop. It is implemented as a combination of integer length delay line and a fractional delay filter.

Fractional delay can be classified into two categories, fixed fractional delay and variable fractional delay. Variable fractional delay filter will provide more flexibility and be used in more interesting areas compared to fixed fractional delay. The desired frequency response of variable fractional delay filter is given by

$$H_d(\omega, p) = e^{-j\omega p} \tag{1}$$

where p represents the value of variable fractional delay, in the range of -0.5 to 0.5 for all $\omega \in \Omega = [0, \omega_p]$.

The transfer function of FIR filter to approximate the frequency response is given by Farrow [5] as

$$H(z, p) = \sum_{r=0}^{N-1} a_r(p) z^{-r} \tag{2}$$

where ar(p) represents the coefficients of filter which depends upon two variables r & p . N is the order of the filter.

Several approaches are available to approximate the desired frequency response [6]-[11]. An approach is available using DCT interpolation method, which is further utilized in this paper for variable fractional delay[11]. Given the discrete-time sequence x(0), x(1), ..., x(N-1), the type-II discrete cosine transform pairs are defined by

$$X(k) = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} c_k \bullet x(n) \bullet \cos\left(\frac{2n+1}{2N} \bullet k\pi \right)$$

$$x(n) = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} c_k \bullet X(k) \bullet \cos\left(\frac{2n+1}{2N} \bullet k\pi \right)$$

where $c_k = \begin{cases} \frac{1}{\sqrt{2}} & k = 0 \\ 1 & others \end{cases}$

In this paper, the discrete cosine transform method based on zero padding in the frequency domain is introduced, which can provide interpolation in the time domain [12-14]. Providing suitable zero padding in the DCT of the given sequence will result in an interpolation in the time domain [11]. This method is then utilized in designing VFD FIR Filter. Realization structure is also proposed and at last design examples are provided to validate its performance through magnitude, group delay and error responses.

II DESIGN METHOD

This section deals with designing of variable fractional delay filter using DCT interpolation. Fig.2 shows the filter H(z,p) with s(n) as input and y(n) as output. As per DCT Interpolation method [11],

$$x(t) \approx \sum_{m=0}^{N-1} x(m) b(m,t) \tag{3}$$

where

x(m) = N-samples of input signals

x(t) = Corresponding continuous signal

b(m,t) = Continuous time interpolation basis and equal to

$$b(m,t) = \frac{2}{N} \sum_{k=0}^{N-1} c_k^2 \bullet \cos\left(\frac{2m+1}{2N} \bullet k\pi \right) \bullet \cos\left(\frac{2t+1}{2N} \bullet k\pi \right) \tag{4-1}$$

$$b(m,t) = \frac{2}{N} \sum_{k=0}^{N-1} a_k \bullet \cos\left(\frac{2m+1}{2N} \bullet k\pi \right) \bullet \cos\left(\frac{2t+1}{2N} \bullet k\pi \right) \tag{4-2}$$

where

$$a_k = \begin{cases} 1/2, k = 0 \\ 1, otherwise \end{cases}$$

This b(m,t) replaces the sinc function[1] in ideal interpolation and approximates the impulse response in (1). Now from Fig.2, according to convolution theorem

$$\mathrm{Y(z) = S(z) * H(z,p)} \tag{5}$$

or

$$y(n) = \sum_{r=0}^{N-1} a_r(\mathrm{p})\, \mathrm{s(n - r)} \tag{6}$$

If H (z, p) approximates the ideal response Hd(w, p), then

$$\mathrm{y(n) = s(n\text{-}p)} \tag{7}$$

Then from (6) and (7)

$$s(n-p) = \sum_{r=0}^{N-1} a_r(\mathrm{p})\, \mathrm{s(n - r)} \tag{8}$$

Now with x(t) = s(n-(N-1) + t), (3) become as

$$s(n-(N-1)+t) \approx \sum_{m=0}^{N-1} s(n-(N-1)+m)\, \mathrm{b(m,t)} \tag{9}$$

Let m = N-1-r and t = N-1-p to make (9) in more appropriate form

$$s(n-p) \approx \sum_{r=0}^{N-1} \mathrm{s}(n-r)\, \mathrm{b(N\text{-}1\text{-}r, N\text{-}1\text{-}p)} \tag{10}$$

On comparing (8) with (10), ar (p) comes out to be

$$\mathrm{b\ (N\text{-}1\text{-}r,N\text{-}1\text{-}p) = ar(p)} \tag{11}$$

Using (4), (11) becomes

$$\begin{aligned} \mathrm{a_r(p)} &= b(N-1-r, N-1-p) \\ &= \frac{2}{N}\sum_{k=0}^{N-1} a_k \bullet \cos\left(\frac{2(N-1-r)+1}{2N} \bullet k\pi\right) \bullet \cos\left(\frac{2(N-1-p)+1}{2N} \bullet k\pi\right) \\ &= \frac{2}{N}\sum_{k=0}^{N-1} a_k \bullet \cos\left(\frac{2N-2r-1}{2N} \bullet k\pi\right) \bullet \cos\left(\frac{2N-2p-1}{2N} \bullet k\pi\right) \end{aligned} \tag{12}$$

$$\mathrm{a_r(p)} = \frac{2}{N}\sum_{k=0}^{N-1} a_k \bullet \cos\left(\frac{(2N-2r-1)}{2} \bullet J_k\right) \bullet \cos\left(\frac{(2N-2p-1)}{2} \bullet J_k\right) \tag{13}$$

Where $J_k = \dfrac{k\pi}{N}$

$$\mathrm{a_r(p)} = \frac{1}{N}\sum_{k=0}^{N-1} a_k\left[\cos\left((2N-r-p-1)J_k\right) + \cos\left((r-p)J_k\right)\right] \tag{14}$$

$$\mathrm{a_r(p)} = \frac{1}{N}\sum_{k=0}^{N-1} a_k\left[\cos\left(2k\pi - J_k r - J_k p - J_k\right) + \cos\left(J_k r - J_k p\right)\right] \tag{15}$$

$$\mathrm{a_r(p)} = \frac{1}{N}\sum_{k=0}^{N-1} a_k\left[\cos\left(J_k r + J_k p + J_k\right) + \cos\left(J_k r - J_k p\right)\right] \tag{16}$$

$$\mathrm{a_r(p)} = \frac{1}{N}\sum_{k=0}^{N-1} a_k[(\cos\left(J_k(r+1)\right) + \cos\left(J_k r\right))\cos\left(J_k p\right) - (\sin\left(J_k(r+1)\right) - \sin\left(J_k r\right))\sin\left(J_k p\right) \tag{17}$$

Turn cos(J_kp) and sin(J_kp) into the taylor series，comes to

$$a_r(p) = \frac{1}{N}\sum_{k=0}^{N-1} a_k[(\cos(J_k(r+1)+\cos(J_k r))\cos(J_k p) - (\sin(J_k(r+1)) - \sin(J_k r))\sin(J_k p)$$

$$\approx \frac{1}{24N}\sum_{k=0}^{N-1} a_k[b_r\,(24 - 12Q_k^{\,2} + Q_k^{\,4}) - d_r(24Q_k - 4Q_k^{\,3})\,] \tag{18}$$

Where

$$br = \cos(J_k(r+1)+\cos(J_k r)$$

$$dr = \sin(J_k(r+1)) - \sin(J_k r)$$

$$Q_k = J_k p = \frac{k\pi}{N} p,$$

$$a_k = \begin{cases} 1/2, k = 0 \\ 1, others \end{cases}$$

The ar (p) in equation (18) represents the coefficient of FIR Filter shown in Fig.2. As it is clear, these coefficients depend on two variables, r and p which define the fractional delay.

III PROPOSED STRUCTURE

In this section we discuss the structure of the VFD FIR Filter designed in the last section. Z-transform coefficients of FIR Filter can be used to realize the VFD Filter.

$$H(z,p) = \frac{1}{12N}\sum_{k=0}^{N-1} a_k[H_1(z,p)\,(24 - 12Q_k^{\,2} + Q_k^{\,4}) - H_2(z,p)(24Q_k - 4Q_k^{\,3})\,] \tag{19}$$

where H1(z,p) and H2(z,p)are the z-transforms of br and dr , and Qk which is a product of Jk & p, gives the control on delay. it is clear from (19) that VFD Filter requires two sub filters H1(z) & H2(z), adders and multipliers.H1(z) &H2(z) are the Direct Form FIR structure and easy to implement.

IV DESIGN EXAMPLE

In this section, a numerical example performed with MATLAB language is presented. Simulation results for magnitude, group delay and mean square error for different parameters are evaluated. The normalized root mean square error is defined as

$$NRMS = \left(\frac{\int_0^{\alpha\pi} |H_d(\omega,p) - H(\omega,p)|^2 d\omega}{\int_0^{\alpha\pi} |H_d(\omega,p)|^2 d\omega} \right)^{\frac{1}{2}} \times 100\%$$

Two different examples are considered, for positive and negative fractional delay.

Example 1: The parameters chosen are N = 100 , p = 0.05. The magnitude and group delay responses shown in Fig.3. Clearly H(z,p) approximate the Hd(w,p) in a wide range (0,0.95Π).

Example 2: The parameters chosen are N = 100, p =-0.1. The magnitude and group delay responses shown in Fig.4. Again H(z,p) approximate the Hd(w,p) in a wide range (0,0.9Π).

Fig. 5 shows the error curves for α =0.8 and α =0.9 with maximum error 0.045% & 0.33% respectively

V. CONCLUSIONS

In this paper, DCT based interpolation has been presented to design the variable fractional delay FIR filter. Essentially, it is designed for the entire range(-0.5 to 0.5), but according to the approximation method we can get a better performance in the range (-0.1 to 0.1) , which is also distinct in the given examples. The structure is also proposed, but the complexity of it can be reduced via various approaches. In the future, the various approaches will be studied to reduce the complexity of the structure and the 2-D VFD FIR Filter will be realized by using DCT.

REFERENCES

[1] T. I. Laakso, V. Valimaki, M.Karjalainen and U. K. laine,"Splitting the unit delay: tools for fractional delay filter design," IEEE Signal Processing Magzine, pp30-60, Jan 1996

[2] P. Kroom and B. S Atal, Pitch prediction and temporal resolution, in Proc. IEEE Intr. conf. Acoust Speech Signal Processing (ICASSP-90), vol-2, pp 661-664, Albuquerque, New Mexico, April 3-6, 1990.

[3] T. Chen, H. P Graf and K. Wang, Speech-assisted video processing: interpolation and low bit rate coding in Proc. 28th Annual Asilomar Conf. on Signals, System and Computers, pp 975-979, Pacific Grove,California, oct-1994.

[4] J. O. Smith and B. Friedlander, Adaptive interpolated time delay estimation, IEEE Trans. Aerospace and Electronic System, vol. AES-21,no-3,pp 180-199, March 1985

[5] C. W. Farrow," A continuously variable digital delay element," in Proc.IEEE Int. Symp. Circuits System (ISCAS-88), vol-3 pp 2641-2645,Espoo, Finland, June 6-9,1988.

[6] C. C. Tseng," Design of fractional delay FIR Filter using differentiator bank," in Proc. Int. Symp. Circuits and Systems vol-4, 2002, pp 412-424.

[7] J. Vesma and T Saramaki," Design and properties of polynomial-based fractional delay filters," in Proc. Int. Symp. Circuits and Systems,2000,pp421-424.

[8] Chien-Cheng Tseng and Su-Ling Lee ,“Design of fractional delay FIR filter using discrete cosine transform ,” 2008 IEEE Asia Pacific Conference on Circuits and Systems, Nov 2008, Macao, China, pp. 858-861.

[9] W. S. Lu and T.B.Deng, "An improved weighted least square design for variable fractional delay FIR Filters," IEEE Trans. Circuits Syst. II vol-46, pp 1035-1040, Aug-1999.

[10] T. B. Deng," Discretization-free design variable fractional delay FIR Filters," IEEE Trans Circuits Syst.II, vol-48, pp 637-644, June -2001.

[11] A. Tarezynski, G. D. Cain, E. Hermanowicz, M. Rojewski, " WLS design of variable fractional response FIR Filters," in Proc. Int. Symp. Circuits and Systems,1997, pp 2244-2247.

[12] D. Fraser ," Interpolation by FFT revisted an experimental investigation." IEEE Trans. on Acoust. , Speech & Signal Processing, vol-37, pp 665-675, May 1989.

[13] Z. W. Wang, J. J. Soltis and W. C. Miller," Improved Approach to Interpolation using the FFT," Electronics Letters, vol. 28, pp 2320-2322 Dec 1992.

[14] M. D. Macleod," Fast Interpolation by FFT with greatly increased accuracy," Electronics Letters, vol-29, pp1200-1201, June 1993.

[15] D. Upadhyay," Design of Variable Fractional Delay Filter Using DFT Interpolation," computer socienty, pp691 - 693, Oct. 2009.
[16] Hai Huyen Dam," Design of Allpass Variable Fractional Delay Filter," IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 59, NO. 12, DECEMBER 2011

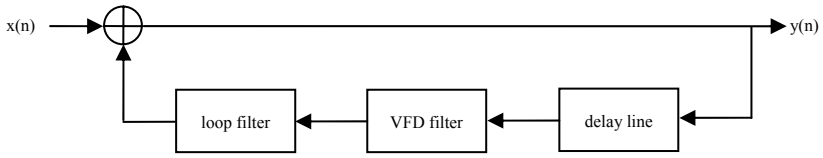


figure 1 A fractional delay filter model for vibrating string

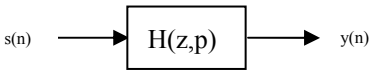


Figure 2 the relationship of variable fractional delay FIR filter between input and output

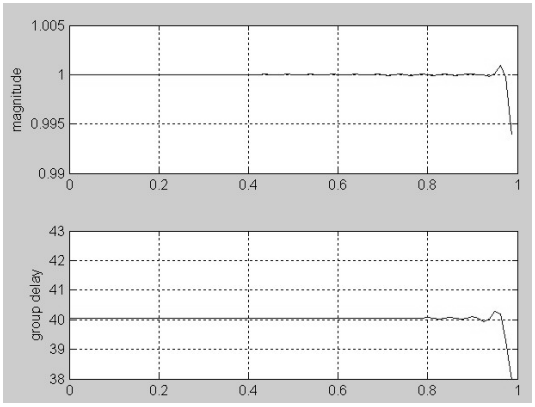


Fig.3 The magnitude and group delay of variable fractional delay FIR filter for p = 0.05

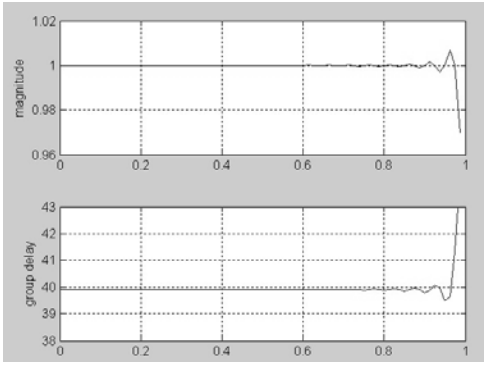


Fig.4 The magnitude and group delay of variable fractional delay FIR filter for p = -0.1

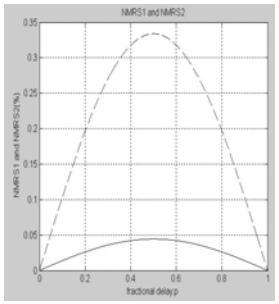


Fig.5 The NRMS error curve of the VFD-DCT method.dash line is α =0.9 when solid line is α =0.8

Research on Precise Remote Time Synchronization Based on Communication Satellite

WANG Haifeng[1,2], ZHANG Shengkang[1,2], YANG Renfu[1,2],WANG Xueyun[1],
WANG Hongbo[1,2], YI Hang[1,2], SHI Fan[1,2], SHANG Huaiying[1,2]

1. Beijing Institute of Radio Metrology and Measurement, 100854, Beijing, China;
2. Science and Technology on Metrology and Calibration Laboratory, 100854, Beijing, China
1. cappuccino_whf@163.com

Abstract: Expatiated the research and application of precise remote time synchronization based on communication satellite all over the world at present. Mainly introduced the research of precise remote time synchronization equipment in Beijing Institute of Radio Metrology and Measurement (BIRMM). Described the comparison experiment between BIRMM modem and SATRE, and accomplished the evaluation and validation on performance of the equipment. The result indicate that the precision of remote time synchronization is less than 0.3ns, and the equipment can satisfy the requirement of high precision time synchronization in the metrology organizations and weapon systems.

Keywords: Communication Satellite; Remote Time Synchronization; Modem

1 INTRODUCTION

With the development of science and technology, the high-precision remote time synchronization in the national economy is becoming more important in the high-tech industry, such as telecommunication, electricity, transportation and high-speed digital networks. In recent years, with the development of national defense and space technology, for high-precision remote time synchronization put forward higher requirement, even though China has established COMPASS navigation system, but its timing precision can only reach a few nanoseconds. It is unable to meet the requirements in the realms of higher precision time synchronization, such as air target detection and interception, deep space exploration, time and frequency measurement and traceability.

Two-way satellite time transfer (TWSTT) can be achieved with high precision and high accuracy time synchronization via a communication satellite and the system contains a modem and a VSAT. It is widely believed that the precision of time synchronization is less than one nanosecond. TWSTT measurement data have become a main part of the international atomic time calculation in current days. The European Galileo program have compared GPS common view (GPS-CV) and TWSTT time synchronization experiments, the comparison result shows that the performance of TWSTT technology is more better, and due to dependence and security considerations for GPS navigation system, the European Galileo system chose TWSTT technology for time synchronization of ground stations.

Although China time-frequency system construction has made great achievements, but for various reasons, reliable and independent time and frequency standard system has not yet built, about 80% of civilian and military areas commonly used GPS time. The relevant state departments have clearly strengthened to establish our independent time and frequency system. At present, time resources are dispersed and underutilized in China, establishing domestic data share links of

keeping clock ensemble, forming a nationwide time keeping system, improving domestic time keeping capacity is imperative.

In addition, as COMPASS navigation system works steadily, positioning and timing function need further improvement. Time synchronization precision of satellites and ground stations, satellites, ground stations is one of the key determining factors. Meanwhile, high-precise clock onboard and ionosphere changes monitoring in real-time all depend on high-precise time synchronization.

Firstly, this paper expatiated the research and application of precise remote time synchronization based on communication satellite all over the world at present. Then, introduced the research of precise remote time synchronization equipment in Beijing Institute of Radio Metrology and Measurement (BIRMM). Finally, described the comparison experiment between BIRMM modem and SATRE (TIMETECH). The result shows that BIRMM modem has reached the international advanced level.

2 STATUS OF DOMESTIC AND ABROAD

From the 1960s, proposed and applied to the time transfer, TWSTT technology have been occurred a number of technical changes, especially with the development of commercial communication satellites which made time synchronization device more simplified and more high performance.

2.1 STATUS OF ABROAD

In terms of the use of TWSTT device, about twenty international time laboratories have established two-way satellite links and took part in the International Atomic Time (TAI) calculation. These links are the Asia-Pacific, the European and the North American time comparison networks, and the core of them are the National Institute of Information and Communication Technology (NICT) in Japan, the United States Naval Observatory (USNO) in America and the Physikalisch-Technische Bundesanstalt (PTB) in German.

In terms of development and production of TWSTT device, TWT-100 developed by ATLANTIC company in America and SATRE developed by TIMETECH in German are the two representative products, the two have completed the commercial production and marketing all over the world, and the major product in use is SATRE.

2.2 STATUS OF DOMESTIC

In terms of the use of TWSTT device, since 1998 National Time Service Center (NTSC) have established TWSTT link between China and Japan, and in 2009 NTSC and National Institute of Metrology(NIM) have joined the newly established Asia-Europe TWSTT network. Since 2006 Beijing Institute of Radio Metrology and Measurement have began to concentrate on the research of TWSTT technology, established TWSTT system based on the SATRE and planed to join the next Asia-Europe TWSTT network.

In terms of development and production of TWSTT device, appearance of a mature commercial one which is comparable with the SATRE is not yet. Since 2006 BIRMM have been committed to the research of the TWSTT modem, have broken through a number of key technology about the production of the device, formed a set of practical technical solutions, and completed a TWSTT modem product prototype.

3 RESEATCH STATUS IN BIRMM

3.1 TIME SYNCHRONIZATION SYSTEM IN BIRMM

The principle of the TWSTT system in BIRMM is similar to the common international method. Using pseudo-random coding technology, the two ground stations which need time synchronization send its second pulse signal to each other via geosynchronous commercial communication satellite. The signal is received through an antenna and a receiver on the other side, the transmission time can be got by the demodulation technology. When the two ground stations exchange the transmission time data, the two can all obtain the time difference and the time synchronization is got. Figure 1 shows the TWSTT system at BIRMM.

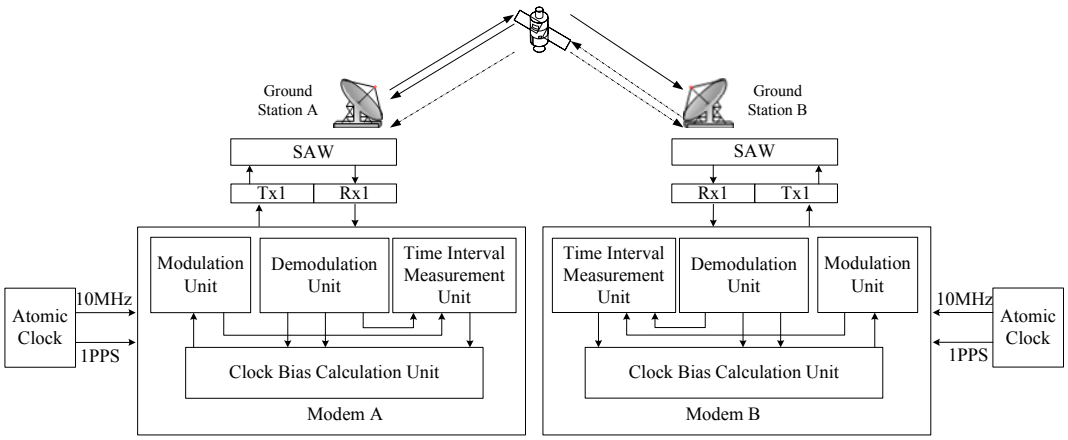


Figure 1. TWSTT System at BIRMM

3.2 RESEARCH STATUS IN BIRMM

The core parts of the time synchronization equipment are logic devices, high-speed processor, DAC and ADC chips. The modulation, demodulation and time interval measurement unit are highly synchronized to an external 10MHz frequency signal input. The architecture of the modulation and the demodulation unit is digital processing. The processing stream[1,3] of intermediate frequency(IF) signal is shown in Figure 2.

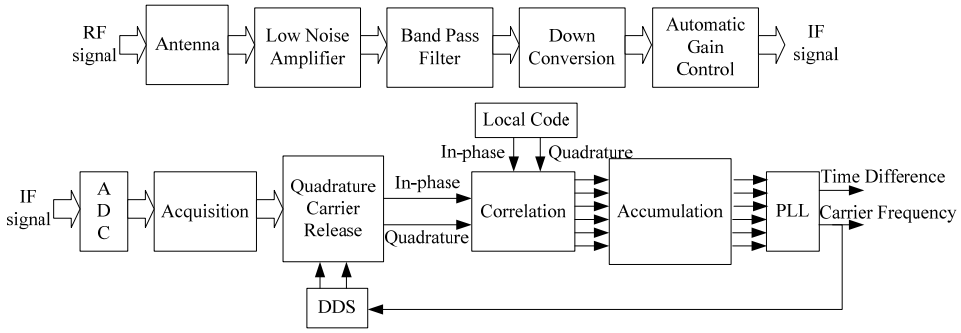


Figure 2. Stream of IF Data

The RF signal become into IF signal through a low noise amplifier (LNA), a band pass filter (BPF) and a down conversion module; IF signal is digitized through an analog to digital conversion(ADC) module; Acquisition module make coarse synchronization to the signal, and the initial code phase and carrier frequency are supplied to the tracking module; The signal is quadrature down converted in quadrature carrier release module; Six correlation values are supplied to the integral summation module; PLL module track the code and carrier in real-time and the error signal is fed back to the local code and carrier generator; The measure value of the time difference is framed for transmission in the host computer.

User can configure the device for their different requirements, the parameters can be configured through the host computer software, and make it possible to work separately at 5MHz, 2.5MHz and 1MHz, and each code rate corresponding to 6 different patterns to meet multi-user and multi-channel high-precision remote time synchronization simultaneously as shown in Figure 3. The device is more flexibility and configurability compared to the traditional ones, and the device is show in Figure 4.

Figure 3. Interface

Figure 4. BIRMM Modem

3.3 Experiment

3.3.1 Experiment Condition

We made zero baseline comparison experiment between BIRMM modem and Germany SATRE, the link is established by communication satellite Star 10 (110.5°E) on March 2015, and the test site is at BIRMM (39.91°N,116.26°E). Both the environment and auxiliary equipment are identical in the comparison experiment as shown in Table 1, and the Outdoor Experiment Scene is shown in Figure 5.

Table 1. Validation Condition

	SATRE	BIRMM MODEM
Chip rate	5MHz/2.5MHz/ 1Mz	5MHz/2.5MHz/ 1Mz
Antenna caliber	1.2m	1.2m
Satellite in use	Star 10	Star 10
Outdoor unit	ANACOM	ANACOM
Environment temperature	20℃	20℃

Figure 5. Outdoor Experiment Scene

3.3.1 EXPERIMENTAL SCHEME

In order to facilitate comparison performance between BIRMM modem and SATRE, we must ensure that the time and frequency source are identical. So the 1PPS and 10MHz signal

from BIRMM modem are supplied to SATRE simultaneously. Meanwhile, the 1PPS signals of both devices input to an external high-precision time interval counter (CNT91R), and choose HP5071 as its external reference. We all know that the data from CNT91R is the real value of time bias, measure data of time bias from BIRMM Modem and SATRE can reflect the performance of two devices. So, comparing time bias data from CNT91R and two devices is a reasonable way to validate the performance of BIRMM modem, the connection of the validation system is shown in Figure 6.

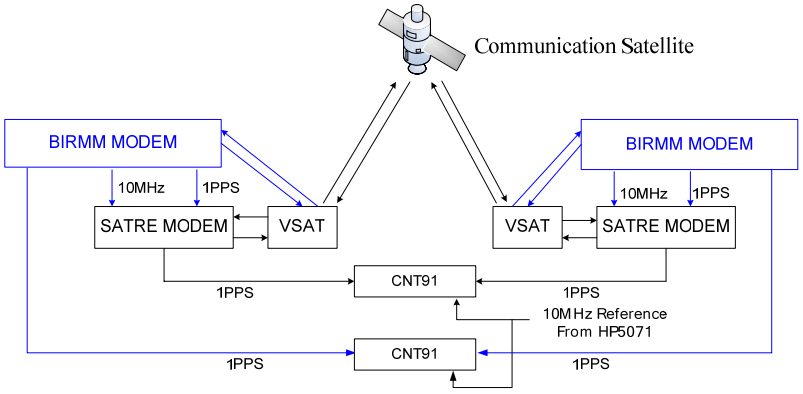


Figure 6. Comparison Experiment

3.3.1 EXPERIMENT RESULT

Generally, satellite navigation signal strength is about -130dBm, while the time synchronization signal strength is much higher so that this value exceed the limit of the peak of ADC. The received signal cannot be quantified accurately which makes the measurement value trembling violently at the beginning. Then we add automatic gain control (AGC) module before ADC, then the measurement results significantly improved. Meanwhile, increasing the non-coherent integration time which is equivalent to measurement results filtering makes the results further improved.

Figure 7, 8, 9 shows the final result of the two devices, wherein each curve is the difference between measurement results of the two devices and external time interval counter.

In addition, the performance of the interval counter inside is another key factor in the decision of the system performance, and it measures the local time difference between the modulator and demodulator. The result of the time interval counter which is self-developed is shown in Figure 10, and the measurement accuracy is better than 0.058ns(1σ).

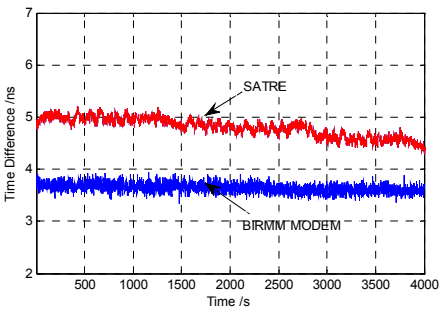


Figure 7. Comparison of 5MHz Chip Rate

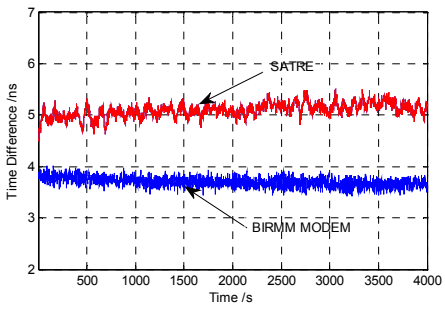


Figure 8. Comparison of 2.5MHz Chip Rate

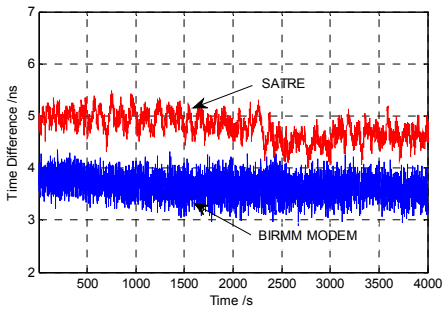


Figure 9. Comparison of 1MHz Chip Rate

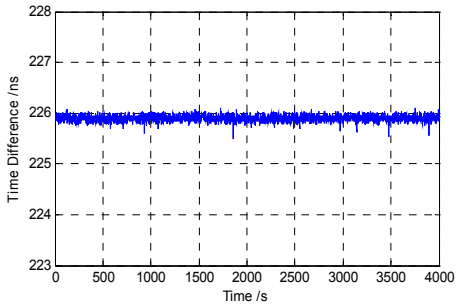


Figure 10. Time Interval Measurement Result

3.3.2 RESULT ANALYSIS

Table 2 shows the result of statistical comparison between BIRMM modem and SATRE, and the overall performance is quite.

Table 2. Standard Deviation of Experiment Result

Chip rate	Standard Deviation of SATRE	Standard Deviation of BIRMM MODEM
5MHz	0.13ns	0.16ns
2.5MHz	0.14ns	0.13ns
1MHz	0.25ns	0.26ns

4 CONCLUSION

Based on communication satellite, this paper introduced structure and feature of the prototype of self-developed high-precision remote time synchronization equipment. Presented the validation Scheme and program. Compared the result between BIRMM modem and SATRE, and the result showed that the accuracy of BIRMM modem have reached a level better than 0.3ns. Long baseline experiment based on GEO satellite will be the focus of future work. BIRMM Modem can meet the requirements of COMPASS navigation system, and is expected to be applied to high-precision remote time synchronization between ground stations and stars.

ACKNOWLEDGMENT

The authors would like to express their gratitude to staff at BIRMM for a great deal of cooperation of this experiment.

REFERENCES

[1] Elliott D.Kaplan， Understanding GPS Principles and Applications, Second Edition[M]. Beijing: Publishing House of Electronics Industry, 2009.

[2] Wen-Hung Tseng, Yi-Jiun Huang, Tadahiro Gotoh, Thomas Hobiger, Miho Fujieda, Masanori Aida, Tingyu Li, Shinn-Yan [2] Lin, First International Two-Way Satellite Time and Frequency Transfer Experiment Employing Dual Pseudo-Random Noise Codes, IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 59, no. 3, March 2012.

[3] Tadahiro Gotoh, Jun Amagai, Thomas Hobiger, Miho Fujieda, and Masanori Aida, Development of a GPU-Based Two-Way Time Transfer Modem, IEEE Transactions On Instrumentation and Measurement, vol. 60, no. 7, July 2011.

[4] Lei Dong, IF GPS Signal Simulator Development and Verification[M], Department of Geometrics Engineering, 2003

Application of User Terminal in the Field of Space-based Internet

Li Haihao, Huang Yin , Chu Jianxiang, and Liu Xiaokuan

（Beijing Institute of Remote Sensing Equipment）
E-mail:lihaihao@sohu.com

Abstract： For the main purpose of accessing space-based Internet, “Fuxing” is designed as a LEO broadband communication system which has very indispensable application prospects in the field of Internet extension. “Fuxing” has broad prospects in the civil communication. The system consists of the space part, the ground part and the user part. Because of user terminal directly provides service for the user, whose technique design directly affects the difficult degree of the satellite communication system commercial operation. “Fuxing” terminal gives attention to both performance and cost advantage, consists of variety of types such as phased array antenna and reflector antenna, which is suitable for high speed motion, low speed motion and static platform, covers user group such as high speed train, plane, ship, automobile, sea islands, drilling platform, the remote areas, office building and personal portable device. Firstly this paper analysis the different user group demand for satellite space-based Internet, then describes the function and use of “Fuxing” terminal, which has the reality guiding purpose for the better use of LEO space-based networks. Finally, this paper concludes that the user terminal will play an important role in user experience of space-based Internet.

KeyWords: “Fuxing” communication system, User terminal, Space-based Internet

1 Introduction

User terminal is an important component of the satellite communication function realization. At the technical level, user terminal has developed from analog system to a variety of digital system. At the business level, user terminal has expanded from the single telephone service to services including data, fax, short message, E-mail, broadband multimedia and Internet application, as well as some new services, such as satellite positioning, distress alert rescue, aviation and maritime security, shipping management. In terms of the application field, it has covered a variety of users, including land, sea and air[1].

The user terminal equipment is structurally shrinking in size and weight, and develops from the fixed mode of ship, vehicle, airborne to suitcase which is suitable for handling type, portable type and multimode hand holding type[2]. We can conclude that user terminal technology of communication satellite is developing to the direction of miniaturization, intelligence, multimode, multi-service and combination with navigation.

2 Domestic and foreign development actuality

Since 1990s, LEO satellite communication system draws high attention. Countries in the world have developed a number of LEO satellite communication systems. Iridium, Globalstar, Orbcomm, Tleldesic are the typical LEO satellite communication systems[3]. However, only Iridium, Globalstar and Orbcomm are operating successfully currently. In this paper, we will take the above

three foreign systems and domestic typical systems as example, and introduce the development of satellite communication terminals[4][5].

a) Iridium user terminal

Iridium system supports several categories and types of user terminals, including telephone, marine, aviation terminal and so on. Data transfer rate of the first generation Iridium telephone was only 2.4kbps. As a result, except for call, it could only send short e-mail or slow fax, which was unable to meet the needs of the increasingly popular Internet. The second generation Iridium can support hand holding terminal, provide L-band (of which the speed is up to 1.5 Mbps), and Ka-band (of which the speed is 8 Mbps) high speed services.

b) Globalstar user terminal

In addition to the north and south poles, "Globalstar" satellite system provides seamless worldwide coverage, and provides low-cost satellite mobile communication services, such as voice(2.4/4.8/9.6kbps), fax, data(7.2kbps), group 3 fax, short message, positioning. And its user terminal includes four types: fixed station, mobile station, personal station and radiolocation terminal. The hand holding terminal has three modes, which include Globalstar single-mode type, Globalstar/GSM dual-mode type and Globalstar/CDMA/AMPS tri-mode type. There are several user terminal manufacturers, such as Qualcomm Incorporated of USA, Ericsson of Sweden and Italy Telital Company. The sigle-mode terminal can only be used within the Globalstar system. The dual-mode terminal system can be used not only within the Globalstar system, but also within the ground cellular system or other mobile communication systems.

c) Orbcomm satellite user terminal

There are two types of Orbcomm satellite terminal. The first one was used in data acquisition and transmission in remote fixed location, the other one was used in bidirectional data and information communication for mobile terminal. Satellite communication terminal which is used in fixed location data service adopts low-cost VHF electron device, of which the antenna design and equipment structure is simple, and it can be flexible installed, consumes low power, and can work for a long time using battery, solar panel and other available power supply device. Satellite communication terminal which is used in mobile carrier bidirectional communication was portable handset, and there are usually alphanumeric keyboard and small screen on that. Users can use the keyboard and screen at any time.

d) User terminal of domestic LEO satellite

Shanghai Microsatellite Engineering Center successfully runs a microsatellite, which is called "Innovation One" whose ground user terminal weigh is less than 300g. It has simple structure and is easy to carry. On the other hand, "Tsinghua University —Xinwei Telecom Aerospace Information Network Technology Joint Research Center" develop smart communications test satellite, of which the hand holding satellite terminal can realize services including the communication between the hand holding satellite terminals, communication between the hand holding satellite terminals with telephone, Internet data transmission, electromagnetic spectrum monitoring and location.

3 "Fuxing" system user terminal

For the main purpose of accessing space-based Internet, "Fuxing" is designed as a LEO broadband communication system which has very indispensable application prospects in the field of Internet extension. It consists of dozens of satellites and covers "One Belt And One Road", Russia and other neighboring regions. The first phase of "Fuxing" will adopt the function of satellite with

ground network and transparent transmittability. It can provide the communication ability which is high rate, low delay, undifferent in wide area and get rid of the traditional optical fiber transmission means.

Various types of user terminal of “Fuxing” system consist of antenna, high-frequency transceiver, signal processing, power supply and other component. To build terminal custom platform, users can customize the terminal according to their need. Custom contents include multiple antennas, high-frequency transceiver and other accessory parts. Among them, the antenna includes two-dimensional phased array antenna, one-dimensional phased array antenna, servo reflector antenna, fixed reflector antenna and portable multi-plane antenna. Other accessory parts will be customized according to installation environment, installation method and power supply mode.

a) User terminal of two-dimensional phased array antenna

The user terminal of two-dimensional phased array antenna is mainly applied to high speed motion carrier, which requires high communication rate. And it can be conformal with external of aircraft and high-speed railway, which will meet the shape requirements of aircraft and high-speed railway, such as aerodynamic characteristics. This type of terminal antenna adopts smart two-directional phased array antenna, uses electronic scanning to track satellites. The antenna standard size is 0.6m diameter (single-sided). The overall structure of the terminal is 0.65m×1.4m×0.1m. The antenna gain is no less than 35dB and its communication elevation is 20°. The theoretical communication rate is 100Mbps. The user terminal can provide data download capability of 16Mbps ~ 66Mbps in fine weather, and can provide data download capability of 2Mbps~8Mbps under the condition of rainfall, and can achieve various services including data and voice. It has built-in WIFI module (802.11 a/b/g/n/ac) and Beidou / GPS satellite positioning module. Mobile phone, PAD, etc. can be connected with the terminal through WIFI. Terminal uses 220V/50Hz power supply. Its appearance and application scenarios are shown in Figure 1.

Figure 1 Appearance and application scenarios of Two-dimensional phased array antenna user terminal

b) User terminal of one-dimensional phased array antenna

The user terminal of one-dimensional phased array antenna is low-cost, portable terminal, and is mainly applied to personal communication, emergency communication, data acquisition, individual combat and other scene which requires relatively low rate communication. It can be also placed on small mobile carriers, and meets the needs of family car and bus. In order to decrease the number of T/R module and cost, this type of terminal antenna uses one-dimensional phased array antenna, another one-dimensional antenna uses mechanical rotation servo. The antenna standard size is 0.4m diameter (single-sided). The overall structure of the terminal is 0.8m ×0.4m×0.06m. The antenna gain is no less than 30dB and its communication elevation is 20°. The theoretical

communication rate is 50Mbps. The user terminal can provide data download capability of 7Mbps ~ 30Mbps in fine weather, and can provide data download capability of 0.9Mbps~3.7Mbps under the condition of rainfall. It can achieve various services including data and voice. It has built-in WIFI module (802.11 a/b/g/n/ac) and Beidou / GPS satellite positioning module. Mobile phone, PAD, etc. can be connected with the terminal through WIFI. Terminal uses two types of power supply, 220V/50Hz power and built-in battery. Its appearance and application scenarios are shown in Figure 2.

Figure 2 Appearance and application scenarios of one-dimensional phased array antenna user terminal

c) Portable user terminal of multi-plane antenna

The user terminal is low-cost, portable terminal, and mainly is applied to personal communication, emergency communication, data acquisition, individual combat and other scene, meets the demand of personal communication. In order to cut the cost, this type of terminal adopts multiple narrow beam antennas, achieving hemispheric pattern covering. The antenna gain is no less than 6dB. The overall structure of the terminal is 0.15m × 0.15m × 0.01m. This type terminal supports 2.4kbps communication rate of voice and SMS. It has built-in WIFI module (802.11 a/b/g/n/ac) and Beidou / GPS satellite positioning module. Mobile phone, PAD, etc. can be connected with the terminal through WIFI. Built-in battery can ensure continuous working time 2h, 30h standby time. Its appearance and application scenarios are shown in Figure 3.

Figure 3 Appearance and application scenarios of multi-plane antenna portable user terminal

d) User terminal of servo reflector antenna

The user terminal is mainly applied to low speed motion or stationary carrier, which requires higher communication rate. It gives attention to both performance and cost advantages, meets the application demand of boats, cars, islands, drilling platforms, villages, office building and other low speed or stationary platform. This type terminal uses reflector antenna, whose standard size is 0.6m diameter. It adopts servo scanning mode to track satellite. The overall structure of the terminal is

0.7m ×1.4m. The antenna gain is no less than 35dB and its communication elevation is 20°. The theoretical communication rate is 100Mbps. The user terminal can provide data download capability of 16Mbps ~ 66Mbps in fine weather, and can provide data download capability of 2Mbps~8Mbps under the condition of rainfall. It can achieve various services including data and voice. It has built-in WIFI module (802.11 a/b/g/n/ac) and Beidou / GPS satellite positioning module. Mobile phone, PAD, etc. can be connected with the terminal through WIFI. Terminal uses 220V/50Hz power supply. Its appearance and application scenarios are shown in Figure 4.

Figure 4 Appearance and application scenarios of servo reflector antenna user terminal

e) User terminal of fixed reflector antenna

The user terminal is mainly applied to low speed motion or stationary carrier, which requires low communication rate. It can achieve data return at very low cost. Since it doesn't have the ability to track satellite, it is mainly used in non-real-time communication and it can be applied to field data acquisition and other scenarios. This type terminal uses reflector antenna, whose standard size is 0.2m diameter. The overall structure of the terminal is Φ0.3m ×0.5m. The antenna gain is no less than 25dB. The theoretical communication rate is 10Mbps. The user terminal can provide data download capability of 1.8Mbps ~ 7.3Mbps in fine weather, and can provide data return capability of 0.2Mbps~0.9Mbps under the condition of rainfall. It has built-in WIFI module (802.11 a/b/g/n/ac) and Beidou / GPS satellite positioning module. Mobile phone, PAD, etc. can be connected with the terminal through WIFI. Terminal uses two types of power supply, 220V/50Hz power and external solar battery. Its appearance and application scenarios are shown in Figure 5.

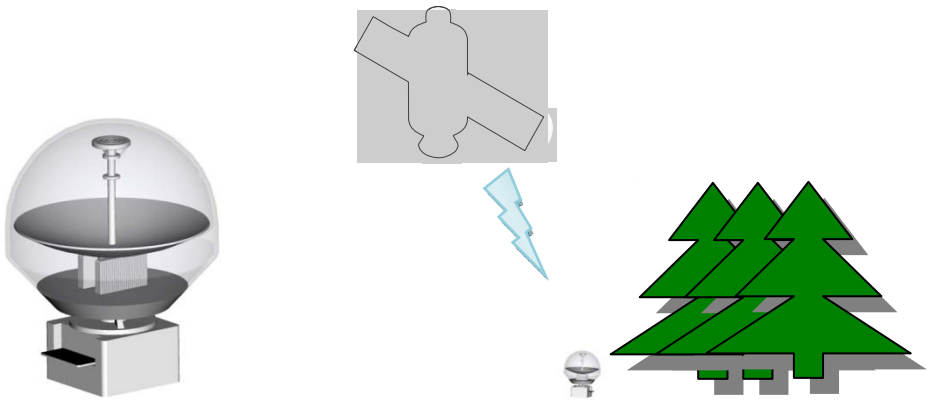


Figure 5 Appearance and application scenarios of fixed reflector antenna user terminal

4 Summary and outlook

From the above we can conclude that, the general trend of the future development of satellite communication terminal is developing from a large volume and weight to being small, flexible and portable, developing from supporting a single type of business to supporting a variety of services[6]. The user terminals finally made up seamless global coverage personal communication system with ground communication network. “Fuxing” satellite communication system, which is designed by CASIC, is an indispensable part to build space-based Internet. Because of its extremely low communication delay, true global coverage, and higher communication rate, “Fuxing” will better serve economic development in “One Belt And One Road” area in which the user terminal of “Fuxing” system will play an important part. Lately it will rely on LTE communication technology, SOC chip technology and other assistant to reduce the cost of user terminal development and production, to increase the richness of the user experience and to boost the promotion of “Fuxing” system. “Fuxing” system will realize new developing application and value-added area of China’s aerospace Industry in the “Internet Plus” era.

Reference:

[1]Pulak K C,Mohammed A.Handover Scheme in Satellite Networks:State-of-the-Art and Future Research Directions[J].IEEE Communications Surveys&Tutorials,2006,8(4):2-14.

[2]Papapetrou E,Pavlidou F-N,QoS Handover Management in LEO/MEO Satellite Systems [J].Wireless Personal Communication,2003,24(2):189–204.

[3]Paolo Barsocchi,Nedo Celandroni,Franco Davoli and Erina Ferro，Radio resource management across multiple protocol layers in satellite networks:A tutorial overview[J]，International Journal of Satellite Communication Network.2005;23,pp.265–305.

[4]Ahmed M.H.,Call admission control in wireless networks:a comprehensive survey [J],IEEE Communications Surveys&Tutorials,2005,7(1),pp.50-69.

[5]Majid Ghaderi,y and Raouf Boutaba,Call admission control in mobile cellular networks:a comprehensive survey[J],Wiley,Wireless.Communication.Mobile Computing,2006;6,pp.69–93.

[6]Hoang Nam Nguyen,Olariu,S.,Todorova,P.,A novel mobility model and resource reservation strategy for multimedia LEO satellite networks,Proceedings of Wireless Communications and Networking Conference,WCNC2002,Vol.2,17-21,March 2002,pp.832-836.

Application of ISLs(Inter-Satellite-Links) in the Field of Space-based Internet

Li Haihao, Chu Jianxiang, Huang Yin and Liu Xiaokuan

（Beijing Institute of Remote Sensing Equipment）

E-mail:lihaihao@sohu.com

Abstract： For the main purpose of accessing space-based Internet, “Fuxing” is designed as a LEO broadband communication system, which has very broad application prospects in the fields of transitional Internet extension. Specially, ISLs are transmission links between adjacent satellites, and make the satellite constellation as a whole. Furthermore, satellite network with ISLs is not entirely dependent on ground network infrastructure, expanding the coverage of the system. Especially, for the LEO satellite, the single coverage is small. Maybe, there are no ground gateway stations visible, such as in the ocean or politically sensitive land, etc. In this case, ISLs used to make the terminals access to the ground backbone network is almost the only one solution. Firstly, this paper reviews the domestic and overseas developing experience of ISLs in the recent ten years. Then, the key technologies are summarized in detailed. Finally, this paper indicates implementation methods in second phase of “Fuxing” system.

KeyWords: “Fuxing” communication system, ISLS(Inter-Satellite-Links), Space-based Internet

1 Introduction

In the 21th century, big data, cloud services and smart home, which is based on the “Internet Plus”, have become labels of the era. However, there are still 3 billion people in the world who can’t use the infrastructure of broadband Internet. As a result, it will be a huge new emerging market to utilize the global satellite constellation, reaches the same speed as the optical fiber and provided satellite Internet services for ground areas which can not be covered by the Internet. For the main purpose of accessing space-based Internet, “Fuxing” is designed as a LEO broadband communication system, which has very broad application prospects in the fields of transitional Internet extension. ISLs(Inter-Satellite-Links) can achieve interoperability between satellites and it can effectively reduce the number of ground gateway stations on the basis of on-board processing and on-board routing. This is particularly important for the construction of satellite communication system which doesn’t have a global distribution capacity, such as “Fuxing” system[1].

ISLs are transmission links between adjacent satellites, which consists of the transmitter, receiver, antenna and other components. Using ISLs, constellation can construct network in the space, make the satellite constellation as a whole syetem[2]. Furthermore, satellite network with ISLs is not entirely dependent on ground network infrastructure, expanding the coverage of the system. Especially, for the LEO satellite, the single coverage is small. Maybe, there are no ground gateway stations visible, such as in the ocean or politically sensitive land, etc[3]. In this case, ISLs used to make the terminals access to the ground backbone network are almost the only one solution. Establishing ISLs is the only way to make every satellite as a node in the communication network and it can be called "space-based Internet" really[4].

It can be drawn from the above that, establishing ISLs is important key technology to improve

the service capability and autonomous operation capacity of satellite communication system. In this paper, based on the literature published at home and abroad, we introduce the domestic and overseas construction situation of ISLs in the famous star network system, and indicate implementation methods in second phase of “Fuxing” system.

2 Domestic and foreign construction situation

Global satellite communications companies carry out demonstration and construction work of ISLs in succession. Among the foreign satellite network systems which have certain commercial and military value, the Iridium system(launched by Motorola), Teledesic system(jointly proposed by Microsoft and McCaw Communications Corporation), Astrolink system and Spaceway system(which located in the high and middle orbit) all have declared that they already have or will achieve on-board processing and ISLs capabilities[5][6].

The most important feature of the Iridium system is to achieve global communications via links between satellites, which is equivalent to moving the cellular mobile telephone system to space. “Iridium” communicates with the neighboring satellites via ISLs in Ka-band. There are 4 ISLs for each satellite: two “Iridium” at the same orbit plane, and two “Iridium” at the adjacent plane. It takes about 100 minutes to orbit from the South Pole to the North Pole. This design of Iridium means good satellite visibility and services coverage in the north and south poles. Since the “seams” satellites are designed according to the polar-orbit, when the satellites run along the track direction of orbit plane, satellites on the other orbit plane are running in the opposite direction. Horizontal ISLs will switch fast to cope with Doppler shift. As a result, Iridium ISLs can only support satellites in the same direction orbit.

Teledesic communication satellite is planning to provide two-way broadband telecommunications services for the whole world, including computer networks, fast broadband Internet access, interactive multimedia, and high-quality voice technology. Teledesic is known as aerial optical fiber network, which is based on satellite networks, utilizes fast packet switching and provide seamless global coverage. The satellites establish fast packet switching and communication with satellites in the same or neighboring orbit pane via cross channels. Teledesic system takes every satellite as a communication node, constructs inter-satellite switching network, has the ability of global coverage, so it is truly “global space Internet”. But due to the huge investment and the low return on investment, its development is greatly restricted.

In middle and high orbit, Astrolink uses on-board processing to improve system efficiency, and uses satellite switching to enhance system flexibility. Each satellite is a node of the communication network, which is different from “bend relay” transmission, and the business needs to be forwarded through satellite nodes. However, Spaceway3 system uses on-board processing and unique fast packet switching technology. And aimed at improving the invulnerability of satellite communication system and high-speed access onto the ground, Spaceway3 system plans to add ISLs in the next step.

LEO satellite constellation communication system which has ISLs is capable to form a true global coverage system. It can reduce dependence of ground facilities, provide broadband real-time service and will be an important component of China’s future space-based network. Since the 10th five-year plan, many domestic research institutions have carried out research about the constellation system design, networking technology, and the key technology of LEO satellite constellation. They have got a certain achievements. However, China has no star network system which is designed and

implemented independently.

3 Application in the second phase of “Fuxing” system

“Fuxing” system adopts uniform Walker constellation. 156 satellites are evenly distributed in 13 orbit planes, and there are 12 in each orbit plane. The maximum bandwidth demand of a Single satellite is 2.5Gbps and the distance between the adjacent satellites ranges from 3000km to 5500km. The first phase of “Fuxing” will adopt the function of satellite with ground network and transparent transmittability. The second phase of “Fuxing” will adopt the work way of ISLs. Satellite constellation is shown in Figure 1.

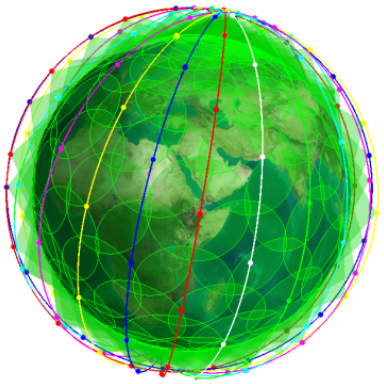
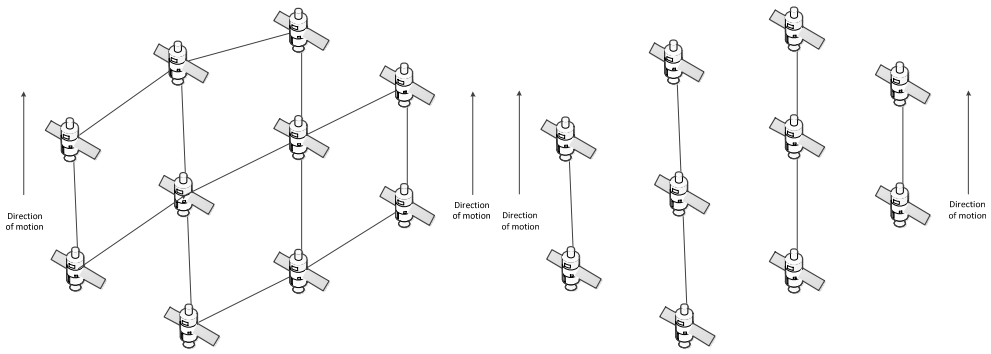


Figure 1 satellite constellation

Network topology can be neither too loose nor too dense. Loose network topology will increase the number of routing hops, thereby will increase the signal transmission delay. Dense network topology will have high capacity requirements for the ISLs, which will increase the difficulty and cost of implementation[7]. In satellite constellation, there are 6 satellites around each satellite. In most sparse topology, the satellite link is established with only one satellite around. In the most intensive topology, the satellite will maintain links with the six satellites around. Two network topologies are proposed in this paper. The first network topology adopts each satellite establishing links with 4 satellites around and its local topology is shown in Figure 2(a). The second ring topology adopts Each satellite connected with two satellites, before and after. Its local topology is shown in Figure 2(b).



(a)network topology (b)ring topology

Figure 2 topology of Inter satellite network

Using the “Fuxing” communication system, user terminal can transfer data, voice and Internet services through its top satellite. The transmission process of information from the user to the core network is called uplink access. And the transmission process of information from the core network to the user is called downlink access. The simplified information transfer diagram is show in Figure 3.

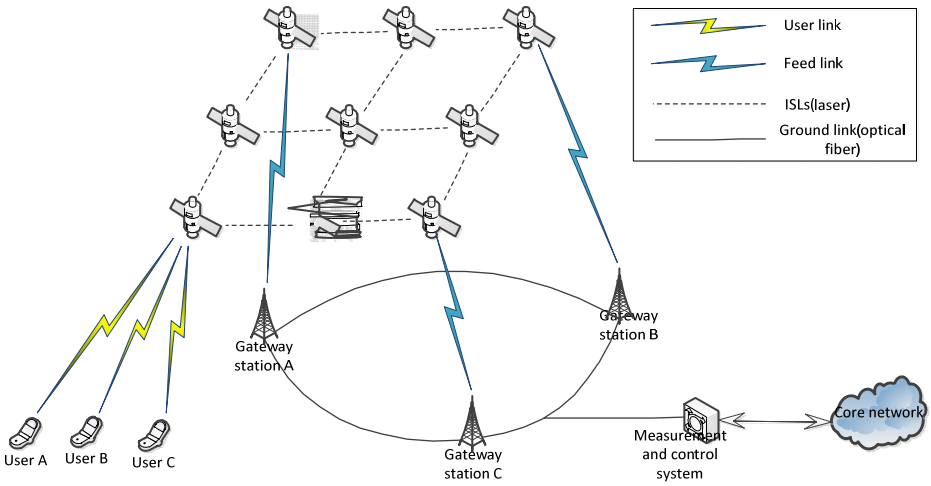


Figure 3 ISLs system diagram

Depending on the purpose of transmission of information, the communication process can be divided into uplink access and downlink access. The aim of uplink access is that user’s request and other information access to the core network via satellite network, gateway station, measurement and control system. When the information is transferred in the space, its purpose is to transfer onto ground as soon as possible. So in the satellite network, the selection of uplink access node is non-unique, and it can be any satellite who has a feed link. The aim of downlink is to transfer information of core network to the users via measurement and control system, gateway station and satellite network. When the information is transferred in the space, there is a clear satellite which covers the ground user, and its destination node is unique. As a result, we will propose routing protocols which are available to reference in accordance with different types of information transmission.

Inter satellite routing is carried out according to the procedure shown in Figure 4.

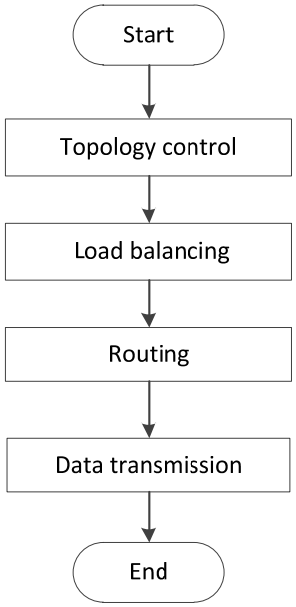


Figure 4 Inter satellite routing flow chart

3.1 Uplink access process

In the uplink access process, destination nodes are not unique. So we recommends dynamic routing in this paper.

Traditional dynamic routing finds the shortest path from the source node to the destination node on the basis of distance vector. Reference to thought of traditional dynamic routing, we can find the best destination node and perform information landing onto the ground. Since the total throughput of the entire system is limited by the capacity of feed links form the satellite to the gateway station, it is essential to balance load between the satellites. When source satellite has the information of user access to ground, group by broadcasting destination request, find the satellite node which can land to ground, and return the destination answer along a reverse path. The source satellite node chooses the satellite with the greatest remaining bandwidth as the destination node according to a certain rule.

This route maintenance algorithm will delete a route entry when its lifetime expires. The flowchart of uplink routing is shown in Figure 5.

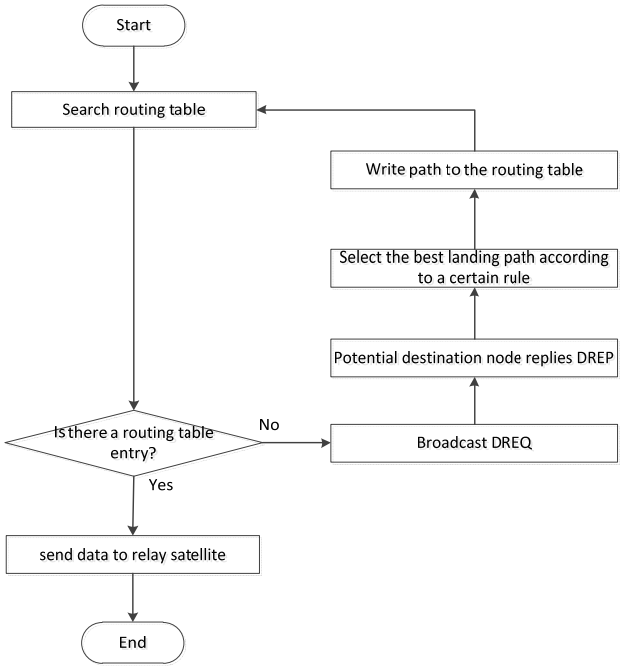


Figure 5 uplink routing flowchart

3.2 Downlink access process

The routing scheme of downlink access is proposed to use static routing, which is the improved OSPF routing protocol.

The traditional OSPF routing protocol adopts Dijsktra algorithm to establish the shortest path from source node to all the nodes in the entire network. The precondition of traditional OSPF algorithm is that it is already known that topological relationships and channel transmission capacity is much greater than demand bandwidth. It generates the routing table according the shortest hop count. But in the “Fuxing” system, channel capacity of ISLs may be lower than the required bandwidth. So we can not simply establish routing table according to the shortest hop count. We need to establish a main routing table according to load balancing, and maintain a backup route. The routing path can not be overlapped.

Routing maintenance includes main routing table maintenance and backup rout maintenance. Main routing table is generated in accordance with Dijsktra algorithm， but is not restricted to the shortest hop count. It will find the shortest path routing on the condition that there is no routing path overlapped. Since the ephemeris map of the satellite network is certainly known, each satellite can foresee the topological relationship of the entire network. The maintenance of the main routing table just needs to be updated according to the operation rules of satellites constellation and the topology control algorithm. The maintenance of backup route is performed by a hop broadcast Hello packet, and the Hello packet carries the backup path information. When the node establish routing transmission and finish routing transmission (or the ISL connected to the node is overloaded), the satellite will broadcast Hello grouping and tell the neighboring satellite nodes.

When the satellite has data to relay, it inquires the routing table and find the relay satellite which is suitable to be the next hop. Once there is a ISL overloaded (greater than 80% of the channel capacity is considered to be overloaded) in the transmission path, the satellite will adopt backup route for diversion. The flowchart of downlink routing is shown in Figure 6.

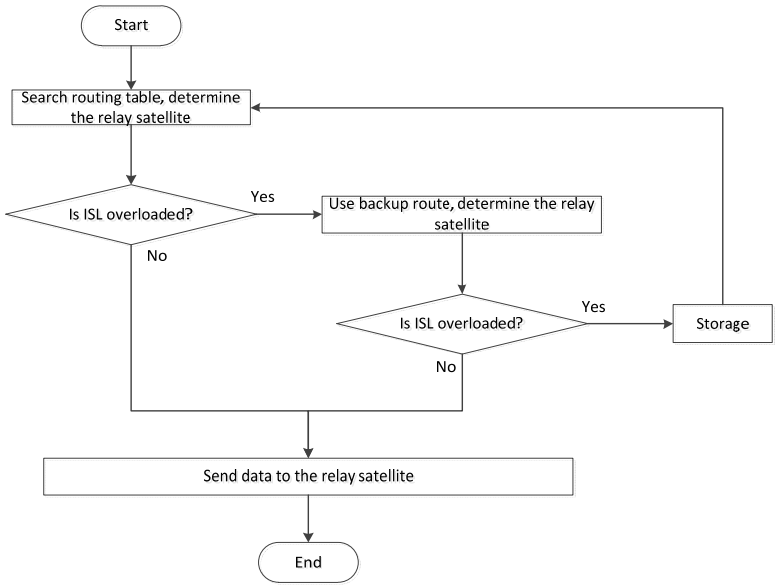


Figure 6 downlink routing flowchart

4 Summary and outlook

Conclusion can be drawn from the above that, the ISLs is an important component of the satellite communication system. Both the Iridium system which is already mature and systems which are still developing by various domestic and foreign companies, establishing ISLs is one of the key contents to be constructed[8]. In this paper, we summarize the construct situation of typical satellite network system present, further carry out the research of ISLs in the second phase of "Fuxing" system, and promote the transformation from research results to actual application. This paper will provide theoretical basis and solid technical support for the construction of ISLs in satellite communication system.

Reference

[1]Patterson D P.Teledesic:A Global Broadband Network[C]//IEEE Aerospace Applications Conference Proceedings,Snowmass at Aspen,CO,USA:IEEE Press,1998:547-552.

[2]Kota S L.Broadband satellite networks:trends and challenges[C]//WirelessCommunications and Networking Conference,New Orleans,LA,United states:IEEE Press,2005:1472-1478.

[3]Zahariadis T B,Vaxevanakis K G,Tsantilas C P,et al.Global Roaming in next-generation networks[J].IEEE Communications Magazine,2002,40(2):145-151.

[4]Lloyd W.Internetworking with satellite constellations[D].University of Surrey,2001.

[5]Kota S L.Broadband satellite networks:trends and challenges[C]//Wireless Communications and Networking Conference,New Orleans,LA,United states:IEEE Press,2005:1472-1478.

[6]Zahariadis T B,Vaxevanakis K G,Tsantilas C P,et al.Global Roaming in next-generation networks[J].IEEE Communications Magazine,2002,40(2):145-151.

[7]Pulak K C,Mohammed A.Handover Scheme in Satellite Networks:State-of-the-Art and Future Research Directions[J].IEEE Communications Surveys&Tutorials,2006,8(4):2-14.

[8]Papapetrou E,Pavlidou F-N,QoS Handover Management in LEO/MEO Satellite Systems [J].Wireless Personal Communication,2003,24(2):189–204.

The Development of Forestry Terminal System Based on Beidou Multimode Navigation

Jinsheng Ni, Jinlin Yang

Beijing Aerospace TITAN Technology Co., LTD.

Abstract: For the needs of forestry information services, develop Beidou multimode terminal hardware for the forestry, with adopting multi-modular, modular design approach, and unifying interface standards. Develop background intelligent monitoring, scheduling and managing technology that matches navigation terminal software, to realize the communication service, map service, data service and monitoring service of the terminal. Based on the Android operating system and mobile GIS platform get on the secondary development of the terminal, and realize spatial data infrastructure management, navigation, positioning and data acquisition functions, to meet the various needs of users of forest resources survey and to promote large-scale industrialization and application of Beidou system, and finally to led chips, devices, terminals, software integration and a series of high-tech industries to develop.

Keywords: Beidou/GPS; forestry; multimode navigation terminal; intelligent monitoring

Beidou satellite navigation system, which China developed and operated independently, is the important space infrastructure of China. Since its operation, Beidou navigation has been successfully used in national defense, mapping, telecommunications, water conservancy, fishery, transportation, meteorological, forest fire prevention, disaster relief and public safety and many other areas, resulting in significant social and economic benefits, especially playing an important role in Wenchuan earthquake relief. In the aspect of civil, however, there is a great distance from industrialization and marketization.

At present, the global satellite navigation and positioning technology research in forestry has concentrated on the GPS system(Chen., 2004). America has begun to develop GPS since the 1980s, and now its technology is very mature, and the application is extremely widespread, accounting for more than 95% of the global market. Therefore, the research on Beidou multimode navigation system for forestry will improve the utilization of the whole system, and promote the Beidou navigation system toward marketization and internationalization.

1. Software architecture

The Beidou multimode navigation terminal for forestry adopts a design of multi-modular and combined optimization. The system combines many functions in one (Fig.1), such as forestry data acquisition recorder, vehicle traveling data recorder, laser rangefinder, GPS, satellite navigation, GIS services, through full use of system platforms, mobile communications network, the Internet network. The system which has light weight, small volume, friendly interaction and high security features, and has an independent operating system and graphical user interface, can accurately locate and secure communication around the clock, and will meet the various needs of users of forest resources survey.

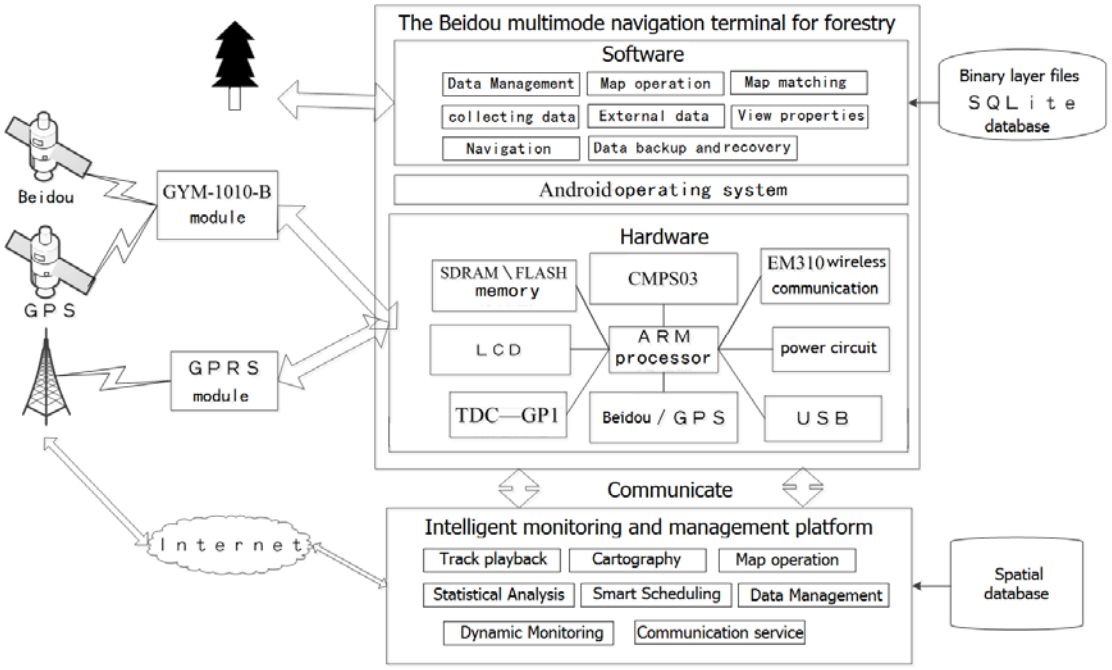


Fig.1. The overall framework

2. System design

The terminal mainly includes Beidou multimode navigation terminal hardware equipment and terminal intelligent monitoring, scheduling and managing platform software. The whole system i suited to the characteristics of the forestry industry, solving technical problems such as the end-product stability, multimode interoperability and rational allocation of memory.

2.1 Hardware design

Choose the appropriate processor and multimode positioning chip type based on the resource requirements of the terminal equipment, order to design and realize the terminal hardware platform. The processor, communication module, positioning module, and display module are integrated seamlessly. The overall power circuit is designed and integrated with the Android operating system, and ultimately forms a mobile communications and navigation terminal meeting the needs of survey of forest resources.(Fig.1)

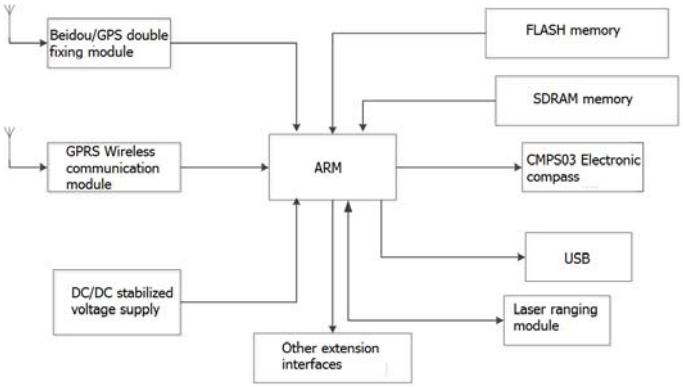


Fig.2. Terminal equipment structure

The key technology of Beidou navigation terminal hardware development for forestry mainly includes the following several aspects.

- The low power design of GPS/BD baseband signal processing

The baseband signal processing chip and low power consumption: use single-chip design, including high-speed satellite signal acquisition circuit and 16 fully independent all-digital channel and interface circuitry, with each channel can be arbitrarily configured as GPS (L1) / Beidou (B1) frequency and with the module using 45nm technology devices and low-power design methodology, average operating power consumption less than 1W.

- GPS / BD baseband signal capture technology

Capturing and tracking with high-precision technology: Capturing circuit adopts the matched filter and Fast Fourier Transform (FFT)， and can finish the two-dimensional search of code phase and Doppler phase simultaneously in a related process, to quickly capture satellite signals with capture accuracy achieving 1/4 to 1/8 chip. Full Digital signal tracking loop uses third-order Jaffe-Rechtin filter, and uses adjustable integration time and adjustable noise bandwidth technology, to make sure that theory tracking accuracy can reach 0.001 weeks.

- Precise multi-constellation positioning solution method

Location solver algorithms: On the basis of the classical least-squares algorithm, use error Revision troposphere, ionosphere model error / dual frequency ionospheric error correction, the carrier phase smoothing pseudorange and differential modification methods, to improve the accuracy of the solution; through adaptive kalman filtering technology, provide the stability of the solution; use autonomous integrity RAIM algorithm to provide availability of positioning results.

- Machine electromagnetic compatibility

Low noise amplifier circuit and the amplifier circuit are shielded by a shield cavity separately, to reduce mutual interference. Use cavity design to make isolation reach 60dB above. The isolator and filter are adopted in at the transmitter, and filters are also added respectively on each receiving end, to avoid the impact on B1 / L1 signal receiving output signal produces.

2.2 Software design

2.2.1 Beidou multimode navigation terminal software

(1) Architecture design

According to functional requirements of the forestry resource survey navigation terminal, design the navigation terminal software architecture shown in the Fig.3.

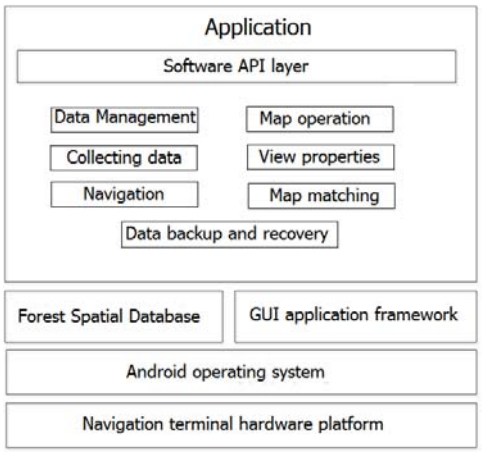


Fig.3. Forestry navigation terminal software frame

(2) The technical route

Based on the Android development, adopting the hypergraph mobile GIS platform for secondary development and management of spatial data, using SQlite as the local database, use the

interact with the server such as GPSR/WCDMA, Beidou navigation module is adopted to improve the navigation, positioning, such as data acquisition, the specific technical route as shown in Fig.4.

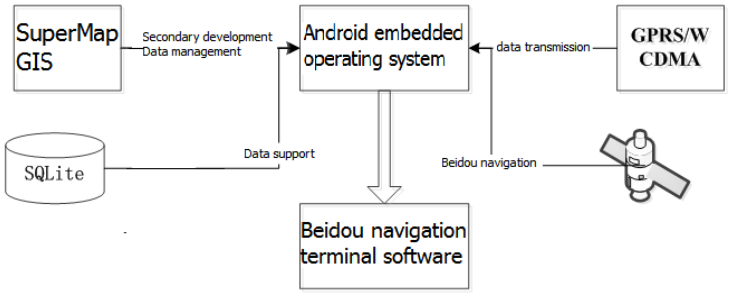


Fig.4. Forestry navigation terminal software technology roadmap

The key technology of Beidou multimode navigation terminal software mainly includes the following aspects.

- Real-time positioning and data transmission technology

Develop Beidou/GPS module drive, realize latitude and longitude information extraction accepted by Beidou/GPS module(Liu., 2012). Then the location information displayed in the mobile terminal.

- Database synchronization technology

After the field investigation, field data, personnel data, tasks, equipment and so forth, are synchronized through artificial net synchronization technology or brake technology (automatic synchronization in a fixed time period) , to ensure the consistency of the data.

- Data backup and recovery technology

Different ways of data management use different methods of data backup, and the common ways are using file data and table data. After the field investigation, it is mainly the file data that is backuped manually on a regular basis using an incremental backup. At the same time, realize rapid recovery of the lost data, to ensure data integrity.

2.2.2 Intelligent monitoring platform software

(1) Architecture design

Develop background intelligent monitoring, scheduling and managing software that matches navigation terminal software, including communication service module, map service module, data service module and monitoring platform and other functional modules. Realize visualization of fieldwork personnel location information, dynamic monitoring of vehicle information, driving record, track playback, operating statistics and query, and other functions(Guo., 2010). The overall structure is shown in Fig.5.

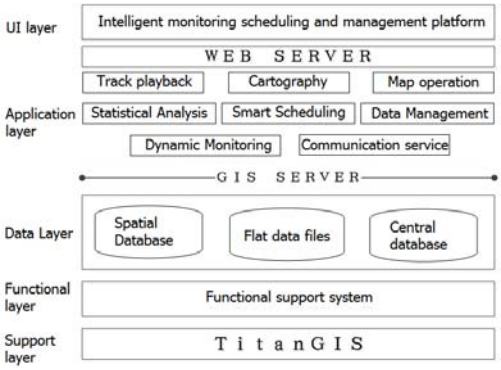


Fig.5. The overall architecture of intelligent monitoring platform software

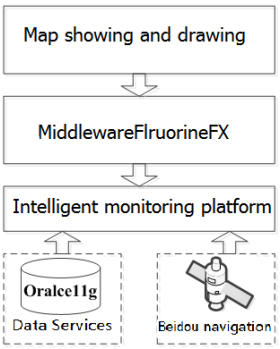


Fig.6. The technology roadmap of intelligent monitoring platform software

(2) Technology roadmap

Intelligent monitoring platform software mainly achieve real-time monitoring, scheduling and managing of field investigation, to provide assistance for decision makers(Sun., 2011). The specific technical route is shown in Fig.6.

The key technology of intelligent monitoring platform software development mainly includes the following several aspects.

- Get real-time location information

The key of the whole technology is to quickly transmit the location information in the database to monitor platform, adopting the way of polling, that front end accesses server every once in a while, look to if there is updated location information, and transmits the updated data to the monitoring platform, enabling real-time display of location information.

- Communication with mobile terminals

In order to make the monitoring platform more human, with the method of server for message delivery, the service side use XMPP agreement, can push some of the information to the mobile terminal in real time. At the same time, the mobile terminal receives the information real-time, and display in the notification bar, such as text messages to the mobile terminal (e.g., scheduling information, weather information, etc.).

3 Summary

The application of Beidou multimode navigation technology to forestry informationization, will promote large-scale industrialization and application of Beidou system, and will led chips, devices, terminals, software integration and a series of high-tech industries to develop. Combined with GPS, Android-based software development will greatly improve the accuracy of satellite positioning, promote the Beidou into the civilian market, and help Beidou navigation system as a new product go into the international market.

Reference

[1] Chen Gaiying. GPS Technology and Its Application in Forestry [J]. Beijing Agricultural College, 2004,19 (3): 45-46.

[2] Guo Hongzhi. Android application development Comments [M]. Beijing: Electronic Industry Press, 2010.

[3] Liu Shengqian, Chen Liding. Android-based vehicle navigation system design and implementation [J]. Thematic studies, 2012 (4): 1-4.

[4] Nong Liping, Wang Lihu, Huang Yiping. Application of Android in embedded vehicle navigation system [J]. Computer Engineering and Design, 2010,3 (11): 73-76.
[5] Sun Weifeng. Development of logistics information platform for forestry based on mobile terminal [J]. Fujian computer, 2011 (6): 28-29.

Research and application of intelligent processing technology of remote sensing data

Jinsheng Ni, Xue Yang, Ting Liu, Wenting Ma

Beijing Aerospace TITAN Technology Co.,LTD

Abstract: In this paper, cloud computing, cloud storage and cloud services and other technologies are used in the intelligent processing of remote sensing data, combined with the professional knowledge and experience of forest fire prevention, establishing the forest fire monitoring and control system.The system has completed the rapid distribution of forest fire information at all levels, real time management of forest fire, real time monitoring and data updating of forest climate dynamics, forest fire prevention and fighting command and other functions, realizing the digital, network, visualization and automation of the forestry fire prevention work, and the standard management and information sharing. It has important significance for the sustainable development of the whole society.

Keywords: Remote sensing data, Cloud computing, Cloud storage, Cloud services, Forest fire monitoring

I. Introduction

With the continuous improvement of space infrastructure and technology development, the number of satellites in orbit around the world has exceeded one thousand. The amount of remote sensing image data show explosive growth, has reached PB level. With the emergence of various sensors, the spectral resolution, spatial resolution and temporal resolution of the satellite are greatly improved. In recent years, the demand for space applications is increasing, and the requirements for acquiring and processing technology of remote sensing information are also increasing, At present, most of the mainstream remote sensing data processing softwares are based on the single machine processing, do not support the parallel data processing, and data throughput, processing precision and automation level is not very high, can not meet demand for large-scale production of spatial information application.

China's forest area is 208000000 hectares, the forest coverage rate is 21.63%, which is below the global average level of 31%, the per capita forest area is only 1/4 l of the global level, per capita forest volume is only the 1/7 of world's per capita level, the total amount of forest resources is relatively low, the quality is not high and the distribution of the situation has not been fundamentally changed [1].In recent years, the system of forest fire prevention function has been relatively mature, which demonstrated a strong function in data acquisition and input, spatial data analysis and processing and data output, etc[2].The technology has been achieved the practical application using the radar monitoring, monitoring of laser, infrared monitoring and satellite remote sensing monitoring, such as laser radar multi spectral satellite forest fire prevention system for detecting forest fire and environmental resources satellite monitoring system[7,8].

In this paper, we apply cloud computing, cloud storage and cloud services to remote sensing data processing, combined with the professional knowledge and experience of forest fire prevention and control, establishing the forest fire monitoring and control system. The system realizes a series

of functions, including the quick circulation of information of forest fire prevention office at all levels, real time management of forest fire safety, real time monitoring and data updating of forest climate dynamics, forest fire prevention and control, reslizing the digital, network, visualization and automation of the forestry fire prevention work, and the standard management and sharing of information. The system has important significance for the sustainable development of the whole society.

II. Parallel cloud computing processing technology of remote sensing data

The basic framework of remote sensing data cloud computing is a kind of effective method to solve the problem of massive remote sensing data using multiple computing resources, and it is an effective method to improve the speed and processing power of the computer. It uses multiple processors to solve the same problem, and the problem is solved by decomposing the problem into several parts, each department has a separate processor to parallel computing. Remote sensing data parallel cloud computing process flow show as figure 1.

The parallel computing system can be a super computer specially designed, containing multiple processors, or a cluster of independent computers interconnected in some way. Through the parallel computing cluster to complete the data processing, and then return the results to the user.

Parallel processing generally has two kinds of forms: Functional parallelism and data parallelism. Functional parallelism is divide a computational task into different parts according to different functions, such as data input, large scale computing, results output, information dissemination, etc.,then these parts are then assigned to different processing units at the same time. Data parallelism is divide the raw data into small pieces, and then to divide the processing units in a balanced way. In specific applications, the need for comprehensive consideration.

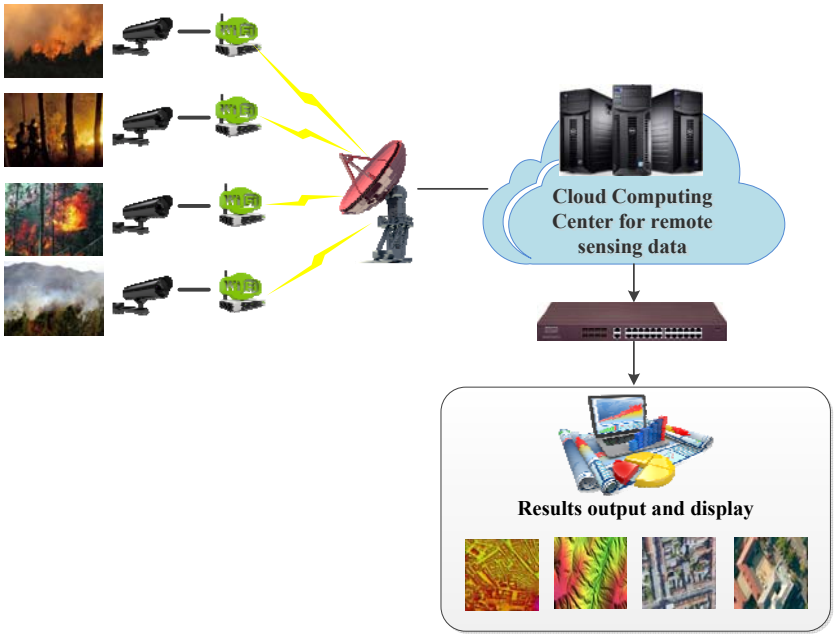


Fig 1 Parallel cloud computing processing flow

The information transmission between the remote sensing satellite data and the parallel algorithm is used in the network parallel computing environment which conclude peer to peer model and master slave mode.Peer to peer model is relative to massive remote sensing data, each of the computer processes processing units generally do the same work. Master slave mode is relative to

parallel processing of remote sensing algorithm and satellite data, parallel processing algorithm between master slave processes is relative to function parallel, Therefore, master slave mode can be regarded as the combination of data parallel and function parallel.

In the process of remote sensing data parallel computing, due to task scheduling monitoring and data transmission between processes, in the parallel processing of remote sensing data, the image data is partitioned and mapped to the process of parallel computing,

It will take different methods to divide in different remote sensing image data processing level. For pixel level processing, the parallel spending is small, data partitioning is the most simple,the dividing method can be selected according to the circumstances;For feature level processing such as line processing, the data can be divided into horizontal strips or vertical strips; In the feature level and target level, some processing methods such as the regional processing is determined by the specific issues and operating environment. On the whole, in order to ensure the premise of parallel, it must make the data communication between the parts is the smallest.

In the process of executing a given algorithm in parallel system, some calculation processes must be accomplished after other processes, at this time synchronization is necessary. Synchronization also has two aspects: on the one hand is to achieve synchronization has to cost time; on the other hand, some processors may be idle, waiting for the message to be calculated.

Three division methods of remote sensing data showed as Figure 2, (a) deviding the image data by the horizontal bar, (b) deviding the image data by the numerical, (c) deviding the image data by rectangular block.

(a) (b) (c)

Fig2 Three division methods of remote sensing data

III. Cloud storage model construction and resource management of large remote sensing data

According to the characteristics of remote sensing data, constructing the storage model of big remote sensing data, combining the big data processing platform and computer storage devices, completing the transformation of computer storage devices to storage services through big data processing platform. Building a remote sening database in cloud environment, the technical architecture of the database is a large image data block, distributed storage and multiple copies, which can automatically adjust the image data and the required computer storage resources according to need, using the redundant storage to ensure reliability of image data and efficiency of access processing. Realizing the multi - scale remote sensing data integrated management, sharing and service through research on the management technology of remote sensing cloud.

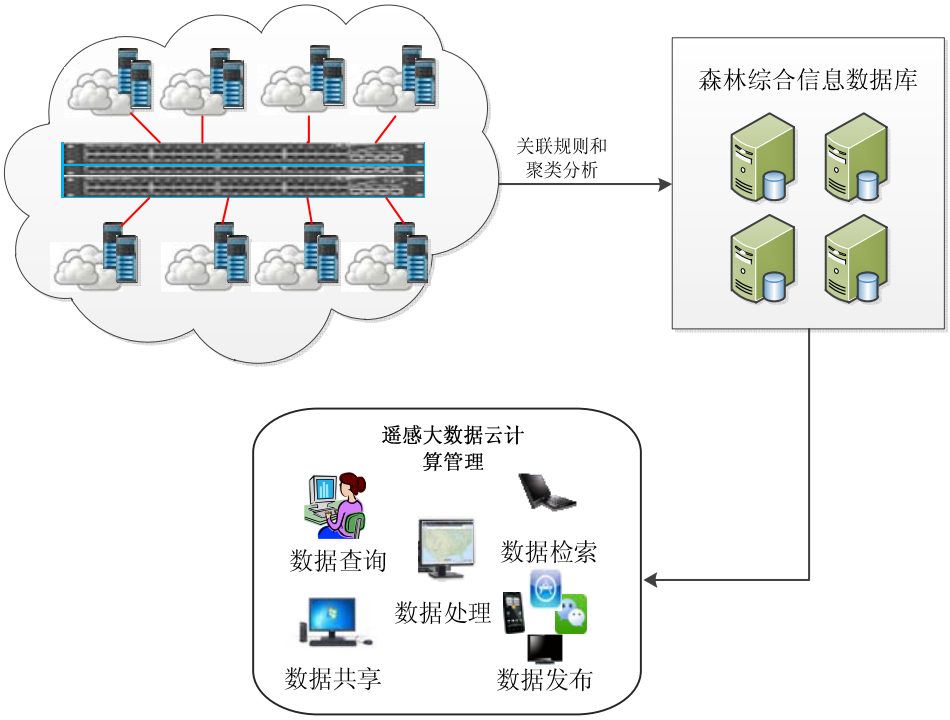


Fig 3 Cloud storage model construction and resource management of large remote sensing data

IV. The application of remote sensing data intelligent processing technology in forest fire monitoring.

Through the research on the remote sensing data intelligent processing technology, changing the previous remote sensing image data storage and processing technology and service pattern, combining positioning system (GPS) technology, satellite remote sensing (RS) technology, modern communication technology, computer network technology, multimedia database technology, real-time monitoring technology, display control technology and high technology in the system, developing a variety of software including fireworks intelligent identification software, fireworks spread the analysis of software, geographic information software, forest fire fighting command management system, forest resource management software, forestry 3D terrain simulation display platform software, ranger position monitoring platform software.

The system improvethe scientific and technological level of firefighting command, achieving the requirements for fireplace quick positioning, rapid deployment command, rapid reporting and convenient, accurate and intuitive management. Its establishment takes forest fire prevention work to a new stage, its use will make the forest fire prevention work change from the traditional experience based management into automation, standardization, standardization of quantitative management, greatly improve the efficiency and modernization level of forest fire management, and further improve the scientific nature and rationality of forest fire prevention.

The combination of traditional video surveillance and forest fire identification has realized unattended in the forest; The combination of forest fire identification and geographic information system (GIS) has achieved precise positioning of fireworks; The balancing combination of traditional video monitoring and network loading has ensured the effective transmission of video and system stability;The combination of video and video viewing can realized the function of play back.

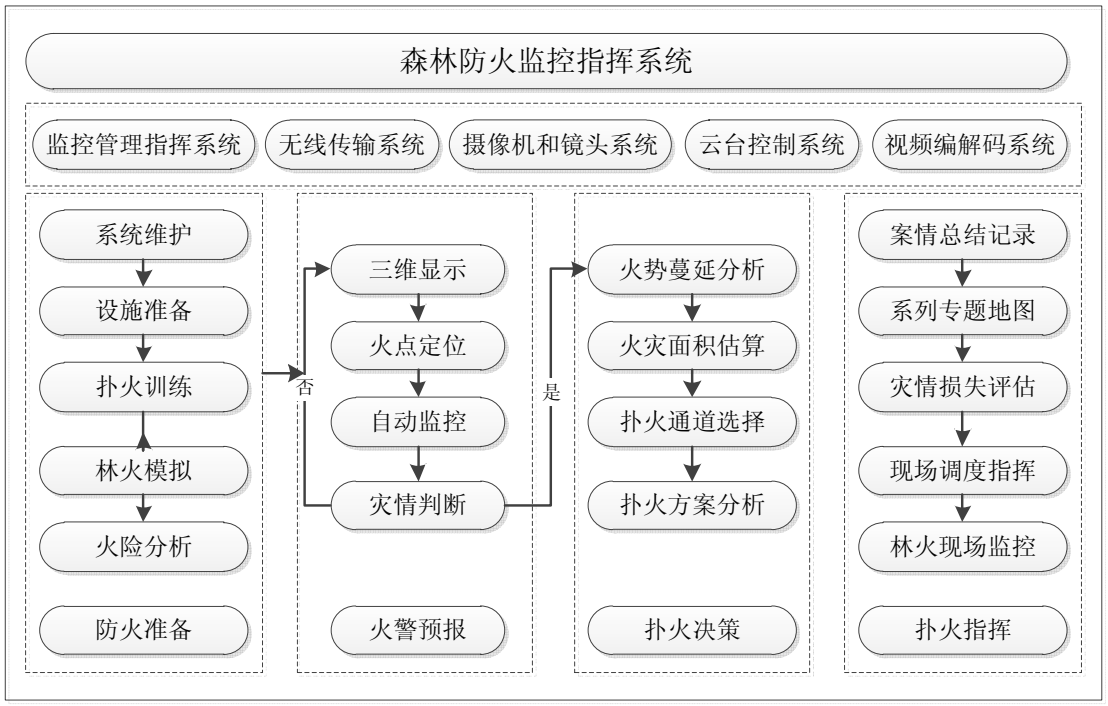


Fig 4 The system of forest fire monitering

V. Conclusion

In this paper, we studied the technology of remote sensing data processing, integrated all kinds of remote sensing information and technology resources,to provide remote sensing applications by using the Internet in a way as customer needed, including remote sensing data cloud storage, cloud processing, remote sensing application cloud services and other functional modules, and to provide the integration and one-stop service of remote sensing data, information, software and computing resources. The development of computer technology has brought a broader space for remote sensing data processing and service, especially under the rapid development of cloud computing technology, people are constantly exploring the combination of remote sensing and cloud computing technology and exploring the advantages of cloud computing technology can bring to remote sensing technology[8-12].The introduction of cloud computing technology has made great changes n the forest fire video monitoring technology.The factors such as the use of network video monitoring technology and wireless networking technology, the vigorous development of Forestry Geographic Information System, the ranger's cultural quality increased year by year have made the digital monitoring equipment to gradually replace the old analog monitoring equipment, and the modernization of forest fire GIS system more and more perfect. Through real-time analysis of monitoring data, monitoring and forecasting accuracy has been greatly improved, the number of forest fires and the affected area have been declining, effectively promoting the economic development and ecological environment construction.

References

[1] The eighth national forest resources inventory results (2009-2013),2014.

[2] Qiang Wu, Hua Xu, Wanfang Zhou. Development of a 3D GIS and its Application to Karst Areas. Environmental geology. 2008, 54 (5): 1037-1045.

[3] Hai Ren, Xiangfeng Gu, Changhai Gao. New technology and equipment for forest fire monitoring in Europe and America. Forestry Machinery & Woodworking Equipment.2005, 33 (5): 8-11.
[4] D. Llewellyn-Jones, M. C. Edwards, C. T. Mutlow, A. R. Birks, L J. Barton, H. Tait.AATSR: Global-Change and Surface-Temperature Measurements From Envisat. ESA bulletin. 2001, 105: 11-21.
[5] Qijian Lin. Introduction to forest fire: University of science and technology of China press, 2003.
[6] Yuanyuan Wang. Research on automatic posititioning techniques for forest fire video monitoring system based on GIS. Master Candidate Thesis of Beijing Forestry University.2008:93.
[7] Quanshan He, Jietai Mao. Micro-Pulse Lidar and Its Applications. Meteorological Science and Technology.2004, 32 (4): 219-224.
[8] Zhihe Yang, Xuhuai Hu, Guanqi Guo. Research on massive remote sensing data processing architecture and parallel computing algorithm. Computer Applications and Software,2014, 31(4):281−284.
[9] Yinghui Zhao, Congfeng Jiang. Research on High Performance Parallel Processing Technology for Remote Sensing Images. Computer Technology and Development,2014,24(7):201−205.
[10] Tianxin Fu, Zhengjun Liu, Haowen Yan. Remote sensing retrieval method for biomass based on MapReduce parallel model. Journal of Arid Land Resources and Environment,2013,27(1):130-136.
[11] Xiantao Li, Zhi Zeng, Feng Zhang, Renyi Liu. Cluster-based ocean remote sensing image fusion parallel computing strategy. Computer Applications and Software,2012,29(1): 84-87.
[12] Haibin Ai, Jianqing Zhang. Distributed orthorectification of high resolution satellite imagery based on middleware. Science of Surveying and Mapping,2009,34(4):158-160.

Error Analysis and Experiment Validation of Height Retrieval for the Interferometric Imaging Radar Altimeter

Li Yin-wei, Lu Hu-lin, Wei Wei-wei, Wang Hai-tao

（Shanghai Radio Equipment Research Institute, Shanghai 200090, China）

Abstract: The accurate and generally applicable interferometric imaging radar altimeter height reconstruction geometry model is established. Then, based on the model, interferometric sensitivity equations of height retrieval error are deduced. And the influence of all the interferometric parameters on height retrieval error is studied using the quantitative simulation analysis method. Finally, the amplitude images, interferometric coefficient images and interferometric phase images of lands are obtained by the way of processing airborne experimental data. After interferometric calibration, height retrieval precision of topography maps is up to 1.5 m. The validity of quantitative simulation analysis is demonstrated.

Keywords: Interferometric imaging；Radar altimeter；Height retrieval；Topography maps

1. Introduction

Radar altimeter is an important microwave remote sensing equipment, whose first-level product measured includes sea level, billow altitude, and anemometry [1]. The retrieval results from the first-level product of radar altimeter could be applied in oceanic geophysics, oceanic dynamics, oceanic climate and environment, sea ice surveillance etc. It supplies an important means of surveillance and detection of the change such as sea level and sea bed, abnormal change of sea level in extreme condition; also it supplies high definition ocean gravity field and seabed terrain. And thus radar altimeter means a great significance not only in military application, but also in civilian application.

Since the launch of Skylab, the first radar altimeter carrier in space, the radar altimeter has experienced a great overall advancement in measurement technology, especially in practicability, technology, and metrical accuracy. Up to now the countries of America and Europe have launched a series of radar altimeter satellites[2], such as Geosat、ERS-1、Topex/Poseidon、ERS-2、GFO、Jason-1、Envisat、Cryosat、Jason-2 etc.. The metrical accuracy has developed from meter level to centimeter level, from science experimentation to practical application. However, the traditional radar altimeter only measures oceanic parameters of nadir point; moreover, it has shortcomings of narrow swath, low resolution, long period etc., far below the requirements of modern oceanic research work and application for the intervals of space and time of oceanic parameter samples.

Interferometric SAR, a combination of radio interferometry and SAR imaging, obtains three dimensional position information via interferometric phase between radar-echo of two channels [3][4]. It works round-the-clock and in all weather without restriction of weather and sunlight. Therefore, interferometric imaging radar altimeter [5] has advantages of wide swath, high resolution in space and time, high accuracy, which resolves the problem of increment of metering point and decrement of time intervals and actualizes observation of oceanic small-size phenomenon. In addition to above we also could get dynamic observation data of ocean wind field, ocean current

field, wave power spectrum etc., and get more precise data of land elevation above sea-level, off-shore tide, real billow height[6].

Based on the interferometric sensitivity equations of height retrieval error, the paper uses the quantitative simulation analysis method to study the influence of all the interferometric parameters on height retrieval error. Finally, the validity of the quantitative simulation analysis is demonstrated through airborne experimental data.

2. DEM height reconstruction model

Figure 1 shows the DEM height reconstruction geometry of interferometric imaging radar altimeter system. H denotes the height of the master antenna phase centre; h is the height of the ground point F; r is the distance between the master antenna A_m and the ground point F and $\hat{\boldsymbol{r}}$ is the unit vector of the line of sight (LOS); β denotes the initial squint angle. $\boldsymbol{b}$ is the baseline vector from the master antenna to the slave antenna; $\boldsymbol{b}_n$ and $\boldsymbol{b}_v$ are vectors, denoting the cross-track component and the along-track component of the baseline vector, respectively. α is the angle of cross-track component of the baseline with respect to the horizontal plane. In order to derive the DEM height reconstruction model with an arbitrary squint angle value, we define three coordinate systems: sensor coordinate system, platform coordinate system and earth-fixed axis system, shown in figure 1 and denoted by (μ,ξ,η), $(x,y,z)_p$ and $(x,y,z)_e$, respectively.

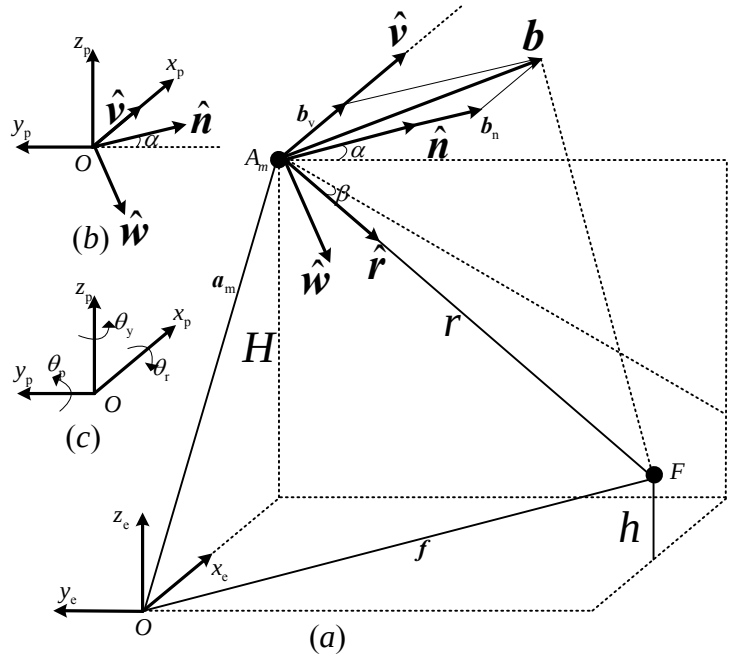


Figure 1. DEM height reconstruction geometry of interferometric imaging radar altimeter system.

In the sensor coordinate system, the unit vector of LOS $\hat{\boldsymbol{r}}$ can be expressed as

$$\hat{\boldsymbol{r}} = \mu\hat{\boldsymbol{v}} + \eta\hat{\boldsymbol{n}} + \zeta\hat{\boldsymbol{w}} \tag{1}$$

where

$$\left.\begin{aligned}
&\mu = <\hat{\boldsymbol{r}},\hat{\boldsymbol{v}}> = \frac{\lambda f_{\mathrm{dc}}}{2v} \\
&\eta = <\hat{\boldsymbol{r}},\hat{\boldsymbol{n}}> = <\hat{\boldsymbol{r}},\frac{\boldsymbol{b}-\boldsymbol{b}_{\mathrm{v}}}{b_{\mathrm{n}}}> = \frac{1}{b_{\mathrm{n}}}<\hat{\boldsymbol{r}},\boldsymbol{b}> - \frac{b_{\mathrm{v}}}{b_{\mathrm{n}}}<\hat{\boldsymbol{r}},\hat{\boldsymbol{v}}> = \frac{b^2}{2rb_{\mathrm{n}}} + s_1 s_2 \frac{\lambda\phi}{2Q\pi b_{\mathrm{n}}} - \frac{\lambda^2\phi^2}{8Q^2\pi^2 rb_{\mathrm{n}}} - \frac{\lambda f_{\mathrm{dc}} b_{\mathrm{v}}}{2vb_{\mathrm{n}}} \\
&\zeta = -s_2\sqrt{1-\mu^2-\eta^2} = -s_2\sqrt{1-\left(\frac{\lambda f_{\mathrm{dc}}}{2v}\right)^2 - \left(\frac{b^2}{2rb_{\mathrm{n}}} + s_1 s_2 \frac{\lambda\phi}{2Q\pi b_{\mathrm{n}}} - \frac{\lambda^2\phi^2}{8Q^2\pi^2 rb_{\mathrm{n}}} - \frac{\lambda f_{\mathrm{dc}} b_{\mathrm{v}}}{2vb_{\mathrm{n}}}\right)^2}
\end{aligned}\right\} \tag{2}$$

Where $<\ ,\ >$ denotes the dot product of vectors; λ is the used wavelength; v is the platform velocity; b is the baseline length; b_{n} and b_{v} denote the cross-track baseline length and the along-track baseline length, respectively; f_{dc} is the Doppler centre frequency; ϕ stands for the absolute interferometric phase of the target point. s_1, s_2, Q are used to denote the operation mode: $s_1 = -1$ if left side-looking, $s_1 = 1$ if right side-looking; $s_2 = -1$ if the master antenna lies in the left, $s_2 = 1$ if the master antenna lies in the right; $Q = 1$ if standard mode, $Q = 2$ if ping-pong mode.

Based on equations (1)-(2), we can obtain that the coordinates of $\hat{\boldsymbol{r}}$ in the sensor coordinate system are as follows:

$$\hat{\boldsymbol{r}}_{\mathrm{vnw}} = \begin{bmatrix} \mu \\ \eta \\ \zeta \end{bmatrix} = \begin{bmatrix} \sin\beta \\ \sin\theta_1 - \dfrac{b_{\mathrm{v}}}{b_{\mathrm{n}}}\sin\beta \\ -s_2\sqrt{1-\sin^2\beta - \left(\sin\theta_1 - \dfrac{b_{\mathrm{v}}}{b_{\mathrm{n}}}\sin\beta\right)^2} \end{bmatrix} \tag{3}$$

where

$$\sin\beta = \lambda f_{\mathrm{dc}}/2v \tag{4}$$

$$\theta_1 = \arcsin\left(\frac{b^2}{2rb_{\mathrm{n}}} + s_1 s_2 \frac{\lambda\phi}{2Q\pi b_{\mathrm{n}}} - \frac{\lambda^2\phi^2}{8Q^2\pi^2 rb_{\mathrm{n}}}\right) \tag{5}$$

Figure 1 shows that the coordinate change from the sensor coordinate system to the earth-fixed axis system can be implemented by the following two steps: Firstly, change the coordinate in the sensor coordinate system to the platform coordinate system through the rotation; secondly, change the coordinate in the platform coordinate system to the earth-fixed axis system through the translation.

So, the position vector $\boldsymbol{f}$ of the target point F in the earth-fixed axis system can be expressed as

$$\boldsymbol{f} = \boldsymbol{a}_{\mathrm{m}} + r\hat{\boldsymbol{r}} = \boldsymbol{a}_{\mathrm{m}} + r \cdot \boldsymbol{\Gamma} \cdot \hat{\boldsymbol{r}}_{\mathrm{vnw}} \tag{6}$$

where $\boldsymbol{a}_{\mathrm{m}}$ represents the position vector of the master antenna A_{m} in the earth-fixed axis system; $\boldsymbol{\Gamma}$ denotes the rotation matrix and has the following expression

$$\boldsymbol{\Gamma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & s_2 \cos\alpha & s_1 s_2 \sin\alpha \\ 0 & -s_1 s_2 \sin\alpha & s_2 \cos\alpha \end{bmatrix} \tag{7}$$

In fact, airplane instability will lead to the difference between the sensor coordinate system and the platform coordinate system. The difference can be expressed by three angles of the platform attitude motion, i.e. roll angle θ_r, pitch angle θ_p and yaw angle θ_y. They are defined as counter-clockwise around the x_p, y_p and z_p axes respectively, as showed in figure 1(c). The rotation matrices of attitude angles are as follows:

$$\boldsymbol{\Gamma}_r = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_r & -\sin\theta_r \\ 0 & \sin\theta_r & \cos\theta_r \end{bmatrix} \quad \boldsymbol{\Gamma}_p = \begin{bmatrix} \cos\theta_p & 0 & -\sin\theta_p \\ 0 & 1 & 0 \\ \sin\theta_p & 0 & \cos\theta_p \end{bmatrix} \quad \boldsymbol{\Gamma}_y = \begin{bmatrix} \cos\theta_y & -\sin\theta_y & 0 \\ \sin\theta_y & \cos\theta_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{8}$$

Therefore, equation (6) becomes

$$\boldsymbol{f} = \boldsymbol{a}_m + r\hat{\boldsymbol{r}} = \boldsymbol{a}_m + r \cdot \boldsymbol{\Gamma}_y \cdot \boldsymbol{\Gamma}_p \cdot \boldsymbol{\Gamma}_r \cdot \boldsymbol{\Gamma} \cdot \hat{\boldsymbol{r}}_{vnw} \tag{9}$$

Substituting equation (3) and equations (7)-(8) into equation (9), we obtain the expression of the target elevation by taking out the height component

$$h = H - r \begin{bmatrix} \cos\theta_p \cos(\alpha - s_1\theta_r)\sqrt{\cos^2\beta - \sin^2\theta_1} \\ -s_1 s_2 \cos\theta_p \sin(\alpha - s_1\theta_r)\sin\theta_1 - \sin\theta_p \sin\beta \end{bmatrix} \tag{10}$$

Equation (10) shows the accurate and generally applicable interferometric imaging radar altimeter DEM height reconstruction geometry. When the initial squint angle and the along-track baseline are equal to zero, equation (10) degrades into the common expression of broadside model. In the paper, we discuss the following situations: $b_v = 0$, $s_1 = 1$ and $s_2 = -1$. Thus, equation (10) becomes

$$h = H - r\cos\theta \tag{11}$$

where

$$\cos\theta = \cos\theta_p \cos(\alpha - \theta_r)\sqrt{\cos^2\beta - \sin^2\theta_1} + \cos\theta_p \sin(\alpha - \theta_r)\sin\theta_1 - \sin\theta_p \sin\beta \tag{12}$$

$$\theta_1 = \arcsin\left(\frac{b}{2r} - \frac{\lambda\phi}{2Q\pi b} - \frac{\lambda^2\phi^2}{8Q^2\pi^2 rb}\right) \tag{13}$$

3. Interferometric sensitivity equations

As shown in equations (11)-(13), the main interferometric parameters affecting the DEM height precision include the platform height H, the slant range r, the baseline length b, the baseline angle α, the interferometric phase ϕ, the initial squint angle β as well as the platform motion attitudes, i.e. the roll angle θ_r, the pitch angle θ_p. Due to the aircraft jitter during the flight, these interferometric parameters are not derived accurately. According to equations (11)-(13), the sensitive relationship between these interferometric parameters and the DEM height can be given by calculating partial derivatives.

$$
\left.\begin{aligned}
&\frac{dh}{dH}=1\\
&\frac{dh}{dr}=-\cos\theta+l\cdot\left(\frac{\lambda^2\phi^2}{8Q^2\pi^2br}-\frac{b}{2r}\right)\\
&\frac{dh}{db}=r\cdot l\cdot\left(\frac{1}{2r}+\frac{\lambda\phi}{2Q\pi b^2}+\frac{\lambda^2\phi^2}{8Q^2\pi^2rb^2}\right)\\
&\frac{dh}{d\alpha}=r\cdot l\cdot\sqrt{\cos^2\beta-\sin^2\theta_{\mathrm{l}}}\\
&\frac{dh}{d\phi}=r\cdot l\cdot\left(-\frac{\lambda}{2Q\pi b}-\frac{\lambda^2\phi}{4Q^2\pi^2rb}\right)\\
&\frac{dh}{d\beta}=r\left[\cos\theta_{\mathrm{p}}\cos\left(\alpha-\theta_{\mathrm{r}}\right)\frac{\sin\beta\cos\beta}{\sqrt{\cos^2\beta-\sin^2\theta_{\mathrm{l}}}}+\sin\theta_{\mathrm{p}}\cos\beta\right]\\
&\frac{dh}{d\theta_{\mathrm{r}}}=-r\cdot l\cdot\sqrt{\cos^2\beta-\sin^2\theta_{\mathrm{l}}}\\
&\frac{dh}{d\theta_{\mathrm{p}}}=r\left[\sin\theta_{\mathrm{p}}\cos\left(\alpha-\theta_{\mathrm{r}}\right)\sqrt{\cos^2\beta-\sin^2\theta_{\mathrm{l}}}-\sin\theta_{\mathrm{p}}\sin\left(\alpha-\theta_{\mathrm{r}}\right)\sin\theta_{\mathrm{l}}+\cos\theta_{\mathrm{p}}\sin\beta\right]
\end{aligned}\right\}\quad(14)
$$

where

$$
l=\cos\theta_{\mathrm{p}}\left(\frac{\cos\left(\alpha-\theta_{\mathrm{r}}\right)\sin\theta_{\mathrm{l}}+\sin\left(\alpha-\theta_{\mathrm{r}}\right)\sqrt{\cos^2\beta-\sin^2\theta_{\mathrm{l}}}}{\sqrt{\cos^2\beta-\sin^2\theta_{\mathrm{l}}}}\right)\quad(15)
$$

Assuming that all of these interferometric parameters sensitive to the DEM height are independent of each other, according to the law of propagation of random errors, the standard variance of DEM height error σ_h can be expressed as

$$
\begin{aligned}
\sigma_h=\sqrt{\left(\frac{dh}{dH}\right)^2\sigma_H^2+\left(\frac{dh}{dr}\right)^2\sigma_\rho^2+\left(\frac{dh}{db}\right)^2\sigma_b^2+\left(\frac{dh}{d\alpha}\right)^2\sigma_\alpha^2+\left(\frac{dh}{d\phi}\right)^2\sigma_\phi^2}\\
\sqrt{+\left(\frac{dh}{d\beta}\right)^2\sigma_\beta^2+\left(\frac{dh}{d\theta_{\mathrm{r}}}\right)^2\sigma_{\theta_{\mathrm{r}}}^2+\left(\frac{dh}{d\theta_{\mathrm{p}}}\right)^2\sigma_{\theta_{\mathrm{p}}}^2}
\end{aligned}\quad(16)
$$

where σ^2 is the variance of parameter.

4. Quantitative simulation analysis

To simplify the quantitative analysis of the DEM height error, we assume that the error factors influencing the DEM height accuracy are independent of each other. Table 1 gives the airborne interferometric imaging radar altimeter system parameters and their accuracy values used in quantitative error analysis.

Figure 2 shows the error relationships between the interferometric parameter (slant range, interferometric phase, baseline length, baseline/roll angle, pitch angle, initial squint angle) error and the DEM height error varying with the look angle. From these figures, we find that with increasing the look angle, the DEM height error due to the slant range decreases, while the DEM height error

increases due to other interferometric parameters except for the pitch angle. Figure 3 shows the height error and its standard variance induced by the order of magnitude of the expected errors listed in table 1 varying with the look angle. It is clear that the height reconstruction accuracy of the system is better than 1.5m.

Table 1. Airborne interferometric imaging radar altimeter system parameters and their accuracy values.

Parameter	Value	Parameter	Value
Band width	200 MHz	Pulse repetition frequency	400 Hz
Wavelength	2.17 cm	Interferometric mode	1
Platform height	2985 m	Platform height accuracy	0.1 m
Slant range	3460 m	Slant range accuracy	0.3 m
Baseline length	0.99 m	Baseline length accuracy	0.3 mm
Baseline angle	0.10 rad	Baseline angle accuracy	0.02°
Initial squint angle	0°	Initial squint angle accuracy	0.01°
Roll angle	-0.41°	Roll angle accuracy	0.02°
Pitch angle	1.32°	Pitch angle accuracy	0.04°
Target height	150 m		

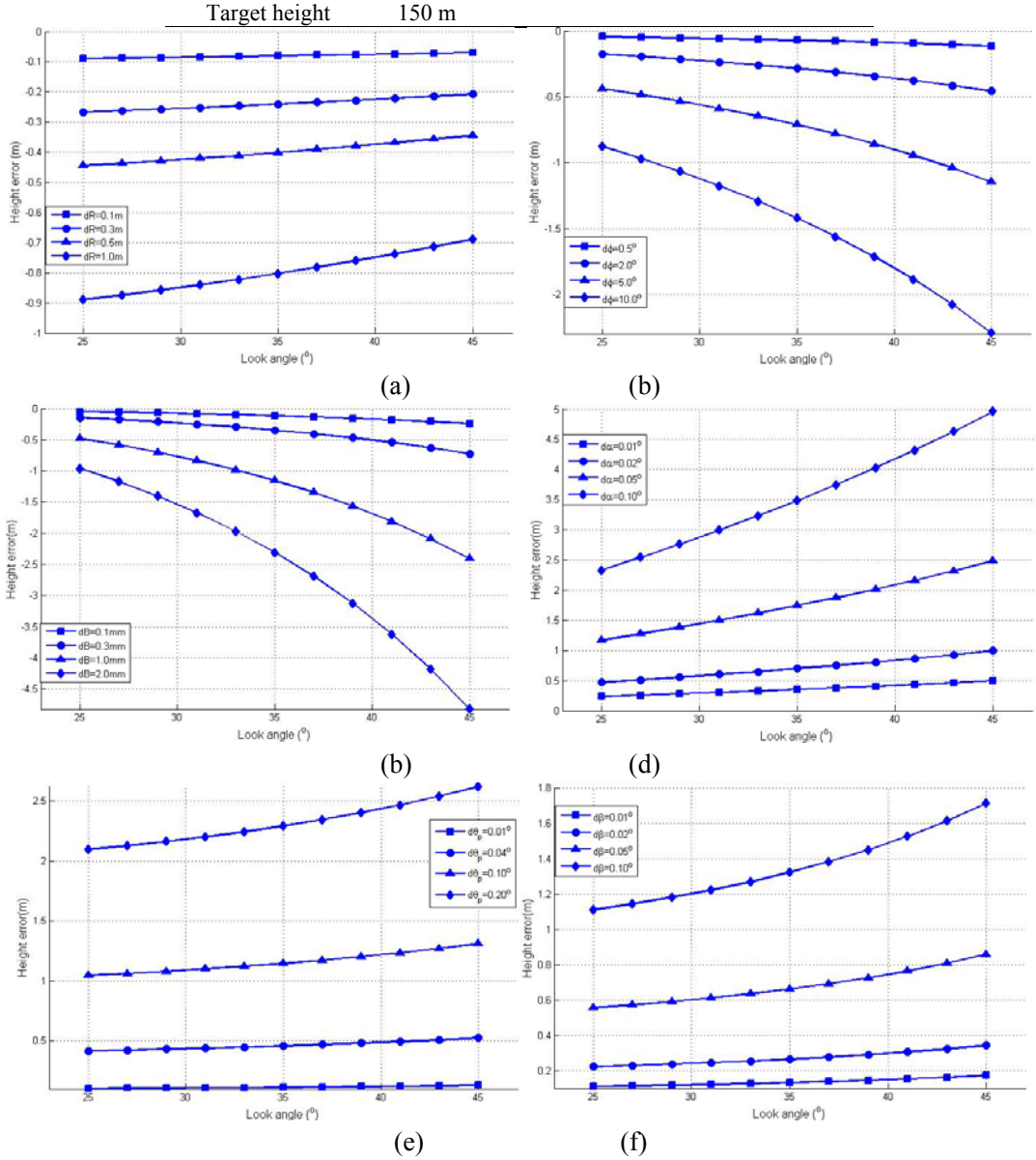


Figure 2. Quantitative relationship between (a) Slant range, (b) Interferometric phase, (c) Baseline length, (d) Baseline\Roll angle, (e) Pitch angle, (f) Initial squint angle and the DEM height error.

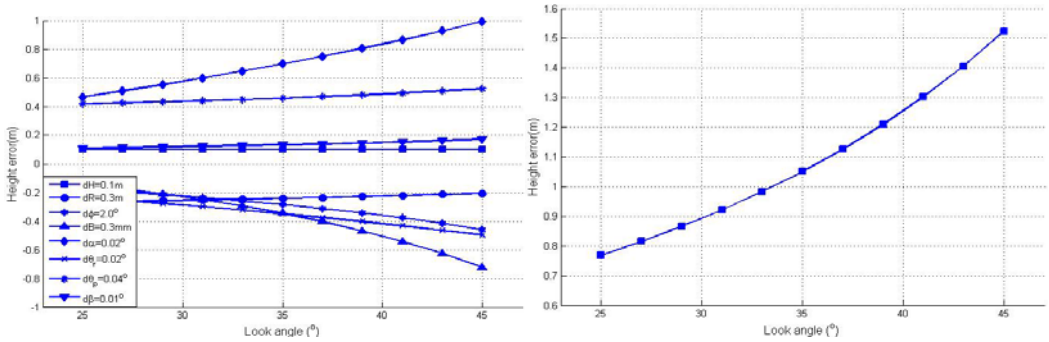
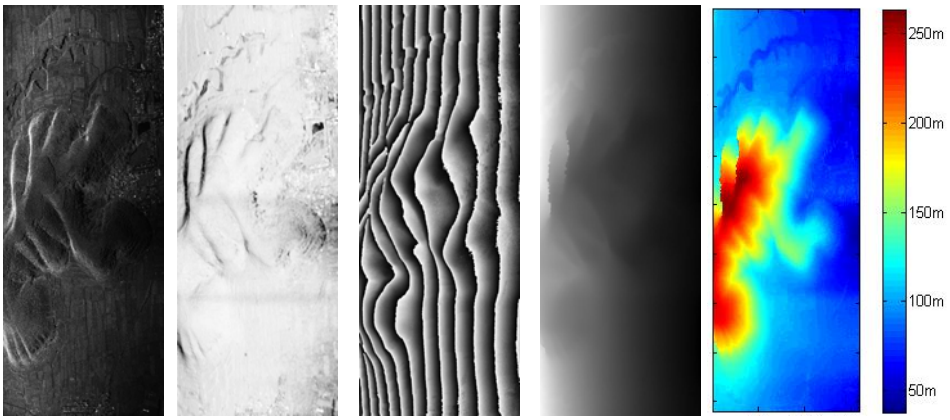


Figure 3. Height error and its standard variance induced by the order of magnitude of the expected errors listed in table 1.

5. Experiment validation

An experiment of airborne prototype was conducted. The test area is located in hilly country, in which corner reflector was set up for interferometric calibration and to validate retrieval accuracy. The detail analysis of experiment result is presented as following.

Figure 4 respectively shows the amplitude, phase and coherent coefficient images of the results obtained through two-dimensional imaging and interferometric processing. Figure 4(a) is the amplitude image, which indicates that the imaging area is mainly plain. Figure 4(b) is the coherent coefficient image of two results. The more correlated that two results are, the brighter the coherent coefficient image will be, of which the pixel value will be closer to 1. Otherwise, the less correlated that two results are, the darker the coherent coefficient image will be, of which the pixel value will be closer to 0. Furthermore, we have the statistical result of coherent coefficient shown in table 2. From table 2, the coherent coefficient larger than 0.8 take 92.01% of the whole, meanwhile its average value is up to 0.90. Figure 4(c) is the filtered interferometric phase and figure 4(d) is unwrapped real interferometric phase. In the center of testing area, we set up 11 corner reflectors, of which five corner reflectors serve as interferometric calibrator and the other 6 corner reflectors for the validation of elevation retrieval accuracy. Figure 4(e) presents DEM of the testing area, which was obtained through interferometric calibration and retrieval based on interferometric geometry of imaging radar altimeter. The real elevations and retrieval elevations of other 6 validation points are shown in table 3. In table 3 the elevation errors within 2m have 1.06m of absolute statistical average value and 0.22m of absolute variance. The results conform to the analysis in aforementioned simulation.



(a) Amplitude image (b) Coherent coefficient (c) Interferometric phase (d) Unwrapped interferometric phase (e) DEM height

Figure 4 Results after two-dimensional imaging and interferometric processing

Table 2 Statistical results of correlation coefficient

Coherent coefficient	>0.5	>0.6	>0.7	>0.8	>0.9	Average value
Distribution (%)	98.84	98.16	96.51	92.01	67.51	0.90

Table 3 Results of elevation retrieval

	Point 1	Point 2	Point 3	Point 4	Point 5	Point 6
Real elevation (m)	72.27	79.47	82.37	84.10	105.47	107.89
Retrieved elevation (m)	71.46	80.66	80.96	85.04	104.35	106.99
Error (m)	-0.81	1.19	-1.41	0.94	-1.12	-0.90

6. Conclusions

This paper first establishes the accurate and general applicable interferometric imaging radar altimeter height reconstruction geometry model. Based on the model, interferometric sensitivity equations of height retrieval error are deduced and the influence of all the interferometric parameters on height retrieval error is studied using the quantitative simulation analysis method. Finally, the amplitude images, interferometric coefficient images and interferometric phase images of lands are obtained via the way of processing the airborne experimental data. The validity of quantitative analysis is demonstrated. The high accuracy interferometric imaging radar altimetry performance validation and DEM retrieval results of land observation pave a feasible technical path for oceanic interferometric imaging radar altimetry and oceanic environment retrieval research.

Acknowledgments

This work was supported by the Shanghai Natural Science Foundation Program (Grant 15ZR1439500).

Reference

[1] Iain H. Woodhouse. Introduction to microwave remote sensing [M]. Beijing: Publishing House of Science, 2014.

[2] Cai Yu-lin, Cheng Xiao, Sun Guo-qing. A Review of Development of Radar Altimeter and its Applications [J]. Remote Sensing Information, 2006, (4): 74-78.

[3] L C Graham. Synthetic Interferometric Radar for Topographic Mapping [J]. IEEE Proceedings, 1974, 62(6): 763-768.

[4] Paul A Rosen, et al. Synthetic Aperture Radar Interferometry [J]. IEEE Proceedings, 2000, 88(3): 333-382.

[5] National Research Council. Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond [R]. Washington, DC: National Academies Press, 2007.

[6] Zhang Rong, Yang Jing-song, Huang Wei-gen, Zhang Hua-guo, Chen Xiao-yan, WANG Juan. Imaging Mechanism and Possible Oceanographic Applications of China Imaging Altimeter [J]. JOURNAL OF MARINE SCIENCES, 2009, 27(3): 99-102.

An Experimental Study on Airborne Polarimetric SAR Data Calibration

Peng Wang[1,2,3], Lei Yang[1,2], Zhi Tan[1,2]

(1) Shanghai Radio Equipment Research Institute, Shanghai 200090, China
(2) Shanghai Academy of Spaceflight Technology, Shanghai, China
(3) Email: wangpeng84@outlook.com

Abstract: An experiment of polarimetric SAR calibration was conducted. The SAR data was acquired using an airborne Ku-band full polarimetric SAR system with single transmitter and single receiver. A polarimetric SAR data calibration procedure including phase calibration, cross-talk calibration, and channel imbalance calibration was applied to the data. Qualitative and quantitative analysis was applied to the calibration results. Quantitative analysis of the backscattering from the trihedral corner reflectors demonstrates that, after calibration, the imbalance between co-polarized channels is reduced into less than 1 dB, while the cross-talk is basically reduced to less than -20dB. After calibration, the polarimetric signatures of these trihedral corner reflectors are closer to the ideal trihedral corner reflector. Freeman-Durden decomposition result of the calibrated data shows that the dominant polarimetric scattering mechanisms are consistent with the land cover, which demonstrates the effectiveness of the polarimetric calibration method.
Keywords: polarimetric SAR calibration, phase calibration, cross-talk, channel imbalance.

I. Introduction

Polarimetric SAR acquires polarimetric information of targets by alternately transmitting and simultaneously receiving of two mutually orthogonal polarimetric waves through a pair of mutually orthogonal polarized antennas. The complete polarimetric information is contained within the four channels of polarimetric SAR data, i.e., two co-polarized channels and two cross-polarized channels. In most cases, polarimetric information is extracted from the relative amplitude as well as the phase difference of these four data channels. Thus the polarimetric SAR data must be calibrated before application.

An airborne polarimetric SAR calibration experiment was conducted in Shandong Province, China, in December 2014. After some pre-processing, such as SAR imaging and removal of spatially varying phase error that caused by displacement between the phase centers of the vertical and horizontal antennas[1], we applied polarimetric SAR data calibration to the acquired SLC (single look complex) data, including phase calibration[2], cross-talk calibration, and channel imbalance calibration[3]. Qualitative and quantitative analysis was applied to the calibration results.

II. Polarimetric SAR Data Calibration Method

Polarimetric SAR data calibration principally consist of phase calibration, cross-talk calibration and channel imbalance calibration.

A. Phase calibration

Assuming that the true polarimetric back scattering matrix for a resolution element is:

$$\boldsymbol{S}=\begin{bmatrix} S_{\mathrm{HH}} & S_{\mathrm{HV}} \\ S_{\mathrm{VH}} & S_{\mathrm{VV}} \end{bmatrix} \tag{1}$$

Due to path difference in the radar hardware, the backscattering matrix actually be measured is

$$\boldsymbol{S}'=\begin{bmatrix} S'_{\mathrm{HH}} & S'_{\mathrm{HV}} \\ S'_{\mathrm{VH}} & S'_{\mathrm{VV}} \end{bmatrix}=\begin{bmatrix} S_{\mathrm{HH}}\exp\left(\mathrm{j}\left(\phi_{\mathrm{T,H}}+\phi_{\mathrm{R,H}}\right)\right) & S_{\mathrm{HV}}\exp\left(\mathrm{j}\left(\phi_{\mathrm{T,V}}+\phi_{\mathrm{R,H}}\right)\right) \\ S_{\mathrm{VH}}\exp\left(\mathrm{j}\left(\phi_{\mathrm{T,H}}+\phi_{\mathrm{R,V}}\right)\right) & S_{\mathrm{VV}}\exp\left(\mathrm{j}\left(\phi_{\mathrm{T,V}}+\phi_{\mathrm{R,V}}\right)\right) \end{bmatrix} \tag{2}$$

For simplicity, the phase on S_{VV} is factored out

$$\boldsymbol{S}'=\exp\left(\mathrm{j}\left(\phi_{\mathrm{T,V}}+\phi_{\mathrm{R,V}}\right)\right)\begin{bmatrix} S_{\mathrm{HH}}\exp\left(\mathrm{j}\left(\phi_{\mathrm{T}}+\phi_{\mathrm{R}}\right)\right) & S_{\mathrm{HV}}\exp\left(\mathrm{j}\phi_{\mathrm{R}}\right) \\ S_{\mathrm{VH}}\exp\left(\mathrm{j}\phi_{\mathrm{T}}\right) & S_{\mathrm{VV}} \end{bmatrix} \tag{3}$$

where $\phi_{\mathrm{T}}=\phi_{\mathrm{T,H}}-\phi_{\mathrm{T,V}}$, and $\phi_{\mathrm{R}}=\phi_{\mathrm{R,H}}-\phi_{\mathrm{R,V}}$.

In the case of reciprocal propagation medium, according to the reciprocity theorem, $S_{\mathrm{HV}}=S_{\mathrm{VH}}$, thus the phase of $S^{*}_{\mathrm{HV}}S_{\mathrm{VH}}$ is 0. By averaging the product $S'^{*}_{\mathrm{HV}}S'_{\mathrm{VH}}=S^{*}_{\mathrm{HV}}S_{\mathrm{VH}}\exp\left(\mathrm{j}\left(\phi_{\mathrm{T}}-\phi_{\mathrm{R}}\right)\right)$ over the whole image, we get $\phi_{\mathrm{T}}-\phi_{\mathrm{R}}$. Find a region in the image, where the theoretic phase difference between S_{HV} and S_{VH} is known. Generally select these areas where this phase difference is near zero, for instance, slightly rough dry bare soil surface. $\phi_{\mathrm{T}}+\phi_{\mathrm{R}}$ could be calculated by averaging $S'_{\mathrm{HH}}S'^{*}_{\mathrm{VV}}=S_{\mathrm{HH}}S^{*}_{\mathrm{VV}}\exp\left(\mathrm{j}\left(\phi_{\mathrm{T}}+\phi_{\mathrm{R}}\right)\right)$ over this region. Since $\phi_{\mathrm{T}}-\phi_{\mathrm{R}}$ and $\phi_{\mathrm{T}}+\phi_{\mathrm{R}}$ are known, ϕ_{T} and ϕ_{R} could be deduced immediately. At last phase calibration could be applied to the measured scattering matrix:

$$\boldsymbol{S}=\begin{bmatrix} S'_{\mathrm{HH}}\exp\left(-\mathrm{j}\left(\phi_{\mathrm{T}}+\phi_{\mathrm{R}}\right)\right) & S'_{\mathrm{HV}}\exp\left(-\mathrm{j}\phi_{\mathrm{R}}\right) \\ S'_{\mathrm{VH}}\exp\left(-\mathrm{j}\phi_{\mathrm{T}}\right) & S'_{\mathrm{VV}} \end{bmatrix} \tag{4}$$

B. Cross-Talk and Channel Imbalance Calibration

If the horizontal and vertical channels are not completely isolated, there will be energy leakage between them during wave transmitting and receiving, thus causes cross-talk. Polarimetric channel imbalance is the relative difference of amplitude and phase that caused by the gain difference between different channels during wave transmitting and receiving.

The cross-talk and channel imbalance could be modeled by a two-stage linear process

$$\boldsymbol{M}=\boldsymbol{R}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{T}=\begin{bmatrix} 1 & \delta_2 \\ \delta_1 & f_1 \end{bmatrix}\begin{bmatrix} S_{\mathrm{HH}} & S_{\mathrm{HV}} \\ S_{\mathrm{VH}} & S_{\mathrm{VV}} \end{bmatrix}\begin{bmatrix} 1 & \delta_3 \\ \delta_4 & f_2 \end{bmatrix} \tag{5}$$

Where $\boldsymbol{M}$ is the observed backscattering matrix, $\boldsymbol{S}$ is true backscattering matrix of target, $\boldsymbol{R}$ and $\boldsymbol{T}$ are 2x2 complex matrix describing the amplitude and phase distortion introduced on receive and transmit respectively. δ_{i} denote cross-talk factors, while f_{i} denote channel imbalance factors. If

these cross-talk factors and channel imbalance factors are obtained, the backscattering matrix could be calibrated with $\boldsymbol{S} = (\boldsymbol{R}^{\mathrm{T}})^{-1}\boldsymbol{M}\boldsymbol{T}^{-1}$.

According to [3], assuming that the transmitting and receiving channels are reciprocal, which means $\boldsymbol{R} = \boldsymbol{T}$, i.e., $\delta_1 = \delta_3$, $\delta_2 = \delta_4$, and $f_1 = f_2$. Introduce $f = f_1 = f_2$, formula (5) could be rewritten as

$$\boldsymbol{M} = \boldsymbol{R}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{T} = \begin{bmatrix} 1 & \delta_2/f \\ \delta_1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & f \end{bmatrix}\begin{bmatrix} S_{\mathrm{HH}} & S_{\mathrm{HV}} \\ S_{\mathrm{VH}} & S_{\mathrm{VV}} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & f \end{bmatrix}\begin{bmatrix} 1 & \delta_1 \\ \delta_2/f & 0 \end{bmatrix} \tag{6}$$

Introduce a transition matrix

$$\boldsymbol{Y} = \begin{bmatrix} Y_{\mathrm{HH}} & Y_{\mathrm{HV}} \\ Y_{\mathrm{VH}} & Y_{\mathrm{VV}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & f \end{bmatrix}\begin{bmatrix} S_{\mathrm{HH}} & S_{\mathrm{HV}} \\ S_{\mathrm{VH}} & S_{\mathrm{VV}} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & f \end{bmatrix} \tag{7}$$

It is widely accepted that

(1) The amplitudes of δ_1, δ_2 are much less than 1, i.e., the system has a good isolation performance between horizontal and vertical polarimetric channels.

(2) For random distributed target that satisfies azimuth symmetry, the co-polarization and cross-polarization channels are uncorrelated, i.e., $\langle S_{\mathrm{HH}}^* S_{\mathrm{HV}} \rangle = 0$, and $\langle S_{\mathrm{VV}}^* S_{\mathrm{HV}} \rangle = 0$. The corresponding components of matrix $\boldsymbol{Y}$ also satisfy this condition.

From above mentioned condition we have following recursion:

$$\begin{aligned} \langle Y_{\mathrm{HV}} Y_{\mathrm{HV}}^* \rangle \approx\ & \langle M_{\mathrm{HV}} M_{\mathrm{HV}}^* \rangle + |\delta_1|^2 \langle M_{\mathrm{HH}} M_{\mathrm{HH}}^* \rangle + |\delta_2/f|^2 \langle M_{\mathrm{VV}} M_{\mathrm{VV}}^* \rangle - \delta_1 \langle M_{\mathrm{HH}} M_{\mathrm{HV}}^* \rangle \\ & - \delta_1^* \langle M_{\mathrm{HV}} M_{\mathrm{HH}}^* \rangle - (\delta_2/f) \langle M_{\mathrm{VV}} M_{\mathrm{HV}}^* \rangle - (\delta_2/f)^* \langle M_{\mathrm{HV}} M_{\mathrm{VV}}^* \rangle \\ & + \delta_1 (\delta_2/f)^* \langle M_{\mathrm{HH}} M_{\mathrm{VV}}^* \rangle + \delta_1^* (\delta_2/f) \langle M_{\mathrm{VV}} M_{\mathrm{HH}}^* \rangle \end{aligned} \tag{8}$$

$$\begin{cases} \delta_1 = \dfrac{\Delta P + 4P^* \langle Y_{\mathrm{HV}} Y_{\mathrm{HV}}^* \rangle \langle M_{\mathrm{VV}} M_{\mathrm{HH}}^* \rangle}{\Delta^2 - 16 \langle Y_{\mathrm{HV}} Y_{\mathrm{HV}}^* \rangle^2 \langle M_{\mathrm{VV}} M_{\mathrm{HH}}^* \rangle \langle M_{\mathrm{HH}} M_{\mathrm{VV}}^* \rangle} \\ \dfrac{\delta_2}{f} = \dfrac{\Delta Q + 4Q^* \langle Y_{\mathrm{HV}} Y_{\mathrm{HV}}^* \rangle \langle M_{\mathrm{HH}} M_{\mathrm{VV}}^* \rangle}{\Delta^2 - 16 \langle Y_{\mathrm{HV}} Y_{\mathrm{HV}}^* \rangle^2 \langle M_{\mathrm{VV}} M_{\mathrm{HH}}^* \rangle \langle M_{\mathrm{HH}} M_{\mathrm{VV}}^* \rangle} \end{cases} \tag{9}$$

Where

$$\Delta = \langle M_{\mathrm{HH}} M_{\mathrm{HH}}^* \rangle \langle M_{\mathrm{VV}} M_{\mathrm{VV}}^* \rangle - \langle M_{\mathrm{VV}} M_{\mathrm{HH}}^* \rangle \langle M_{\mathrm{HH}} M_{\mathrm{VV}}^* \rangle - 4 \langle Y_{\mathrm{HV}} Y_{\mathrm{HV}}^* \rangle \langle Y_{\mathrm{HV}} Y_{\mathrm{HV}}^* \rangle$$

$$P = \langle M_{\mathrm{HV}} M_{\mathrm{HH}}^* \rangle \langle M_{\mathrm{VV}} M_{\mathrm{VV}}^* \rangle - \langle M_{\mathrm{VV}} M_{\mathrm{HH}}^* \rangle \langle M_{\mathrm{HV}} M_{\mathrm{VV}}^* \rangle - 2 \langle Y_{\mathrm{HV}} Y_{\mathrm{HV}}^* \rangle \langle M_{\mathrm{VV}} M_{\mathrm{HV}}^* \rangle$$

$$Q = \langle M_{\mathrm{HH}} M_{\mathrm{HH}}^* \rangle \langle M_{\mathrm{HV}} M_{\mathrm{VV}}^* \rangle - \langle M_{\mathrm{HV}} M_{\mathrm{HH}}^* \rangle \langle M_{\mathrm{HH}} M_{\mathrm{VV}}^* \rangle - 2 \langle Y_{\mathrm{HV}} Y_{\mathrm{HV}}^* \rangle \langle M_{\mathrm{HH}} M_{\mathrm{HV}}^* \rangle$$

Start with the initial value $\delta_1 = 0$ and $\delta_2 / f = 0$, and repeat the process of expression (8) and (9) until $\left\langle Y_{\mathrm{HV}} Y_{\mathrm{HV}}^{*} \right\rangle$, δ_1, and δ_2 / f are convergent. Note that channel imbalance factor f varies with range and incident angle, therefore, the statistical averages in above formulas should be calculated from pixels that along with a straight line in the azimuth direction (thus with the same range and incidence angle).

Finally, the channel imbalance factor f could be estimated from the backscattering of the trihedral corner reflectors:

$$|f| = \left| \frac{Y_{\mathrm{VV}}^{*} Y_{\mathrm{VV}}}{Y_{\mathrm{HH}}^{*} Y_{\mathrm{HH}}} \right|^{\frac{1}{4}}, \quad \arg(f) = \frac{1}{2} \arg\left(Y_{\mathrm{HH}}^{*} Y_{\mathrm{VV}} \right) \tag{10}$$

As channel imbalance factor f varies with range, several trihedral corner reflectors should be set along the range direction during data acquiring. After calculate out the factor f from these reflectors, factor f of all other range lines could be obtained through interpolation.

III. Data Acquisition and Processing

The experimental data was acquired around the Pingyin airfield area in Shandong Provience, China, in December 2014, using an airborne Ku-band full polarimetric SAR system. The system resolution is 1 m. The major land covers in this observed scene are airfield runway, road, lake, cropland, bare soil, building, etc. Picture of the observed scene is shown in Fig.1.

Trihedral corner reflectors were set on the concrete airfield runway. Considering that the calibration parameters vary with range, the corner reflectors were placed dispersedly along range direction, which was the direction of runway in this experiment. To avoid interference among corner reflectors, the interval between adjacent reflectors were set above 10 resolution units.

Fig.1 Picture of the observed scene

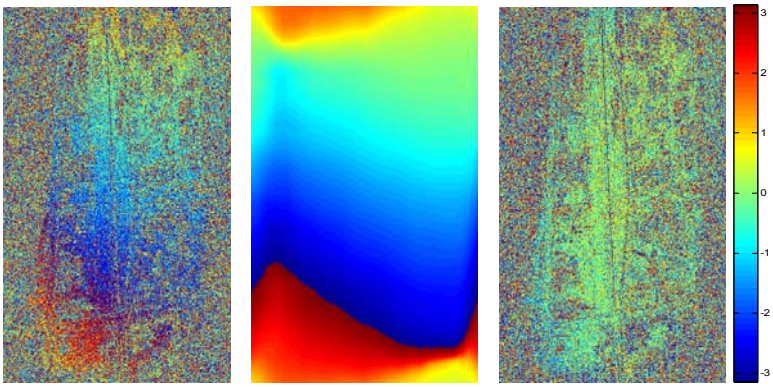
The experimental system is a single transmitter and single receiver polarimetric system, and raw data of four polarimetric channels were alternately recorded by a single channel data collector. Four SLC (single look complex) images were obtained after SAR imaging processing including

matched filtering in the range dimension and dechirp processing in the azimuth dimension, as shown in Fig.2.

(a) (b) (c) (d)

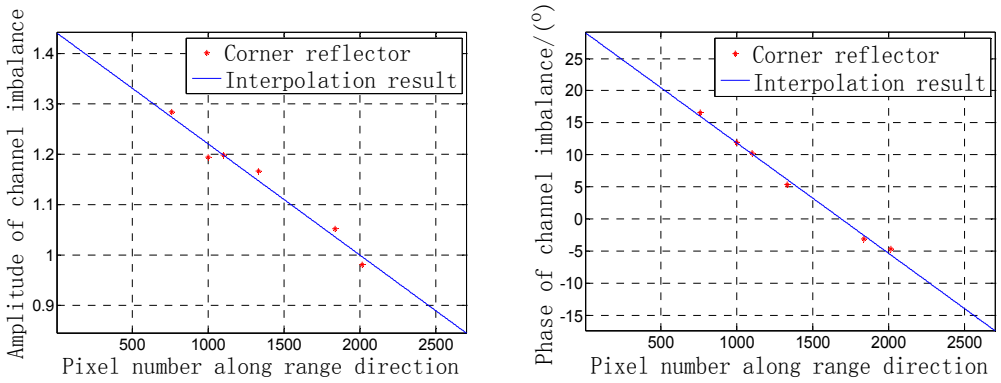
Fig.2 SLC image of each polarimetric channel: (a) HH; (b) HV; (c) VH; (d) VV

Due to displacement between the phase centers of the vertical and horizontal antennas, the Pauli image synthesized using four polarimetric channels shows a spatial varied phase error. Fig.3 (a) shows the phase difference between HH and VV polarization channels, clearly it reveals a phase modulation. This phase error must be eliminated during the data preprocessing before polarimetric SAR calibration. We estimated the phase error through sliding window averaging and then compensated it to the SLC data. As shown in Fig.3 (c), this phase error was eliminated. The phase error between other data channels were eliminated by the same way.



(a) (b) (c)

Fig.3 Removal of the spatial varied phase error. (a) before compensation; (b) the estimated value of phase error; (c) after phase compensation



(a) (b)

Fig.4 The interpolation result of channel imbalance factor: (a) Amplitude; (b) Phase.

After removal of the phase error that caused by displacement between the phase centers of the vertical and horizontal antennas, phase calibration, cross-talk calibration, and channel imbalance calibration that introduced in section 2 were applied to the data. Cross-talk factors calculated from six trihedral corner reflectors and azimuth symmetric distributed targets are $\delta_1 = -21.82\text{dB}\angle -64.23^\circ$ and $\delta_2 = -15.71\text{dB}\angle 35.05^\circ$. Channel imbalance factors corresponding to six trihedral corner reflectors are $2.17\text{dB}\angle 16.60^\circ$, $1.54\text{dB}\angle 11.90^\circ$, $1.57\text{dB}\angle 10.25^\circ$, $1.34\text{dB}\angle 5.37^\circ$, $0.44\text{dB}\angle -3.03^\circ$, $-0.17\text{dB}\angle -4.61^\circ$, respectively. Channel imbalance factors corresponding to each azimuthal line could be derived through interpolation of amplitude and phase of these six factors along the range direction, as shown in Fig.4. Finally the data was calibrated using these cross-talk factors and channel imbalance factors.

Backscattering matrix corresponding to those trihedral corner reflectors before and after calibration are listed in Table 1. As a reference, the backscattering matrix of ideal trihedral corner reflector is also listed. After calibration, the channel imbalance between co-polarization channels are improved to less than 1dB, besides, data amplitude of cross-polarization channels are reduced to around -20dB. Without loss of generality, take the 1st and 6th corner reflectors as example, the co-/cross-polarization signatures of these two reflectors before and after calibration are shown in Fig.5 and Fig.6. After calibration, the polarization signatures are more close to the signatures of ideal corner reflector, which demonstrates the effectiveness of the calibration processing.

Table 1 Backscattering matrix of trihedral corner reflectors before and after calibration

Corner reflector No.	Backscattering matrix of trihedral corner reflectors (ideal value: $\begin{bmatrix} 0\text{dB} & -\infty\text{dB} \\ -\infty\text{dB} & 0\text{dB} \end{bmatrix}$)	
	Before calibration	After calibration
1	$\begin{bmatrix} 0.00\text{dB}\angle 0.00^\circ & -14.09\text{dB}\angle 75.13^\circ \\ -14.10\text{dB}\angle 79.66^\circ & 4.47\text{dB}\angle 44.49^\circ \end{bmatrix}$	$\begin{bmatrix} 0.00\text{dB}\angle 0.00^\circ & -18.10\text{dB}\angle 115.63^\circ \\ -17.35\text{dB}\angle 118.63^\circ & -0.38\text{dB}\angle 1.90^\circ \end{bmatrix}$
2	$\begin{bmatrix} 0.00\text{dB}\angle 0.00^\circ & -18.75\text{dB}\angle 58.72^\circ \\ -19.51\text{dB}\angle 59.07^\circ & 2.80\text{dB}\angle 35.04^\circ \end{bmatrix}$	$\begin{bmatrix} 0.00\text{dB}\angle 0.00^\circ & -20.60\text{dB}\angle 136.95^\circ \\ -20.36\text{dB}\angle 141.63^\circ & -0.81\text{dB}\angle -0.50^\circ \end{bmatrix}$
3	$\begin{bmatrix} 0.00\text{dB}\angle 0.00^\circ & -20.67\text{dB}\angle 49.05^\circ \\ -21.37\text{dB}\angle 48.77^\circ & 2.87\text{dB}\angle 31.80^\circ \end{bmatrix}$	$\begin{bmatrix} 0.00\text{dB}\angle 0.00^\circ & -20.69\text{dB}\angle 152.71^\circ \\ -20.36\text{dB}\angle 155.64^\circ & -0.48\text{dB}\angle -0.76^\circ \end{bmatrix}$
4	$\begin{bmatrix} 0.00\text{dB}\angle 0.00^\circ & -25.01\text{dB}\angle 42.11^\circ \\ -25.69\text{dB}\angle 41.24^\circ & 2.83\text{dB}\angle 22.27^\circ \end{bmatrix}$	$\begin{bmatrix} 0.00\text{dB}\angle 0.00^\circ & -18.24\text{dB}\angle 166.87^\circ \\ -18.06\text{dB}\angle 167.86^\circ & -0.22\text{dB}\angle -3.50^\circ \end{bmatrix}$
5	$\begin{bmatrix} 0.00\text{dB}\angle 0.00^\circ & -16.19\text{dB}\angle -13.36^\circ \\ -15.69\text{dB}\angle -8.80^\circ & 1.03\text{dB}\angle 5.61^\circ \end{bmatrix}$	$\begin{bmatrix} 0.00\text{dB}\angle 0.00^\circ & -25.23\text{dB}\angle -134.33^\circ \\ -28.44\text{dB}\angle -130.64^\circ & -0.14\text{dB}\angle -3.12^\circ \end{bmatrix}$
6	$\begin{bmatrix} 0.00\text{dB}\angle 0.00^\circ & -15.74\text{dB}\angle 22.07^\circ \\ -14.57\text{dB}\angle 22.45^\circ & -0.46\text{dB}\angle 1.85^\circ \end{bmatrix}$	$\begin{bmatrix} 0.00\text{dB}\angle 0.00^\circ & -23.85\text{dB}\angle 104.11^\circ \\ -22.52\text{dB}\angle 86.15^\circ & -0.52\text{dB}\angle 2.29^\circ \end{bmatrix}$

Fig.7 (a) is the color coded image of the Freeman-Durden target decomposition[4] result of the calibrated polarimetric SAR data observed over the calibration field. Fig.7 (b) (c) are the images

acquired from the same flight, and processed through the same procedure as Fig.7 (a). The Freeman-Durden decomposition is a technique for fitting a physically based, three-component scattering mechanism model to the polarimetric SAR observations[5]. These three components are single-bounce scattering, double-bounce scattering, and volume scattering, respectively. In the color coded image, these three components are coded with blue, red, and green color components, respectively. In Fig.7, blue component occupies most part of the image, which denotes single-bounce scattering from the bare ground. Red denotes double-bounce scattering from dihedral structures that composed of buildings, airplanes or other architectures together with ground surface. The formation of volume scattering (coded with green) is relatively complicated. It might be scattered from forest, vegetation, or some complex structures; buildings not aligned along the azimuthal direction also have the characteristics of volume scattering. According to our analysis, the scattering mechanisms of the calibrated data are in accord with the land cover types.

IV. Conclusion

An experiment of airborne polarimetric SAR calibration was conducted. After SAR imaging processing and removal of phase error that caused by displacement between the phase centers of the vertical and horizontal antennas, a polarimetric SAR data calibration procedure including phase calibration, cross-talk calibration, and channel imbalance calibration was applied to the data. Both qualitative and quantitative analysis of the calibrated data demonstrate the effectiveness of the polarimetric calibration method.

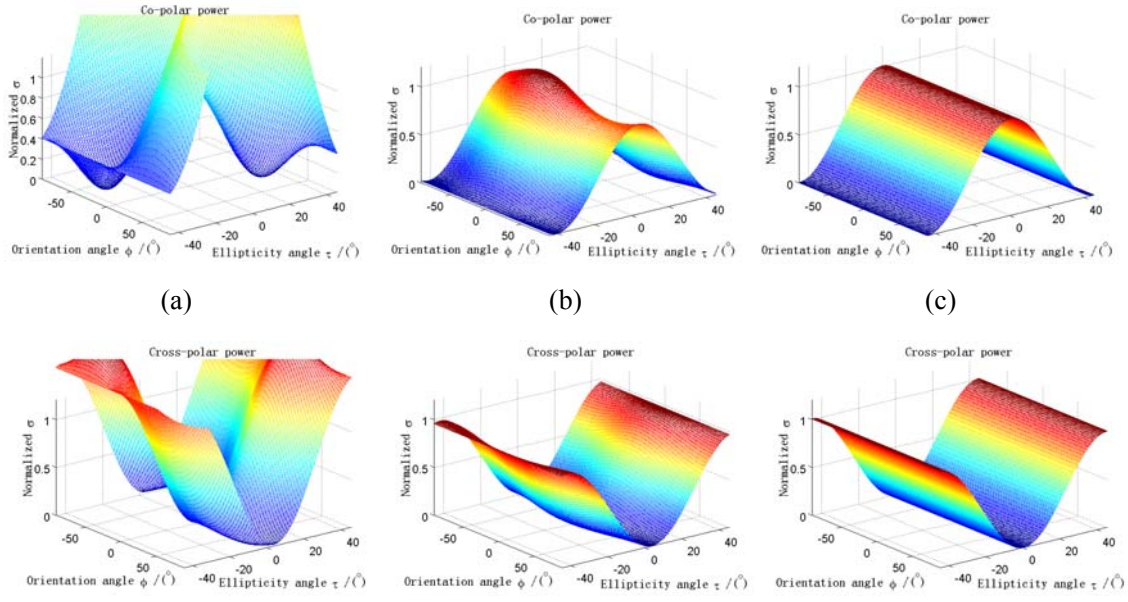
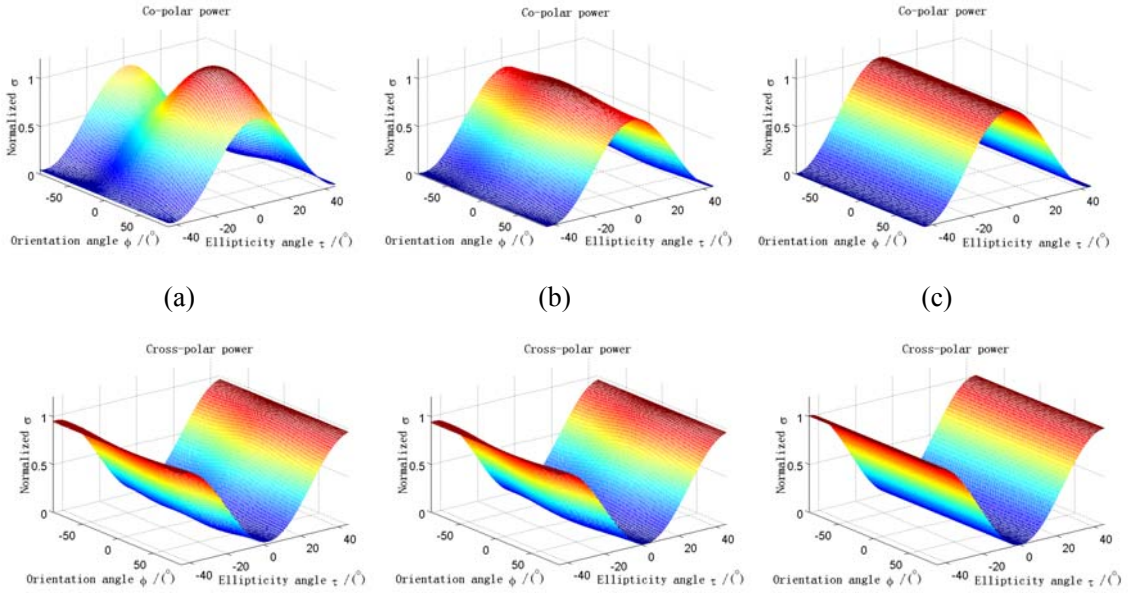


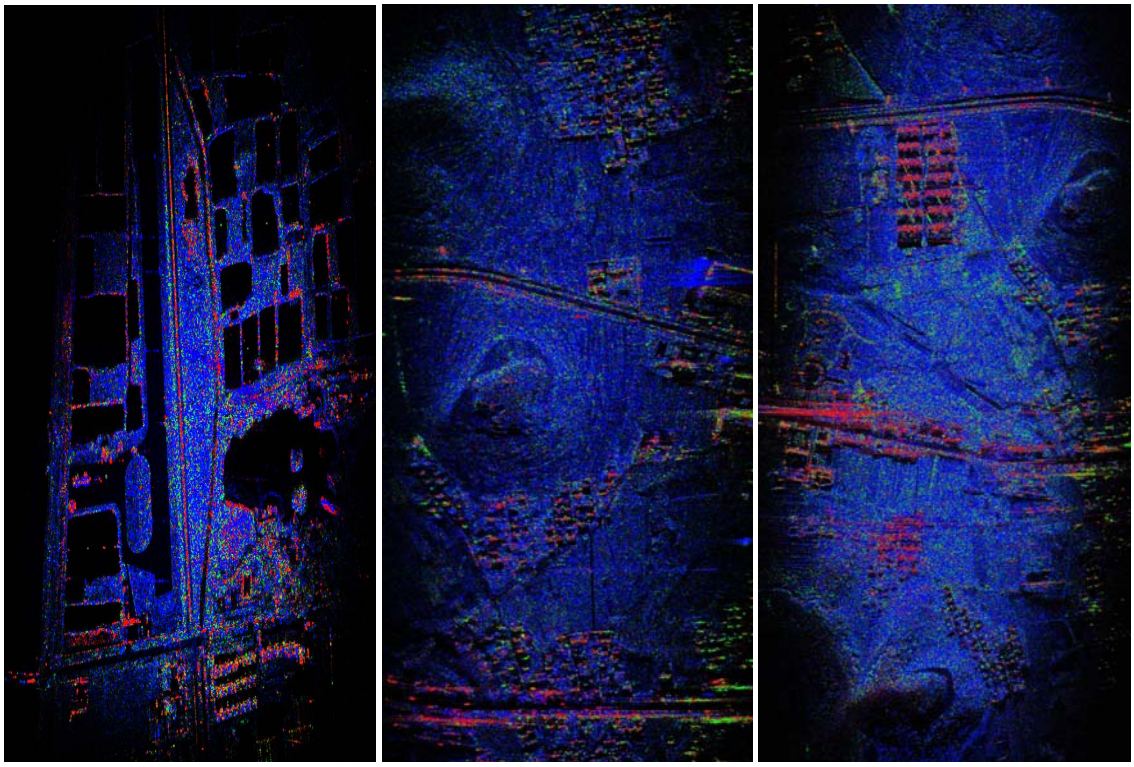
Fig.5 Polarimetric signature of the ideal trihedral corner reflector and the 1st trihedral corner reflector before and after calibration: (a) co-polarization signature before calibration; (b) co-polarization signature after calibration; (c) co-polarization signature of ideal trihedral corner reflector; (d) cross-polarization signature before calibration; (e) cross-polarization signature after calibration; (f) cross-polarization signature of ideal trihedral corner reflector.



(a) (b) (c)

(d) (e) (f)

Fig.6 Polarimetric signature of the ideal trihedral corner reflector and the 6th trihedral corner reflector before and after calibration: (a) co-polarization signature before calibration; (b) co-polarization signature after calibration; (c) co-polarization signature of ideal trihedral corner reflector; (d) cross-polarization signature before calibration; (e) cross-polarization signature after calibration; (f) cross-polarization signature of ideal trihedral corner reflector.



(a) (b) (c)

Fig.7 Color coded image of the Freeman-Durden target decomposition results of calibrated polarimetric SAR data

Reference

[1] D. R. Sheen, A. Freeman, and E. S. Kasischke. Phase Calibration of Polarimetric Radar Images. IEEE Transactions on Geoscience and Remote Sensing, 1989, Vol. 27, No. 6, pp. 719-731.

[2] Zebker, H. A. and Y. L. Lou. Phase calibration of imaging radar polarimeter Stokes Matrices. IEEE Transactions on Geoscience and Remote Sensing, 1990, Vol. 28, No. 2, pp. 246-252.

[3] J. J. van Zyl. Calibration of polarimetric radar images using only image parameters and trihedral corner reflector responses. IEEE Transactions on Geoscience and Remote Sensing, 1990, Vol. 28, No. 3, pp. 337-348.

[4] A. Freeman and S. L. Durden. A Three-component Scattering Model for Polarimetric SAR Data. IEEE Transactions on Geoscience and Remote Sensing, 1998, Vol. 36, pp. 963-973.

[5] J. S. Lee and E. Pottier. Polarimetric radar imaging: from basics to applications. New York: Taylor & Francis Group, 2009.

The Research on the Implementation and Transformation of GPS Technology Patents

Li Ping[2], Zhang Yue[1], Zhao Pengfei[1], Zhou Ziyan[1]

Abstract: GPS technology has been gradually used in transportation, disaster relief, telecommunications and other professional areas and consumption of human lives. Therefore, to promote the implementation of GPS patented technologies and transformation of human society can effectively benefit, for the benefit of the people. However, in the process and transformation of domestic GPS technology patent application, there are still a large gap with foreign countries, the system is imperfect, institutions are not perfect and a series of problems exist, it is necessary to further study the implementation of GPS technology patents and Transformation in China. Through the implementation of the conversion application and the status quo at home and abroad and GPS technology research with the United States examples in GPS technology implementation and conversion applications, analysis of China's GPS technology, especially the "Beidou" system of intellectual property management and implementation of patent transformation, the gap with foreign countries, as well as practical implementation and conversion process problems and obstacles encountered. Finally, the implementation of GPS technology patents and conversion applications next Intellectual Property management and related recommendations, provide a reference for the future transformation of the successful implementation and use of domestic GPS technology patents.

Keywords: patent conversion, GPS, Beidou Navigation Positioning System, Military and civilian general technology

1. GPS technology used in civilian areas

Since the advent of GPS technology, it has fully demonstrated its dominance in navigation, positioning field. Many areas also changed because of the emergence of GPS revolutionary. Currently, almost all of the world need to navigate, locate the user, are GPS precision, all-weather, global coverage, and convenient and flexible to attract. Our GPS application development momentum of rapid, just a few years, GPS application in China has grown from a small number of research institutions and military departments quickly extended to all civilian areas, widely used GPS to change the way people work, improve work efficiency, bring a huge economic benefits. It can be said, GPS applications in our country is unlimited. GPS technology in civilian areas is mainly reflected in the following aspects.

1.1 GPS used to measure

GPS technology for surveying and mapping industry has brought a revolution. Use carrier phase differential technology (RTK), on the basis of real-time processing of two stations on the carrier phase can achieve centimeter-level accuracy. Compared with the traditional manual measurements, GPS technology has enormous advantages: high accuracy; easy operation, equipment is small, easy to carry; all-weather operation; no communication between the observation points of view; measurements unity in WGS84 coordinates, information automatically receive, store, reducing the

tedious process of intermediate links. Current, GPS technology has been widely used in geodesy, resources exploration, crusted movement, cadastral and other fields.

1.2 GPS used in transportation

In the transportation industry to use GPS technology to track vehicles, scheduling management, equitable distribution of vehicles, the fastest response to user's car or send a request to reduce energy consumption and save operating costs. GPS vehicle navigation plays an important role in the establishment of digital traffic radio station in the city, real-time traffic information broadcasts the city, automotive equipment for precise positioning through GPS, combined with electronic maps and real-time traffic conditions and automatically matches the optimal path, and implement autonomous navigation of the vehicle. Civil aviation through the GPS receiving device, so that the driver can be accurately aligned with the runway during landing, while the aircraft compaction, improve airport efficiency, safety guide aircraft approach and departure.

1.3 GPS applied to the Rescue

Rescue applied to rescue Using GPS technology, for fire, ambulance, police emergency dispatch, improving the efficiency of the emergency services to respond to fires, crime scenes, accidents, traffic jams and other emergencies. Special vehicles (such as armored cars), etc., can alarm emergencies, positioning, to limit the damage. With the help of GPS, rescue workers can in inaccessible, harsh conditions of the sea, the mountains, the desert, for the implementation of effective search of missing persons, rescue. Fishing boats equipped with GPS devices, in the event of danger, timely positioning, alarm, making it faster and more timely access to rescue.

1.4 GPS used in agriculture

Currently, developed countries have started to GPS technology into agricultural production, the so-called "precision farming." The method using GPS positioning to obtain field information, including production monitoring, soil samples were collected and other computer systems through the analysis of the data processing, decision of the management practices of agricultural land, the yield and soil status information is loaded with GPS devices spraying device, thereby accurately to fertilize farmland plots, spraying. Through the implementation of precision farming, in the case of try not to cut production, reduce agricultural production costs, avoid waste of resources, reduce pollution caused by fertilizer insecticides on the environment.

1.5 GPS applied in recreation

Recreation applied to entertainment with entertainment miniaturization of GPS receivers and lower prices, GPS gradually into people's daily life, as people travel, adventure is a good helper. Through GPS, one can find in a strange city destination quickly and can be the optimal path travel; campers carry a GPS receiver, can quickly and easily find the right camping place, do not worry about getting lost; even some high-end video games also use the GPS simulation technology.

1.6 GPS applied in navigation

It is mainly for positioning and navigation of ships, automobiles, aircraft and other moving objects. For example: In the ship into port ocean navigation and water diversion, route guidance and approach landing aircraft, automobile independent navigation, ground vehicle tracking and urban intelligent traffic management, emergency rescue, personal travel and outdoor adventure, personal communications terminal (cell phone, PDA, electronic maps are integrated) and the like.

1.7 GPS frequency applied to time service

GPS time can be used for electric power, telecommunications, communications and other network synchronization, timing the exact time and the exact frequency of the award.

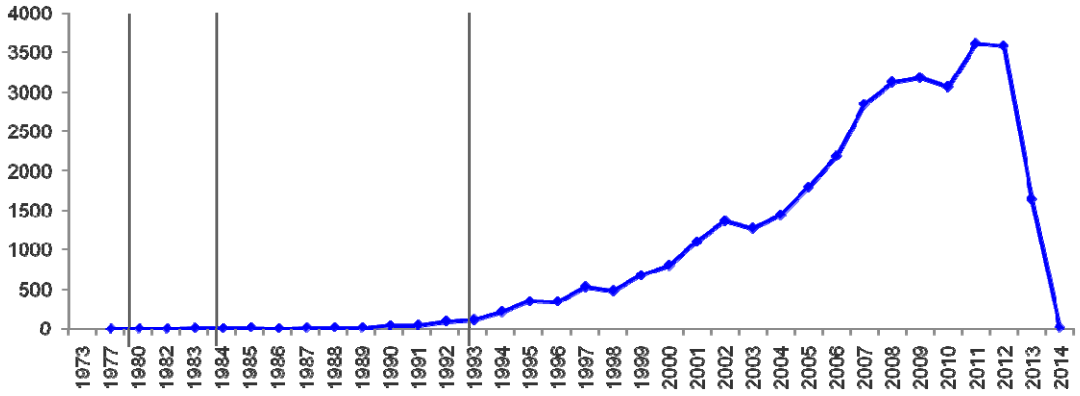
1.8 GPS applied to precision measurement

GPS can be used for geodetic control measurements, stakeout road and the various lines, underwater topography, crusted deformation measurement, dams and large-scale building deformation monitoring.

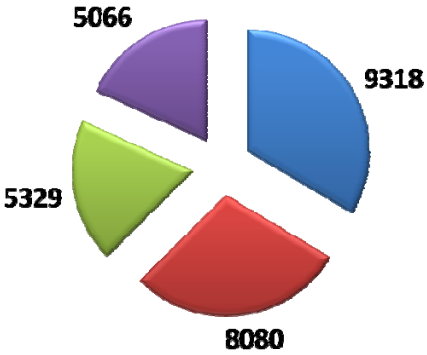
2. Implementation and Transformation and Application Status US GPS technology patents

2.1 Stages of GPS project development

American GPS research project can be divided into four stages: demonstration program phase (1973-1979), research and experimental phase (1979-1984), GPS network construction phase (1985-1993) and the use of GPS conversion stage (1994-present).So far, the total global patent application 130000, Derwent duplicates removed, the result is 34,000 patent families. While 98.8 percent of the total patent applications focused on the use of phase transformation.



The figure reflects the beginning to the present, the number of global trends on GPS patent applications. From the national perspective, the US, China, Japan and South Korea the most patent applications, GPS applications is the world's most dynamic markets. The following figure performance in several key countries on the proportion of case filings:



Blue: USA
Red: China
Green: Jap.
Purple: Kor.

(1) GPS Project Feasibility

First, the US Department of Defense approved the GPS system plan, and fired four test satellites, ground receiving equipment was developed. 1977, the first GPS application for patent was field by NASA.

(2) Development and test phase GPS

At this stage, the United States has launched seven test satellites, in order to conduct a comprehensive research and testing. The stage, has generated a patent has a military background of six, including two US Air Force, the Navy 2, NASA2 member.

(3) GPS network construction phase

Launch 24 satellites, the formal completion of the GPS network, the stage has a military background nine patents, four of which the US Air Force, Navy five.

(4) GPS conversion application stage

At this stage, the US-based GPS technology advanced countries began the conversion of the use of related technologies, including:

a)GPS technology into the United States to promote the use of military and civilian establishment of a sound management system;

b)GPS technology into the United States to promote the use of a range of policies:

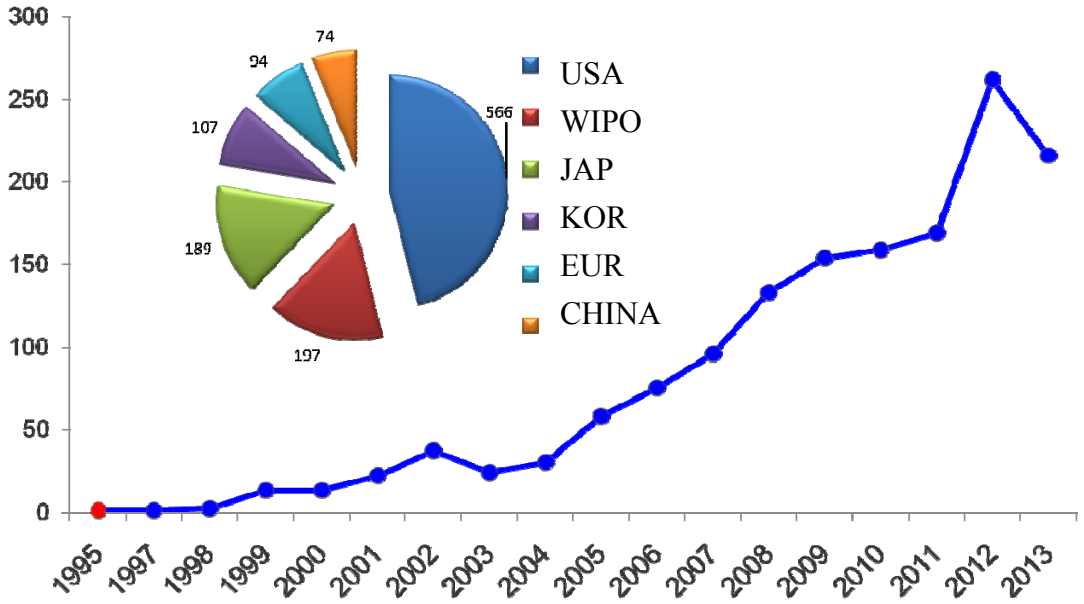
- 1996,White House announced that everyone can use the GPS
- 1996, FCC (Federal Communications Commission) administrative command E911-build public safety wireless network
- 1997, the United States military and civilian dual-use GPS Modernization Program
- 1999 US Vice President announced an increase in the second and three civil frequency
- 2000 US President Bill Clinton ordered a halt SA (ie reduce civil signal performance), so accuracy of civilian into a 6.3 m from the 100 m
- 2004 "American Space-Based PNT Policy"
- Implementation of the new GPS modernization program, GPS capabilities ensure continuous improvement

c)In support of the US government's 428 patent applications:

- Air Force 20
- NASA 32
- Navy 83
- Boeing 193
- Lockheed Martin 100

2.2 Case Study

This report is well-known US satellite navigation technology R & D institutions - Qualcomm Incorporated, for example, analyze its main mode of transformation in terms of navigation technology patents.



Qualcomm satellite navigation technology has been patented layout in 22 countries worldwide, and its patent portfolio in line with the market. The figure reflects the situation.

QUALCOMM conducted the main mode navigation technology patents conversion, there are three, independent research and development, joint research and development and the purchase of patents. Around specific conversion carried out by way of four patents: patent licensing, patent transfer, and patent litigation and investment shares.

Take the patent (Patent No. US20090298511A) for example, tap the existing Qualcomm patents market prospects, by way of purchase of the franchise. The change in ownership of the patent as follows:

US20090298511

2008	2009	2010	2013	2014
patentee	PALM	BANK OF N.Y	HP	PALM	Qualcomm
PALM company has satellite navigation technology on your handheld (PDA) related technology applications and related patent application in 2008	2009, mortgage-related patents are to Bank of New York	$ 1.2 billion acquisition of PALM, and the name for sale Hewlett-Packard	2013 HP will re-transfer of patent rights to PALM	purchase of related patents from Qualcomm hands PALM company 12

We can cite an example (Patent No. US20110292871A), Qualcomm collect relevant market demand, and then according to the actual needs of the development of related technologies and generate patents. 2010, developed by CHIN personal and patent the patentee; by the end of 2010 the transfer of patent rights to Qualcomm, but Qualcomm CHIN I shall researchers.

3. The status of implementation and transformation of domestic GPS technology patents

3.1 Development Status "Beidou"

China's development of GPS technology mainly to "Beidou" project for traction, the current development of China's Beidou satellite navigation system has gone through five stages: to carry out the research phase of satellite navigation (1960), proposes a variety of contemplated phase of the global system (20 1970), proposed independently developed implementation of the "Beidou" construction phase (20th century 80--90 years), "Beidou plan officially approved" phase (1994) and "Beidou" system basic construction is completed (2012). At present, China has now entered the technological achievements of information integration, conversion and use of specific implementation phase.

3.2 "Beidou" technology implementation and Transformation of Patent

At present, the domestic bear the primary development tasks in the field of satellite navigation technology unit of military enterprises, including the China Academy of Space Technology, China Academy of Launch Vehicle Technology and its affiliated research institutes and companies; development of the entire project has been managed as a state secret, and a large number of scientific and technological achievements in various stages of development are technical secrets, the way the national defense patent protection of intellectual property; ordinary basic patent application originated in 2006, and most of them belong to the civilian areas extending Beidou navigation system. In China Academy of Space Technology, for example, general patent directly related to the Beidou satellite navigation technology only 8.

4. Domestic GPS technology patent and the problems in the conversion process

4.1 Technological achievements that involve state secrets, conversion and use of the work carried out difficult

As most of the intellectual property rights based on national defense patent, in the form of technical secrets be protected, it involves state secrets; the defense of intellectual property which can be transformed, when transformed, how the conversion, there is no top-level development of relevant national policies and regulations, various R & D units dare touch the red secrecy, a direct result of the defense of intellectual property rights on the shelf, and technological achievements of information integration, conversion and use of work can not be carried out.

4.2 Satellite navigation at the national level industrial policy inadequate

Although China has introduced a part of the policy, but support is not enough. Our country in 2010, by the Central Military Commission Chairman Hu Jintao signed an order issued the "People's Liberation Army Standard Time Regulations." , Has released its first "military standard timetable" be granted based on the second generation Beidou satellite navigation system in 2011, in 2012 the Ministry of Transport issued a notice requiring nine provinces of Shandong, Jiangsu, Anhui, Tianjin to ensure completion before the end of March 2013 more than 80 percent of the province Beidou satellite navigation system Car terminal installation tasks. However, China still lacks a comprehensive national level satellite navigation industry policy, various professional fields policies and regulations, industry policy and industrial policy in key areas.

4.3 Core intellectual property and not enough to master the key technology, industrial development disorder

In this regard, the most typical example is the large number of Beidou Industrial Park "blind construction," Compass Enterprise "rush" into. Some local governments and businesses around the

country are in the dividend policy to seek their own interests' foothold, based on the demands of their own interests rather than the real market demand to do the Big Dipper. China is currently engaged in the Beidou chip, systems, terminals, applications and other satellite navigation application industry vendors institutions over 2500, specializing in satellite navigation industry, there are about 1000, but smaller in scale, weak and not yet form a joint force, corporate lack of innovation, international competitiveness is low. On the other hand, domestic base band and RF chip in terms of power consumption, sensitivity, product stability, application solutions are still a wide gap between the core intellectual property. In the area of national security of civil aviation, transport, telecommunications, finance, electricity, logistics and other heavily dependent on GPS, Beidou system in the domestic market share of only 1%.

4.4 Lack of market-oriented technological achievements into use mechanisms

Mainly in the sharing of resources, demand docking, intellectual property assessment, and the lack of distribution of benefits and other support mechanisms. Therefore, we should pay attention to intellectual property rights and ordered to create a competitive environment and achieve seamless link "top-down" government action and the "bottom-up" market behavior.

5. Further implementation of the recommendations of domestic GPS technology patents and Transformation

Recommendations for the implementation of intellectual property management, improve efficiency, "Compass" patented technology into implementation, the report further proposed the "Compass" and patent applications related to the conversion proposal.

5.1 States should introduce relevant policies to promote satellite navigation system to promote conversion

The United States since 1996 transformed into the GPS application stage, twice presidential decree, to promote GPS's military conversion, the development of GPS industry has benefited from the full support of the US government;

On the contrary, Russia's GLONASS almost the same period to establish and GPS, and to achieve full coverage of the time even earlier than GPS, but now can not even achieve full coverage of the Russian regions. While this is partly due to the disintegration of the former Soviet Union, but more in the process of industrialization is not enough government support.

The Russian government in recognizing this aspect after:

In December 2006, the Government of the Russian Federation First Deputy Prime Minister Sergei Ivanov said that Russia will take effect January 1, 2007 to lift restrictions GLONASS accuracy, resulting in global navigation satellite system of arms control for precise, unlimited business use, and for the Russian Economic Development and Transport provide high-precision service.

In October 2008, Russian Deputy Prime Minister Sergei Ivanov said that "the Russian Federation navigation Business Act" will be submitted for discussion, the provisions in the bill, "all funded from the government budget system or related to the national security system must first use GLONASS system. "" In some special circumstances, may even use more extreme way, prohibit the use of other countries positioning device. "

(1) improve the national level of the satellite navigation industry policy

According to China's National Defense Science and Technology Information Center for domestic enterprises, universities, and other groups of domestic and international users survey results show

that four of the most talked about policy: technology innovation policies, intellectual property rights policies, market regulation policies and enforce policies in key industries.

(2) Industrial policies of recommendation at the national level of Satellite navigation.

According to China's National Defense Science and Technology Information Center for domestic enterprises, universities, and other groups of domestic and international users survey results show that most suggestions for China's satellite navigation industry development countermeasure attention are:

- Encourage and foster technological innovation, the protection of intellectual property rights
- Accelerate the construction of the Beidou system, improve system performance, the development of top-level industrial development policies
- Long-term development plan for satellite navigation industry
- Government to strengthen coordination, standardizing market order, and create a favorable market environment

5.2 The State shall establish and improve the military and civilian project management system

The United States attaches great importance to GPS, placing height management and policy development of national strategies, strong leadership by the senior officials responsible for sectors, increased coordination efforts, to ensure that policy planning departure from the national interests, to avoid the various departments fragmentation and duplication;

American establishment and management of GPS running in a perfect military and civilian management system, which the national space-based PNT Executive Committee for the overall policy, institutional demand, resource allocation and other integration and strategic decision-making and coordination department, aerospace commander of US Space section as a military technical support, space combat organization, the Department of Transportation, the Department of Homeland Security, the Coast Guard navigation Center, the Federal aviation Administration, and other departments as a civilian requirements development, civil service and other regulatory bodies.

5.3 Conversion of military technology to use should be based on the patent as a breakthrough

98.9% GPS total patent applications in the conversion application stage, because the patent is technology or products to enter the market umbrella, there is no patent protection, any technological innovation, inventions outcome would not be the commercialization and industrialization. Satellite navigation technology from military projects, should be converted to civilian use in the field of pay more attention to patent protection.

Patent application development stage of basic GPS with a military background, including a US government supports background 428 patent applications.

Can the following key points as a breakthrough, the conversion of military technology to achieve GPS application:

(1) Through professional assessment and nurturing a unique process to produce high quality patented invention.

(2) By analyzing the scope of the claims, writing and management of domestic and foreign patent text responses to improve the quality of patent applications.

(3) Enhance the value of patents by means of specialized marketing and business negotiation skills.

(4) The adoption of internationally accepted practice of intellectual property into income allocated to mobilize the enthusiasm of all parties

5.4 Establishing an efficient communication platform between military technology and civilian needs

American Civil GPS Service Liaison Committee (CGSIC) as an important military and civilian communication platform has played an important role in the industrialization of the GPS civilian areas. CGSIC is effective communication channels between the GPS service providers and users, is an important link between global civil GPS users and GPS among service providers, whose mission is to tap the full potential of GPS civil part. Implementation requirements and market, specific ways to communicate the following information:

First, the military group of companies, universities as the main platform for the establishment of industry intellectual property information resources, the effective elimination of trade barriers, norms of confidentiality issues, and enhance the operability of innovation resources, to achieve the centralized management of scientific intellectual property to address research Information Circulation units and production units, military and civilian industry between;

Second, carding military and civilian technology research directory, IPR combing, classification, and valuation and market analysis, use of intellectual property as the main line to promote the conversion of military technology to civilian;

Third, the use of intellectual property into the promotion phase to solve the actual needs of trading and industrialization of scientific and technological achievements, should make full use of existing information exchange platform, to achieve the purpose to show results and technology needs, promote technology to promote the greatest degree of conversion.

5.5 Beidou navigation system patent information application integration and transformation principles

(1) Task-oriented, demand-driven: the dual mission requirements Beidou system as a guideline, the market demand for the driving force of technological innovation, identify and then develop ideas and direction of technology in the research and production tasks and market demand environment.

(2) "Innovative mechanisms, integration of resources": Exploring the ordinary patents, patent achievements defense effective information integration model, methods and techniques and efficient transfer of operational mechanism, and fully integrate existing patents results, optimize resource allocation.

(3) "Government guidance, multi-stakeholder involvement": to have a large number of technological achievements of aerospace enterprises as the mainstay, played patent achievements transformation and application of information integration and guiding role of the government in the same time the absorption of other different technologies involved in innovation subjects and related resources, improve the efficiency of the overall outcome of the patent application;

(4) "Clear interests, joint development": the transformation around the patent application rights and obligations between the different stakeholders should be defined in the contract or other forms, making the benefits of what they want to participate in different subjects of technological innovation and achieve win-win cooperation objectives.

5.6 Beidou navigation system patent information application integration and transformation solutions

(1) Technical Category

Beidou system for dual-use R & D needs, carded patent the results of each business to technology, key technology, research and study needs classification, patent achievements over classified information integration will sort out.

(2) Data Integration

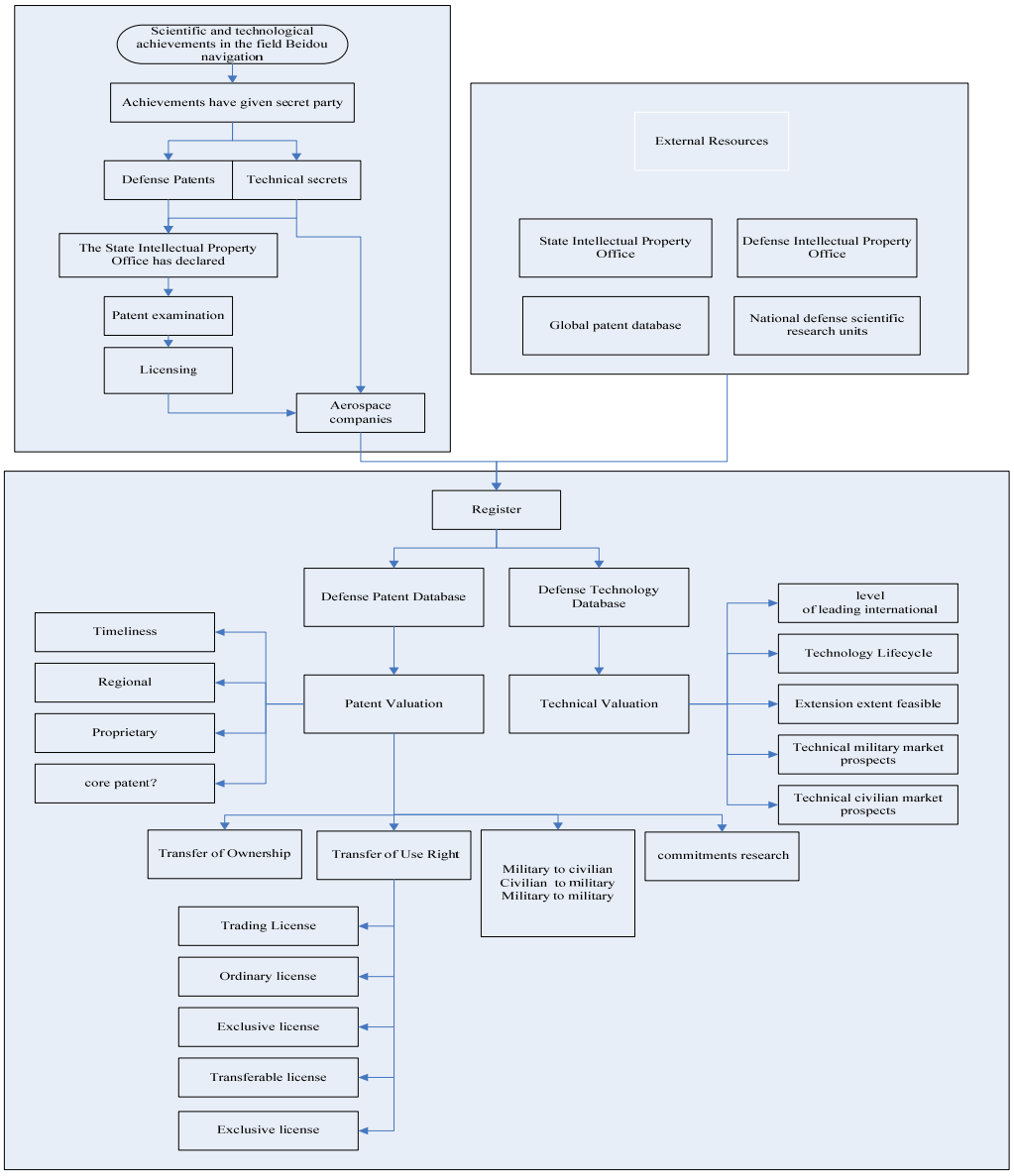
For dual-use research and aerospace companies Beidou navigation system in the professional and technical development needs, according to the technical integration of national defense patent system diagram results, and other information.

(3) System

Induction and screening arrangement to collect patent results obtained condition, the demand for information aggregated statistical analysis will be matching demand and supply side of information technology to integrate, establish patent achievements into the application database.

(4) The outcome of the patent application to convert Demonstration of market mechanisms

In aerospace enterprises as the mainstay, form a highly efficient transfer of patent achievements sharing mechanisms and internal technical market program.



The figure below shows the results of the Beidou navigation area of intellectual property information integration and application platform conversion diagram:

5.7 Training an intellectual property into use professionals

One important use of intellectual property into the smooth development of security conditions is talent, intellectual property management and transformation are extremely complex business practices, expertise professionals need complex, where both have the patent field, but also to understand the law, business, in order to better complete identification patents, training, protection and transformation range of business activities. Professional guidance is designed to help inventors claim to expand the required experiments, can become valuable worthless invention, the low value of the invention into high-value, create value for the unit and inventors play a significant role.

Design of flywheel configurations for fast maneuver on single-axis of spacecraft

Ming Qi, Yi Yang, Dengyun Wu, Shaowei Zhang, Pengbo Zhang

Beijing Institute of Control Engineering

Abstract: The fast maneuver problem of the spacecraft is one of the research hotspots recently. The flywheel is usually adopted as attitude execution mechanism of the spacecraft for its characteristics of long life and high reliability. Nevertheless, limited by the quality and power consumption, the torque produced by the general flywheel cannot meet the rapid maneuvering requirements of the spacecraft. It needs to develop the high torque flywheel for its better capabilities. The research status of the typical high torque flywheel in the world is given. From the investigation it can be known that the power consumptions of the high torque flywheels are generally great, so the maneuver project should take into account that just laying the high torque flywheel on the fast maneuver axis (defines as X axis) while the other axes are equipped with the general flywheels for decreasing the power consumption of the system. At this point, the flywheel configuration can be defined as follows: There are n flywheels to be used in total($4\leq n\leq 7$). One of them is a high torque flywheel and its angular momentum axis is coincident to the maneuvering axis and the others are all same general flywheels slanting at particular angles. The angles of each momentum axis of the general flywheels in relation to the maneuvering axis (defines as α) are equal and the lines which be produced by projecting the momentum axes on the non-maneuvering plane (defines as Y-Z plane) are distributed equally. Thereinto, the angle between the projection line of the first general flywheel and the axis Y is defined as β. The process of configuration design can be defined as follows: Firstly, to keep α and β as variables and give the location and torque allocation matrices. Then to calculate the energy optimal solution of the variables according to the maximum torque demands of the three axes and the maximum torque capability of the general flywheel when n is minimum. Finally, to calculate the optimal torque of the high torque flywheel based on the above. And, using our conventional product (Output torque $\geq$ 0.075 N·m) and high torque product (Output torque $\geq$ 0.5 N·m) as the baseline, a design example is given. The proposed flywheel configurations and the line of thinking is useful to the design of flywheel configurations for fast maneuver on single-axis of spacecraft and it can be developed to be used in conditions which need more maneuvering axes if necessary.

Key words: Spacecraft; High torque flywheel; Fast maneuver; Flywheel configuration; Attitude control; Energy optimal

1 Introduction

Flywheels can generate control torques via the internal speed adjustable inertia rotors. When they are used, the magnitude change of the angular momentum can provide an output torque, which will change the angular velocity of the spacecraft. As the important parts in the attitude control system for spacecraft, the flywheels have the advantage of low power consumption, light weight and high reliability. The continuous output torque can accommodate the precise attitude control requirements of the spacecraft. So the space wheels are the ideal choice for a variety of high quality spacecraft[1-4].

Nevertheless, limited by the quality and power consumption, the torque produced by the general flywheel cannot meet the rapid maneuvering requirements of the spacecraft. Recently, the high torque flywheel became one of the research focuses for its better capabilities with the greater requirements for the maneuver capability[5-8].

The research status of the typical high torque flywheel in the world is given. The one produced by Beijing Institute of Control Engineering which can produce no less than 0.5 N·m torque is introduced. From the investigation it can be known that the power consumptions of the high torque flywheels are generally great, so the maneuver project should take into account that just laying the high torque flywheel on the fast maneuver axis (defines as X axis) while the other axes are equipped with the general flywheels for decreasing the power consumption of the system. The process of configuration design is defined and a design example is given using our conventional product (Output torque ≥ 0.075 N·m) and high torque product (Output torque ≥ 0.5 N·m) as the baseline.

2 General situations

From the investigation of the research status of the high torque flywheel we find that the representative companies are German Teldix company and American ITHACO company. If only taking into account of the reaction torque, German Teldix company has a kind of high torque flywheel which can produce no less than 0.6 N·m torque. But the mass of that product is more than 60 kg. Thus that product isn't a meaningful high torque flywheel and it is a general flywheel in nature which has heavier mass and bigger volume. About decadal years lately, those companies are all produced some high torque flywheels which can produce 0.2~0.3 N·m torque. And it can account for that the high torque flywheels are accord with the needs of the developments of international spaceflight in future. The mostly capabilities of the high torque flywheels of Teldix are shown in the following Table.

Table 1 The mostly capabilities of the high torque flywheels of Teldix

Characteristics	RSI 50-220/45	RSI 30-280/30
Working mode	Momentum/Reaction	Momentum/Reaction
Angular momentum at nominal speed	50N·m·s	30N·m·s
Operation speed range	±4500r/min	±3000r/min
Motor Torque	0.22N·m	0.28N·m
Loss torque	<0.025N·m	<0.025N·m
Power consumption	Maximum torque at nominal speed：<150W Steady state at nominal speed：<20W	Maximum torque at nominal speed：<150W Steady state at nominal speed：<20W
Mass	<9.2kg	<9.0kg
Dimensions	Φ347×124mm	Φ347×124mm
Life Time	≥15 years	≥15 years

The mostly capabilities of the high torque flywheels of ITHACO are shown in the following Table.

Table 2 The mostly capabilities of the high torque flywheels of ITHACO

Characteristics	TW-16B200	TW-26E300
Working mode	Momentum/Reaction	Momentum/Reaction
Angular momentum at nominal speed	16.6N·m·s	26N·m·s
Operation speed range	±5100r/min	±2000r/min
Reaction Torque	0.2N·m	0.3N·m
Power consumption	Maximum torque at nominal speed：<250W Steady state at nominal speed：<20W	Maximum torque at nominal speed：<120W Steady state at nominal speed：<28W

Beijing Institute of Control Engineering developed a high torque flywheel which can produce no less than 0.5 N·m torque. The 0.5 N·m flywheel is a non-mechatronic product operated by external Wheel Drive Electronics (WDE) respectively shown in Figure 1 and 2. The flywheel has many advantages such as high torque, high control accuracy, light weight and long lifetime and so on. The mostly capabilities of the 0.5N·m high torque flywheel is shown in Table 3.

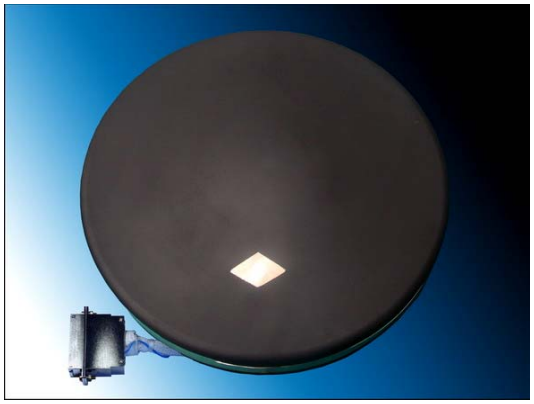


Figure 1 0.5N·m high torque flywheel

Figure 2 0.5N·m high torque flywheel WDE

Table 3 The mostly capabilities of the 0.5N·m high torque flywheel

Characteristics	TRMW-350-25B TRMWE-1-42B
Control type	Torque control
Working mode	Momentum/Reaction
Angular momentum at nominal speed	25N·m·s
Nominal speed	2000r/min
Operation speed range	±2000r/min
Reaction Torque	0.5N·m
Loss torque	<0.02N·m
Static unbalance	≤2g·cm
Dynamic unbalance	≤30g·cm^2
Power consumption	Standby：<8W Maximum torque at nominal speed：<200W Steady state at nominal speed：<25W
Mass	Wheel：9.25kg WDE：4.35kg
Dimensions	Wheel：Φ343×120mm WDE： 267×183×117mm
Life Time	≥15 years

3 Flywheel configurations and optimal design

From the investigation it can be known that the power consumptions of the high torque flywheels are generally great, so the maneuver project should take into account that just laying the high torque flywheel on the fast maneuver axis (defines as X axis) while the other axes are equipped with the general flywheels for decreasing the power consumption of the system. Therefore, it's significant to study the flywheel configurations and optimal design in this situation.

For the problem of fast maneuver on single-axis of spacecraft, literature[9] had putted forward an optimization method of flywheel configurations in case of n=5. All of the five flywheels are same. One of them is laying on the fast maneuver axis while the others are slanting at particular angles. The angles of each momentum axis of the slanting flywheels in relation to the maneuvering axis are equal and the lines which be produced by projecting the momentum axes on the non-maneuvering plane are distributed equally. Based on the above, the flywheel configurations and design process are proposed as follows using the high torque flywheel.

3.1 Flywheel configurations

At this point, the flywheel configuration can be defined as follows: There are n flywheels to be used in total. One of them is a high torque flywheel and its angular momentum axis is coincident to the maneuvering axis and the others are all same general flywheels slanting at particular angles. The angles of each momentum axis of the general flywheels in relation to the maneuvering axis (defines as α) are equal. The value range of α is

$$0 \le \alpha \le \frac{\pi}{2} \qquad [1]$$

The lines which be produced by projecting the momentum axes on the non-maneuvering plane (defines as Y-Z plane) are distributed equally. Thereinto, the angle between the projection line of the first general flywheel and the axis Y is defined as β. The value range of β is

$$0 \le \beta \le \frac{2\pi}{n} \qquad [2]$$

Considering the requirements of three axes stabilization， we define n≥4. And we can define n≤7 taking into account that the number of flywheels will increase the layout complexness of the spacecraft. Commonly, it needs to replace the initial general flywheel with the other kind of general flywheel and the high torque flywheel if it can't satisfy the using requirements in the instance of n=7.

For giving a demonstration, the layout of n=4 is given in Figure 3.

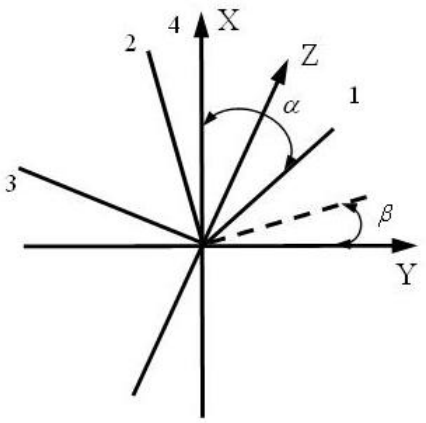


Figure 3 The layout of n=4

3.2 Optimal design

For the sake of adequate utilizing the capacities of the actuators in the energy optimal condition, it needs to distribute the locations of the flywheels reasonably when n is minimum.

Based on the above, we can keep α and β as variables and give the location matrix. For an example, the location matrix of n=4 is shown as follows:

$$U = \begin{bmatrix} 1 & \cos\alpha & \cos\alpha & \cos\alpha \\ 0 & \sin\alpha\cos\beta & \sin\alpha\cos(\beta + 2\pi/3) & \sin\alpha\cos(\beta + 4\pi/3) \\ 0 & \sin\alpha\sin\beta & \sin\alpha\sin(\beta + 2\pi/3) & \sin\alpha\sin(\beta + 4\pi/3) \end{bmatrix} \qquad [3]$$

The torque allocation matrix on condition that energy optimal is defined as follows:

$$D = U^{\mathrm{T}}\left(UU^{\mathrm{T}}\right)^{-1} \qquad [4]$$

Define the maximum of the output torque of the high torque flywheel and the general flywheel is Thmax and Tmax respectively, thus the matrix of the maximum torque can be obtained as

$$T_{\mathrm{I}} = \left[\mathrm{T_{h\,max}} \quad \mathrm{T_{max}} \quad \cdots \quad \mathrm{T_{max}}\right]^{\mathrm{T}}_{1\times n} \qquad [5]$$

The torque demands matrix distributed to the three perpendicular axes is

$$T_{\mathrm{C}} = \begin{bmatrix} T_{\mathrm{X}} & T_{\mathrm{Y}} & T_{\mathrm{Z}} \end{bmatrix}^{\mathrm{T}} \quad [6]$$

Then, define

$$T_{\mathrm{R}} = DT_{\mathrm{C}\max} = \begin{bmatrix} T_1 & T_2 & \cdots & T_n \end{bmatrix}^{\mathrm{T}} \quad [7]$$

Finally, the optimal objective function can be defined considering the restrictions of the non-maneuvering axes.

$$\varepsilon = \left(T_{\max} - \left|T_2\right|\right) + \cdots + \left(T_{\max} - \left|T_n\right|\right) \to \min \quad [8]$$

and

$$T_{\max} - \left|T_1\right| \geq 0 \ , \ \cdots \ , \ T_{\max} - \left|T_n\right| \geq 0 \quad [9]$$

Then to calculate the energy optimal solution of the variables according to the maximum torque demands of the three axes and the maximum torque capability of the general flywheel when n is minimum.

Finally, the optimal torque of the high torque flywheel can be obtained based on the above. It needs to point out that it can increase the angular momentum of the high torque flywheel by increasing the rotational speed and decreasing the maximum output torque if the calculation value of the torque is less than the actual capabilities. It will help to decrease the unload frequency in the instance that the energy is about the same.

4 Calculating example

A design example is given using our conventional product (Output torque ≥ 0.075 N·m) and high torque product (Output torque ≥ 0.5 N·m) as the baseline.

The matrix of the maximum torque of the flywheels is obtained as

$$T_{\mathrm{I}} = \begin{bmatrix} 0.5 & 0.075 & \cdots & 0.075 \end{bmatrix}^{\mathrm{T}}_{1\times n}$$

Here, the maximum torque demands of the non-maneuvering axes and the maneuvering axis are 0.3N·m and 0.1N·m respectively. Thus, the maximum torque demands matrix distributed to the three perpendicular axes is

$$T_{\mathrm{C}\max} = \begin{bmatrix} 0.3 & 0.1 & 0.1 \end{bmatrix}^{\mathrm{T}}$$

From the section 3 it can be known that the number n of the flywheels can be verified according to the exceptive example of $\alpha = \pi / 2$. If the flywheels can't provide adequate torque for the demand, the number n of the flywheels shall be increased.

Figure 4 shows the calculating values of the maximum torque capability that the general flywheel needs to have when n=4.

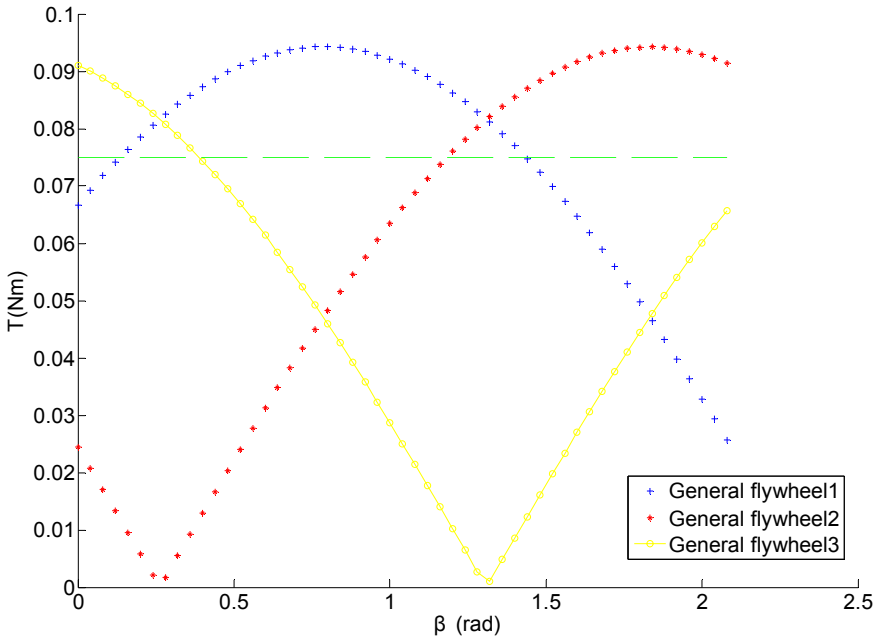


Figure 4 The calculating values of the maximum torque capability that the general flywheel needs to have when n=4

From the Figure 4 it is shown that the 0.075N·m flywheel can't satisfy the torque requirements no matter what the value of β is chosen when n=4.

Then we get the calculating values of the maximum torque capability that the general flywheel needs to have when n=5, as shown in Figure 5.

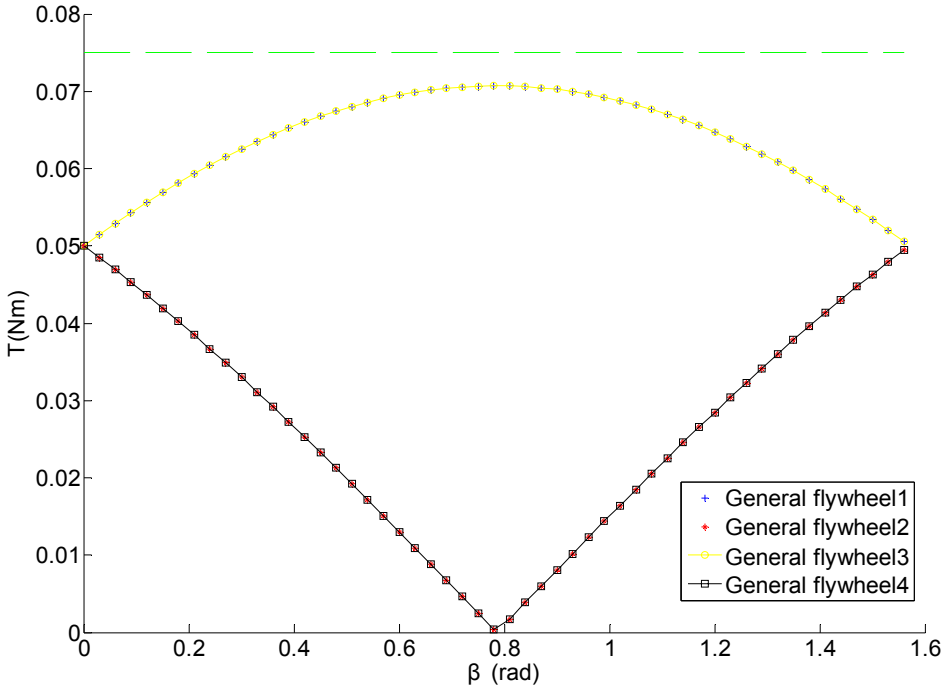


Figure 5 The calculating values of the maximum torque capability that the general flywheel needs to have when n=5

Figure 5 shows that it has value range for torque requirements when n=5.

Thus, the optimal objective function can be solved keeping α and β as variables. The solution distribution of ε is shown in Figure 6.

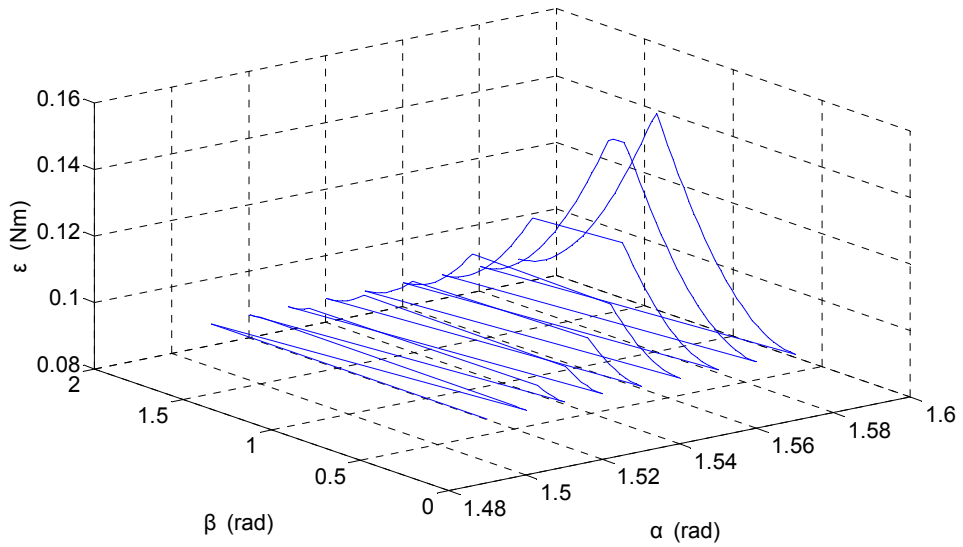


Figure 6 The solution distribution of ε when n=5

Then the optimal solution can be obtained according to Figure 6 as

$$\alpha = 1.49rad, \beta = 0rad$$

Here, the output matrix of the flywheels is

$$T_{\mathrm{I}} = \begin{bmatrix} 0.2924 & 0.0738 & 0.0738 & -0.0266 & -0.0266 \end{bmatrix}^{\mathrm{T}}$$

The torque capability of the high torque flywheel is 0.5N·m. It is greater than the practical torque 0.2924N·m. Therefore，the angular momentum of the high torque flywheel can be properly increased by increasing the rotational speed and decreasing the maximum output torque according to the unload frequency.

And it needs to point out that the reason that the practical torque is similar to the torque demand of the maneuvering axis in the above instance is the torque capability of the general flywheel is relatively smaller than the torque demand of the non-maneuvering axis. If the torque capability of the general flywheel is greater, the conditions will be improved.

5 Conclusions

The design of flywheel configurations for fast maneuver on single-axis of spacecraft is investigated in this paper. The conclusions can be made as follows:

1) A maneuver project should take into account that just laying the high torque flywheel on the fast maneuver axis (defines as X axis) while the other axes are equipped with the general flywheels for decreasing the power consumption of the system. The flywheel configuration can be defined as follows: There are n flywheels to be used in total(4≤n≤7). One of them is a high torque flywheel and its angular momentum axis is coincident to the maneuvering axis and the others are all same general flywheels slanting at particular angles. The angles of each momentum axis of the general flywheels in relation to the maneuvering axis (defines as α) are equal and the lines which be produced by projecting the momentum axes on the non-maneuvering plane (defines as Y-Z plane) are distributed

equally. Thereinto, the angle between the projection line of the first general flywheel and the axis Y is defined as β.

2) The process of configuration design can be defined as follows: Firstly, to keep α and β as variables and give the location and torque allocation matrices. Then to calculate the energy optimal solution of the variables according to the maximum torque demands of the three axes and the maximum torque capability of the general flywheel when n is minimum. Finally, to calculate the optimal torque of the high torque flywheel based on the above. It needs to point out that it can increase the angular momentum of the high torque flywheel by increasing the rotational speed and decreasing the maximum output torque if the calculation value of the torque is less than the actual capabilities. It will help to decrease the unload frequency in the instance that the energy is about the same.

3) A design example is given using our conventional product (Output torque ≥ 0.075 N·m) and high torque product (Output torque ≥ 0.5 N·m) as the baseline. The results show that 4 flywheels can't fulfill the 0.1N·m demands of the non-maneuvering axis. The optimal solution when n=5 is α=1.49rad，β=0rad.

4) The proposed flywheel configurations and the line of thinking is useful to the design of flywheel configurations for fast maneuver on single-axis of spacecraft and it can be developed to be used in conditions which need more maneuvering axes if necessary.

References

[1] AVISHAI W，ILYA K，DENNIS S B．Inertia-free spacecraft attitude control using reaction wheels [J]. Journal of Guidance, Control, and Dynamics, 2013, 36（5）: 1425-1439.

[2] YANG D D，YUE B Z．Attitude maneuver of spacecraft with long cantilever beam appendage by momentum wheel [J]. International Journal of Control, 2013, 86（2）: 360-368.

[3] KIM Y M, KIM B K. Energy-efficient trajectory generation for space manipulators with reaction wheels under a fixed base orientation [J]. Journal of Intelligent & Robotic Systems: Theory & Application, 2014, 76（2）: 219-237.

[4] CHOU M C，LIAW C M．PMSM-driven satellite reaction wheel system with adjustable DC-link voltage [J]. IEEE Transactions on Aerospace and Electronic Systems, 2014, 50（2）: 1359-1373.

[5] VIT BABUSKA, SCOTT M BEATTY, BRETT J DEBLONK, et al. A review of technology developments in flywheel attitude control and energy transmission systems[C]. IEEE Aerospace Conference Proceedings, Piscataway, NJ, USA, 2004, (4): 2784-2800.

[6] BLALKE B. A combined earth and momentum wheel for attitude determination and control of small spacecraft[R]. N92-25078.

[7] GU S, JIA J Q, JIN G. Research on momentum management of fast slew maneuver small satellite[J]. Opt．Precision Eng, 2008, 16(8)：1540-1545

[8] ZHOU W M, ZU L Y, ZHU Q H. An attitude maneuver control method based on path planning for agile satellite with large torque flywheels[J]. Aerospace Control and Application, 2014, 40(2)：37-41,46

[9] HOU Z L，GENG Y H. Optimization method of flywheel configurations for fast maneuver on single-axis of spacecraft：China，CN102616386B[P]. 2014-04-02

An Improved Linearization Calibration Method of Space-borne Three-axial Magnetometer Based on Mathematical Platform

WANG Jiabo* CHEN Xi

(*Space Center, Tsinghua University)
Room 106, Floor 10, Main Building, Tsinghua University,100084
wangjiabo13@mails.tsinghua.edu.cn, chenxiee@tsinghua.edu.cn

Abstract: The 3-axial magnetometer is an essential component of modern small/micro satellites, which is usually used as geomagnetism sensor in geomagnetism-based attitude determinations. The errors source inherent to magnetometers including zero point offset error, 3-axis orthogonal error and calibration factor error are main measurement error sources in small/micro satellites attitude determinations. In this work, a modified 3-axial magnetometer calibration algorithm using linear model is proposed. Different from existing algorithms, it performs zero point offset estimation before figuring out other parameters. Simulation results show that the proposed algorithm outperforms existing algorithm especially in terms of 3-axis orthogonal residue error when zero point offset error is larger than 50nT, which makes it better in meeting the low-cost and fast-to-market requirements of modern space-borne magnetometers.
Keywords: Magnetometer, Three-axial error; Calibration; Linear Model

1. Introduction

Three-axial magnetometers, as are used for spacecraft's attitude measurement by geomagnetism method, are essential components on the microsatellites [1]. Magnetometers have zero point offset, three-axial orthogonal errors, and three-axial calibration factor errors due to the restriction of material technology and manufacturing process. These errors bring measurement deviation by decades to thousands of nTs, which is unbearable in microsatellite applications. Thence, calibration is indispensable before putting magnetometers into use.

In order to enhance the calibration capability, specific mechanical structures used for the calibration are set on the probe, and resistors used for zero setting and sensitivity setting are reserved on the circuit. However, in the experiment, it is difficult to set a magnetometer in a certain magnetic field strictly parallel to the direction of one axis [2]. The calibration method based on the mechanical structure on the probe increases the manufacturing cost, manufacturing periodicity and weight of the probe. Also, the calibration structure on the probe introduces problems to reliability, especially when the satellite is launched by a solid rocket. Therefore, substituting the mechanical structure by mathematical methods (so called mathematical platform) to calibrate the measurement errors is a better way to reduce the cost, weight, manufacturing periodicity, and enhance the reliability. Mathematical platform comprises in-orbit calibration method and off-line calibration method, and they are both adopted. In-orbit calibration, by filter method, can eliminate the deviation of IGRF model and the true geomagnetic field as well as the dynamic errors of the satellite [3-5]. Off-line

calibration is responsible for calibrating the inherent measurement errors and the body static magnetic errors [6-9].

The existing work to revise the inherent measurement errors can be categorized into two types. One is concluding the calibration of three-axial measurement as a single objective optimization problem without constraint, which can be solved by simulated annealing algorithm, ant colony algorithm, genetic algorithm, and artificial neural network algorithm[6-7]. These algorithms have the advantage that they can constrain themselves when the errors are rather small. But their defect is that they may work out a second-best solution. Another category is establishing equations by linearization method and solving the equations by least square method [8-9] or kalman filter [10-12]. In the debugging stage, the sensitivity resistors and zero setting resistors interact with each other so that it is difficult to get the expected sensitivity as well as the zero point offset. The existing linearization method (e.g. the method in [9]) ignores the zero point offset when calculating the orthogonal degree, which results in performance reduction when the zero point offset is big enough. In this work, an improved linearization calibration method is proposed, by which zero point offset is figured out first, and then other parameters are calculated. The simulation result shows that the proposed method outperformed the existing method when there exists a big zero point offset.

2. The error model of a three-axial magnetometer

Given a reference coordinate $O-xyz$ and actual polarity coordinate $O–x'y'z'$ of a magnetometer, their relationship is given in Fig.1[13-14]. The originals of the two coordinates coincide, z-axis and z'-axis also coincide, x' lies in xOz plane, α □ $\angle x'Ox$, γ is the angle of y' regarding with plane-xOy, β is the angle between the projection line of y' on xOy and y axis.

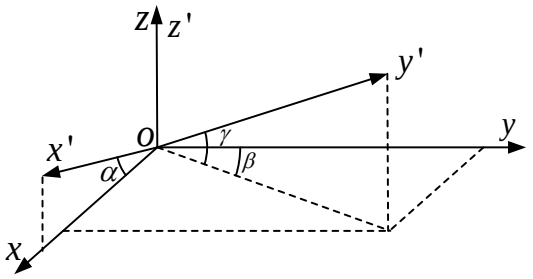


Fig.1. The reference coordinate and actual polarity coordinate

With the above definitions, we have

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos\alpha & \cos\gamma\sin\beta & 0 \\ 0 & \cos\gamma\cos\beta & 0 \\ \sin\alpha & \sin\gamma & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \tag{1}$$

Assume that the actual measurement output of a magnetometer in $O-x'y'z'$ is $Hm' = [Hm'_x, Hm'_y, Hm'_z]^{\mathrm{T}}$, the true value in $O-xyz$ is $H_p = [H_{px}, H_{py}, H_{pz}]^{\mathrm{T}}$, the magnetometer zero point offset is $E = [E_x, E_y, E_z]^{\mathrm{T}}$ and a calibration factor is $K = [k_x, k_y, k_z]^{\mathrm{T}}$, then we have

$$H_p = \begin{bmatrix} H_{px} \\ H_{py} \\ H_{pz} \end{bmatrix} = \begin{bmatrix} k_x & 0 & 0 \\ 0 & k_y & 0 \\ 0 & 0 & k_z \end{bmatrix} \times \begin{bmatrix} \cos\alpha & \cos\gamma\sin\beta & 0 \\ 0 & \cos\gamma\cos\beta & 0 \\ \sin\alpha & \sin\gamma & 1 \end{bmatrix} \times \begin{bmatrix} Hm'_x - E_x \\ Hm'_y - E_y \\ Hm'_z - E_z \end{bmatrix} = \begin{bmatrix} k_x\cos\alpha & k_y\cos\gamma\sin\beta & 0 \\ 0 & k_y\cos\gamma\cos\beta & 0 \\ k_x\sin\alpha & k_y\sin\gamma & k_z \end{bmatrix} \times \begin{bmatrix} Hm'_x - E_x \\ Hm'_y - E_y \\ Hm'_z - E_z \end{bmatrix} \tag{2}$$

Define

$$S = \begin{bmatrix} k_x \cos\alpha & k_y \cos\gamma \sin\beta & 0 \\ 0 & k_y \cos\gamma \cos\beta & 0 \\ k_x \sin\alpha & k_y \sin\gamma & k_z \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} & 0 \\ 0 & s_{22} & 0 \\ s_{31} & s_{32} & s_{33} \end{bmatrix} \tag{3}$$

and

$$H_m = \begin{bmatrix} Hm_x \\ Hm_y \\ Hm_z \end{bmatrix} = \begin{bmatrix} Hm'_x - E_x \\ Hm'_y - E_y \\ Hm'_z - E_z \end{bmatrix} \tag{4}$$

Then we have

$$|H|^2 = H^T H = Hm^T S^T S Hm \tag{5}$$

and

$$|H| = \sqrt{H_x^2 + H_y^2 + H_z^2} \tag{6}$$

which is a constant in uniform magnetic field. Considering the effect non-orthogonality has on magnetometers, the measurement results of the three axis varies when the attitude of the magnetometer changes. Hence, the output of the magnetometer will change because of the non-orthogonality of the coordinate.

There exist errors between the actual output values of the three-axis magnetometer and the ideal sensor output values, and these errors will not change after the sensor is made. Such errors become inherent errors, which includes zero point offset errors, three-axial axis orthogonal errors, and axis calibration factor errors.

3. The proposed calibration method

$|H|^2$ is a constant, and $S^T S$ can be formulated as

$$S^T S = \begin{bmatrix} s_{11}^2 + s_{31}^2 & s_{11}s_{12} + s_{31}s_{32} & s_{31}s_{33} \\ s_{11}s_{12} + s_{31}s_{32} & s_{12}^2 + s_{22}^2 + s_{32}^2 & s_{32}s_{33} \\ s_{31}s_{33} & s_{32}s_{33} & s_{33}^2 \end{bmatrix} \tag{7}$$

Then $|H|^2$ can be formulated as

$$\begin{aligned} |H|^2 &= H^T H = Hm^T S^T S Hm \\ &= (s_{11}^2 + s_{31}^2)Hm_x^2 + (s_{12}^2 + s_{22}^2 + s_{32}^2)Hm_y^2 + s_{33}^2 Hm_z^2 + 2(s_{11}s_{12} + s_{31}s_{32})Hm_x Hm_y \\ &\quad + 2s_{32}s_{33}Hm_y Hm_z + 2s_{31}s_{33}Hm_z Hm_x \end{aligned} \tag{8}$$

Define

$$HM = [Hm_x^2, Hm_y^2, Hm_z^2, Hm_x Hm_y, Hm_y Hm_z, Hm_z Hm_x] \tag{9}$$

$$RM = \begin{bmatrix} Rm_1 \\ Rm_2 \\ Rm_3 \\ Rm_4 \\ Rm_5 \\ Rm_6 \end{bmatrix} = \begin{bmatrix} s_{11}^2 + s_{31}^2 \\ s_{12}^2 + s_{22}^2 + s_{32}^2 \\ s_{33}^2 \\ 2(s_{11}s_{12} + s_{31}s_{32}) \\ sa_{32}s_{33} \\ 2s_{31}s_{33} \end{bmatrix} \tag{10}$$

We can get

$$|H|^2 = HM \cdot RM \tag{11}$$

Suppose the magnetometer rotates K groups of attitudes, resulting in K groups of observed values, we can get

$$(|H|^2)_{K\times 1} = HM_{K\times 6} \cdot RM_{6\times 1} \tag{12}$$

It can be deduced that

$$RM_{6\times 1} = ((HM_{K\times 6})^{\mathrm{T}} HM_{K\times 6})^{-1} \cdot (HM_{K\times 6})^{\mathrm{T}} \cdot (|H|^2)_{K\times 1} \tag{13}$$

By solving the equations, we can get the solutions as

$$\begin{aligned} s_{11} &= \sqrt{Rm_1 - s_{31}^2} \\ s_{12} &= (0.5Rm_4 - s_{31}s_{32})/s_{11} \\ s_{22} &= \sqrt{Rm_2 - s_{12}^2 - s_{32}^2} \\ s_{31} &= Rm_6/2s_{33} \\ s_{32} &= Rm_5/2s_{33} \\ s_{33} &= \sqrt{Rm_3} \end{aligned} \tag{14}$$

Furthermore, we get the nine parameters α、β、γ、k_x、k_y、k_z、E_x、E_y、E_z and eliminate the errors of the magnetometer.

From the above, the calibration of the three-axial magnetometer can be concluded as the following steps:

(1) Measure the magnetic field L times in a uniform magnetic field and work out the mean value $Hl1'$. Then place the magnetometer in the opposite direction, repeat the steps above, and get a mean value $Hl2'$. We have

$$E = [Hl1' + Hl2']/2 \tag{15}$$

(2) Rotate the magnetometer K (K>5) times homogeneously in a uniform magnetic field [15]. Measure the magnetic field K times and work out the mean value at each attitude. We can get the measurement matrix $Hm'_{3\times K}$.

(3) According to equations (4) and (9), we can get $HM_{K\times 6}$.

(4) Then we can get the evaluation value of the magnetic field according to the formula

$$|H| = \frac{1}{K}\sum |Hm' - E| \tag{16}$$

(5) According to equations (13) and (14), we can work out S.

(6) Finally, according to equation (3), we can get α、β、γ、k_x、k_y、k_z.

4. Performance analysis

To begin with, simulation results are employed to analyze the performance of the proposed algorithm. Assume that the intensity of magnetic field is 41628nT. The mean squared measurement noise is 6nT, K=10, L=5. The orthogonal error of axis X, Y, Z is 0.5°, 1° and 1.5° respectively. Assume that the three axis's zero point offset errors have the same absolute value. |E| represents the

total zero point offset error. The relationship between the orthogonal error after calibration and |E| is shown in Fig. 2.

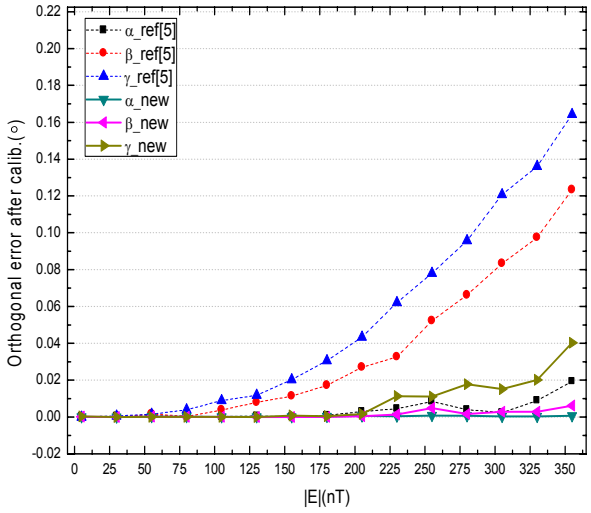


Fig. 2. The relationship between the orthogonal error after calibration and |E|

The relationship between calibration factor after calibration and total zero point offset error |E| is shown in Fig. 3.

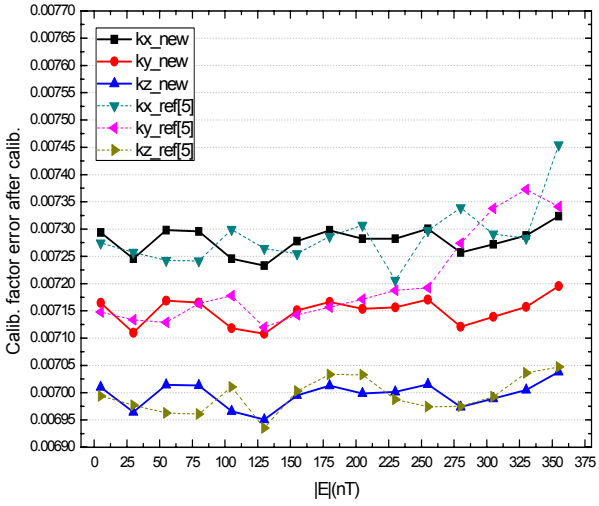


Fig. 3. The relationship between calibration factor after calibration and total zero point offset error |E|

According to Fig. 2, the proposed algorithm has the same performance as existing algorithms when the zero point offset error is small. However, when the total zero point offset error |E| is big, the proposed algorithm has an obvious advantage over the existing algorithms. As to calibration factor error, it is shown in Fig. 3 that the proposed algorithm also outperforms the existing algorithms.

5. Conclusion

The three-axial orthogonal calibration structure increases the weight of the probe, while the probe cannot avoid the three-axial orthogonal error. By replacing the mechanical platform with mathematical platform, we can reduce the volume, weight and manufacturing periodicity of the

satellite borne magnetometer, and enhance the reliability. In this work, the measurement model of vector fluxgate magnetometer is analyzed. The measurement errors and its calibration method are concluded. To solve the problems that the existing linearization method cannot bear big zero point offset, we proposed an improved linearization calibration method. By this means, the zero point offset error is figured out first, after that, other parameters are worked out. The simulation results show that the proposed method has an obvious advantage over the existing methods when the zero point offset is big.

Acknowledgement

This work is supported by National Nature Science Foundation of China (NSFC) with Project No. 61132002 and 61231011.

References

[1] QIAN Shan; ZHANG Shifeng; CAI Hong, "Study on Attitude Acquisition of Micro-Satellite with Only Magnetic Actuators", Journal of Astronautics,Vol. 32, No. 1, Pp. 72-80,Jan. 2011.

[2] YAN Hui, XIAO Changhan, ZHANG Chaoyang, "HorizontalCalibrationtoTriaxialFluxgateMagnetomete", Transducer andMicrosystem Technologies, Vol. 20, NO. 6, Pp. 2006.

[3] TAKAYA I, NOBUTADA S, et al. "Compensation of time-variable magnetic moments for a precise attitude control in nano and micro-satellite missions". Advances in Space Research. Vol. 48, Pp. 432-440, 2011.

[4] INAMORI T, SAKO N, NAKASUKA S. "Strategy of magnetometer calibration for nanosatellite missions and in-orbit performance". AIAA Guidance. Navigation and Control Conference. AIAA-2010-7598, Toronto Canada, August, 2010.

[5] Huang Lin, Jing Wuxing."Spacecraft attitude determination and three-axis magnetometer calibration". Journal of Astronautics, Vol. 29, No. 3, Pp. 854-859, 2008.

[6] ZHOU Jianjun, LIN Chunsheng, ZHANG Jian. "Correction of non-orthonormality of three-axis magnetometer based on ACO arithmetic", Ship Science and Technology, Vol. 33, No, 2, Pp. 106-109, 2011.

[7] WU Dehui, HUANG Songling, ZHAO Wei. "Research on Correction of tri-axial magnetometer based on FLANN", Chinese Journal of Science Instrument. Vol. 30, No. 3, Pp. 449-453, 2009.

[8] YANG Zhaoxin ,DU Hongmian, FAN Jingbiao, ZU Jing. "Dynamic Modeling of Acceleration Sensor Based on Generalized Least Square Method ", Chinese Journal of Sensors and Actuators. Vol. 22, No. 2, Pp. 195-198.

[9] ZHANG Qi, PAN Mengchun, CHEN Dixiang, PANG Hongfeng, "Scalar Calibration of Tri-Axial Magnetometer with Linearized Parameter Model", Chinese Journal of Sensors and Actuators. Vol. 25, No. 2, Pp. 215-219.

[10] PANG Hongfeng, Pan Mengchun, CHEN Dixiang, LUO Shitu, LUO Feilu, "Error Calibration of Three Axis Magnetometer Based on UKF and Equipment" Chinese Journal of Instruments, Vol. 33, No. 8, Pp. 1800-1805, 2012.

[11] TIAN Jing, WU Mei-ping, HU Xiao-ping,"Three-Axis Magnetometer Attitude Determination", Journal of National University of Defense Technology, Vol. 23, No. 5, Pp. 17-21, 2001.

[12] Zuo Wenji, Song Fuxiang, "An Approach to Autonomous Navigation Using Magnetic measurements for Small Satellite", Journal of Astron Autics. Vol. 21, No. 2. Pp. 100-104. 2000.

[13] WONMO KOO,SANGKYUNG SUNG,YOUNG JA-E LEE, "Error Calibration of Magnetometer Using Nonlinear Integrated Filter Model With Inertial Sensors", IEEE Transaction on Magnetics, Vol.45, No. 6, Pp. 2740-2743, 2009.
[14] Zhang Xiaoming; Gao Lizhen, "A Novel Auto-Calibr- tion Method of the Vector Magnetometer" The NinthInternational Conference on Electronic Measure-ment & Instruments, Vol. 1, Pp. 145-150, 2009.
[15] Vojtech, P.;Peter, K.; Merayo, J. M. G.,et al, "Non-Magnetic Platform for Scalar Calibration of Magnetometers", In Proc. of IEEE Sensors 2008 Conference, Pp. 1444-1447, 2008.

High-precision Control of Mechanical Flywheel Based on RBF-PID Algorithm with Disturbance Compensator

Zhang Cong, Ma Jun, Han Shi-yu, Wei Hui, Tang Xu-dong, Zhong Qi-liang

Beijing Mechanical Equipment Institute, Beijing, China
e-mail: haibou571@163.com

Abstract: Mechanical flywheel is widely used for satellite attitude-control in space that requires fine pointing accuracy and stability nowadays. But the inner disturbance torque and noises produced by the mechanical flywheel itself seriously deteriorate its output torque precision. Accordingly, this paper adopts the RBF-PID algorithm with a simple disturbance compensator for the torque mode control system of mechanical flywheel. Based on the dynamics of flywheel and the analysis of disturbance torque, the RBF-PID(Radial-basis-function, Proportional-integral-derivative) controller with a straightforward disturbance compensator are designed. A performance comparisons of the proposed methodology with conventional PID algorithm are performed to show the advantages of the former over the latter on many aspects, such as response time, overshot, anti-disturbance and so on.

Keywords: mechanical flywheel; satellite attitude-control; inner disturbance torque; RBF-PID algorithm; disturbance compensator; torque mode

I. Introduction

The development of space technology has improved the demands on the characteristics of satellite attitude-control system. Three-axis stabilization has been paid more attention in the field with mechanical flywheel as a significant actuator of satellite attitude-control system[1-2]. Permanent magnet brushless DC motor is favored in mechanical flywheel due to its superior efficiency, simple structure, no commutation spark, no excitation loss, high power density, torque-to-inertia and etc, that make it an adaptive choice for variable-speed direct-drive applications[3-4].

The flywheel system is in essence a momentum transfer and storage device, absorbing momentum form satellite. The steady-state behavior of the system in presence of a constant satellite's external disturbance torque is a contstant acceleration of flywheel, whose reaction torque on satellite is equal and opposite to satellite's external disturbance torque. In the absence of disturbances, constant speed is maintained. There are usually two types of work modes for mechanical flywheel: torque mode and speed mode. Nowadays, the existing mechanical flywheel systems adopt torque mode in a great measure, only realizing equivalency between instruction torque and electromagnetism torque[5-6]. However, there are multifarious disturbances and noises inside mechanical flywheel system, consisting of three parts mostly[7-9]: the first is inner damping torque, proportional to the rotation speed of mechanical flywheel approximately, such as air resistance torque, oil-velum resistance torque and magnetism resistance torque, etc; the second is Coulomb friction torque, whose symbol changes with the rotation direction; the third is noises torque of mechanical bearings, from holding framework, balls and rolling passage for ball bearing, that are related with lubrication mode and operation state. Among these, the damping torque and noises torque are small relatively, and constant or continuous variables with less effection on system.

The Coulomb friction torque is the most and reverses abruptly with amplitude doubled or more as mechanical flywheel changes from clockwise to anticlockwise rotation direction[10]. These disturbances and noises, especially friction disturbance torque inside flywheel, produces error for flywheel's output torque and disturbances for satellite attitude-control system.

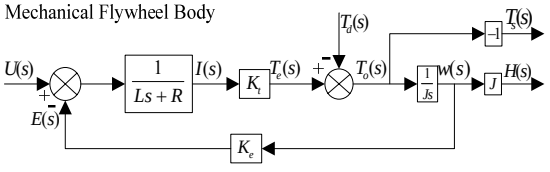
With the torque mode, conventional means obtains the feedback value of DC-link current for motor by a sampling resistance, and then adopts PID controller in the current closed-loop to regulate the error between the feedback value and the reference value. Essentially, the means imports the inner disturbances torque of mechanical flywheel on satellite attitude-control system, resulting in the deterioration of satellite's attitude control precision directly. Moreover, the classical PID controller is simple, reliable and convenient to engineering applications with outstanding control effect in less parametric variation[11]. But, permanent magnet motor applied in mechanical flywheel is a multi-vairable non-linear system, difficult to describe the electromagnetism relationship with accurate mathematical expression[12]. Hence, it's hard to obtain ideal effectiveness using the conventional control scheme.

Based on the above considerations, the anti disturbance capability is significant for mechanical flywheel torque mode control system. Therefore, this paper is attempting to adopt a method which can provides the satisfied performance of both command torque tracking and disturbance compensation ability, as well as keeps the control scheme to be simple and straightforward. In the presented methodology, a disturbance compensator with simple structure and convenient implementation is adopted to counteract the disturbance torque of mechanical flywheel. Moreover, PID controller parameters of mechanical flywheel are regulated automatically by a RBF neural network, that overcomes the disadvantages of conventional fixed parameters PID algorithms.

II. DYNAMICS OF MECHANICAL FLYWHEEL

In order to establish the dynamics model of mechanical flywheel, some assumptions are made as follows[13]:

1. Stator armature windings join with three-phase star type and three-phase windings are complete symmetry.
2. Neglecting the magnetism presses and the effect of eddy current and the magnetic hysteresis in core.
3. Don't consider the armature reaction and the distribution of air-gap magnetic field approximates trapezoidal wave with electrical angle 120° of flat top width.



Dynamics of mechanical flywheel

The complete dynamics of mechanical flywheel body can be described by the above Fig.1, where U(s): armature voltage; E(s): back EMF; I(s): armature current; $T_e(s)$: electromagnetism torque; $T_d(s)$: disturbance torque; $T_o(s)$: output torque of mechanical flywheel; $T_s(s)$: reaction torque to satellite; w(s): rotation speed; H(s): angular momentum; R: armature resistance; L: armature inductance; K_t: torque coefficient; K_e: back EMF coefficient and J: moment of inertia.

According to the Angular Momentum Theorem, the output torque of mechanical flywheel is expressed as:

$$T_o(s) = sH(s) = Jsw(s) = T_e(s) - T_d(s).$$

According to the Newton's Law of Action and Reaction, the reaction torque of mechanical flywheel on satellite is equal to the torque action on flywheel's rotor with opposite direction:

$$T_s(s) = -T_o(s) = -Jsw(s)$$

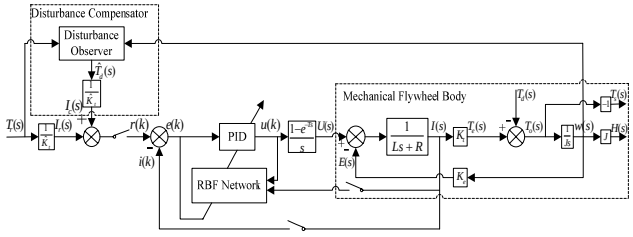
Based on the motor operation principle and voltage balance equation, we can obtain:

$$\begin{cases} T_e = K_t I(s) \\ E(s) = K_e w(s) \\ U(s) - E(s) = (Ls + R)I(s) \end{cases}.$$

The controlled voltage of motor is generated by controller to change the rotation speed of motor and store rotation momentum. Simultaneously, the reaction torque is operating on satellite so as to change the attitude of satellite.

III. CONTROL SYSTEM DESIGN

The torque mode control system diagram for mechanical flywheel adopted in this paper is displayed in the Fig.2, that is made up of three main parts: mechanical flywheel body, a disturbance compensator and RBF-PID controller, where, $T_r(s)$ is the reference instruction torque; r(k) is the reference input of controller; k_t^ is the nominal value of torque coefficient K_t; u is the output control variable of controller; i(k) is the DC-link current of mechanical flywheel motor and the reference input r(k) is sum of the reference current $I_r(s)$ converted by the instruction torque and disturbance compensation current $I_c(s)$ generated by the disturbance compensator, namely $r(k)= i_r(k)+ i_c(k)$.



Mechanical flywheel torque mode control systeml

Disturbance Observer Design

According to the equations1, the output torque is not only the error between the electromagnetism torque and the disturbance torque, but also proportional to the differential of the rotation speed. In other words, the error between actual measuring rotation speed and ideal rotation speed reflects the error between the output torque and the reference insturction torque. So, the designed disturbance observer obtains the ideal rotation speed by integral of reference instruction torque, compares it with the measured value of actual rotation speed, and then estimates the disturbance torque by a simple PD regulator that is strongly robust and performed easily[6]. The disturbance observer can be expressed as Fig.3 or the following equation.

$$\hat{T}_d(s) = \left[\frac{T_r(s)}{\hat{J}s} - w(s) \right](k_{cp} + k_{cd}s) \quad . \quad .$$

Where, J^ is the nominal value of moment of inertia J; k_{cp}, k_{cd} are proportional and differential coefficients of PD regulator, respectively.

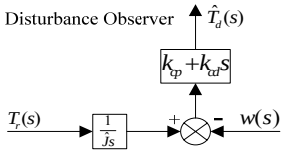


Diagram of disturbance observer

On-line RBF-PID Controller Design

The control system carries on on-line training and on-line control which can weaken the perturbation influence and improve the stability and reliability of the system. Conventional increment PID controller can be showed as:

$$u(k) = u(k\text{-}1) + k_p x(1) + k_i x(2) + k_d x(3) \quad . \quad .$$

Where, k_p, k_i, k_d are proportional, integral, differential coefficients respectively which can be optimized by RBF neural network in real time based on the operation state of control system, and

$$\begin{aligned} e(k) &= r(k) - i(k) \\ x(1) &= e(k) - e(k-1) \\ x(2) &= e(k) \\ x(3) &= e(k) - 2e(k-1) + e(k-2) \end{aligned} \quad . \quad .$$

The quadratic performance index is:

$$E(k) = \frac{1}{2}e^2(k) = \frac{1}{2}\left[r(k) - i(k)\right]^2 \quad . \quad .$$

The regulation of k_p, k_i, k_d adopts the following gradient descent method:

$$\Delta k_p = -\eta \frac{\partial E}{\partial k_p} = -\eta \frac{\partial E}{\partial i}\frac{\partial i}{\partial u}\frac{\partial u}{\partial k_p} = -\eta e(k)\frac{\partial i}{\partial u}x(1)$$

$$\Delta k_p = -\eta \frac{\partial E}{\partial k_i} = -\eta \frac{\partial E}{\partial i}\frac{\partial i}{\partial u}\frac{\partial u}{\partial k_i} = -\eta e(k)\frac{\partial i}{\partial u}x(2) \quad . \quad .$$

$$\Delta k_p = -\eta \frac{\partial E}{\partial k_d} = -\eta \frac{\partial E}{\partial i}\frac{\partial i}{\partial u}\frac{\partial u}{\partial k_d} = -\eta e(k)\frac{\partial i}{\partial u}x(3)$$

Where, $\frac{\partial i(k)}{\partial u(k)}$ is the unknown Jacobian information of controlled mechanical flywheel. In this paper, RBF neural network is used to identify the Jacobian information.

Identification of Mechanical Fflywheel's Jacobian Information

The RBF neural network is a three-layer network which reflects the input via the function of the hidden layer, namely the radial basis function[14]. The output of the whole network is the weight sum of hidden nodes outputs. In order to obtain better control effect, we need to know the approximate Jacobian information $\frac{\partial i_m(k)}{\partial u(k)}$ through RBF neural network to substitute the exact Jacobian information $\frac{\partial i(k)}{\partial u(k)}$. In the adopted 2-4-1 type RBF neural network structure,

$X=[u(k),i(k)]^T$ is the input vector of network, $H=[h_1,h_2,h_3,h_4]^T$ is the radial basis vector of the hiddent layer, where h_j is selected as Gauss basis function of the jth node, and denoted as:

$$h_j = \exp(-\frac{\left\|X - C_j\right\|^2}{2b_j^2}), (j = 1,2,3,4) \qquad . \quad .$$

Where, $C_j=[c_{j1},c_{j2}]^T$ is the center vector of the jth hidden node. Suppose $B=[b_1,b_2,b_3,b_4]^T$ is the basis width vector of the hidden layer, and b_j is the basis width of the jth node, $b_j>0$, while $W=[w_1,w_2,w_3,w_4]^T$ is the weight values vector of the output layer. Then, the output of identification network can be denoted as:

$$i_m(k) = w_1h_1 + w_2h_2 + w_3h_3 + w_4h_4 \qquad .\tilde{1}$$

The function of quadratic performance index for identifier can be expressed as:

$$J = \frac{1}{2}\left[i(k) - i_m(k)\right]^2 \qquad .11$$

According to gradient descent method, the iterative algorithm of w_j, b_j and c_{ji} can be calculated as:

$$\begin{aligned}
&w_j(k) = w_j(k-1) + \eta'(i(k) - i_m(k))h_j + \alpha(w_j(k-1) - w_j(k-2)) \\
&\Delta b_j = (i(k) - i_m(k))w_j h_j \frac{\left\|X - C_j\right\|^2}{b_j^3} \\
&b_j(k) = b_j(k-1) + \eta'\Delta b_j + \alpha(b_j(k-1) - b_j(k-2)) \\
&\Delta c_{ji} = (i(k) - i_m(k))w_j \frac{x_j - c_{ji}}{b_j^2} \\
&c_{ji}(k) = c_{ji}(k-1) + \eta'\Delta c_{ji} + \alpha(c_{ji}(k-1) - c_{ji}(k-2))
\end{aligned} \qquad .1 \quad .$$

Where, η' is learning rate, α is momentum factor.

So, the approximate Jacobian information can be expressed as:

$$\frac{\partial i(k)}{\partial u(k)} \approx \frac{\partial i_m(k)}{\partial u(k)} = \sum_{j=1}^{4} w_j h_j \frac{c_{ji} - u(k)}{b_j^2} \qquad .1 \quad .$$

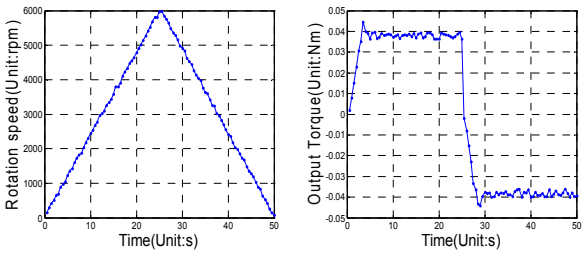
IV. SIMULATIONS OF MECHANICAL FLYWHEEL PROTOTYPE

To verify the effectiveness of the adopted methodology, the prototype simulationss of mechanical flywheel are implementd. For the motor adopted by the prototype, with rated power 60W, maximum speed 6000rpm, maximum output torque 0.06Nm and rated DC-link voltage 28V, the specifications are listed as follows: R=0.164Ω, L=53μH, J=0.00273kgm^2, kt=0.00172Nm/A and ke=0.00165V/(rpm).

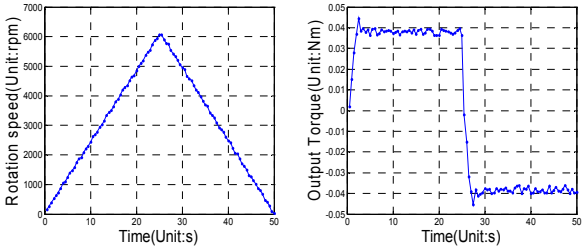
Conventional PID controller is compared with the RBF-PID controller with disturbance compensator in the speed response and torque response. PID controller parameters are set as: k_p=0.14, k_i=0.35, k_d=0.00018. In the RBF-PID controller with disturbance compensator, the PD regulator parameters of disturbance observer are set as: k_{cp}=0.65, k_{cd}=0.42, while the initial parameters of RBF-PID controller are set as: η=0.01, η'=0.02, α=0.005, $C_j=\begin{bmatrix} 0.2, & 0.2, & 0.2, & 0.2 \\ 0.2, & 0.2, & 0.2, & 0.2 \end{bmatrix}^T$, $bj=[1.2,1.2,1.2,1.2]^T$, $w_j=[0.1,0.1,0.1,0.1]^T$.

First of all, the reference instruction torque is set as Tr=0.04Nm, the mechanical flywheel begins to accelerate in the anticlockwise direction from the rest. And then, when the rotation speed reaches the maximum 6000rpm, the reference instruction torque changes from 0.04Nm to -0.04Nm. So, the mechanical flywheel decelerates until 0rpm in the anticlockwise direction from 6000rpm.

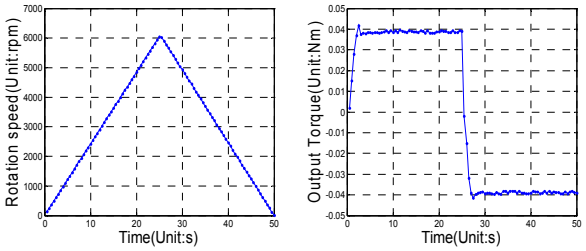
Fig.4 shows the speed and output torque response curves adopting traditional PID controller, and Fig.5 displays the speed response and output torque response curves adopting the proposed RBF-PID controller without the disturbance compensator, while Fig.6 reveals the speed response and output torque response curves adopting the proposed RBF-PID controller with the disturbance compensator.



Speed & output torque response curves with PID controller



Speed & output torque response curves with RBF-PID controller without disturbance compensator



Speed & output torque response curves with RBF-PID controller with disturbance compensator

As seen from the comparison of Fig.4 and Fig.5, the RBF-PID controller posseses less response time, less overshot than the conventional PID controller. From Fig.6 and Fig.6, the RBF-PID controller with the disturbance compensator is provided with better stability, stronger anti disarrangement capacity than the RBF-PID controller without the disturbance compensator.

V. conclusions

This paper adopts the RBF-PID algorithm with disturbance compensation as the current regulation arithmetic of mechanical flywheel torque mode control system. By comparison with conventional PID controller, experimental results demonstrate that the methodology can improve

the dynamic performance of the mechanical flywheel system with on-line tuning of the PID parameters and anti disturbance ability with the disturbance compensator.

References

[1] S. R. Valadi, H. S. Oh. Space station attitude control and momentum management: a nonlinear look[J]. Journal of Guidance, Control and Dynamics, 1992, 15:577-586.

[2] William Bialke. Low cost attitude control system reaction wheel development final report[R]. NASA-CR-191332, ITHACO, Ithaca, New York, 1991:230-240.

[3] D. C. Hanselman. Minimum torque ripple, maximum efficiency excitation of brushless permanent magnet motors[J]. IEEE Trans. On IE, 1994 , 41(3):292-300.

[4] Ishak D. Zhu Z Q, Howe. Permanent brushless machines with unequal tooth widths and similar slot and pole numbers[C]. IEEE Transactions on Industry Applications Conference, 2005, 2(7):1055-1061.

[5] He Linlin, Dou Manfeng. The reaction wheel control mode in the azimuth control of nacelle system[J]. Small & Special Electrical Machines, 2006 , 34(7):17-19,44.

[6] Chen Hao, Ge Shengmin, Liu Fuchen, et al. The design of torque mode control for reaction wheel system[J]. Journal of Astronautics, 2006, 27(6):1248-1253.

[7] Sun Xianjun, et al. Effects of reaction-wheel-induced disturbance on the attitude of three-axis-stabilized geostationary satellite[J]. Computer and Digital Engineering, 2005, 33(12):55-59.

[8] Masterson R A, Miller D W, Grogan R L. Development of empirical and analysis reaction wheel disturbance models. AIAA99-1204.

[9] Li Lianjun, Dai Jinhai. Inner disturbance modeling and simulation analysis of reaction wheel system[J]. Journal of System Simulation, 2005, 17(8):1855-1863.

[10] Wang Bingquan, Li Ying, Yang Di. A compensation observer method for the improvement of performance of low-speed reaction wheel[J]. Chinese Journal of Space Science, 1999, 19(4):362-367.

[11] Xia Changliang, Wang Juan, Shi Tingna, et al. Direct control of currents based on adaptive RBF neural network for brushless DC motors[J]. Proceedings of the CSEE, 2003, 23(6): 123-127.

[12] Niu Xinglin, Zhang Xinghua. Current tracking control of brushless DC motor[J]. Microcomputer Information, 2006, 22(11-1):97-98.

[13] ZABALAWI, S.A. and NASIRI A. State space modeling and simulation of sensorless control of brushless dc motors using instantaneous rotor position tracking[C]. Proceedings of the 2007 IEEE Vehicle Power and Propulsion Conference, 2007, : 90-94.

[14] Xia Changliang, Qi Wenya, Yang Rong. Identification and model reference adaptive control for ultrasonic motor based on RBF neural network[J]. Proceedings of the CSEE, 2004, 24(7):117-121.

SPACESTAR STAR TRACKER: Designed to be The Choice of Future Recurrent Platforms

F. Boldrini, G. Berrighi, D. Procopio

Selex ES S.p.A. – A Finmeccanica Company, Italy

Star Trackers experienced a strong evolution in the recent years, passing from configurations which required on-ground processing to obtain quaternion measurements to "fully autonomous" equipment like our AA-STR product, which is based on an Active Pixel Sensor detector (APS) and features a robust and accurate three axis attitude determination, offering customers the benefits of very low mass and power consumption with respect to conventional CCD units belonging to the same "class".

SPACESTAR represents an innovative, medium Field of View star tracker architecture - originally developed for a high volume constellation program - through an integrated configuration to Attitude Control System computers.

The SPACESTAR sensor consists of up to three Optical Heads, each one of which containing a baffle, an optical system, a focal plane and proximity electronics, and of a Software running in the Attitude Control System Computer of the spacecraft.

The new architecture was conceived to provide significant value to new space programs, including optimized hardware and the elimination of unnecessary redundant hardware (DC/DC conversion, star catalogs, processing elements, etc.) to: minimize recurring costs, significantly reduce size, weight and power, and to offer the advantages of system level efficiencies by providing a centralized control of multiple Star Tracker Optical Heads in the single OH units. The SPACESTAR system leverages standard interfaces whenever possible.

The SPACESTAR OH delivers "compressed sky images" to be processed via a dedicated SW routine that is hosted in the AOCS Computer. The SW performs all the computation necessary to convert the information coming from the Optical Heads, up to the quaternion solution delivery.

SPACESTAR's SW is optimized to match CPU time allocation (sharing resources with all the other satellite AOCS activities), without penalizing robustness and reliability of the system - through simultaneous management of up to three OHs per satellite - and introduces additional features for simultaneous management of up to three SPACESTAR Optical Heads per satellite.

The SPACESTAR product is an evolution of Selex ES's AA-STR Attitude Sensor with the necessary adaptations to minimize production costs. This strategy allowed preserving some very key design items of the already qualified AA-STR unit, reversing the electro-optical performance of the AA-STR to the new program while minimizing development risks.

SPACESTAR's integrated configuration with AOCS computers allows Customers to take advantage of significant benefits, including high resistance to EMC/EMI thanks to the complete elimination of analog electronics, significant size and mass reductions.

Optical Head HW on Ground Qualification and SPACESTAR SW integration in spacecraft Attitude Control Computer SW was achieved in 2013. This paper describes the challenges and lessons learned during the development of this product

1. INTRODUCTION

Star trackers are a central part of satellite attitude determination and control systems. Their ubiquitous use means that optimizations of these systems in terms of cost, size, weight, and power result in benefits realizable throughout the industry. Most star trackers today employ a standalone architecture, whereby each star tracker contains all functions necessary for their use, including DC/DC conversion, star catalogs and processing elements. Since most satellites use two or three star trackers, this standalone architecture drives inefficiencies as each star tracker essentially contains redundant hardware that could be integrated into a centralized controller. Few star trackers use a single electronics box, dedicated to manage multiple Optical Heads, but this configuration may be further optimized in case the ACS Computer directly interfaces the Optical Heads and runs "Multi-Head" star tracker software as a specific task within the attitude control system software. This further step of Star Trackers' configuration improvement (called "SPACESTAR") has been implemented within a contract for a high volume commercial constellation, where Selex ES (SES) has already delivered flight products.

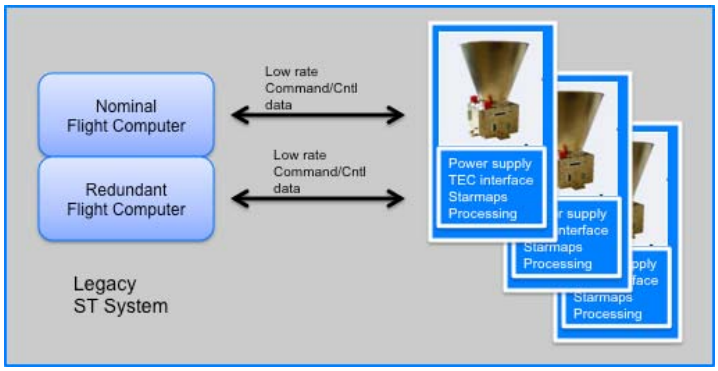


Figure 26. Legacy Star Trackers' Architecture

The SPACESTAR Optical Head is realized around the detection unit (optics plus focal plane) already utilized in the AA-STR star tracker (SES star tracker standard product), with the necessary adaptations to minimize production costs. This strategy allowed the preservation of some very key design items of the in-flight qualified AA-STR unit, allowing the electro-optical performance of the AA-STR to new programs, while minimizing development risks.

The SPACESTAR Optical Head delivers "compressed sky images" and is used with a dedicated SW routine, delivered by Selex ES, hosted in the spacecraft computer. The SPACESTAR SW is based on the AA-STR SW already qualified in-flight on the AlphaBus and other Telecommunication programs, and performs all the computation necessary to convert the information coming from the Optical Heads, up to the quaternion solution delivery. This approach has been used for a high volume commercial constellation contract where SES has already delivered more than one hundred flight units in support of the constellation. This experience in Star Tracking SW "porting" into the Spacecraft Computer (external processor) can be directly applied to new projects.

Figure 27. SPACESTAR Star Tracker Optical Head

2. SYSTEM OVERVIEW

The SPACESTAR sensor consists of up to three Optical Heads (each one containing a baffle, optical system, focal plane and proximity electronics) and a SW hosted and running in the Spacecraft Computer.

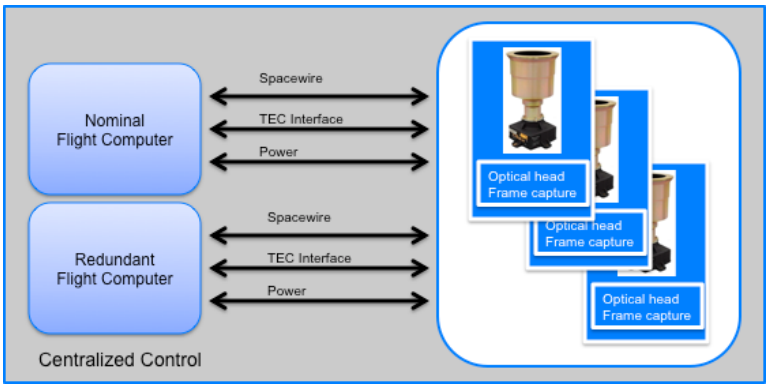


Figure 28. Centralized Star-Tracker Architecture.

SPACESTAR system has been designed to be robust to natural environments at LEO, MEO and GEO orbits, for typical commercial, GNSS and telecommunication satellites. Components are selected to be tolerant to 100 Krad exposure except for HAS detector (tested up to 42 Krads), which is physically located in the nucleus of the optical head, and therefore well shielded by structure and parts (typically the detector is shielded by about 15 mm of equivalent Aluminium).

Transient events induced by particles are filtered out by dedicated SW algorithms, making the tracking mode robust to worst peak solar flares. The acquisition mode, where the whole detector must be explored and no a-priori information is available, works with a SEU rate compatible with a worst day solar peak model (CREME96).

3. SYSTEM ARCHITECTURE

The SPACESTAR system offers system level efficiencies by providing a centralized control of multiple star tracker optical heads. This eliminates the standalone electronics provided by each star tracker when used in a 3 OHs configuration. The redundant features are transitioned to the spacecraft flight computer where the inherent redundancy of the flight computer system can be utilized as the redundant electronics for the star trackers. Though the SPACESTAR system leverages standard interfaces when possible, spacecraft flight computers might need to be modified to accommodate the SPACESTAR's electrical interfaces.

In fact the SPACESTAR Optical Head includes only the detection module (optics plus focal plane), the SPACESTAR OH electronics, and the baffle (attached to the Optical Head Structure through a dedicated thermally isolating spacer).

The SPACESTAR OH electronics is limited to 2 PCBs, one directly connected to the focal plane assembly and supporting only some bias and filter circuitry for the APS detector, the other supporting the digital electronics in a radiation hardened FPGA. Therefore the SPACESTAR OH interface to the "external World" is through two Spacewire connectors, one power connector (including Thermo-Electric Cooler control) and a test connector. These connectors constitute the interface to the spacecraft AOCS flight computer.

Spacecraft flight computer requirements for accommodating the SPACESTAR configuration

In order to accommodate the SPACESTAR Star Tracker Configuration, the spacecraft flight computer shall provide three SpaceWire ports, to provide a direct connection to the ST Optical Heads. A diagram of the SpaceWire interface (including cross-strapping) is shown in Figure 29.

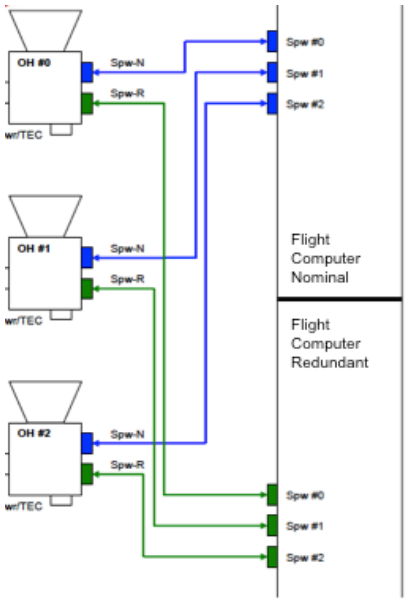


Figure 29. SpaceWire Interface Diagram

The spacecraft flight computer also needs to accommodate powering of three SPACESTAR Optical Heads. Three OH power interfaces capable of supporting 2 watts each shall be foreseen, including spacecraft cabling that supports the interfaces between the three OHs and the primary and redundant interface cards. A diagram of the power interface is shown in Figure 30.

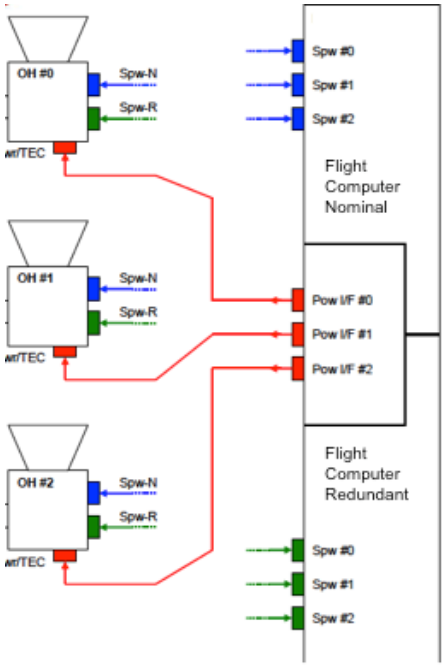


Figure 30. Power Interface

Finally, a Thermo Electric Cooler (TEC) control interface is required in the spacecraft for each OH. The TECs' interface shall contain three analog control interfaces to control up to three OHs. A diagram of each of the required TEC interface is shown in Figure 31.

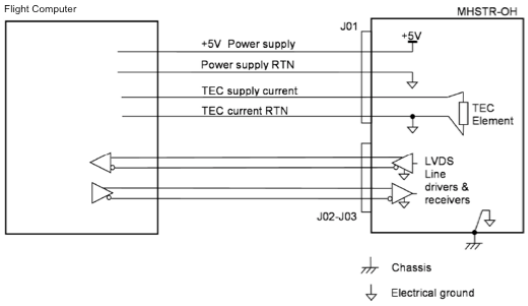


Figure 31. TEC Interface Diagram

4. SPACESTAR SW

The SPACESTAR SW provided by SES together with the SPACESTAR product is a real time SW. It has to be managed as a task in the frame of the spacecraft central computer AOCS SW.

SPACESTAR SW will be executed cyclically by the AOCS computer every 100 ms (or another lower rate, to be agreed with the customer) in case of tracking, providing an attitude refresh rate of up to 10 Hz.

In acquisition, which is the operative mode devoted to the initialisation of tracking, more than one cycle are necessary for the pattern recognition (a few seconds are overall necessary). Moreover, acquisition is efficient when longer exposure time is possible (200ms), for this reason a relaxed call rate is required during this operative mode (e.g. 5 Hz instead of 10 Hz). In other words, AOCS computer shall be able to manage the calls to the SPACESTAR SW task at 5Hz instead of 10Hz and leave it work for consecutive 50ms instead of 25ms when it is in acquisition mode (reported in the communication interface).

In tracking mode the SW is able to conveniently manage the data available from the OHs, swapping each other and avoiding outage of information. It is also possible to let the system working with one OH at the time, i.e. in the "traditional" star tracker usage, thus deciding at avionics computer level the "switching policy".

The SPACESTAR SW monitors at each AOCS cycle the status of the OHs. Thus in case one OH is no longer capable to provide good attitude measurement (for example because it gets blinded by the Sun/Earth entrance in the FOV), the SPACESTAR SW autonomously reacts in order to guarantee the maximum performance at "star tracker system" level. In this way outage of attitude measurement is avoided (the platform of the high volume constellation program where the SPACESTAR is used is "gyroless"), and reacquisition of attitude is not needed, also minimizing CPU load, since attitude acquisition from "lost in space" requires a computational capability significantly higher than maintaining tracking.

The SPACESTAR SW includes also as calibration parameters (from the on ground AIT campaign) the transformation matrices that describe the Optical Heads orientations with respect to an external reference frame, defined by customer. SPACESTAR system will always deliver the attitude data (w.r.t. J2000 frame) referred to the same MHBRF frame, irrespective to the OH currently in use. The matrices computed during the AIT phase and implemented in the SW are used for this reference frame change.

During the first cycles of flight activity, the direct tracking entrance of "secondary" OHs will be managed using the AIT matrices (affected by the on ground measurements and the ground to orbit transfer mechanical stimuli). In order to increase the robustness of the system, the attitude data measured by the primary and secondary OH during flight will be used also to update the misalignment matrices among the OH, allowing a refinement of the AIT parameters.

The SPACESTAR SW (provided by SES) manages the available optical heads based on two main operative modes:

- Normal Mode (for moderate angular rate < 1.5 deg/sec)
- High Angular Rate Mode.

A third mode may be defined together with the Customer targeting the minimum spacecraft flight computer CPU usage, thus fully tailored to the specific mission/platform needs. An example is provided as reference in the following as "minimum CPU usage" mode.

Normal Mode

In Normal Mode, two (or three) OH with 9/8 stars per OH can be used. Exact number of tracked stars per OH is defined on the basis of the CPU budget allocated to the SPACESTAR SW execution.

For each OH, a separated quaternion is computed. Data fusion is done at quaternion level, with autonomous calibration of misalignment among optical heads (i.e. recovery of initial offsets and Thermo Elastic Deformations). The following figure shows the key blocks of the Normal Mode function.

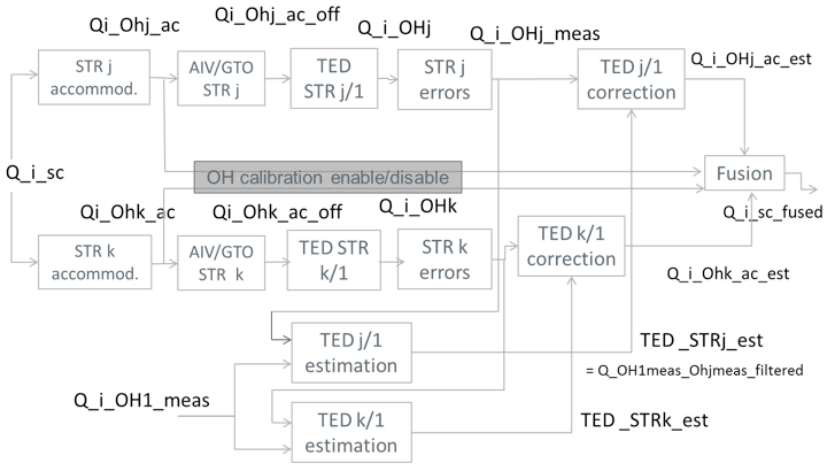


Figure 32. Normal Mode Block Diagram

As previously mentioned, the SPACESTAR SW includes also as calibration parameters (from the on ground AIT campaign) the transformation matrices describing the Optical Heads orientations with respect to an external reference frame, defined by customer (called MHBRF). Starting from this initial alignment information of the OHs on the spacecraft, the SPACESTAR SW refines the misalignments removing the offset errors (due to ground AIV and GTO effects) and the thermo elastic effects.

SPACESTAR system will always deliver the attitude data (w.r.t. J2000 frame) referred to the same MHBRF frame, irrespective to the OH(s) currently in use. The matrices computed during the AIT phase and implemented in the SW are used for this reference frame change.

Since misalignment measurement is performed using measured quaternions from the OHs, the misalignment knowledge is affected by star tracker measurement errors. For this reason, the misalignment estimation is performed using a Kalman Filter, so that high frequency and temporal noise effects are minimized.

This Normal Mode will be used when the SPACESTAR is operating at low spacecraft angular rates.

High Angular Rate Mode

The High Angular Rate Mode (HARM) can be either autonomously entered by the SPACESTAR SW when the measured angular rate is higher than a threshold, or be commanded by the spacecraft avionics.

In this mode the SPACESTAR SW acquires star data measurements from each available OH and performs data fusion at star level. In this way, the robustness against angular rate is strongly improved.

In fact, in the normal mode, a quaternion is computed for each optical head. This means that each OH has to detect at least three (depending on SW setting) stars. In the HARM mode it is sufficient to have three stars, however distributed over the available OHs, to maintain the SPACESTAR system fully functioning.

In the HARM mode the misalignment estimation among OHs is kept frozen (the last values computed in the Normal Mode or the default values are used in this phase). Misalignments estimation is recovered as soon as the Normal Mode is entered.

Minimum CPU Usage Configuration

This mode is suitable for spacecraft configurations where 2 OHs are present (or even 3, for further redundancy), minimizing the spacecraft central computer CPU load but at the same time providing good accuracy and continuous attitude measurement capability.

One example of operating principle, which might be a good approach for telecommunication platforms, is given in the following. Clearly specific software tailoring to the mission/platform needs is possible.

In case two SPACESTAR OHs are available, the SPACESTAR SW manages them setting a "primary" and a "secondary" OH. The SPACESTAR SW computes the attitude quaternion from the Inertial reference frame (IRF) to the OH Boresight Reference Frame (OHBRF) of the primary OH using up to 10 stars, to achieve full accuracy, while the secondary OH will use only 5 stars in tracking, to minimise CPU load. The SPACESTAR SW monitors at each frame the misalignment between the primary and secondary OH, so that in case for any reasons the primary OH is no longer capable to provide attitude measurement (for example because blinded by the Sun/Earth entrance in the FOV), the SPACESTAR SW autonomously switches the secondary OH as primary OH, increasing the number of stars in tracking from 5 to 10. In this way, outage of attitude measurement is avoided, and reacquisition of attitude is not needed, minimising also CPU load, as attitude acquisition requires a computational capability significantly higher than tracking. In case also a third OH hardware is available on the satellite, it may be kept in stand-by and, when the secondary OH is set by the SPACESTAR SW as primary OH, the third OH may be set as "new" secondary OH.

During the first cycles of flight activity, the direct tracking entrance of secondary OH will be managed using the AIT matrices (calibration parameters described above) affected by the on ground measurements and the ground to orbit transfer mechanical stimuli. In order to increase the robustness of the system, the attitude data measured by the primary and secondary OH during flight will be used also to update the misalignment matrices among the OH and to allow a refinement of the AIT parameters.

Moreover, the system is able to estimate the performance of the OHs that are in tracking and detect a degradation of accuracy higher than expected. When that happens, the SPACESTAR autonomously decides to switch the primary OH to the other with better performance, if available (for example if the sun is entering in the FOV of the primary OH, the SPACESTAR SW realizes this degradation and reacts before losing the tracking).

As ancillary information, also the attitude and star measurement data in the OH reference systems are provided in telemetry.

In this Minimum CPU Usage configuration, assuming a Leon 3 μP based spacecraft central computer, the tracking mode (AT) needs 20-25ms of CPU availability, while 50 ms of CPU time availability every 200 ms are needed during acquisition mode (AAD).

5. OPTICAL HEAD HARDWARE DESCRIPTION

The SPACESTAR optical head is built around the optics and focal plane already developed for the ALPHABUS AA-STR star tracker product (successfully operating in space since July 2013).

The optical system is assembled in a titanium structure, to match for thermal expansion coefficient, allowing optics performance maintenance over a wide temperature range.

The optical barrel supports also the focal plane assembly, realising the electro-optical module, following a well consolidated design criteria, adopted in the past in all the SES STR designs.

In this way a steady and rigid detection module is obtained, allowing for high stability of the sensor performance over a large temperature range.

The detection module (optics plus focal plane) is included in an OH structure that support also the SPACESTAR OH electronics and, through a dedicated thermally isolating spacer, the baffle.

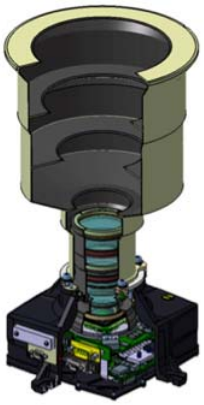


Figure 33. SPACESTAR OH Cutaway View

The electronics is limited into 2 PCBs, one directly connected to the focal plane assembly and supporting only some bias and filter circuitry for the APS detector, the other supporting the digital electronics in an RAD-HARD FPGA.

Two Spacewire connectors, one power connector and a test connector, connected by means of flex connections to the PCB are then directly assembled on the external case.

All the connectors are positioned on the same OH side, to simplify accessibility to the OH during integration on the Spacecraft.

The sensor layout and parts design is made in order to minimise the manufacturing and assembly cost, maintaining low mass and size.

With respect to the configuration under manufacturing for the large constellation contract, the following customizations are available:

1. An alignment cube is included, minimizing bias error due to alignment measurements
2. Envelope structure implementing an increased thickness, thus improving tolerance to radiation of the electronic components. Each SPACESTAR OH for the large constellation contract has a mass of about 1400g; the modified configuration has a mass of about 1600g
3. “Grade 1” EEE components suitable for 18 years GEO missions

Optical system

The optical system is manufactured with radiation resistance glasses, hard mounted in the titanium structure, without use of adhesives or optical cements. The same design already applied in the Stellar Reference Unit for the NASA Cassini mission, in the Star Tracker for the ESA Rosetta mission and in the CCD based A-STR product of SES, was adapted to the AA-STR and reused for the SPACESTAR. The focal length has been selected to produce a 20° full cone FOV on the HAS APS detector, and the pupil area guarantees, with margin, the sensitivity necessary to ensure that enough stars are always visible in the FOV. In addition, the internal layouts of the barrel and the rings have been optimized to improve the stray light performance.

Baffle

The sensor structure supports the baffle to prevent from stray light produced by the Sun, Earth and Moon. The SPACESTAR baffle has been designed for the following exclusions angles:

Table 2. Exclusion Angles.

Item	Value
Sun Exclusion Angle	26 Deg, 1 AU
Earth limb exclusion angle	20 Deg, 36000 Km
Moon limb exclusion angle	15 Deg, 380000 Km

Measurements confirm that the SPACESTAR baffle design is able to work correctly in worst case (Sun at 26deg) reducing the light to a PST (Point Source Transmittance) in the order of 10^{-7}. That leads to a maximum straylight signal of 30 Ke/sec/pixel, included in the accuracy figures presented in the next sections.

The SPACESTAR can operate with the Moon in FOV, with reduced accuracy (less than 30% of attitude error increase when Moon is present). No memory effect is induced by Moon presence, full performance are available as soon the Moon leaves the FOV and reaches its designed exclusion angle.

The baffle is produced with the standard process widely used in the past by SES, and with already large flight proven heritage.

Focal Plane

The APS detector, the thermoelectric cooler and the baseplate for the attachment to the main structure constitute the focal plane assembly.

The HAS2 APS detector used in the SPACESTAR OH is a 1024x1024 pixels detector, especially developed by On Semi (Belgium) for star tracker applications in severe radiation environments. The characterisation and evaluation tests, performed in the frame of an ESA contract, confirmed the expected improvement of electro-optical performance of the HAS2 APS with respect to the earlier STAR1000 APS: all the targeted requirements have been achieved.

The “hot pixels” and relevant dark current non uniformity greatly reduces as APS is cooled.

This reduction, more than the reduction of the average dark current which is already very low in the HAS2 APS, dominates the selection of the APS operating temperature. In addition to APS cooling, presence of “hot pixels” is managed at SW level to minimise the effects of the residual hot pixels after cooling is performed.

Peltier Control Loop

The APS cooling is achieved by means of a two stage Thermo Electric Cooler (TEC). The power dissipated by the TEC is then sunk towards the sensor feet, through a dedicated aluminium thermal strap.

The APS temperature is controlled by means of the SPACESTAR SW running in the spacecraft Computer.

At each cycle the SPACESTAR SW reads the OH focal plane temperature and calculates the percentage of the full input voltage (3.3V) that the OH shall receive at the TEC input in order to keep the correct focal plane temperature. This value is provided at SW interface.

Moreover, the Peltier Control Loop design implements a SW limitation mechanism, through an estimation of the TEC power consumption performed by the SPACESTAR SW itself on the basis of the voltage measurement across the TEC. That makes possible to limit the maximum power consumption in the TEC line, in agreement with the thermal interface and analysis for the project environment, to a level sufficient to maintain (also in the worst hot case) high efficiency of the thermal control loop.

6. THERMO-MECHANICAL INTERFACES

The SPACESTAR OH thermal design is based on the concept that the sensor is conductively coupled to the S/C baseplate, the baffle is thermally decoupled from the sensor structure and the internally produced heat is dissipated trough the mounting feet to the baseplate, acting also as heat sink. The sensor is assumed to be covered by MLI, for thermal control. The sensor housing can be provided also with the external housing black painted in case some radiative exchange is requested with a surrounding cavity.

The SPACESTAR OH envelope dimensions are:

- Length: 290.5 mm
- Housing: 164 mm x 164 mm (with feet protrusions)
- Baffle diameter: 163 mm

The mass budget for each Optical Head is less than 1600g (GEO configuration).

7. POWER INTERFACES

The overall SPACESTAR OH power consumption comes from two separate lines:

- Electronics (on 5V DC line)
- Peltier (on 0-3.3V DC line)

Power consumption of each OH electronics is 1.5 W (peltier power consumption is not included).

The Peltier power consumption depends on the mission thermal environment and target temperature of APS (recommended to be kept lower than +15°C). Usually, a trade-off between thermal and performance analyses dedicated to the project leads the right sizing of this parameter; however the order of magnitude of the maximum power consumption to be added to the OH electronics one is about 3 Watts.

8. SW INTERFACE

The SPACESTAR SW exchanges data with the AOCS SW through data structures stored in the spacecraft central computer RAM (see Figure 34. SW interfaces). The SPACESTAR SW receives as inputs:

-the calibrations parameters (related to the on ground test campaign, for example the optics distortion parameters and the misalignments among the OH and the MHBRF measured during the AIT phase at satellite level)

-the TC from the AOCS COMPUTER (among them, the spacecraft linear velocity has been separated to be updated frequently)

-the data from each OH

-the status of the peltier power supply that signals out the overcurrent event reached on the OH

-the on board time (for time tagging and propagation purposes)

and generates as outputs:

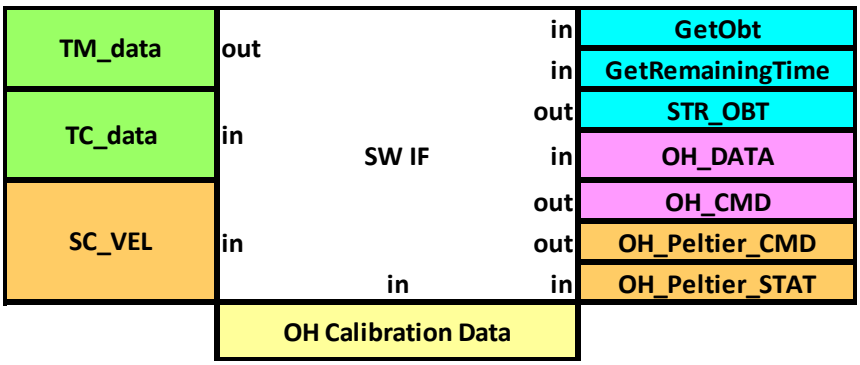
-data related to each OH including all the measurement data in the OHBRF (attitude and star information) and information related to the attitude measurement on the primary OH to be used for attitude control (referred to the MHBRF system)

-the command stream for each OH; the AOCS computer will send it to the OH in order to command the FPGA according the SPACESTAR SW requirements

-data related to the necessary electrical voltage required by each OH for the TEC control

-the remaining time till the end of processing

Particular attention is paid in the misalignment matrixes that allow to command the direct tracking from one OH to the other. This matrix is updated autonomously by the SW when the attitude data on the two OH are "sufficiently" good; that represents an in-flight refinement of the measurements performed on ground. This matrix can periodically be saved by the AOCS computer and, in case of system restart; the AOCS computer itself will initialise the right structures in order to provide the last in flight measurements to the SPACESTAR SW.



"in" areas are used to provides inputs from the AOCS SW to the SPACESTAR SW;
"out" areas are used to provide outputs from the SPACESTAR SW to the AOCS SW.

Figure 34. SW interfaces

The AOCS SW will be in charge to receive the OH's data from the SpaceWire interfaces and to reverse them in the dedicated OH_data according to a defined format. Additionally, the AOCS SW is in charge to transmit to the OH's, via SpaceWire interfaces, the command according to the contents of the OH_CMD memory area.

Timing and Synchronization

Each cycle the SPACESTAR SW prepares, on the basis of the previous results and current command from the AOCS COMPUTER, a command stream for each OH. The AOCS COMPUTER has the task to synchronise the delivery of the streams at the beginning of the cycle. Each FPGA starts the operations related to the current cycle at the end of the stream command receiving (End of Command).

During the APS readout time, the FPGA processes the pixel data and send (in not continuously way) the results of computation (segment and energy data related to overthreshold pixels) to the AOCS COMPUTER through the spacewire links. These operations take few milliseconds (12ms at maximum) in case of tracking cycles, while they take about 35 ms during the acquisition mode. It is important that the SPACESTAR SW is called after the complete receiving of data from each OH, in order to have available all the information to process.

The picture below is showing an example of tracking cycles assuming a 100 ms cycle, with the 25 ms of CPU time, the exposure times, and the time slots related to the readout and storage.

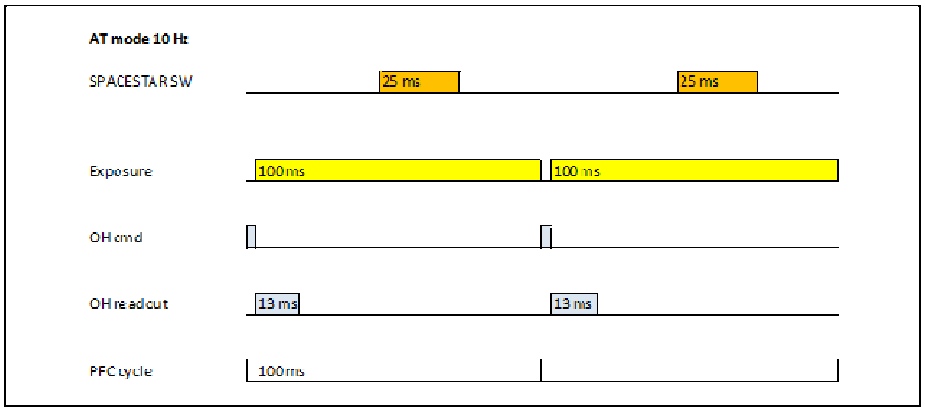


Figure 35. Timing representation during tracking cycles

In acquisition mode, the cycles are longer (200 ms instead of 100 ms) and the time slot dedicated to SPACESTAR SW is 50 ms instead of 25 ms. Information about duration of cycle are present in the OH command data coming from SW to OH (through AOCS computer) at the beginning of each cycle.

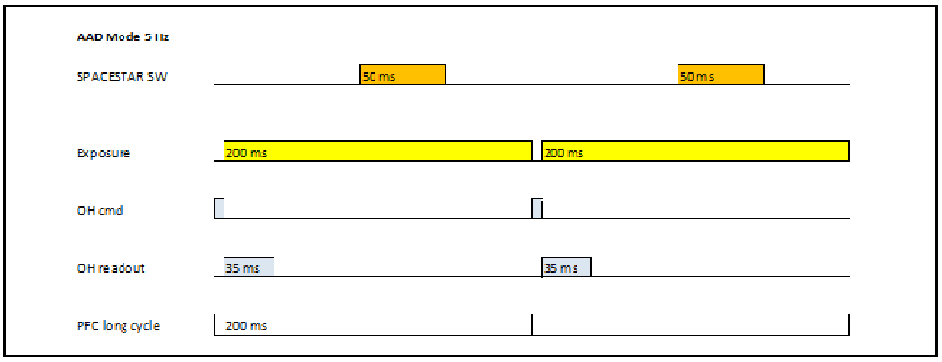


Figure 36. Timing representation during searching cycles

Remark: as shown in the pictures above, time necessary for transmission between OHs and memory where SPACESTAR SW store and read data is not included in the CPU slots (of 25ms in tracking and 50ms in acquisition) required for SPACESTAR task execution.

The SPACESTAR SW refers to the On Board Time (data provided by the AOCS COMPUTER) to compute calculations aimed to a precise time stamping of attitude data and to the internal propagation of expected star positions of the next cycles. That allows also to be robust to a lack of one cycle call of SPACESTAR SW (given for example by a higher priority task for AOCS COMPUTER).

9. SPACESTAR PERFORMANCE

BIAS ERROR

The attitude measurement is performed with respect to the reference frame of the one SPACESTAR Optical Head, previously defined as master reference (BRF_STR1).

The attitude measurements of the other Optical Heads are projected in the BRF_STR1, based on (on ground or in flight) alignment matrices estimation. The computed fused attitude is then projected in MSBRF (aligned to the S/C reference frame) using constant matrix.

The uncertainties in the knowledge of the alignment frame with respect to the functional frame of the master head will remain therefore also in the bias error of the fusion. The following bias values are applicable also to the case of data fusion (Normal Mode):

- Pitch/Yaw: 8 arcsec (3 σ)
- Roll: 11 arcsec (3 σ)

LOW FREQUENCY ERROR

The low frequency error in the data fusion (Normal Mode) depends on the OH mutual orientation and on the misalignment stability among the OH. This stability is affected by two contributions:

- Boresight stability of the star trackers (stability of measurement frame w.r.t. sensor feet)
- Mounting plates stability (stability of each mounting plate w.r.t other OH plates)

The second contribution is not under the SPACESTAR responsibility, therefore is not included in the budget. The data fusion accuracy depends on the relative alignment between the OHs.

Table 3. Low frequency error

Baselate Temperature Range	Error (3σ) Pitch/Yaw	Error (3σ) Roll
±5°C around T_{nom}:	2.5 arcsec	4.5 arcsec
From -25°C to +50°C	10.2 arcsec	4.9 arcsec

Values reported above assume that SPACESTAR OHs are oriented on the spacecraft in a "tulip" configuration, with boresights at about 90 degrees one with respect to the other. The low frequency error is computed assuming to perform the data fusion using data from 2 OHs only (worst case).

HIGH FREQUENCY ERROR

The High frequency error resulting from the data fusion of two OH is reported in the following. The table assumes that raw data are used, without filtering.

Table 4. High Frequency Error

Spacecraft Rate	X (arcsec, 3σ)	Y (arcsec, 3σ)	Z (arcsec, 3σ)
0.1 deg/sec	7.5	8.4	7.2
0.5 deg/sec	10.8	11.6	9.9
1 deg/sec	17.3	18.6	15.9
4 deg/sec	45	45	45

10. PRODUCT DEVELOPMENT AND IMPLEMENTATION ON THE PLATFORM

Thanks to the fact that the SPACESTAR product has its roots in the space qualified AA-STR, both for core hardware elements and SW routines, the relevant development went very smoothly and the qualification of the hardware and software was achieved already in 2013. At this time the production phase is running, with more than 100 Flight Models already delivered to the high volume constellation customer.

The SPACESTAR has been qualified to the following environmental levels (which are not necessarily the design limits):

Table 5. Environmental Levels used during qualification

Temperature:

Temperature limits	Min Op	Max Op	Min Non-Op	Max Non-Op	Min Start up
Qualification limits	-30	+60	-35	+65	-35
Acceptance limits	-25	+55	-30	+60	-30

Sine Vibrations:

	⊥ Mounting Plane		// Mounting Plane	
	Range (Hz)	Level	Range (Hz)	Level
Zone STD	Standard equipement			
Sine	5 - 22 22 - 100	+/-10mm +/-20 g	5 - 22 22 - 100	+/-10mm +/-20 g

Random Vibrations:

Mounting Plane		// Mounting Plane	
Range(Hz)	Level	Range (Hz)	Level
STAR TRACKERS			
20 - 100 100 - 500 500 - 2000	+ 6 dB/oct 0.5 g²/Hz -6 dB/oct	20 - 100 100 - 500 500 - 2000	+ 6 dB/oct 0.5 g²/Hz -6 dB/oct
Global:	20.1 g RMS	Global:	20.1 g RMS

Shock:

	IN PLANE		OUT OF PLANE
Freq. [Hz]	Acc. [g]	Freq. [Hz]	Acc. [g]
100	20	100	20
1000	1200	1000	1200
10000	1200	10000	1200

Hardware Validation

New set ups have been developed for acceptance test of SPACESTAR star trackers. The design of set up derives from the already available facilities used on our standard A-STR and AA-STR products, fulfilling the necessary accuracy suitable for the application and introducing some modifications aimed to optimize, and therefore shorten, the test campaign duration.

In particular the new facilities are ready to accommodate an high rate of production without impacting the SES capability to test and delivery star trackers for single programs demanding higher performance, for which a longer test campaign and calibration activity are usually required.

The goal of implementing a specific validation line, suitable for SPACESTAR Optical Heads, has been achieved dedicating a clean room area that hosts all the necessary items for the final tests (see Figure 37and Figure 38). This clean room area has benches and ground support equipment designed to group more Optical Heads in the same test facility, or offering the opportunity for parallel execution of tests on more benches.

Manufacturing (performed in the central Selex ES mechanical and optical workshops, which can support high volumes productions) and integration of the SPACESTAR Optical Heads hardware was also studied not to conflict with other products of the Selex ES Space Line of Business. Finally, the achieved automatisms in tests execution and reports generation allow an effective testing time for an optimal costs to performance ratio.

Figure 37. TVac test bench (for acceptance test)

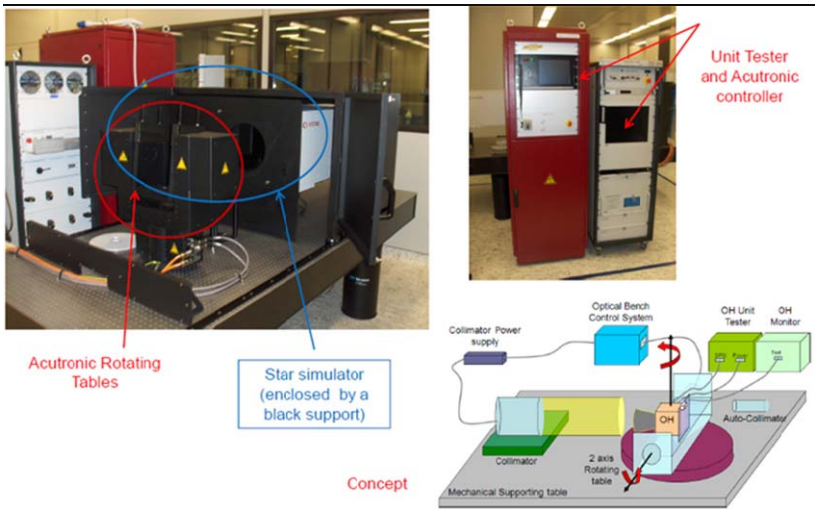


Figure 38. Optical bench (for calibration)

Software validation

The verification and validation of SPACESTAR SW have been executed in three main steps, in joint collaboration with the spacecraft Prime.

- In the first step the SPACESTAR SW "task" has been developed and validated in "stand alone" at SES premises, using a high fidelity emulator of the spacecraft computer that runs the AOCS SW;
- In the second step the SPACESTAR SW "task" has been integrated in the "real" spacecraft computer and further debugged at Prime premises, with the support of the SES experts;
- In the third and final step the SPACESTAR SW "task" has been embedded into the AOCS software and final verifications in the "real" spacecraft computer environment have been carried out, again with the support of the SES experts at Prime premises.

It is extremely important that since the very beginning all the activities (including ICDs definition) are executed in close collaboration with the Prime. For example "incremental" SPACESTAR SW deliveries were agreed among the Parties, allowing parallel developments at SES and at satellite Prime level:

- A SW version that only reproduced the “external interfaces” of the SPACESTAR SW was delivered first, allowing the Prime to work on the preparation of data structures to be shared in the spacecraft central computer RAM;
- SW emulators of the “basic” SPACESTAR OHs data exchanges (in the real case via SpaceWire links) were delivered in the initial stages (exchanging static data) to support the early phases of the AOCS SW development.
- Afterwards high fidelity SW emulators of the SPACESTAR OHs replaced the simplified ones to fully replicate the SPACESTAR behaviour before the EM hardware was shipped to the Prime. At that point the EMs availability allowed a full hardware in the loop verification. It is worthy to highlight that SES have now developed a High Fidelity HW emulator of the SPACESTAR OHs, which may be delivered to customers (even on loan basis) to allow a full “hardware in the loop” SPACESTAR system emulation at Prime level.

In more details, during the initial SW development phases at SES, an evaluation board (GR-UT699) of the Leon 3 based Platform Computer was used in conjunction with some Ground Support Equipments, to be able to reconstruct an environment similar to the one present in the spacecraft central computer. Optical Heads were instead replaced by HW representative breadboards plus an Electrical Stimuli Generator, the latter for efficiently simulating all the sky and satellite dynamics conditions (a duplicate of this HW Breadboards+ESG set-up may also be delivered to the Prime in support of their AOCS development activities).

The set-up was completed by one equipment able to communicate with the SW interface as well as with the AOCS computer, plus one equipment able to enter in SW variables and put break points in the running task.

In conclusion, the representativeness of the set up and the robustness of the interfaces, based on a dedicated slot of CPU time and a shared RAM area with well-defined read/write permissions, strongly facilitated the integration of the SPACESTAR SW into the spacecraft AOCS SW.

Moreover, thanks the approach here above summarized, the process of porting and integrating the SPACESTAR SW into the spacecraft AOCS computer went quite smoothly and both parties (i.e. SES and the Satellite Prime) finally agreed that it was “easier than expected”. Only few ordinary modifications resulted necessary after the first integration phase, and occurred in a second release of the SW.

11. CONCLUSION

An opportunity exists to utilize a cost optimized Star Tracker architecture developed for a high volume commercial constellation program. This concept, pioneered by Selex ES in the SPACESTAR product, has the potential to improve Star Trackers’ cost, size, weight, and power. Selex ES is proud to offer it to the GNC community as a fully qualified commodity, confident that its future spreading will be further enabled by the constant growth in the spacecraft central computers micro-processing power.

This concept has very high potentials to become the new standard for the AOCS of the future. Many people have contributed to its success, both at Selex ES and at our Customer, and the authors wish to thank all of them.

Sun Sensor on-a-chip: progress and final product capabilities

F. Boldrini[1], S. Brogi[1], D. Procopio[1]

[1]Selex ES S.p.A. – A Finmeccanica Company, Italy

The Sun Sensor on a Chip (SSoC) is a fully digital Sun sensor that provides the pixel coordinates of the Sun spot imaged through a pinhole on an image detector in the visible range. This concept has been already demonstrated and flight proven by Selex ES through its current product "Smart Sun Sensor" (SSS).

The SSoC is the evolution of the Smart Sun Sensor SSS where, taking benefit of recent achievements in CMOS and MEMS technology, a high level of integration is reached since all the processing logic is embedded on the same Silicon chip, together with the sensitive area (an Active Pixel Sensor, APS, array). Moreover, pinhole optics is attached directly on the Silicon to form a unique assembly enclosed in the same package, allowing full miniaturization of the unit.

At this time, taking benefits of the results and lesson learnt of previous technological development contracts let by ESA, Selex ES is running the final detailed design and qualification of the SSoC product. This will be achieved at unit level through the development of an Engineering Model (EM) to be submitted to and environmental test campaign, and at component level directly on the sensor chip.

In this paper the challenges encountered in the miniaturization will be described, the general SSoC design concepts will be outlined and the final performance capabilities and interface data of the product will be presented to the users' community, on the basis of the product on-ground qualification activities currently running.

I. Introduction

In the frame of a number of ESA study contracts the use of MEMS and CMOS technologies was proved to be suitable, attractive and mature for use in space, with good market opportunities for the miniaturized sensors they enable. The Sun sensor miniaturization is, among different applications like star tracker or navigation cameras, the most promising one. Thanks to the high optical radiation provided by the Sun, it is possible to realize a sun sensor equipment using a very simple optics, i.e. a pinhole. It is also possible to hard code all the required image processing thus avoiding the use of microprocessor and storage memory, as on board software is not required.

The development of a totally integrated system-on-chip digital sun sensor started in February 2009. The sensor's requirements specification and architecture were derived from work in earlier studies, from breadboard tests at ESA and from Selex ES's long experience in the attitude sensors field. This was augmented with insights and developments originating in CMOSIS' machine vision sensors, from knowledge on MEMS technology from BAE Systems and from Thales Alenia Space AOCS expertise. The Sun Sensor on Chip was also to be used to de-risk technologies and circuits for future single-chip autonomous star trackers.

MEMS optics - with pinhole - and pixel dimensions were selected to obtain a broad FoV of 64 deg half cone and high accuracy in the Sun position determination. Such a FoV retains backward compatibility with many Sun sensors currently on the market today.

The chip, packaged in a standard JLCC84 package for the prototype (but will be in a much smaller custom package for the final product), is hosted in a miniaturized mechanical housing that guarantees adequate radiation shielding and proper fixing / alignment to the satellite.

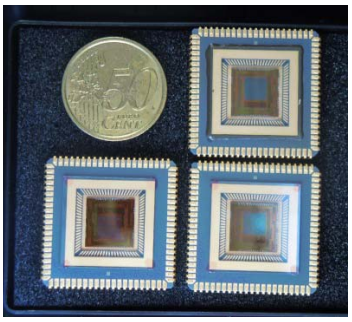


Fig. 1: First SSoC prototype in the standard package

II. SUMMARY OF SSOC DESIGN AND PROTOTYPE DEVELOPMENT

The Sun sensor design is based on the pin hole structure depicted in Fig. 2, bonded on the sensor core chip.

The sensor chip includes:

- Analogue section: the 512x512 10 µm pitch 4 T rolling shutter pixel array, 10-bit column ADCs, and readout circuits, all rad-hard
- Digital section: the Processing Logic with image filtering implementation and the SpaceWire/UART Telemetry/Telecommand (TM/TLC) interface
- Power section: the 3.3V and 1.8V power regulator, the 5V input voltage
- Oscillator

Only few components (quartz and other passive components) and are still needed off chip.

Thanks to the MEMS technology, the direct integration of the optics (pinhole) on the chip allowed to reduce the sensor size and assembly/integration complexity, while the on chip integration of the electronics completes the miniaturization process.

The geometrical dimensioning of the pin hole structure (pin hole diameter and height of the glass under the pin hole) was studied in order to obtain a large FoV (64° half cone) and an apparent hole diameter independent from the wavelength. A 160 micron pin hole diameter was then selected.

This miniaturization is made possible with the present technology considering also that:

- The unit is autonomous with only two operating modes, the Sun acquisition and the tracking modes. No telecommands are required for nominal in orbit operation, but for the use in interplanetary missions a change of integration time with the change of the astronomical unit (AU) is possible, as is the grabbing and downloading of full images.
- The processing for Sun position determination is simply a thresholding and barycentre, thus implemented without the use of any microprocessor. In any case an improvement in the tolerance to SEU has been achieved by using filtering and image processing algorithms.
- The power conditioning, being the overall power consumption limited to few hundreds of mW, is realized with a linear regulator.
- Full operating temperature range (-40° C. to +70° C., with the goal to increase it) can be achieved without the use of any active cooling system

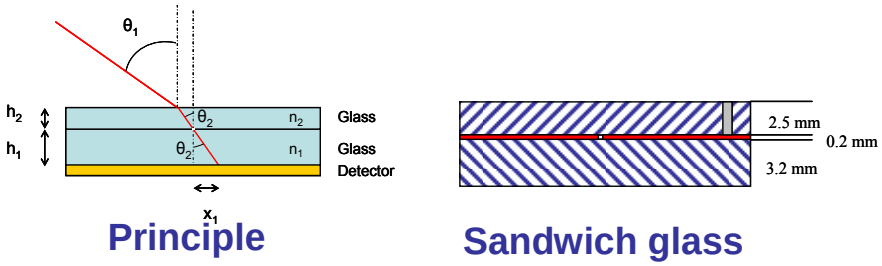


Fig. 2: Pinhole structure (left) and sandwich glass dimensioning (right)

III. Integrated optical chip design

The basic design of the optical chip consists of a metal pinhole layer, spaced above the sensor chip by a thick glass spacer layer, creating an optically immersed chip. The FoV is defined by the spacer thickness. A top glass layer is added above the pinhole layer, providing immunity to assembly dust and additional radiation blocking. Between the top layer and spacer and between the spacer and sensor chip, fully immersed bonds are made with BCB polymer (BenzoCycloButene). Details are shown diagrammatically in Fig. 3.

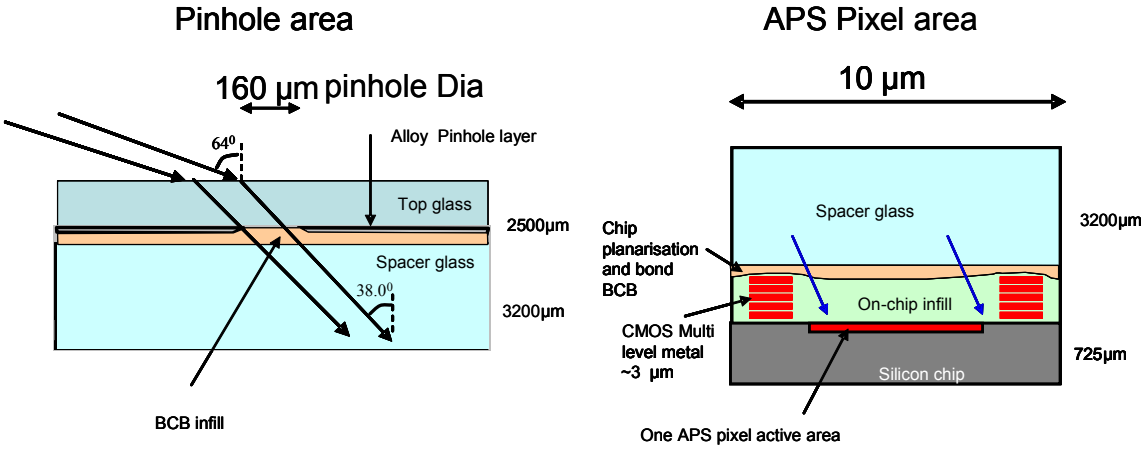


Fig. 3: Integrated optical design concept: chip (left) and pixel detail (right)

IV. Materials selection

The materials' selection has resulted in the following choices.

- Glass Material: Fused silica is chosen as it has good radiation hardness, is available to good tolerances and finish, is conveniently bondable and has a low TCE, which was shown to be an advantage in modelling. Wafers of fused silica can be handled easily in a MEMS facility in a similar way to silicon wafers.
- Pinhole material: Developments have taken place which favour the use of an aluminium alloy as the pinhole material. This material is deposited relatively thick (~1 µm) and given a very short etch, equivalent to removing 10nm of the layer. This operation converts the underside of the pinhole layer into a poorly reflecting layer, as required for good ghost image rejection in the sun sensor operation.
- Bond/fill materials: There are two bond layers, first within the optical chip between the two glass layers, and second at integration of the optics and electronic chips where it is used between the two chips. For both purposes, BCB has proved satisfactory in prototype demonstration testing. BCB has been radiation tested with good results, at least as far as TID is concerned (300 kRad).

V. Fabrication and integration

The optical chips are fabricated on a wafer scale, currently 44 chips on a 100mm wafer. First a wafer of fused silica is prepared from the design of top glass thickness and double-side polished. Then the pinhole layer is deposited, the pinholes defined and its surface treated for low reflection. A fused silica wafer of the spacer glass design thickness is then bonded to the top glass wafer using a spun-on liquid BCB layer, which is subsequently cured to a cross-linked solid.

For this current electrical prototype stage, the optical chips are then sawn into 44 individuals for integration with already packaged electronic sensor chips. In the above operations, because this is an unusually thick chip, at 5.7mm, it has required special developments to regular semiconductor/MEMS handling equipment.

For this chip scale integration, the optical chips are individually coated on their underside with a BCB layer, and then the chip is positioned over a packaged sensor chip for alignment and bonding of the two chips. Curing of the BCB takes place on a low temperature/long time profile, in order not to affect the packaged electronic chip. A cover is then applied to the integrated chip to restrict light entry to the desired field of view, and protect the unit from mechanical damage.

An uncovered integrated prototype unit is shown in Fig. 4 on the left and with a straylight cover on the right.

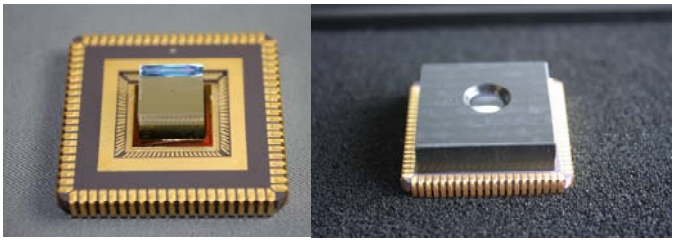


Fig. 4: Integrated prototype in the standard package: uncovered (left), with simple straylight protection cover (right)

VI. Overall packaged chip design

The overall design consists of a conventional custom electronic package with pins underneath, routing layers between the pins and bond pads aligned with the chip pads for wire bonding. The package lid, shown in Fig. 5 will be special, however, being relatively tall to accommodate the integrated chip, and having two seals, one to the electronic package and one to the top of the optical chip, allowing a window for the field of view.

Packaging prototypes of the integrated chip and packaged unit are shown below. The following Fig. 6 shows the package dimensions shrinking between standard and custom packages.

Fig. 5: Packaging prototype of the integrated final product chip (left) and packaged unit with dedicated lid (right)

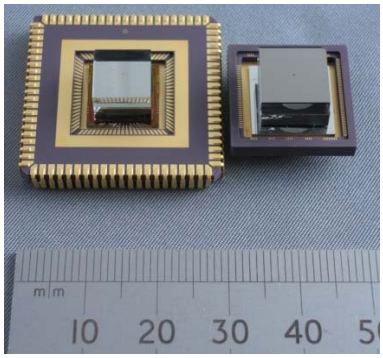


Fig. 6: Packaging dimension shrinking: standard package (left) and custom package (right)

VII. Sensor chip design

Chip design of all analogue circuits and logic was performed by CMOSIS. A strong desire for physical robustness led to a fully radiation hardened design. For the analogue circuits this was done by layout. For the logic the hardness was obtained using IMEC's DARE library of rad-hard core and IO standard cells for the UMC 0.18µm CMOS foundry and process. In support of this the DARE libraries were ported to the specific UMC .18 CMOS Image Sensor process variant, which has less metal layers and a thinner dielectric stack than the standard logic CMOS process. Tape out to UMC was done in January 2010. First silicon returned in June 2010 and was tested and characterised through the rest of the year.

The sensor chip comprises of a mixed signal imager core, a logic core, a 1.8V voltage regulator, Power-On-Reset block (POR), and input/output interfaces - including CMOS and LVDS drivers.

The focal plane array (FPA) has 512x512 4-Transistor buried photodiode pixels, at a pitch of 10 µm, with an electronic rolling shutter. The pixels are read out through fixed gain column amplifiers and 512 10 bit column ADCs. Thermal drift of internal reference levels is, by design, extremely low. There is also a temperature sensor with ADC.

The logic contains the FPA readout sequencer, image processing, sun detection and tracking application, and telecommand/telemetry (TC/TM) interfaces. The latter include SpaceWire and UART. The SpaceWire interface operates at a fixed 10 Mbps rate and allows the output of sun coordinate telemetry and raw or processed images. The UART is limited to 115200 Baud and does not support image readout. The chip is configured by strapping a small number of pins to power or ground. The TC/TM protocol is packet based and does not support Remote Memory Access Protocol (RMAP). The sensor supports SpaceWire point-to-point operation as well as networks with up to 3 additional nodes between the SSoC and its host

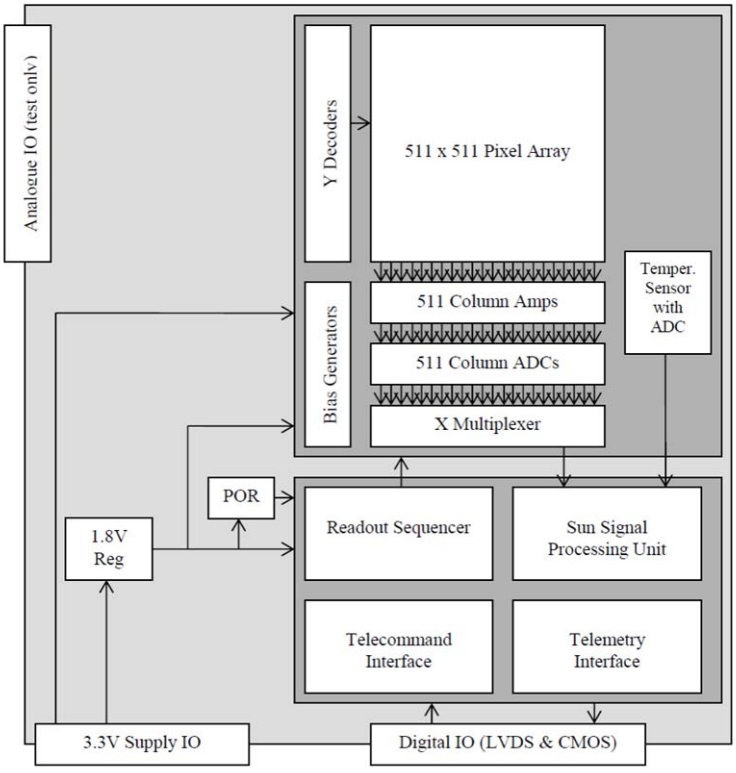


Fig. 7: Sensor chip block diagram

Upon power-on the sensor is intended to autonomously acquire and then track the Sun. No telemetry data are generated: the host must query the sensor for data. The chip can be used interactively as well: modes, configurations, and parameters are stored in registers that are user accessible over the TC interface.

The sensor is highly immune to the effects of radiation. Total dose hardening of analogue and digital circuits was done, including the pixels. There is extensive protection against SEUs: all flip flops and latches in the logic are HIT cells, and all long term storage nodes in the ADC employ Triple Mode Redundancy (TMR).

The sensor die measures less than 12 x 12 mm^2. Some of the silicon area is inactive, serving only as a mechanical support for the MEMS optical stack. The chip has 50 active pins but, as mentioned before, the prototypes were housed in large (but convenient) JLCC84 packages. The future flight models will come in a much smaller custom made ceramic µPGA package.

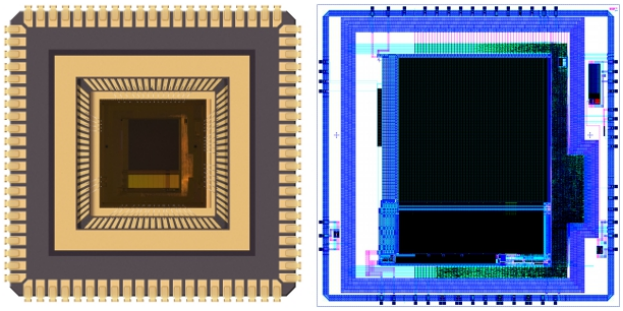


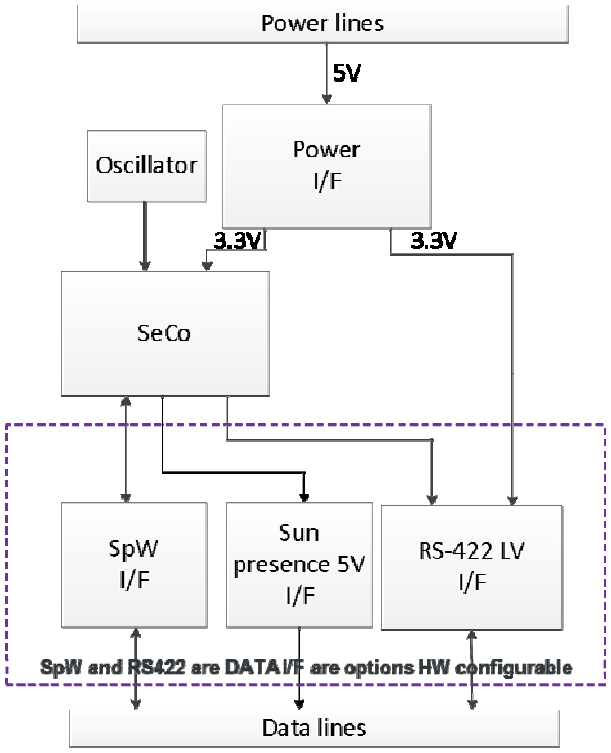
Fig. 8: Prototype sensor chip in its package (left) and die layout (right)

VIII.OFF-chip COMPONENTS

The sensor chip (named “SeCo” in the following diagram) circuitry requires 3.3V and 1.8V power supplies. The 3.3V is generated outside the chip (from the +5 Volt power supply) while an on-chip linear voltage regulator generates the 1.8V. The Voltage regulator – outside the chip - is rated for 5V input (5.5V maximum), although short term (i.e. hours) tolerance to a raw supply input of up to 6.7V over full temperature range is possible.

The off-chip quartz oscillator operates with a specific 10MHz external crystal. A direct clock input is also available.

Block	Description
Power I/F	5V to 3.3V traslator (RHFL4913S33-04V) + bulk capacitors
SeCO	CMOSIS Chip
SpW I/F	Implemented inside SeCO, only 100Ω resistor are needed
RS422LV I/F	26CLV31R and 26CLV32RH line driver.
Oscillator	10MHz CMOS 1118R10M00000BD by Vectron
Sun presence 5V I/F	3.3V to 5V npn transistor based level translator



IX. Sun detection and tracking

The SSoC has two main operating modes: Sun Acquisition Mode (SAM) and Sun Tracking Mode (STM), as shown in Fig. 9 and Fig. 10.

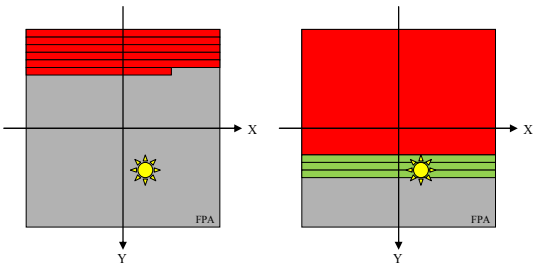


Fig. 9: Sun Acquisition Mode: full frame scan until the sun image is detected

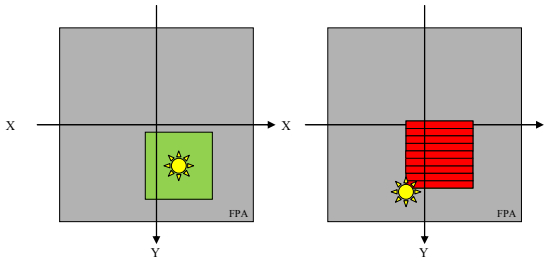


Fig. 10: Sun Tracking Mode with window scan: tracking (right), loss of Sun (right)

Sun Acquisition Mode is entered upon power-on or reset. In this mode the SSoC scans a user defined field of view (normally the full 511 x 511 pixel frame) at user defined rate (normally 5Hz) and searches it for the occurrence of a bright object that matches the dimensions of the sun image. Once such an object has been found, the SSoC determines its exact position, interrupts the full frame scan and transitions to Sun Tracking Mode.

Sun Tracking Mode positions and maintains a user defined readout window (by default 80 x 80 pixels) around the sun image, at a user defined rate (by default 10Hz, but up to 60Hz and even more). Each window is scanned completely, after which the position of the bright object that is most likely the sun is extracted. The window position is updated and the cycle repeats. If no sun image is encountered, the SSoC transitions back to Sun Acquisition Mode.

At all times the coordinates of the sun image's photometric centre of gravity are extracted and made available on the TM interface. Furthermore, the presence of the sun in the field of view is always flagged on the SunPresent chip pin as a logic output.

The output of the FPA is a stream of raw images. These images are subjected to pre processing, sun detection, and sun tracking. The robustness and reliability of these processing is paramount for the sensor's intended application.

Image pre processing contains the following steps, all with user programmable parameters:

- Defect pixel and spike signal filtering
- Average background level calculation
- Background level subtraction
- Pixel value thresholding

The output of this pre processing is a stream of black images, with in it a number of bright objects. These objects can have any size, but their individual pixel signal strengths all exceed a user defined threshold.

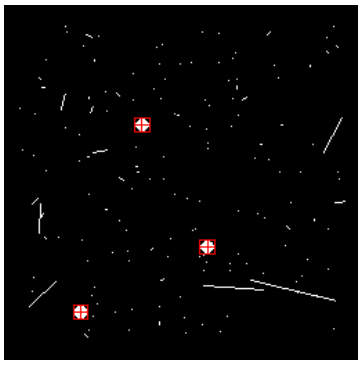
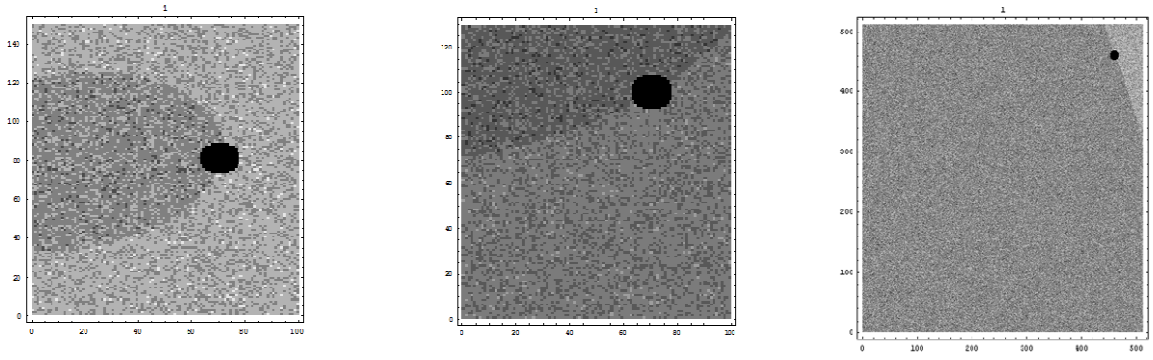


Fig. 11: Simulated demonstration of sun detection and centroiding on artificial input image

In a next processing step, isolated bright pixels are rejected and two dimensional clusters of bright pixels are grown. Whenever a bright cluster has been isolated, its X and Y dimensions and its number of bright pixels are compared to a user defined model of the sun image. Undersized clusters are immediately rejected. Oversized clusters are rejected, but their presence is logged and potentially communicated to the host. If one cluster matches the sun image's dimensions it is tracked, and its photometric centroid is calculated with 16 bit precision. If more clusters match the sun, then the one nearest to the last known sun position is tracked, and the occurrence of extra sun like objects is flagged.

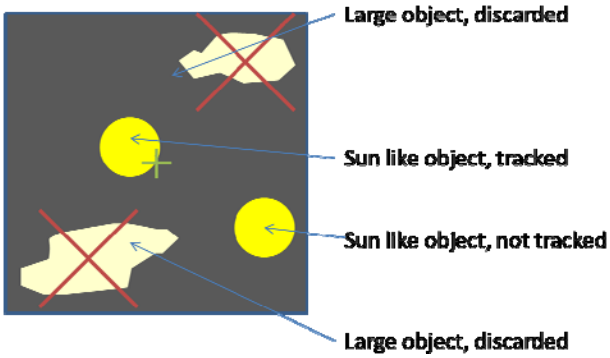
X. SPECIAL FEATURES: Immunity to Earth Albedo and Reflections Rejection

The SSoC is “by design” immune to Earth Albedo. In fact, only the direct light of the Sun is sufficient to generate a signal on the detector above the threshold. As a result, the Earth Albedo DOES NOT affect Sun position detection and centroid estimation.



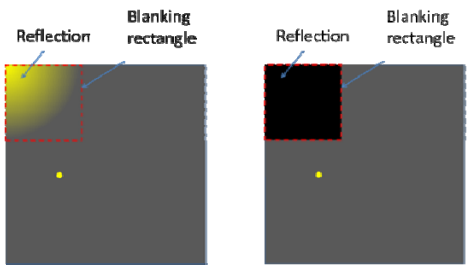
h=50000 km h=20000 km h=100 km

The SSoC can operate in presence of objects in the field of view (like antennas, appendages and booms) and reflections from spacecraft. The logic can autonomously detect and identify up to 4 clusters of over threshold pixels. These 4 clusters can be filtered out according to specific dimensions and energy criteria or maintained and tracked, with relevant information reported in telemetry, while the real Sun is correctly identified and tracked.



- "Large objects" are, by default, discarded
- In presence of multiple Sun-like objects:
- The one closest to the last known Sun position is automatically tracked
- A flag is set (solution not to be trusted)

In addition the user can define - and dynamically update through telecommand - two "blind" rectangles in the detector array. All the pixels contained in the area inside these rectangles are automatically excluded by the SSoC logic, both in terms of background calculation and over threshold pixels determination. The blind rectangles may be defined based on the telemetry data (size and position of the identified clusters) or on the raw image that can be dumped through the SpaceWire interface.



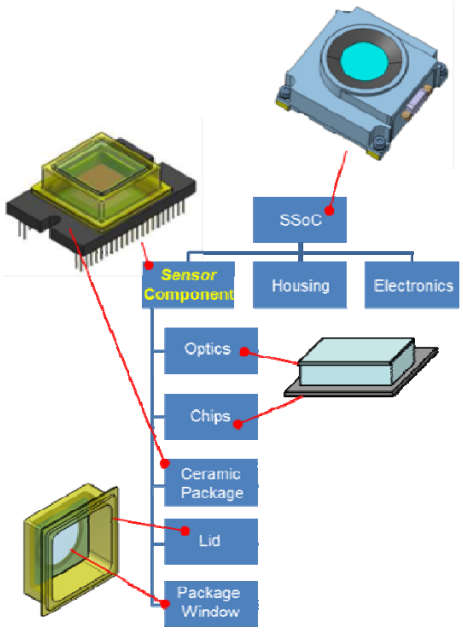
XI. Final Configuration

The Sensor Component (SeCo) is the main device of the unit, where the Sensing and the Logic are implemented.

The Sensor Component (SeCo) is made of:

- **Chip**: the piece of Silicon integrating the pixel array detector, the Processing Logic and electronics, also called "detector", "core chip"
- **Optics**: the glass stack that holds the pinhole assembled on the Chip
- **Package**: the ceramic housing into which the Chip and the Optics are assembled
- **Metallic Lid**: the SeCo mechanical housing assembled on the ceramic package
- **Package Window**: The SeCo closure window assembled on the Metallic Lid

The SSoC product tree can be simply summarized as depicted in the figure reported on the right:



XII. Mechanical housing

The assembly concept (Fig. 12) is based on a main structure constituted by a titanium alloy frame, whose Thermal Coefficient of Expansion (TCE) matches the assembly with glass (closure window) and with the ceramic package of the detector. This choice also offers better radiation shielding than Aluminium within the same volume. A circuit board, hosting the off chip components, will be fixed to the base of the sensor structure, also increasing its stiffness. The upper cover supports internally an attenuation filter. If needed, the attenuation filter can be tailored to the mission needs, and selected on a "mission by mission" case.

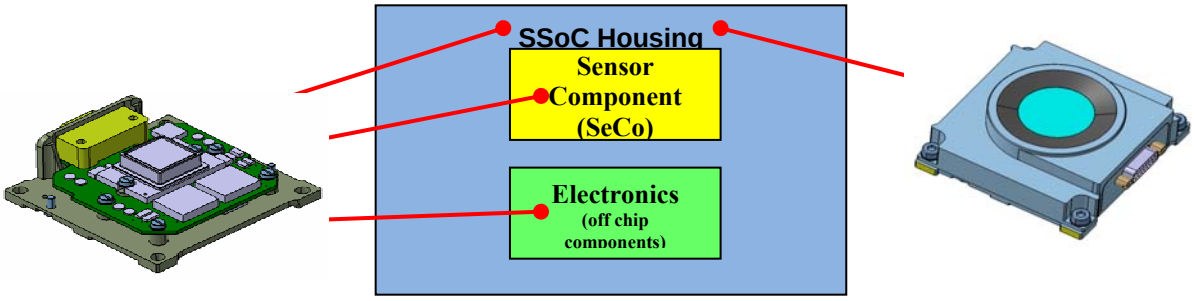


Fig. 12: Sun Sensor Assembly

XIII.FUTURE PLAN AND DEVELOPMENT

In the frame of an ESA Contract started in late 2013 we have already completed the following tasks:

- ✓ Sensor Component (SeCo) Chip (and its packaging) Design
- ✓ Chip evaluation
- ✓ EQMs chip components Design, Manufacturing and Qualification
- ✓ Prepare the tools to decrease the manufacturing costs by introducing batch manufacturing and testing.
- ✓ Sun Sensor on Chip (SSoC) Unit and Packaging Design

and the following activities are on-going:

- Evaluate the SSoC
- Design, build and test EM SSoC
- Design, build and qualify EQM SSoC

As a result of all the activities carried out until now by Selex ES and its partners on this product, the following milestones have been achieved:

- Digital Sun Sensor on Chip designed and developed
- Challenging design goals and specifications met
- First prototype chips manufactured
- Sensor is fully working
- Choice of:
  - Coarse version with no dedicated calibration
  - Accurate version with dedicated calibration

and the following milestones are planned in the near future:

- All Risks retired by Q2 2106
- Environmental test campaign completion by Q4 2016
- Follow up activities mainly dedicated to ESA formal evaluation campaign on the SeCo and cost optimization of SSoC production at Selex ES
- First FM Orders from Q4 2016
- FMs processing will run in parallel to SSoC follow up contract phase

XIV. Conclusion

At the beginning of this ESA contract, the basic idea was to keep the unit architecture defined in the previous ESA contract activity, with the integration of the main electrical and processing features inside the chip. The Sensor Component (SeCo) sensor optics is composed of a glass stack

containing the pinhole, realised with a metallic layer. The bonding of the optics directly on silicon chip is also derived from the previous studies.

During the first trade-off analysis of this contract some modifications have been introduced for the mechanical housing of the SeCo and the final SSoC unit configuration has been defined.

The final product characteristics are shown in the following table.

Tab. 1: SSoC compared with SSS product of Selex ES

	SSS	SSoC
APS Detector	1024x1024 15µm	512x512 105µm
FOV	from ±64° to ±45°	64° half cone
Accuracy (Bias 1σ)	Simple calibration: 0.08 deg (whole FoV) 0.01 deg (±10 deg FoV) Iterative calibration: 0.02 deg (whole FoV) 0.01 deg (±30 deg FoV)	No calibration: 0.5 deg (whole FoV) 0.4 deg (α, β axes) With dedicated calibration: 0.02 deg (whole FoV)
Accuracy (NEA 1σ)	< 15 arcsec	< 5 arcsec
Resolution	< 0.03 deg.	< 0.02 deg.
Power	Bus 50V / pre-regulated +5V±10%	pre-regulated +5V±10%
Power Consumption	<1.5W	< 200 mW
Radiation Hardness	100 Krad with an extension to 300 Krad	> 300 Krad
Mass	< 330 gr	102 gr
Dimensions	112 x 110 x 43 mm	58 x 60 x 23.6 mm
Data Interface	Serial Digital	SpaceWire or UART

XV. ACKNOWLEDGMENTS

This activity is led by the European Space Agency under the Contract Number 4000108300/13/NL/MH, financed by the GSTP (General Support Technology Programme) budget.

The authors of the paper wish to thank the Italian Space Agency (ASI), who recognized the importance of this miniaturized sun sensor development for improving the competitiveness of the Selex ES attitude sensors product line, thus has funded this activity through the GSTP budget of ESA.

Studies on Fast Attitude Maneuver for Spacecraft

Li Yingbo [1,2] Sun Jingru [3] Zhang Zilong [1]

（1、 Shanghai Institute of Spaceflight Control Technology；
2、 Shanghai Key Laboratory of Aerospace Intelligent Control Technology，Shanghai, China；
3、Shijiazhuang Tiedao University，Shijiazhuang, Hebei, China）

Abstract: This paper deals with the studies on the attitude maneuver from rest to rest for both rigid and flexible spacecrafts. Research result for a rigid spacecraft's time-optimal three-axis maneuver is firstly presented. Then it sums up theoretical achievement about time-optimal attitude maneuver for a flexible spacecraft. Engineering methods for fast attitude maneuver is presented based on some typical application. To fast attitude maneuver for a flexible spacecraft, questions to be solved in the future are put forward.

Keywords: spacecraft; fast attitude maneuver; time-optimal control

1 Introduction

Fast attitude maneuver or reorientation of spacecraft is mainly used for the space missions like observation for different areas on the ground, three-dimensional imaging for the same object or tracking a moving target and so on. It can reduce the re-visiting cycle , increase the observed range or carries out a three-dimensional imaging. It is generally required that the attitude maneuver time should be as short as possible. Consequently, research in fast attitude maneuver has been a consistently strong field of study.

Time-optimal control problems are two-point boundary value problem in optimal control theory. Only in very simple configuration such as single-axis attitude maneuver of inertially symmetric rigid body can the analytical solution be found out [1]. The solution to the general three-dimensional reorientation of a rigid body (even inertially symmetric) had to be achieved by means of numerical method. Bilimoria and Wie [2] obtained the time-optimal solution of tree-axis reorientation for a inertially symmetric rigid body under independent three-axis controls (cube constraint) is bang-bang control around non eigenaxis. The switch numbers depend on the angle magnitude. They gave the explanation that rotation around non eigenaxis can provide more torque along the reorientation axis. Xiaoli Bai and John L. Junkins [3] proved that the eigenaxis maneuver is the time-optimal solution for a inertially symmetric rigid body when the magnitude of the control is constrained while the control direction is left free (i.e. sphere constraint). Andrew Flemng [4] studied the effect to the time-optimal maneuver solution by parameter uncertainties of a generic rigid body.

The problem for three-axis maneuver of a flexible spacecraft is more complex. At present, studies on the time-optimal maneuver control of a flexible body are focused on single-axis maneuver by means of numerical method [5]～[14]. They all proved the time-optimal is bang-bang control. HE Haifeng and ZENG Haibo [5] studied the problem of single-axis time-optimal attitude maneuver for a flexible spacecraft. The conclusion was given which maneuver time increases as the flexible mode orders. G. Singh [12] and Joseph Ben-Asher [13] reached the conclusion that the controls are antisymmetric through simulation results. The influence to the maneuver time by the mode frequencies and damp coefficient was investigated [14].

Neither the time-optimal solution of a rigid body attitude maneuver nor that of a flexible one's be directly applied to actual space mission due to its complexity. However, the results obtained for the time-optimal can guide us to achieve a fast attitude maneuver through a feasible approach in practice. JIN Jin [15] optimized controller parameters by using hybrid genetic algorithm to shorten the maneuver time. ZHANG Meihua [16] used the cascade saturation method to eliminate the attitude error successively for large angle attitude maneuver of micro satellite using reaction wheels. WANG Shuyi [17] and LIU Yiwu [18] studied quick attitude maneuver using single-gimbal control moment gyros based on trajectory planning. For attitude maneuver of a flexible spacecraft , unsymmetrical soften trajectory planning was put forward to excite less vibration [19] . Recent years , various methods such as variable structure, input shaping, and adaptive control etc. are used to suppress the residual vibration to the attitude maneuver for a flexible spacecrafts [21]～[23].

This paper summarizes the time-optimal attitude maneuver results respectively for a rigid body and a flexible body. Feasible fast attitude maneuver approaches are put forward for quasi-rigid spacecrafts. For flexible spacecrafts, techniques are presented to suppress residual vibration. The problem to be solved in the future is proposed.

2 Time-optimal Attitude Maneuver for a Rigid Spacecraft

Some Spacecraft can be considered approximately as rigid body. It can help to find out technique approach to fast attitude maneuver by comprehending characteristic of time-optimal attitude maneuver for rigid body.

2.1 Problem Formulation

The Euler's rotational equations of motion for a rigid body, referenced to body-fixed principal axes with origin at the center of mass, are

$$\begin{aligned} I_1\dot{\omega}_1-(I_2-I_3)\omega_2\omega_3&=u_1\\ I_2\dot{\omega}_2-(I_3-I_1)\omega_3\omega_1&=u_2\\ I_3\dot{\omega}_3-(I_1-I_2)\omega_1\omega_2&=u_3 \end{aligned} \qquad (1)$$

In Eqs.(1),ω_1, ω_2, ω_3 are angular velocity components along the orthogonal body-fixed axes; u_1, u_2, u_3 are bounded independent control torques about these axes; and I_i (i=1,2,3) are the moments of inertia about the body fixed axes.

Using the quaternion q_0,q_1, q_2 , q_3 to describe the attitude of the spacecraft.

$$\begin{aligned} q_0&=\cos\frac{\phi}{2}\\ q_i&=\cos\beta_i\sin\frac{\phi}{2}\qquad i=1,2,3 \end{aligned} \qquad (2)$$

where β_i (i=1,2,3) are the orientation of the Euler axis (i.e. equivalent rotation axis or eigenaxis) relative to the body-fixed reference frame. and φ is the principal rotation angle (i.e., the magnitude of the Euler axis rotation), as shown in Fig.1.

The kinematic equations of angular motion are

$$\begin{bmatrix}\dot{q}_0\\ \dot{q}_1\\ \dot{q}_2\\ \dot{q}_3\end{bmatrix}=\frac{1}{2}\begin{bmatrix}0&\omega_3&-\omega_2&\omega_1\\ -\omega_3&0&\omega_1&\omega_2\\ \omega_2&-\omega_1&0&\omega_3\\ -\omega_1&-\omega_2&-\omega_3&0\end{bmatrix}\begin{bmatrix}q_0\\ q_1\\ q_2\\ q_3\end{bmatrix} \qquad (3)$$

`

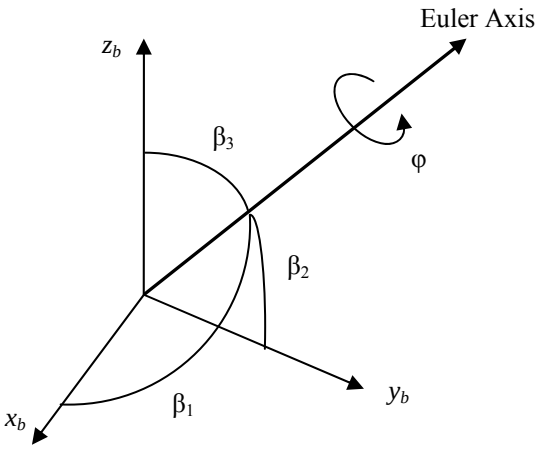


Fig.1 Euler axis and principal rotation angle

Eqs. (1) and (3) form the state equation driving the dynamic system. The minimum-time rest-to -rest reorientation problem may be formally stated as follows: Determining the controls (u_1*, u_2*, u_3*) that driving the dynamic system describe by Eqs. (1) and (3) from rest at its initial orientation described by the quaternion $(q_0, q_1, q_2, q_3)_0$ to rest at its final orientation described by the quaternion $(q_0, q_1, q_2, q_3)_f$, while minimizing the cost function

$$J = \int_0^{t_f} dt = t_f \tag{4}$$

subject to the control constraint

$$|u| \le u_{\max}$$

The problem described above is two-point boundary problem according to optimal control theory. No rigorous analytical solution for a general (inertially unsymmetric body) time-optimal three-axis maneuver has been obtained so far either for a rigid or a flexible body. Only for inertially symmetric or single-dimenssional maneuver problem, some solution (including numerical) be obtained.

2.2 Time-optimal Attitude Maneuver for Inertially Symmetric Spacecraft

Equation (1) can be simplified for the special case of an inertially symmetric body by setting I_1= I_2= I_3= I_0. A time-scale factor equal to $\sqrt{I_0 / u_{\max}}$ is introduced, where I_0 is the moment of inertial of the inertially symmetric body about any principal axis, and u_{max} is the maximum available torque about each control axis. Dimensionless angular velocity components $\dot{\Omega}_i = \omega_i \sqrt{I_0 / u_{\max}}$ and dimensionless control components U_i=u_i/u_{max} (i=1,2,3) are defined. The Euler's rotational equations of motion are now written as

$$\begin{aligned} \dot{\Omega}_1 &= U_1 \\ \dot{\Omega}_2 &= U_2 \\ \dot{\Omega}_3 &= U_3 \end{aligned} \tag{5}$$

subject to the control constraint $|u_i| \le 1$

Eqs. (3) and (5) describe the dynamic motion for a inertially symmetric spacecraft. The optimal control logic for the inertially symmetric spacecraft is then given by the following relationships:

$$\begin{aligned} U_i^* &= +1 \qquad if \quad S_i < 0 \\ U_i^* &= -1 \qquad if \quad S_i > 0 \\ U_i^* &= U_{is} \qquad if \quad S_i = 0 \end{aligned} \tag{6}$$

where S_i is called switching function which is to be found out in the optimal control. The Eqs. (6) shows that the time-optimal attitude maneuver control is either bang-bang control (S_i <0 or S_i >0, also called saturation control) or singular control (S_i =0). It can be proved that all three control components cannot be simultaneously singular. Karl D. Bilimoria and Bong Wie found out the numerical solution about the time-optimal attitude maneuver for inertially symmetric spacecraft under independent three axis control. The results are as follows.

1) The time-optimal attitude maneuver control is bang-bang in all three control components, i.e. three control components either maximum in positive direction or maximum in the other direction.

2) The eigenaxis rotation maneuver is not time optimal one that has a significant nutational components of motion.

3) The switch numbers for time-optimal attitude maneuver bang-bang control are related to the magnitude of attitude maneuver. The switches are 5 times if the maneuver magnitude is between 73°～180°.

Karl D.Bilimoria and Bong Wie obtained the time-optimal solution for a case in which 180° yaw reorientation is considered . Fig2 shows the bang-bang control with 5 switches. Fig 3 shows the dimensionless angular velocity histories.

It is count-intuitive that the eigenaxis rotation maneuver is not time optimal. Reference [2] gives a explanation by converting quaternion into intuitionistic orientation angle α,β,γ as shown in Fig.4. Then during 180 time optimal yaw maneuver process, the orientation angle is shown in the Fig.5. Fig.5 shows that time optimal attitude maneuver is significantly different from the eigenaxis rotation. During the time optimal maneuver , orientation angle γ once increases to a value more than 30°. Converting the optimal control components from body-fixed to inertial reference frame, Fig.6 shows the optimal control components in the inertial frame (only the control component along inertial Z axis is given in the Fig.6). Fig.6 shows that this control component u_z achieves significantly in excess of 1.0 over a substantial portion of the total time interval (region A). Although there is a small time interval where u_z is less than 1.0 (region B), it is clear that region B is significantly smaller than region A. This feature is sufficient (but not necessary) to insure a shorter maneuver time despite of that the maneuver path is not shortest.

The conclusions about time optimal attitude maneuver are obtained in the constraint that three axis independent control components are available ($|u_i| \le 1$,for example in the case of three orthogonal body-fixed wheels), that is the constraint possessing cube feature mathematically. It is due to the cube constraint that the control torque along inertial Z axis can reach a value in excess of 1.0. That is why optimal maneuver faster than the eigenaxis rotation maneuver. If the constraint is ball not cube, it is impossible to provide a torque more than 1.0 along inertial Z axis. In that case the conclusion is not available that optimal attitude maneuver is not around eigenaxis rotation. That is what the Ref. [3] studied on. In the Ref.[3], Xiaoli Bai and John L.Junkins have proved analytically that time optimal attitude maneuver is about eigenaxis rotation maneuver when the control constraint is

sphere form (magnitude not more than 1.0). Ref.[3] gives the optimal solution comparison between sphere constraint and cube constraint.

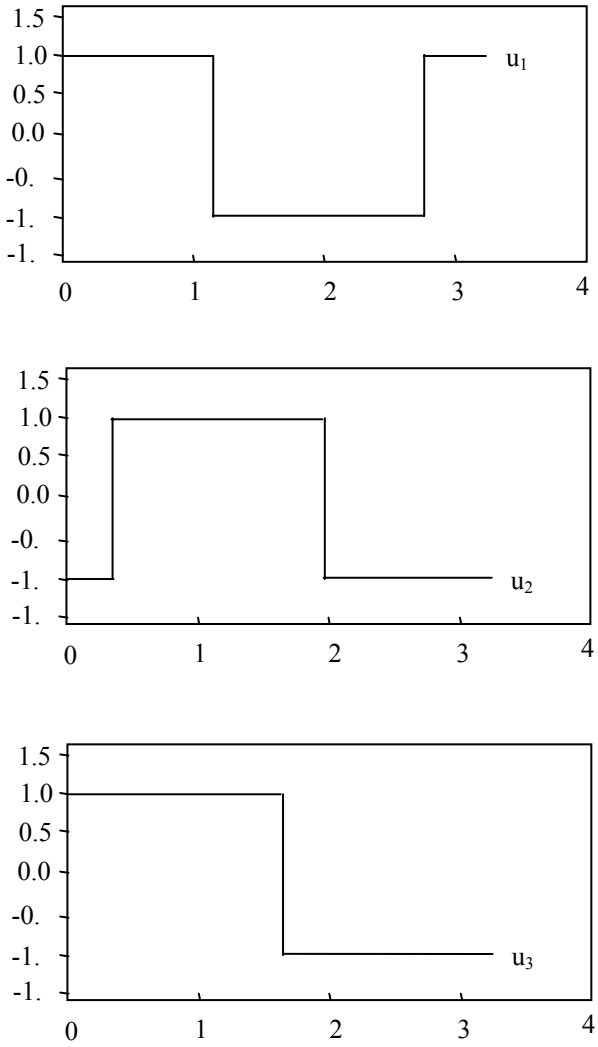


Fig.2 Time-optimal control histories along body axes for □=180°

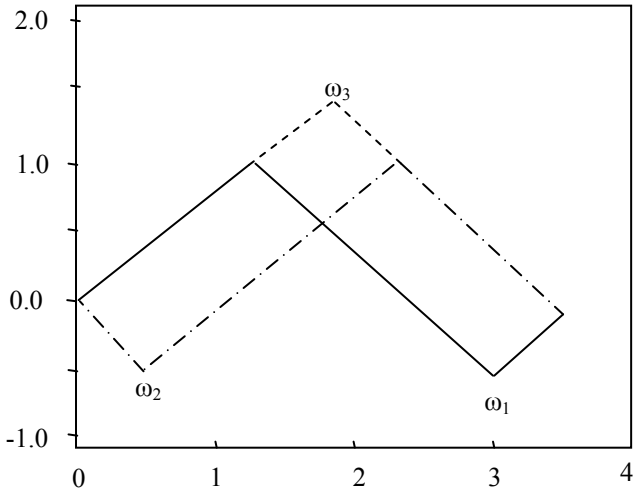


Fig.3 Dimensionless angular velocity history for □=180°

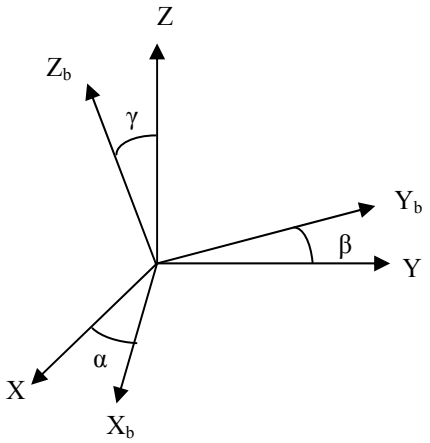


Fig. 4 defination of orientation angles α,β,and γ

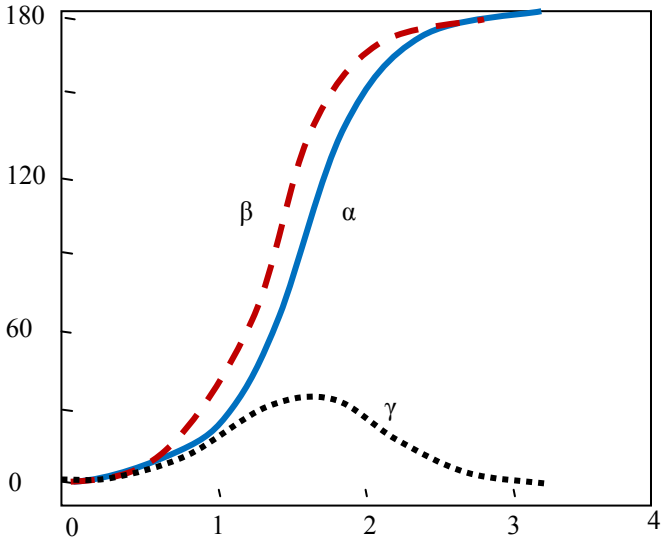


Fig. 5 Orientation angle histories for □=180°

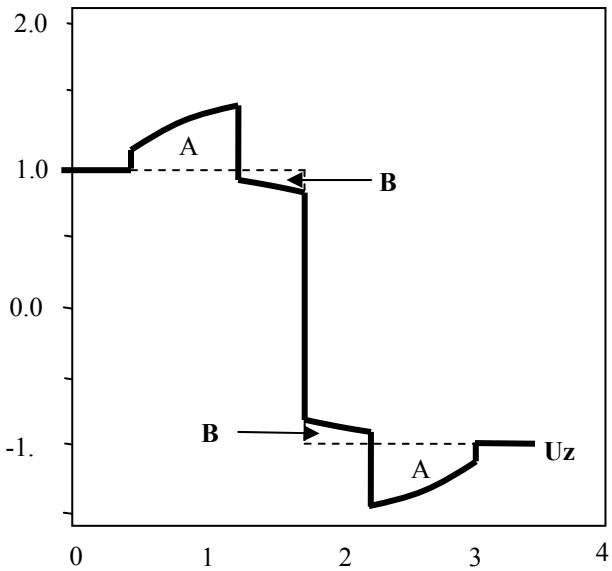


Fig.6 Time-optimal control history along inertial Z axes for □=180°

2.3 Time-optimal Single-axis Maneuver of a Spacecraft

X axis maneuver is discussed as follows. By setting $\omega_2=\omega_3$, and omitting suffix, Eq.(1) can be simplified.

$$I\ddot{\varphi} = u \qquad |u| \le u_{\max} \tag{7}$$

The problem of time-optimal attitude maneuver is to find out the control u* driving the system from rest at its initial angle $\varphi(0)= \varphi_0$ to rest at a final orientation $\varphi(t_f)= 0$ as describe as follows.

$$\begin{aligned} &\varphi(0) = \varphi_0 \qquad \dot{\varphi}(0) = 0 \\ &\varphi(t_f) = 0 \qquad \dot{\varphi}(t_f) = 0 \end{aligned} \tag{8}$$

The time-optimal solution for above problem is a bang-bang control with at most one switch.

$$u^* = -u_{\max}\,\mathrm{sgn}(x) \tag{9}$$

$$S = \begin{cases} \varphi(t) + I\dot{\varphi}(t)|\dot{\varphi}| / 2u & 0 \le t \le t_1 \\ \dot{\varphi} & t_1 \le t \le t_f \end{cases} \tag{10}$$

where S is switch function; t_1 is switch time; t_f is maneuver time

$$t_f = 2t_1 = 2\sqrt{\frac{I}{u_{\max}}|\varphi_0|} \tag{11}$$

Time-optimal control block diagram for single axis attitude maneuver is illustrated in Fig.7.

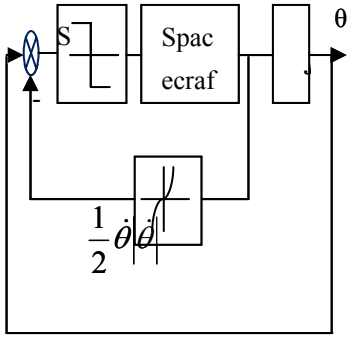


Fig.7 Time-optimal control block diagram for single axis attitude maneuver

3 Time-optimal Attitude Maneuver of a Flexible Spacecraft

The equation for three-axis maneuver of a flexible spacecraft include not only nonlinear term about rigid main body but also nonlinear term about flexible modes. These feature bring trouble to find out the solution for the time-optimal maneuver of a flexible spacecraft. At present, time-optimal solution of a flexible spacecraft is focused on single-axis maneuver by means of numerical method.

HE Haifeng and ZENG Haibo [5] gave the sufficient and necessary condition which time-optimal control should meet through minimum value theory for a flexible spacecraft without damp. They explained why the control has a antisymmetric feature through simulation curve. The Ref. [6] and [8] studied on how the flexible modes frequencies and damp coefficient influence the maneuver time. Ref. [8] gave the tenable condition and attestation for the antisymmetric control. Conclusions are obtained about the correspondence for the switches and flexible modes in general

by means of numerical simulation and rule for the maneuver time increasing adapt to flexible modes order.

3.1 Problem Formulation

Taking pitch axis maneuver as a example, dynamic equation of a flexible spacecraft is as follow

$$\begin{aligned} &I\ddot{\theta}+\mathbf{C}\ddot{\mathbf{p}}=u \\ &\ddot{\mathbf{p}}+2\boldsymbol{\xi}\boldsymbol{\lambda}\dot{\mathbf{p}}+\boldsymbol{\lambda}^{2}\mathbf{p}+\mathbf{C}^{T}\ddot{\theta}=0 \end{aligned} \tag{12}$$

where p is mode coordinates (order n is considered) $\mathbf{p}=[p_1, p_2,..., p_n]^T$; $\mathbf{C}$ is coupled coefficient $\mathbf{C}=[C_1, C_2,..., C_n]$; $\boldsymbol{\xi}$ is mode damping diagonal matrix $\boldsymbol{\xi}=\mathrm{diag}[\xi_1, \xi_2,..., \xi_n]$; $\boldsymbol{\lambda}$ is mode frequency diagonal matrix $\boldsymbol{\lambda}=\mathrm{diag}[\lambda_1, \lambda_2,..., \lambda_n]$

By setting $\mathbf{p}=\mathbf{V}\boldsymbol{\eta}$, $\boldsymbol{\eta}=[\eta_1\ \eta_2...\ \eta_n]^T$,$\mathbf{V}\leftarrow R^{n\times n}$,

$$\mathbf{V}^{T}(\mathbf{E}-\mathbf{C}^{T}I^{-1}\mathbf{C})\mathbf{V}=\mathbf{E}\ ;\mathbf{V}^{T}\boldsymbol{\lambda}^{2}\mathbf{V}=\boldsymbol{\omega}^{2}\quad \boldsymbol{\omega}=diag\{\omega_1,\omega_2,...,\omega_n\}\ ;$$

$$\mathbf{V}^{T}\boldsymbol{\xi}\boldsymbol{\lambda}\mathbf{V}=\boldsymbol{\xi}\boldsymbol{\omega}_{n}\quad \boldsymbol{\xi}=diag\{\xi_1,\xi_2,...,\xi_n\}$$

EQ. (12) can be converted to

$$\begin{aligned} &\ddot{\theta}+I^{-1}\mathbf{C}\mathbf{V}\ddot{\boldsymbol{\eta}}=I^{-1}u \\ &\ddot{\boldsymbol{\eta}}+2\boldsymbol{\xi}\boldsymbol{\omega}\dot{\boldsymbol{\eta}}+\boldsymbol{\omega}^{2}\boldsymbol{\eta}=-\mathbf{V}^{T}\mathbf{C}^{T}\mathbf{I}^{-1}u \end{aligned} \tag{13}$$

Defining state and control variable:

$$\begin{aligned} &x_1=\theta+I^{-1}\mathbf{C}\mathbf{V}\boldsymbol{\eta}\ ,\quad x_2=\dot{x}_1 \\ &x_{3-i}=\eta_i\ ,\ x_{4-i}=\dot{\eta}_i\ ,\ i=1,2,...,n \\ &U(t)=I^{-1}u \\ &b_0=1,\ [b_1\ \ b_2...b_n]^T=-\mathbf{V}^{T}\mathbf{C}^{T} \\ &\mathbf{x}=[x_1\ \ x_2\ \ x_{3-i}\ \ x_{4-i}...x_{m-n}\ \ x_{4-n}]^T \end{aligned}$$

then Eq. (13) can be re-written

$$\dot{\mathbf{X}}(t)=\mathbf{A}\mathbf{X}(t)+\mathbf{B}\mathbf{U}(t) \tag{14}$$

where $\mathbf{A}$=block diag$\{\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_2,..., \mathbf{A}_n\}$, block diag $\{\}$ is element matrix which have the element $\mathbf{A}_i$(i=0,1,2,,n).

$$\mathbf{A}_0=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix},\mathbf{A}_i=\begin{bmatrix}0 & 1\\ -\omega_i^2 & -2\xi_i\omega_i\end{bmatrix},i=1,2...,n$$

$$\mathbf{B}=[0\ \ b_0\ \ 0\ \ b_1\ \ 0\ \ b_2\ \ ...\ \ 0\ \ b_n]^T$$

The problem of rest-rest time-optimal attitude maneuver is to find a control function which is subject to constraint

$$\left|U(t)\right|\le U_0\ ,\ U_0=I^{-1}u_{\max} \tag{15}$$

driving the system described in Eq. (14) from the initial state

$$\mathbf{X}(0)=[-\theta_f\ \ 0\ \ \ 0\ \ ...\ \ 0\ \ \ 0]^T$$

to final state

$$\mathbf{X}(t_f) = [0 \quad 0 \quad 0 \quad 0 \ldots 0 \quad 0]^T$$

with minimum time t_f.

3.2 Time-optimal Control Solution and Its Characteristic

3.2.1 Bang-bang of Time-optimal Solution

It was investigated for time-optimal single-axis attitude maneuver of a flexible spacecraft without damp [5]. It was also investigated for the time-optimal single-axis attitude maneuver of a flexible spacecraft with damp [6][8]. They all have found out that time optimal solution is bang-bang control.

$$U^*(t) = -U_0 \operatorname{sgn}[\mathbf{B}^T \mathbf{Q}^*(t)], \quad t \in [0 \quad t_f^*] \qquad (16)$$

Q*(t) is the vector to be found meeting time-optimal control, in general, by means of numerical approach.

3.2.2 Antisymmetry of Time-optimal Control

In Eq.(16), setting □ □_ □_D_Dd_____□ĝ□□_____

$$U(t) = \begin{cases} \alpha U_0 \quad , \quad t \in [0 \quad t_1] \\ -\alpha U_0 \quad , \quad t \in [t_1 \quad t_2] \\ \vdots \\ (-1)^{k-1} \alpha U_0 \quad , \quad t \in [t_{k-1} \quad t_k] \\ (-1)^k \alpha U_0 \quad , \quad t \in [t_k \quad t_f] \end{cases} \qquad (17)$$

where $t_1, t_2, \ldots, t_k$ shows switch time with total k switches. The conclusion [5][6][8] is antisymmetric if the damp coefficient is zero, i.e. $\xi_i=0, i=1,2,\ldots,n$. In this case, k is add, $t_1=t_f-t_k$, $t_2-t_1=t_k-t_{k-1}$,as shown in Fig.8. In addition, mathematic simulation result show that optimal controls have 2n+1 switches if the system model is the type of rigid+n flexible modes without damp.

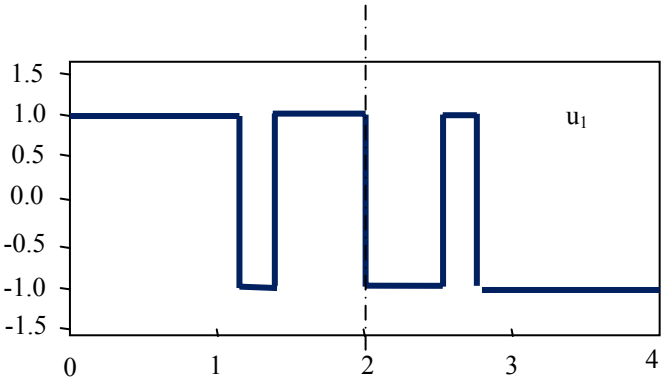


Fig.8 Antisymmetry of time-optimal maneuver control for flexible spacecraft

However, if the system damp is not zero, time-optimal control is not antisymmetric.

4. Implementation of Fast Attitude Maneuver

In engineering application for attitude maneuver of a spacecraft, not only attitude time should be as short as possible, but the feasibility, simplicity and reliability should be considered sufficiently. As a result, the attitude time is generally near-minimum not optimal in the fast attitude maneuver mission.

4.1 Choice for Attitude Maneuver Trajectory

The time-optimal control is not about eigenaxis rotation through the analysis of time-optimal attitude maneuver of a rigid spacecraft in chapter 2. As the time-optimal attitude maneuver controls have the characteristic of numbers of switches, the implementation will be very complicated. Switches has a trend to induce flexible modes. Rotation along the eigenaxis should be a better choice as its easy implementation in engineering application although it is not time-minimum but near-minimum.

Orbit reference coordinate is generally taken as a reference frame for ground observation spacecraft. If initial attitude quaternions in orbit reference frame are $\mathbf{q}_0$, final attitude quaternions are $\mathbf{q}_f$ in the same reference frame, then the attitude eigenaxis quaternions [17] is

$$\mathbf{q}_{bfo} = \mathbf{q}_o^{-1}\mathbf{q}_f = (q_{o0} - q_{o1}i - q_{o2}j - q_{o3}k)(q_{f0} + q_{f1}i + q_{f2}j + q_{f3}k) \tag{18}$$

eigenaxis vector is $\vec{q}_{bf0}$, Euler rotation angle along eigenaxis is

$$\alpha_m = 2\arccos(q_{bf00}) \tag{19}$$

The unit vector (i_x,i_y,i_z) of the eigenaxis vector $\vec{q}_{bf0}$ in the body-fixed frame is

$$\begin{cases} i_x = q_{bfo1} / \sin(\alpha_m) \\ i_y = q_{bfo2} / \sin(\alpha_m) \\ i_z = q_{bfo3} / \sin(\alpha_m) \end{cases} \tag{20}$$

4.2 Control Methods

4.2.1 Program maneuver + Stabilization

In the case of attitude maneuver axis determined, fast attitude maneuver is implemented by simplifying the time-optimal control of single axis maneuver described in the settion 2.3. Theoretically attitude maneuver control around a axis can be achieved by state feedback. However, time-optimal control cannot in itself provide a practical time-optimal orientation without taking into consideration some practical constrains, such as control time delay, uncertainty in the levels of the maximum control acceleration and deceleration, additional time constraints existing in the hardware, and so on. At the final end, $\dot{\theta}(t_f) = 0, \theta(t_f) = 0$ (or some value) cannot be achieve exactly.

However, we can use only one switch for simplicity. A stabilization process is introduced to stabilizing the system to rest at final attitude following the maneuver. Fig.9 illustrates the method "program maneuver +stabilization. The whole process includes two phases. The first phase is attitude maneuver which consists of acceleration and deceleration with a transition at the half maneuver angle $\varphi_0/2$. The second phase is a stabilizing the system to a near rest state at a approximate final attitude acceptable target attitude. The first phase can be implemented in a program control method. Transition point can be achieved by angle measurement. The second phase

is achieved by linear control method. The advantage of this method is only angle information is used. It has a drawback that stabilization time largely depends on the error at the transition point which in turn depends on the hardware delay, measurement error , and so on.

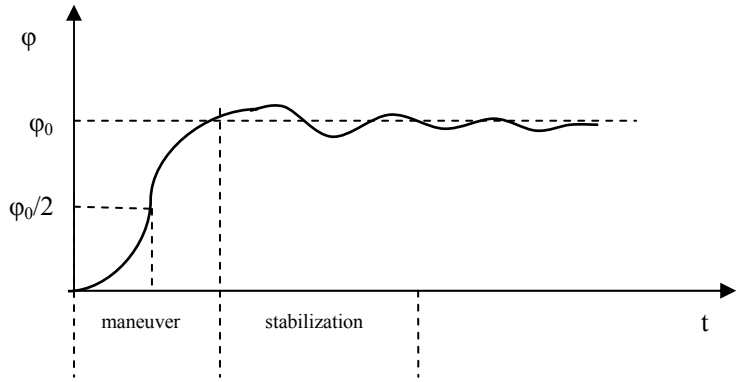


Fig. 9 Program maneuver + stabilization

4.2.2 Trajectory Planning Attitude Maneuver

The attitude maneuver around eigenaxis is easily implemented in practice even though it is not time-optimal path but a shortest one. The Euler angle φ is planned as a trajectory which rotation approaches as close as possible to. The trajectory is planned as follow

`

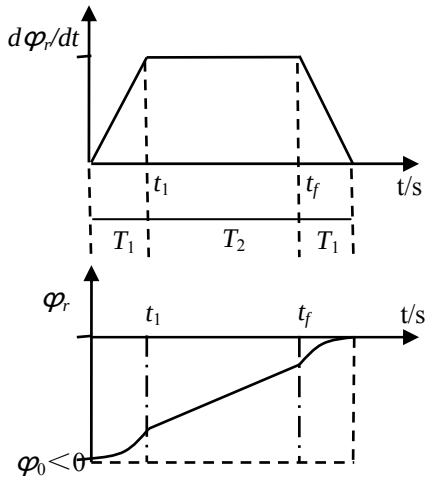


Fig.10 Trajectory planning

Let us also define e=φ_r-φ, $\dot{e} = \dot{\varphi}_r - \dot{\varphi}$

where φ_r and $\dot{\varphi}_r$ are the Euler command angle and its derivative. The torque command T_c will have the following expression:

$$T_c = ke + k_d\dot{e} \qquad (21)$$

where k,k_d are respectively proportional factor and derivative factor. The maneuver can be divided into 4 processes including acceleration, constant velocity, deceleration and stabilization period. The T_2 constant period can be canceled if the maneuver angle is small or the actuator capacity (for example maximum moment for a wheel) is enough. Acceleration time T_1 and maneuver time T ($T=2T_1+T_3$) are important parameter the trajectory planning should focus on. The consideration are dynamic performance, acturator capacity, suppression of vibration, and so on.

4.3 Actuator for Attitude Maneuver

Actuator for attitude maneuver includes thrusters, moment wheels and control torque gyros. Thruster can provide large force so as to large torque, but it cannot put out high accuracy small width pulse as restricted by its minimum impulse. Spacecrafts cannot achieve high performance of stabilization at its target attitude through thrusters. Moreover, attitude maneuver by thrusters consume propellants. It is not adapt to frequent attitude maneuver task.

4.3.1 Attitude Maneuver Using Reaction Wheels

Reaction wheels are used not only for attitude maneuver but also for the attitude stabilization of a spacecraft. Setting the maximum angle is α_{max}, the maximun output torque is u_{max}, the moment increment of the wheel along maneuver axis during the period of rest-to-rest attitude maneuver is

$$h - h_0 = \sqrt{I u_{\max} \alpha_{\max}} \qquad (22)$$

where h_0 is the moment at the initial attitude.

When the wheel's maximum moment is less than the h described in above Eq., the attitude process needs to have a constant velocity phase, otherwise it depends on the angular velocity measurement capacity. A reaction wheel is used for small-scale spacecraft because of its small capacity of torque.

4.3.2 Attitude Maneuver Using Control Moment Gyros (CMG)

Control moment gyros (CMG) has a large capacity of output torque. CMG with a certain type of configuration (pyramid or five edges of cone) can be used for attitude maneuver or stabilization for large scale spacecraft (for example space station etc.). This paper give only a introduction of attitude maneuver using a mix of CMG and wheels, not deal with the CMGs group control. For a fixed axis attitude maneuver, two CMGs are used to quicken the maneuver on the basis of wheels control only for stabilization. These two CMGs are fixed with their gimbals perpendicular to the rotation axis as illustrated in the Fig .11. When two CMGs are driven in synchronization at reverse direction each other, a torque is produced along the fixed axis.

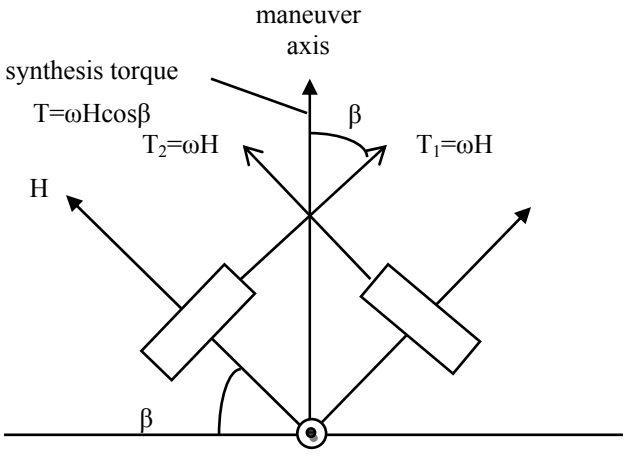


Fig.11 maneuver using two CMGs

In Fig.9, the gimbals rotation axis of CMG A and CMG B are all perpendicular to the paper. The angular velocity vector of CMG A's gimbal is towards to readers, that of the CMG B's gimbal is in the reverse direction. The torque provided along the maneuver axis is $2\omega H\cos\beta$.

5 Problem to be Focused on for the Fast Attitude Maneuver of a Flexible Spacecraft

In some missions, the spacecrafts have complicated structure with appendages such as solar panel, various antennas, and so on. These appendages have obvious flexible modes with low

rigidness, weak damp and centralized modes. The pointing accuracy of spacecrafts will become worse because of appendages vibration coupling with the main body of spacecrafts. It is significant that not only the fast attitude maneuver is implemented but also the residual vibration can be suppressed to a degree as low as possible. The suppression of vibration for flexible spacecraft can be classified into two categories which are consequently active control and passive control. Active control is rarely used in practice task because it needs vibration measurement devices and actuators producing displacement. Passive control can also be divided into two classes. The first one is frequency separation which designs a low system bandwidth away from base mode frequency of flexible modes. It weakens the residual vibration in the price of decreasing attitude maneuver velocity. The second is that at the final attitude the residual vibration is suppressed to a low degree regardless its process has vibration. The frequency separation is widely used in the practice. The second one has become a research attracting more and more investigators to study on. The typical method is input shaping and combination of input shaping and control methods. Different from input shaping, ZHANG Honghua [20] studied vibration suppression of attitude maneuver in the case of sequence square wave. Based on the time-optimal of a flexible spacecraft, the problem to be concerned and possible scheme are put forward for the flexible spacecraft fast attitude maneuver.

5.1 Planning Command Filter

In the chapter 3, we achieve the conclusion that the solution of time-optimal attitude maneuver of a flexible spacecraft (taking single-axis maneuver as an example) is still bang-bang control. Even though the solution is difficult to find out, and difficult to be used in practice, we can use bang-bang control for a rigid body to plan the trajectory to achieve a fast maneuver. As a result, the trajectory generally contains at least one switch which probably excite the vibration of flexible spacecraft. A filter to planning trajectory should be used to eliminate unwanted frequencies near the resonant frequencies of the system. When the dynamic parameter of a spacecraft is known, Fourier function F(ω) should be constructed as illustrated in Fig.12.

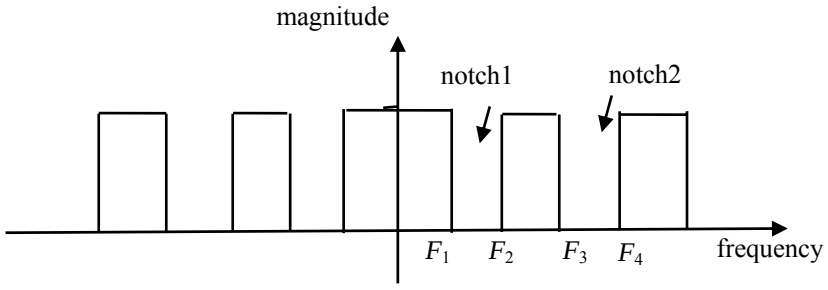


Fig. 12 Notch filter magnitude and frequency

where F1 and F3 are the lower frequency edges of the notch 1 and notch 3; F2 and F4 are the upper frequency edges of the notch 1 and notch 2 . F5 is the cutoff at the Nyquist frequency. F1 and F2 are determined by first mode frequency of the system in order to eliminate frequencies near to the resonance frequencies. Similarly, F3 and F4 are determined according to the second mode.

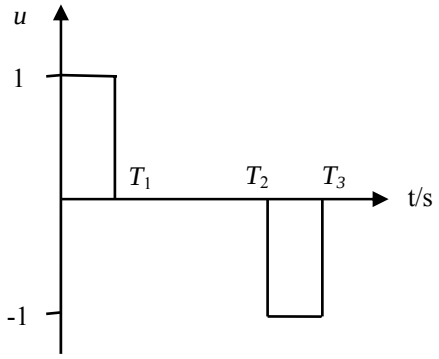


Fig. 13 Planning command function

Setting the planning command function is as follows. The corresponding time filter for the Fourier function F(ω)in Fig .12 is given by f(t). Then planning trajectory through the filter is

$$g(t) = \int_{-\infty}^{\infty} f(t-\tau)u(\tau)d\tau \qquad (23)$$

In general, this filtering technique belong to input shaping. It is used to remove the undesired energy input which can excite resonant vibration.

5.2 Residual Vibration Suppression Using Input Shaping

Input shaping is one of the methods suppressing residual vibration. It has advantage of needing no extra hardware, and shaping input commands through convolving impulses simply basis on the feedback system. That is we produce new input commands to drive the system by convolving the original ones with impulse sequences. These impulse sequences are what we should find out based upon prior knowledge of the closed loop system--to chose frequencies at which to eliminate vibration.

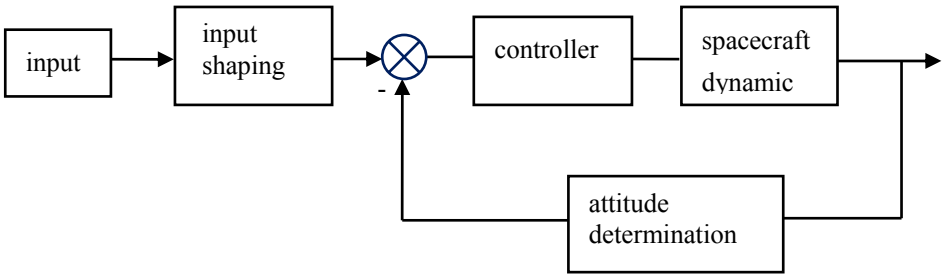


Fig.14 Input shaping control system block diagram

The original input commands are found according to the rigid main body motion. The impulse sequences are design out according to the mode frequencies and damp factor. Therefore, by using input shaping the rigid main body and flexible appendages can be independently account for. The achievement obtained before about rigid body can be used thoroughly. As the input shaper is located outside the closed system, it cannot change the original system stability.

The basic principle of input shaping suppression of vibration is that the system total response are the sum up of the individual response if two or more impulse applied to the system. Taking the second-order system as an example, giving the system an impulse will cause it to vibrate. If we apply a second impulse to the system, we can cancel the vibration induced by the first impulse in the case of the magnitude of the two impulse and the time interval between these two impulse being designed carefully.

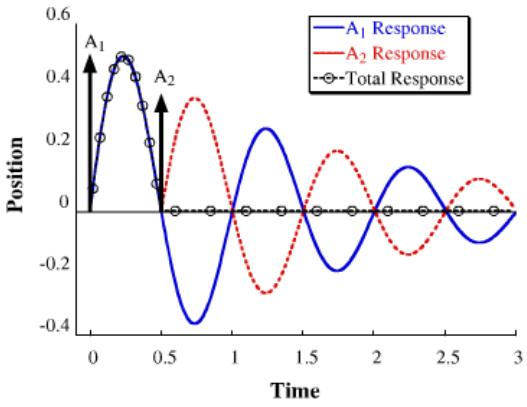


Fig.15 Two impulse response

For two impulses sequence, the response of the system with natural frequency ω and damping ratio ξ will be zero vibration (ZV) after the time the second impulse applied if the parameters of the impulse are chosen as follows.

first impulse time $t_1=0$ magnitude $A_1 = \frac{1}{D}$

second impulse time $t_2 = \frac{\pi}{\omega\sqrt{1-\xi^2}}$ magnitude $A_2 = \frac{K}{D}$

where $K = e^{-\frac{\xi\pi}{\sqrt{1-\xi^2}}}$, $D = 1+K$

The shaper of these two impulses sequence is called ZV shaper. The effect of eliminating vibration suing ZV shaper largely depends on the accuracy of the system parameters (ω,ξ). If there are errors in these values (and there always are), then the impulse sequence will not result zero vibration. Sensitivity of the shaping sequence to modeling errors is an important issue. In order to increase the robustness of the input shaping process, several kinds of shapers are developed. such as three impulse sequence (also called ZVD shaper) and four impulse sequence (ZVDD). In addition, the system is always not a second-order system, so multiple mode shaper are developed.

Following constraints should be account for in the design for the shaper.

a) amplitude constraint $\sum_{m} A_{Im} = 1, m = 2, 3, 1, m,$ m is numbers of impulse sequence. The sum of all impulses magnitude is 1 to guarantee the shaped commands driving the system to the same position as the command before shaped.

b) minimum time constraints. Time length (the final impulse applied to) should be as small as possible in order to bring the system a less time delay.

c) residual vibration constraint. The residual vibration should be under a desired degree after the final impulse applying on.

Actual spacecrafts are multiple mode system. Not only the flexible factors but also the liquid sloshing factor should be considered. However, the more modes be considered, the more impulses sequence can be obtained. This in turn leads to a large time delay. Therefore, a compromise should be account for between the effect of vibration suppression and feasibility of implementation in practice.

5.3 Active Vibration Control of Attitude Maneuver

Although input shaping has advantage of needing no extra hardware, it has shortcoming of time delay, bad effect for big model error. Moreover, input shaping is helpless for the task which require the little vibration during the maneuver because input shaping can only cancel the vibration after the time the final impulse applying.

Active vibration control uses vibration sensors measuring displacement of different places and find out control using algorithm to drive actuator to deform against the displacement of vibration. As the development of smart material and smart structure technology, vibration sensors and actuators has already begun in use in the practice. That provides hardware products for the active vibration control.

Although study and experiment on the active vibration control has developed theoretically, it has not widely used in space mission restricted by sensor and actuator technology. The lightweight of sensor and actuators, stick technique, and multi-dimension vibration control are still need to develop for the future application.

6 Conclusions

Time-optimal control problems, particularly as applied to spacecraft attitude maneuver, are complex and generally difficult to solve. Researcher continue to explore different aspects of the problem, particularly the areas of flexibility effects, specific spacecraft configurations, and solutions to be general three-dimensional reorientation of a rigid body. By numerical method, Ref.[2] obtained the conclusion that time-optimal attitude maneuver of a rigid body is bang-bang control around noneigenaxis under the cube constraint (independent three-axis controls). Ref.[5] have found the bang-bang control solution for single-axis maneuver of a flexible spacecraft. Time-optimal solution can not directly be used to space task. In space application, eigenaxis is generally choosed to be the maneuver axis, and either program maneuver + stabilization or trajectory planning maneuver method is used to get a fast attitude maneuver for quasi-rigid spacecrafts. For flexible spacecrafts, planning command filter or input shaping should be adopted to guarantee eliminating residual vibration. These approaches have advantage of needing no extra hardware. As the development of smart material and smart structure, active vibration control with sensors and actuators during the attitude maneuver will be used more and more widely in the future.

References

[1] XIE Xueshu. Optimal control theory and application [M]. Beijing: Tsinghua University Press，1986，167-189

[2] Bilimoria K D, Wie B. Time-optimal three-axis reorientation of a rigid spacecraft [J]. Journal of Guidance, Control, and Dynamics, 1993,16(3): 446-452

[3] Bai X L, Junklins J L. New results for time-optimal reorientation of a rigid spacecraft [J]. Journal of Guidance, Control, and Dynamics, 2009,32(4): 1071-1076

[4] Andrew F, Pooya S, Ross I M. Minimum-time reorientation of a rigid body [J]. Journal of Guidance, Control, and Dynamics, 2010,33(1): 160-170

[5] He Haifeng, ZENG Haibo. Time-optimal control for flexible spacecraft attitude maneuver [J]. Aerospace Control and Application, 2012,38(1):10-17

[6] Tuner J D, Junkins J L. Optimal large-angle single-axis rotational maneuvers of flexible spacecraft [J]. Journal of Guidance, Control, and Dynamics, 1980,3(6):578-585

[7] Tuner J D, Chun H M. Optimal distributed control of a flexible spacecraft during a large angle maneuver [J]. Journal of Guidance, Control, and Dynamics, 1984,7(3): 257-264
[9] Thompson R C, Junkins J L, Vadali S R. Near-minimum time open-loop slewing of flexible vehicles [J]. Journal of Guidance, Control, and Dynamics, 1987,12(1): 82-88
[10] Byers R M, Vadali S R， Junkins J L. Near-minimum time ,closed-loop slewing of flexible spacecraft [J]. Journal of Guidance, Control, and Dynamics, 1990,13(1): 57-65
[11] Barbieri E. Rest-to-rest slewing of flexible structures in minimum time [C]. The 27th Conference on Decision and Control, Texas, USA, December 1988, 1633-1638
[12] Singh G, Kabamba P T, McClamroch N H. Planar, time-optimal, rest-to rest slewing maneuvers of flexible spacecraft [J]. Journal of Guidance, Control, and Dynamics, 1989,12(1): 71-81
[13] Asher B J, Burns J A, Cliff E M. Time-optimal slewing of flexible spacecraft [J]. Journal of Guidance, Control, and Dynamics, 1992,15(2): 360-367
[14] Pao L Y. Minimum-time control characteristics of flexible structures [J]. Journal of Guidance, Control, and Dynamics, 1996,19(1): 123-129
[15] JIN Jin, ZHANG Jingri, LIU Zaozhen. Optimized design of controller parameters for large angle spacecraft attitude maneuvers [J]. Journal of Tsinghua University, 2009,49(2): 289-292
[16] ZHANG Meihua,ZHANG Bingwen. Cascade saturation controller for micro-satellite large angle attitude maneuver using reaction wheels [J]. Aerospace Shanghai, 2005,3: 15-18
[17] WANG Shuyi, WEI Chunling, LIU Qirui. A quick attitude maneuver algorithm for agile satellites [J]. Aerospace Control and Application, 2011,37(4): 36-40
[18] LIU Yiwu,ZHANG Jun. A quick attitude maneuver control method for agile satellites [C]. The 13th national conference on space and moving object control
[19] ZHOU Rui,GUO Yu,CHEN Qingwei. Study on unsymmetrical trajectory of attitude maneuver for a flexible spacecraft [C] . The 14th national conference on space and moving object control
[20] ZHANG Honghua, LI Kang, ZHAO Yu. Fast maneuver control flexible satellite system [J]. Chinese Space Science and Technology. 2005,2: 53-59
[21] Dimitry Gorinevsky， Georgeukovich。 Nonlinear input shaping control of flexible spacecraft reorientation maneuver [C]. Proceedings of the 1996 IEEE International Conference on Control Applications, Dearborn,MI, 1996
[22] QINGLEI HU, GUANGFU MA. Active vibration suppression in flexible spacecraft attitude maneuver using variable structure control and input shaping technique [C]. AIAA Guidance,Navigation, and Control Conference and Exhibit, San Francisco, California, 2005
[23] ZHANG Guoqi,DING Jianzhao. A daptive control method for attitude maneuver of flexible spacecraft [J]. Aerospace Control and Application, 2008,34(4): 23-27

QBITO DEVELOPMENT AND LESSONS LEARNED

Ignacio Barrios Tascón, Ana Laverón Simavilla , Efrén Moreno Benavides

Universidad Politécnica de Madrid, E-USOC, Spanish User Support and Operations Centre
Centro de Investigación y Desarrollo Aeroespacial, Campus de Montegancedo, UPM,
28223, Pozuelo de Alarcón, Madrid, SPAIN, +34914521626, ibarrios@eusoc.upm.es

ABSTRACT: QBITO is the first CubeSat developed by Universidad Politécnica de Madrid (UPM). It is a 2U CubeSat and is one of the satellites that compound the project QB50 lead by Von Karman Institute (VKI) in Belgium. The main task of QBITO will be to operate the Ion Neutral Mass Spectrometer (INMS) that is the primary payload on-board the CubeSat and that will study the properties of the lower thermosphere. The most outstanding feature of QBITO is the amount of new in-house developments that are present in the design. These include the Structural Subsystem, the Electrical Power Subsystem, the Communications Subsystem and a novel Antenna Deployment Mechanism. These new developments are complemented with Commercial Off-The-Shelf units in order to reach a robust, yet innovative, architecture.

Apart from the INMS, QBITO will carry three other payloads in order to take advantage of the mission as much as possible. The first is an experiment that will assess the performance of the Docosane as a Phase Change Material. It is being developed in collaboration with the University of Liège. The second experiment, the Medium Wave Infrared Detector, is developed by the Spanish company New Infrared Technologies and aims at testing a novel kind of uncooled detectors, in space conditions. Finally, the third additional payload, developed at UPM is the Experimental Software. It implements an attitude determination and control algorithm based on fuzzy control theory. The purpose is to test the suitability of this kind of control technique for spacecraft attitude control applications.

The CubeSat has been integrated and tested by QBITO team which includes UPM staff, professors and students and it will be ready for the launch scheduled in 2016. The team has also the support of the Spanish company SENER, which provides technical support along the project and the access to its environmental test facilities.

This paper will present an overview of the mission together with the main lessons learned in the design and AIT phases.

I. INTRODUCTION

QB50 Project

QB50 is a European 7FP project aiming to put into orbit an international network of 50 CubeSats for multi-point, in-situ measurements in the lower thermosphere and re-entry research. This concept has been under discussion for several years but, so far, no university, institution or space agency has taken the initiative to set up and coordinate such a powerful network.

QB50 mission consists of a network of 50 CubeSats in a ‘string-of-pearls’ configuration. The CubeSats will be launched from the International Space Station. The starting point of the mission will be a circular orbit at around 400 km altitude and 52º inclination, and due to atmospheric drag, the CubeSat orbits will decay until the spacecraft burn in the atmosphere. The separation distance between CubeSats will increase during the mission, eventually leading to a non-uniform distribution

of satellites all the way around the Earth. This happens due to atmosphere density variations along the orbit and variations in the individual CubeSat orientations, mass distributions, etc. The expected lifetime of individual CubeSats is between three and six months, depending on the launch schedule and sun activity.

The characterization of the lower thermosphere and ionosphere properties has been studied before by means of highly elliptical orbit satellites and sounding rocket flights, providing measurements during very small times and in small regions at a time. Also, powerful remote-sensing instruments have been used, which are not ideally suited for exploring the lower thermosphere because the atmosphere is so rarefied that the return signal is weak. QB50 idea is to acquire in-situ and long-duration measurements of key parameters and constituents of the lower thermosphere and ionosphere. This implementation is more complicated than the first two solutions, but results to be the best approach to study these atmospheric layers. The combination of a large number of CubeSats with identical sensors launched at the same time into a network will provide multi-point measurements that allow the separation of spatial and temporal variations of lower thermosphere/ionosphere characteristics.

The main reasons that lead the definition of such special mission, instead of other more common higher orbit missions, are:

- Scientific interest in the very low orbit atmospheric layers.
- The lifetime of the CubeSats in a very low-Earth orbit will be a few months only, which is much less than the 25 years stipulated by international requirements related to space debris. In addition, the educational objectives will be also met with a short mission duration.
- A very low-Earth orbit allows high data rates because of the short communication distances involved.
- In very low-Earth orbits, the CubeSats will be below the Earth's radiation belts, which is advantageous because CubeSats use low-cost Commercial-off-the-shelf (COTS) components which are not radiation hardened.

The 50 CubeSats will comprise about 40 atmospheric double CubeSats with standardized sensors and about 10 double or triple CubeSats for science and technology demonstration.

QBITO's objective is to accomplish the design, development, building, testing, launch and operation of one of the 40 double CubeSat carrying as main payload the Ion and Neutral Mass Spectrometer (INMS).

This project development has a high educational interest for universities that, along with the scientific and technological results, provides an exciting opportunity to be part of an international group of over 90 universities from all over the world working and sharing knowledge to achieve a successful mission.

Schedule

November 2011 was the kick-off date for QB50 project. At that time the call for proposals was made public for the interested institutions to apply. QBITO proposal was submitted by the team in April 2012. After being selected, the team was required to pass the Preliminary Design Review (PDR) in May 2013. The design made a successful PDR. The Critical Design Review (CDR) for QBITO was passed in September 2014. From that moment on, the development activities continued with the emphasis put on testing and integration activities. The Assembly, Integration and Test (AIT) Readiness Review was accepted in February 2015. Currently these testing activities are being carried out in order to pass the Flight Readiness Review which is due February 2016.

After being accepted for flight, QBITO will be sent to Innovative Solutions in Space (ISIS), in the Netherlands, in order for the main payload, the INMS, to be integrated. This will happen in April 2016. Payload integration activities will be done at ISIS premises, before the CubeSats are delivered to Nanoracks, QB50 launch provider, in the USA. Launch window will be opened in July 2016 and will last for several months. The main milestones of the project are presented in Table 1.

Table 6. Schedule for QB50 project.

Call for Proposals	November 2011
Submission of Proposal	April 2012
PDR	May 2013
CDR	September 2014
AIT Readiness Review	February 2015
FRR	February 2016
Shipment for Main Payload Integration	April 2016
Launch window start	July 2016

In May 2014, two precursor CubeSats were launched in order to test critical equipment that will have a great impact on QB50 performance. The precursor flight flew, among other equipment, an INMS unit and a Surrey Space Centre ADCS Bundle. These two units, as shown below will be part of QBITO, so a valuable feedback for the mission is being currently analysed.

Mission goals

QBITO mission encloses different goals related to both the QB50 project and university related advantages. Among the principal goals, the next can be found:

- Comply with QB50 mission requirements.
- Operate secondary payloads included by QBITO team.
- Gain spacecraft engineering experience for staff at Universidad Politécnica de Madrid.
- Involve university students in the design and development of the CubeSat.
- Test in-house developed equipment in real space mission conditions.

QB50 mission requirements were identified by QB50 consortium as the baseline to be followed by the teams. These requirements are based on the CubeSat standard and the main payload restrictions as presented in [1], [2] and [3].

The CubeSat is being developed by QBITO team which includes UPM staff, professors and students. This project is providing a valuable experience for everyone involved in it. It has given as well the opportunity to apply new design techniques. It is worth mentioning the use of the Object-Process Methodology, an emerging ISO standard that has been used for QBITO as the conceptual modelling framework at system engineering level, and concepts derived from modern Design Science methodologies such as Entropy-Based Design.

II. QBITO OVERVIEW

System Configuration

QBITO has been designed in order to comply with the requirements presented in 1.3. Design choices have been based on the search of functional robustness. QBITO is a two-unit (2U) CubeSat and redundancy has not been pursued due to constraints of volume and mass. The INMS was requested to be placed in –Z face and the nominal attitude for its operation is requested to be such that the –Z axis points towards ram direction.

Figure 1 shows two external views of QBITO where the appendages such as the antennas and the magnetometer boom can be observed.

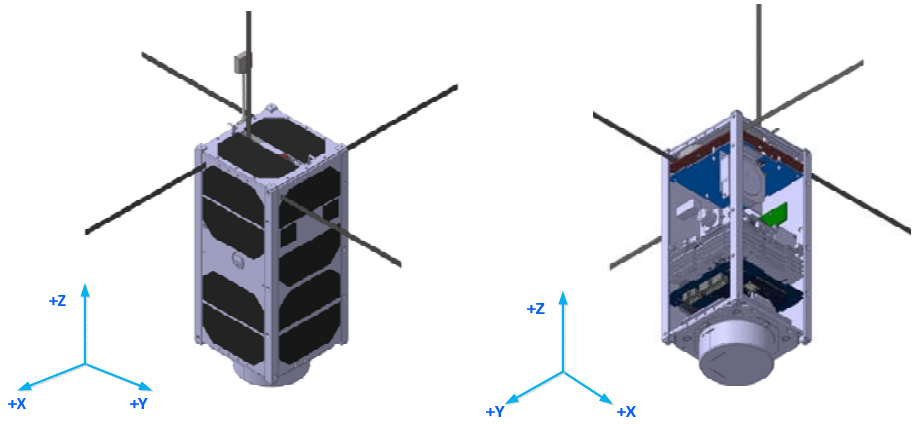


Figure 39. QBITO physical configuration

Figure 2 shows the block diagram where the general architecture of QBITO's subsystems with regards to power and data is presented. Indications are also presented for both analogue and digital lines.

The Electrical Power System (EPS) is in charge of setting the power supply commanded by the On-Board Computer (OBC). The INMS, the Attitude Determination and Control System (ADCS), the Communications (COM) subsystem and the Phase Change Material (PCM) payload are supplied by 5V and 3V3 buses. The GPS receiver only makes use of the 3V3 bus. The power switches offer protection against overcurrent events. A permanent regulated line of 3V3 is available for the OBC to be always ON. Another permanent regulated 5V line is available to supply power to the Medium Wave Infrared Detector (MWIRD) experiment. Finally the unregulated voltage bus provides power to the antenna and magnetometer deployment mechanisms and the momentum wheel.

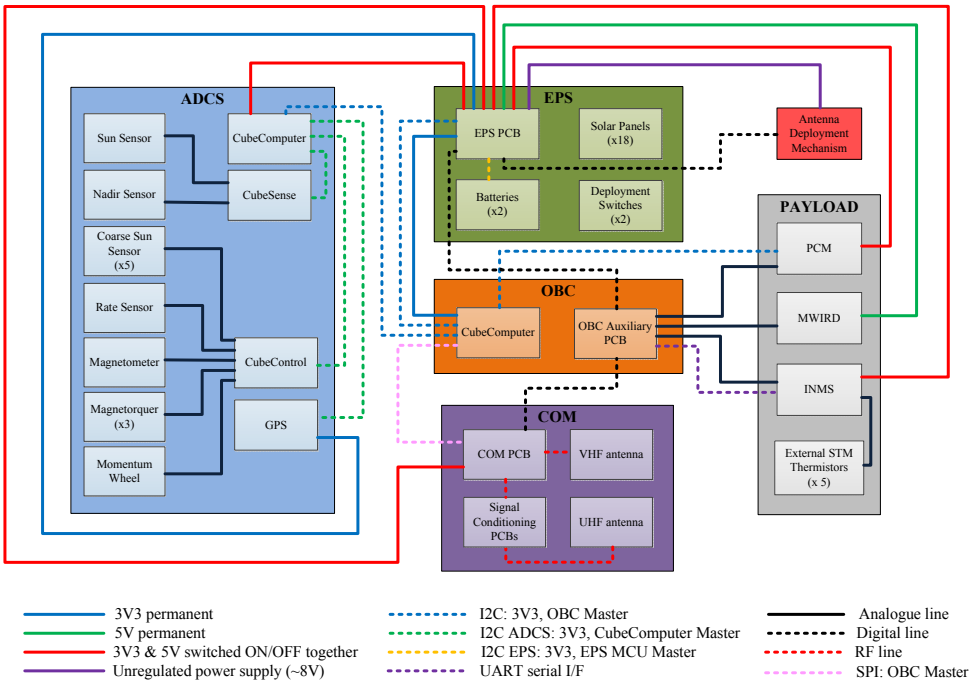


Figure 40. QBITO configuration diagram

Three I2C data buses are present in QBITO architecture:

System I2C: this is the main internal communications bus in QBITO. The OBC is the master in this bus. The slaves will be different subsystems that need communication with the OBC.

ADCS I2C: the ADCS includes its own internal communication bus. In here, the ADCS CubeComputer is the master and the rest of the ADCS equipment work as slaves.

EPS I2C: the EPS includes an internal I2C that is used to communicate the EPS microprocessor (master) with the batteries board (slave).

Other communication buses are used. The OBC is required to use UART communication protocol in order to communicate with the primary payload, the INMS, and SPI bus has been selected for communication between the OBC and the COM subsystem.

Apart from these communication buses there are certain lines that are used to get or monitor certain information. These are:

Analogue lines: the OBC acquires analogue measures of certain parameters by means of these connections.

Digital lines: different digital lines are used as interruption lines for the OBC and as enable lines for the antenna deployment mechanism.

RF lines: these lines transmit the RF signals from the COM PCB to the UHF and VHF antennas.

Most of the data and the power lines are routed through the PC104 bus that goes across the main PCB stack inside the CubeSat.

Modes of operation

Four different modes of operation have been defined for QBITO:

Start-up mode: this is the first mode of operation in which the CubeSat will enter after being deployed. The main purpose of this mode is to initialize the basic functions of the satellite and to deploy the antennas.

ADCS mode: this mode will recover the CubeSat from an uncontrolled tumbling state and put it in a controlled and defined attitude configuration.

Nominal mode: this is the intended mode for scientific operations. All the platform subsystems will be operational but, depending on the activities, different payloads will be working. Downlink will take place in this mode, as well.

Safe mode: this mode has been designed in order to provide with a solution for a case in which the battery charge level falls down under a threshold. Certain failure situations will also lead the satellite to this mode. A passive safe mode with no attitude control has been designed where all the payloads and the ADCS will be turned off.

These modes will cope with all the different scenarios where the CubeSat is required to operate and all the contingencies it might encounter during the whole mission. Due to the limited access time between the CubeSat and the ground stations, the automation of the transitions between the different modes is pursued. Nevertheless, in order to have a higher control of the system, ground commanded transitions have been enabled.

Mass and power budgets

Table 2 presents QBITO mass budget. Efforts have been made in order to reduce the total mass of the CubeSat from the beginning of the project.

Table 2. QBITO mass budget

Subsystem	*Mass with margins (Kg)*
Structure	0.45
EPS	0.25
OBC	0.09
ADCS	0.42
COM	0.05
Mechanisms	0.18
TCS	0.02
Payloads	0.25
Harness and others	0.07
Total	**1.78**

The maximum expected orbit average power consumption is around 2.2 W for the Nominal INMS mode. For this mode the nominal attitude provides the CubeSat with an orbit average power generated of around 2.3 W, coping with the demand. The set of two batteries provide a storage capacity of 20 Wh, and the estimated maximum depth of discharge (DoD) in eclipse shows values no higher than 14%.

The peak power presents values of 3.3 W, 2.1 W and 1.4 W for 3V3, 5V and unregulated power lines respectively. The EPS has been dimensioned to provide these levels of power although the expected ones will be lower.

III. SUBSYSTEMS

Structure

The structure has been developed in-house with the purpose of getting a lightweight multi-mission configurable structure. It was conceived for the mission when QBITO was selected to be part of QB50 project in April 2012. The concept was delivered as part of the QBITO PDR and was presented in the 5th European CubeSat Symposium held in Brussels in June 2013, in [6].

QBITO structure can be divided into primary and secondary structures. The first one comprises an aluminium alloy skeleton which includes the rails that interface with the deployer system, two main squared frames at the top and bottom of the CubeSat (Z axis), and FR4 laminate side panels. The skeleton and the side panels act as a whole in order to provide global stiffness and load-bearing capacity. The secondary structure includes the subsystems PCBs and the spacers that connect these PCBs to the main structure.

The technique used to do the conceptual design was based on applying the design theory "Entropy-Based Design" described in [7], which aims at selecting a concept that keeps as low as possible the number of internal dependencies. The result of applying this theory led to a solution that has the following characteristics:

- Maximum number of degrees of freedom, which is related to the possibility of changing the design.
- Maximum flexibility, which is related to the possibility of adding new components.
- Maximum number of translational and rotational symmetries, which is related to the possibility of moving components.
- Maximum modularity which is related to the possibility of changing a component.
- Minimum number of components, which is related to improvements in mass, cost and reliability..
- Maximum number of equal components, which is related to the reduction of cost.

The result can be seen in Figure 41 where arrows remember that it allows three different 90º rotations, one translation along Z-axis and one symmetry around Z-axis. The result is a modular configuration with enough flexibility to satisfy the requirements of quite different missions.

Finally, the side panels act not only as load-bearing structural members, but are an integral part of the electrical power system. They support the solar cells that provide the power to the satellite and have copper tracks that act as power buses. The FR4 laminate employed is highly-available in the market and is heavily used as a substrate for printed circuit boards because of its excellent electrical, mechanical and thermal properties.

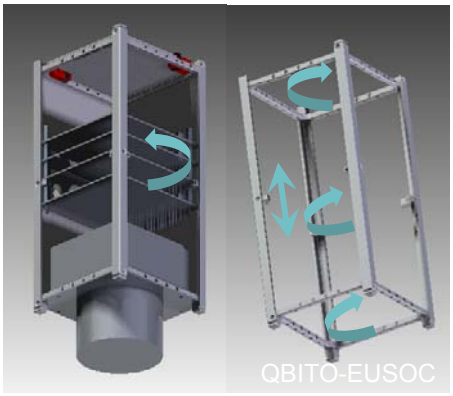


Figure 41. QBITO structure concept

EPS

The EPS is in charge of managing the power supply on-board. The power is obtained from the solar radiation by means of solar cells that are installed in the exterior lateral panels. Five faces of the satellite are covered by a total of 16 solar cells, leaving one of the faces free to be used by the INMS. Triple junction GaAs AzurSpace's 3G30C Advanced solar cells have been selected due to their efficiency of 30%.

The EPS optimizes the power received using five Maximum Power Point Trackers (MPPT). This power is regulated to the proper voltage levels required by the different equipment (3V3 and 5V) and distributed using dedicated switches. For some units such as the antenna deployment mechanism this voltage level will not be regulated. Finally, the EPS will be in charge of managing the charge and discharge processes of the set of two batteries (2.8 Ah, 3.8V each) selected for the mission.

The EPS has five pairs of controlled switches. Each pair is able of switching on/off a 3V3 line and a 5V line, both protected by a dedicated thermal fuse. This design is easily adaptable to different needs that might appear in future missions. The EPS SW is an in-house development that runs in an ultralow consumption microcontroller manufactured by Texas Instruments.

The EPS comprises two PCBs. The first one has the main components for the power control and distribution (MCU, MPPTs, DC/DC converters, etc.) and the second one protects and controls the charge/discharge processes of the batteries. The batteries are placed inside an aluminium protection housing on top of this PCB. Two heaters are in charge of controlling the temperature of the batteries so that cold environments do not decrease the batteries performance. The EPS weight 205 grams.

OBC

The OBC will manage all the processes that must take place during the operational lifetime of the satellite. It will manage the different modes of operation and the logic that is underneath them. In

order to do so, it will monitor all the housekeeping parameters coming from the different subsystems. It will operate the different payloads and will be in charge of storing and packing the relevant data that will be downloaded to ground. The OBC is the core of the communications on board, acting as the main I2C bus master.

The CubeComputerV3 developed by Electronic Systems Laboratory, was selected to act as the OBC. It provides with a low power 32-bit ARM Cortex-M3 based microcontroller. It operates at 4-48 MHz delivering 1.25 DMIPS/MHz. The CubeComputer has a memory subsystem which includes 256 KB EEPROM, 4 MB FLASH, 2 x 1 MB SRAM and 2GB microSD card storage.

The software has been developed in-house. It has been mainly developed in C using the Free-RTOS libraries. Free-RTOS is a Real Time Operating System fully compatible with the CubeComputerV3.

ADCS

The ADCS is in charge of controlling the attitude dynamic state of the CubeSat. The QB50 ADCS Bundle developed by Surrey Space Center is used for this task. It is based on the use of three magnetorquers (one per axis) and a momentum wheel. The processing unit will be a dedicated CubeComputer, equal to the one acting as OBC. Among the sensors are the sun sensors (1 fine and 5 coarse), 1 Earth sensor, 1 rate sensor, 1 magnetometer and 1 GPS. The bundle has been specifically designed for this mission and has been verified against the attitude requirements imposed for QB50 project. The two precursor CubeSats carry this unit and its performance is currently being assessed.

COM

QBITO COM is based on a full duplex architecture with UHF for downlink and Beacon, and VHF for reception. Transmission is based on FSK modulation with rates of 9600 and 1200 bps for downlink and uplink, respectively. The subsystem makes use of programmable transceivers which are configured by the OBC. The COM board is divided into two main parts, transmission (it includes the beacon) and reception. The COM status is monitored and controlled by the OBC. The data interface is based on SPI bus, with power lines of 3V3 and 5 V.

Regarding the antennas, a four monopole turnstile is used for UHF achieving circular polarization and omnidirectional pattern, whereas a monopole is used for the VHF.

Antenna Deployment Mechanisms

The antennas will be deployed with a novel deployment mechanism designed and developed in-house. Four equal mechanisms are used for the deployment of the four monopoles that compound the UHF antenna. Another mechanism scaled to a larger dimension is used for the deployment of the VHF antenna.

This mechanism is an optimized antenna deployment mechanism that fulfils the next requirements:

- Minimum impact in the lateral faces.
- Minimum impact in volume, mass and cost.
- Minimum number of parts in order to have maximum reliability.
- Minimum number of manufacture tolerance.

The mechanism is based on a torsional spring which is previously rolled inside a cylinder. This results in the antenna being pushed out and deployed. Special care has been taken into account for mitigating the problems derived from the friction between the antenna and the walls of the cylinder.

TCS

The thermal control is carried out by a control system based on thermal coatings and paints.

The thermal control design also includes two heaters that are in charge of maintaining the batteries within their operational temperature range. This will cope with all the thermal loads which have been identified for the different operating scenarios for QBITO mission.

The thermal design was made using the software ESATAN, analysing all the extreme scenarios where QBITO could operate given the initial uncertainties related to the launch and final orbit.

IV. PAYLOADS

INMS

This is the main payload of QBITO. It will be in charge of taking in-situ measurements to characterize the composition of ions and neutral particles and their spatial variation. This will provide key measurements that will improve significantly the determination of atmospheric variables and will provide with key inputs for atmospheric models.

The INMS is a miniaturised analyser designed for sampling of low mass ionised and neutral particles in the spacecraft ram direction with the instrument resolutions optimised for resolving the major constituents in the lower thermosphere, i.e., O, O2, N2. The INMS development is being carried out by Mullard Space Science Laboratory (MSSL).

MWIRD

The MWIRD is an experiment that will test the behaviour in space of an uncooled medium wave infrared detector based on the new Vapour Phase Deposited (VPD) PbSe technology. The experiment has been designed by New Infrared Technologies, a Spanish company developer of this technology. Two detectors will be tested in QBITO. One will point towards nadir and the other towards zenith direction.

PCM

The PCM experiment aims at testing the performance of Docosane as a useful phase change material for future use in space thermal control systems. The experiment has been developed together with the Belgian company Walopt. A flight experiment is mandatory to increase the TRL of this technology. A differential test experiment has been designed. An identical thermal environment will be imposed to two experimental cells, one as a witness cell (in fact a thermal spreader) and another one as a cell equipped with a PCM. The thermal evolution of the cells will be recorded and compared to each other.

ESW

The experimental software (ESW) is an attitude determination and control software that will be tested on QBITO taking advantage of the hardware included in Surrey's QB50 ADCS Bundle. It will aim at demonstrating novel fuzzy control techniques in orbit. It is capable of performing full attitude determination and control, tailored for QBITO mission.

It is expected to validate separately the orbit and attitude determination module and the attitude control module. Among the benefits that can be expected from the ESW development there is the consecution of a more powerful controller, which deals easily with inaccuracies of the satellite model and that can be quite easily adaptable to other missions. Also, it is expected that it will improve the performances while reducing the power consumption of the whole subsystem.

Payloads operation

The operation of the INMS has been required to be managed by the OBC by means of uploaded scripts generated by the mission scientific team. These scripts include both direct payload commands and time tagged commands which will define the functioning of the instrument. This

philosophy has been also adopted for the operation of the secondary payloads. Therefore, the PCM and ESW experiments will be managed with the same kind of commands scripts.

The MWIRD experiment will not follow this method and will be active at all times, since the power required to operate is very low (~1 mW).

V. GROUND SEGMENT

The ground segment for the QBITO mission is composed of the following elements:

- Four ground stations, which will be the interface to communicate with the spacecraft. Two of them are placed in Madrid at E-USOC (ETSIA, UPM) and ETSIT (UPM) and the other two are placed in Lima, Peru (UNI) and Mexico City, Mexico (UNAM).
- The Mission Control System (MCS), to support the operations and AIV phases. The MCS comprises two different software units and a database:
  - A TMTC Packet Manager (PM).
  - The Satellite Control Software (SCS).
  - The Mission Database (MDB).
- Human Control Interface (HCI), used to monitor and command the spacecraft.
- One File Server for the mission data archival, located at E-USOC.

Additional features of the QBITO Ground Segment include a planning tool to optimize the uplink and downlink concept, and the possibility to add supplementary ground stations to the network.

Figure 4 shows an overview of these elements and the relations between them.

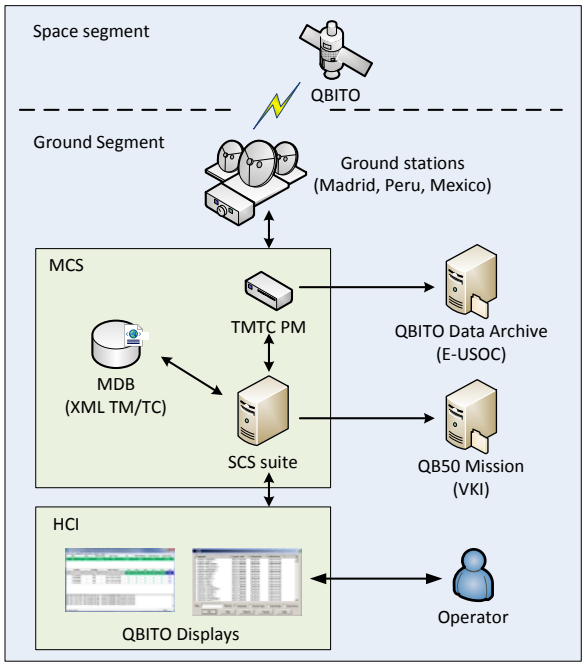


Figure 42. QBITO Ground Segment architecture

VI. AIT PLAN

A hybrid model philosophy has been followed for QBITO. Two models have been used for verification purposes:

- QBITO Structural Model (SM): this model has been used to subject the structure design to structural tests that permit to validate the numeric models. This was considered appropriate since the design were completely new and quite different from other designs present in the CubeSat field.
- QBITO Proto-Flight Model (PFM): this is the final model that will eventually fly. A PFM philosophy has been chosen due to programmatic and cost reasons. Proto-flight qualification tests were carried out on the PFM in order to include in one test campaign both qualification and acceptance requirements verification. This way the test activities are reduced while the verification level is still valid.

Apart from these system models, development models (DM) and engineering qualification models (EQM) were used for different subsystems. New in-house developments underwent qualification and acceptance tests according to ECSS standards.

Most of integration and test activities have been mainly developed at the Spanish User Support and Operations Centre, E-USOC premises where a clean room and several laboratories were used.

The environmental testing has been done in the Spanish company SENER, that has given technical support along the project and that has provided access to their test facilities, and AIRBUS Space and Defence, where the shock test was performed.

Figure 43 summarizes the main test activities for QBITO SM and PFM. QBITO SM has already been successfully qualified against vibration and shock tests. The PFM has undergone the shock test with success. Next tests are planned to be run during the last months of 2015.

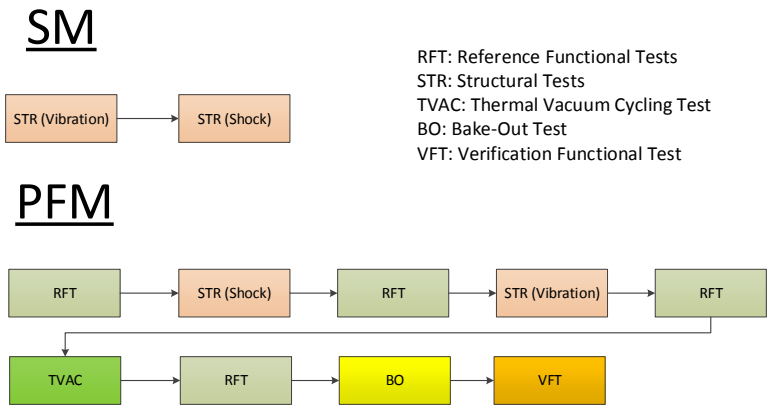


Figure 43. QBITO test activities

VII. CONCLUSIONS

This paper has introduced QBITO mission as part of the QB50 project. The main features of this project have been presented followed by a review of the main characteristics of QBITO design.

The main task of QBITO will be to operate the four payloads that will carry on-board: the INMS, the PCM, the MWIRD and the ESW. The mission will also aim at validating the new in-house developments such as the structure, the EPS, the COM subsystem and the antenna deployment mechanisms.

QBITO is not just the CubeSat itself. It also comprises the ground segment needed to communicate with the spacecraft. At the moment, four ground stations will be in charge of the

downlink of all the valuable data generated on-board and the uplink of instructions necessary for the mission completion.

A hybrid model philosophy has been followed for QBITO. Two models have been used for verification purposes: the structural model and the protoflight model. The first has already been qualified for vibration and shock requirements. The PFM has been tested against shock and the main functional requirements. Future work includes the refinement of the PFM functional performance and the vibration and thermal vacuum test campaigns.

QBITO team expects to receive more valuable experience and lessons learnt as the project evolves. The team includes around fifteen students, most of which are making their contribution to the project the main topic of their final thesis. This number is expected to evolve, due to the duration of the project, and there will be a rotation in which new students will become part of the team while others will leave. The opportunity for the students to take part of a real space project is one of the main motivations of the university to participate in QB50.

REFERENCES

[1] A. Denis, *QB50 System Requirements and Recommendations*, Issue 7, 13th February, 2015.

[2] R. Munakata, *CubeSat Design Specification*, Rev. 12, California Polytechnic State University, 8th January, 2009.

[3] R.A. Chaudery, *QB50 INMS Science Unit Interface Control Document*, Issue 12, University College London, 27th August, 2015.

[4] Electronic Systems Laboratory, *CubeComputerV3 Datasheet*, Revision F.

[5] L. Visagie, *ADCS Interface Control Document*, version 3.0, Surrey Space Centre, 18th June, 2014.

[6] E. Moreno, et al. *QBITO EPS and Structure as Technology Demonstrators for Students* , Oral presentation, 5th European CubeSat Symposium, Brussels, June, 2013.

[7] E. Moreno, *Advanced Engineering Design: An integrated approach*, 1st Edition, Woodhead Publishing, 2012.

[8] QB50 Project Website: www.qb50.eu

[9] Universidad Politécnica de Madrid Website: www.upm.es

[10] E-USOC Website: www.eusoc.upm.es

KM5B – A LARGE SPACE ENVIRONMENT TEST FACILITY OF CHINA

Zhang Shiyi Chen Li Qi Xiaojun Peng Guangdong Kang Shuai Li Hongyu

Shanghai Institute of Spacecraft Equipment, Shanghai 200240,China

ABSTRACT: This paper introduces a new large space environment test facility, named KM5B, which is the second largest space environment test facility in use of China. The KM5B is capable of simulating space vacuum, cold black background and space thermal irradiation. There are three vacuum chambers with the main one being 10m in diameter by 22.4m high, and a vacuum chamber of 7m in diameter by 8m long is in plan. The chambers can be pumped down to the level of 10^{-6}Pa. The thermal shroud can be cooling down below 95K by liquid nitrogen (LN_2) circulation system. The KM5B is used for thermal balance test and thermal vacuum test of the satellites belonging to SAST5000 Satellite Platform and SAR series.

0. INTRODUCTION

The space environment test facility (also called space environment simulator) is capable of simulating space vacuum, cold black background and space thermal irradiation for thermal balance test and thermal vacuum test of the spacecrafts[1].

To meet the test requirement of the satellites belonging to SAST5000 Satellite Platform and SAR series, a large space environment test facility, named KM5B, was built in 2013 in Shanghai of China. The KM5B facility, which includes two large vacuum chambers of 10m in diameter and 7m in diameter, is the second largest space environment test facility in use of China [1-3]. It performed perfectly in the first real test of large satellite in the auxiliary chamber in June 2014 and in the main chamber in October 2014.

The whole facility includes the vacuum chamber, the thermal shroud, the vacuum pumping system, the nitrogen system, the thermal simulation system, the central control system. In addition, another vacuum chamber for extend is in plan.

1. MAIN DESIGN PARAMETER

Main design parameters of KM5B are listed below:

a. Chambers size

·Effective size of the main chamber: Φ10m×11m;

·Effective size of the auxiliary chamber: Φ7m×8m;

b. Thermal shrouds size

·Effective size of the main chamber: Φ9m×11m;

·Effective size of the auxiliary chamber: Φ6.2m×8m;

c. temperature: <100 K;

d. Refrigerateing capacity: 200 kW;

e. Vacuum pressure: $<5\times10^{-4}$Pa;

f. Thermal simulation: infrared simulator;

g. Temperature measurement channel: 2000 (extensible).

2. THE VACUUM CHAMBER

There are three vacuum chambers: the main one is 10m in diameter and 11m high, the auxiliary one is 7m in diameter and 8m long, another one for extend in further is 7m in diameter and 1m long. And a vacuum chamber of 7m in diameter and 8m long for larger SAR is in plan, and then the effective length of whole chamber will be 26m.

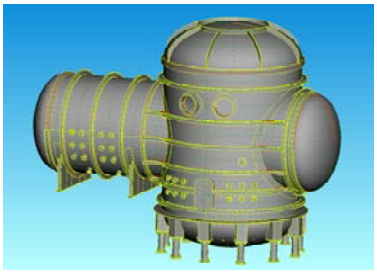


Fig. 1 The chamber shape at present （front）

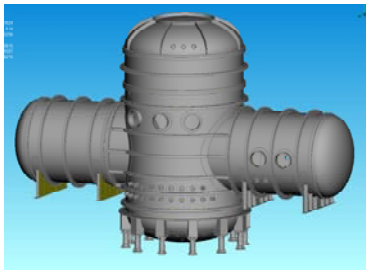


Fig. 2 The chamber shape in final completion (back)

The door of the auxiliary chamber opens in rotation way to improve the use efficiency of hall space, with the maximum opening angle of 135°.

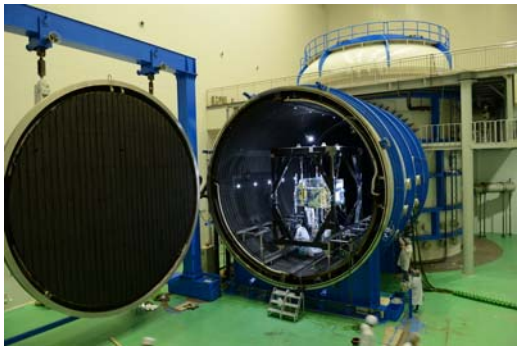


Fig. 3 The photo of KM5B

The door of the auxiliary chamber opens in rotation way to improve the use efficiency of hall space, with the maximum opening angle of 135°.

It is so difficult to make two holes of 7m in diameter in a chamber of 10m in diameter. Locally reinforcement is applied in the intersection of chambers to make sure the strength and stability of the whole chamber. It is confirmed to meet the requirements perfectly through FEM simulation and vacuum test. The Fig. 4 and Fig. 5 are the FEM simulation results of KM5B in effect of chamber gravity, specimen gravity and external pressure.

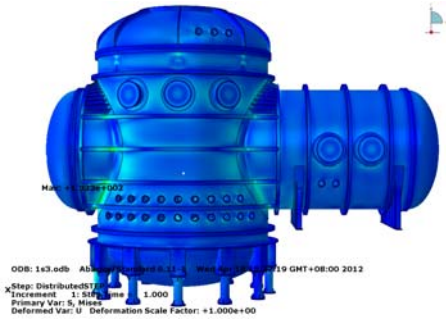


Fig. 4 The stress distribution of chamber

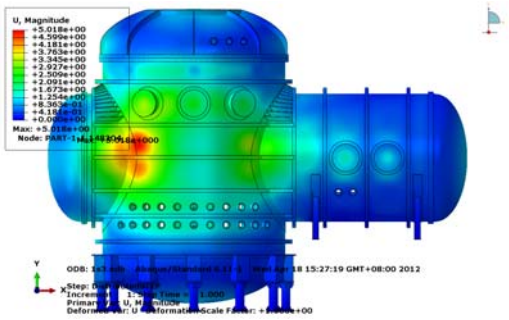


Fig. 5 The displacement distribution of chamber

3. THE THERMAL SHROUD

All the chambers are shrouded inside. The effective volume of the thermal shroud in the main chamber is 9m in diameter by 11m high, and the auxiliary one is 6.2m in diameter by 8m long. The thermal shroud is made of stainless steel (1Cr18Ni9Ti) tubes welded with copper (T2) wing piece, with the maximum heat load of 200kW. The inner surface of the shroud is painted with black paint with a hemisphere emissivity of larger than 0.9.

The removable thermal shroud is set in the intersection of chambers, which make it possible to separate the main chamber and the auxiliary chamber. The chambers can be used individual or together with the removable thermal shroud close or open, so the use efficiency and economy of the facility is improved.

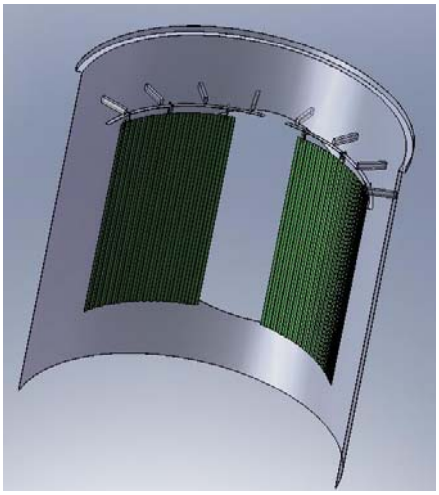


Fig. 6 The removable thermal shroud

4. THE VACUUM PUMPING SYSTEM

The KM5B facility has a three-stage vacuum pumping system, which includes rough pumping system, middle pumping system and high pumping system, shown in Fig. 7.

The rough pumping system includes six dry compressing/ roots vacuum pumps set, which can pump the chambers from 1atm to 5Pa in 2.5 hours. The middle pumping system includes five molecular pumps each of 3200 l/s (N_2). The high pumping system includes seven cryopumps each of 50000 l/s (N_2). The chambers can be pumped down to the level of 10^{-6}Pa by the three-stage pumping system with no load.

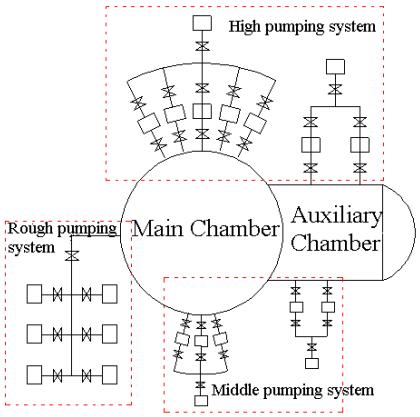


Fig. 7 The principle diagram of vacuum pumping system

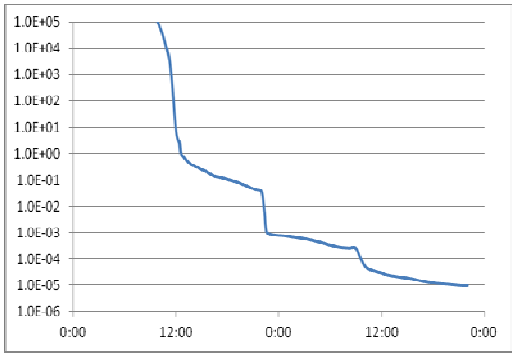


Fig. 8 The curve of pressure

5. THE NITROGEN SYSTEM

The nitrogen system has two parts: liquid nitrogen system and gaseous nitrogen system.

The liquid nitrogen system is mainly used for supplying LN_2 to thermal shroud to make sure the temperature of thermal shroud below 100K. Also, the liquid nitrogen system is able to supply LN_2 to the cryopumps. There are two running ways of the system: monophase close-loop circle drove by LN_2 pumps and open-loop circle, with refrigerating capacity of 260kW.

The gaseous nitrogen system is mainly used for thermal shroud warming from 100K to 300K at the end of test. The system includes two compressors, one heat exchanger, one electric heater, two-level filter, and etc. It is able to run in close-loop circle drove by compressor, and run in open-loop circle warmed by electric heater.

6. THE THERMAL SIMULATION SYSTEM

There are more than 300 program-control DC power supplies and an accurate control program, which supply steady and accurate power to the IR heater (such as infrared lamp array and infrared heating cage) and the heating chip. For different test, the close-loop control mode, open-loop control mode and switch control mode can be chose to simulate thermal flow accurately under steady condition and transient condition. The thermal simulation system can be extended to 2000 heating channels if the test need.

7. THE CENTRAL CONTROL SYSTEM

All of KM5B subsystems are monitored by the central control system remotely and centralized, which includes central control part and local control part, shown in Fig.9. The information transfers between the two parts through network servicing devices. All subsystems can be controlled either on local or on remote computer, with the switching control between local and remote without interference. Lots of logical interlock is set in the control program to prevent misoperation and make the facility safely. A quadrupole mass spectrometer and a quart crystal microbalance are used to monitor the gas components and molecular contamination condensation.

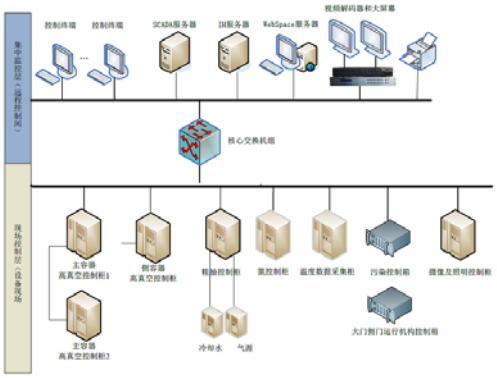


Fig. 9 The structure configuration of central control system

8. SUMMARY

Many tests of satellites have been carried out successfully in KM5B after the facility put into use, in which KM5B is steady and reliable. And a vacuum chamber of 7m in diameter and 8m long for larger SAR will be built in further. The success of KM5B, which is the second largest space environment test facility in use of China, will play a powerful role in the development of Chinese aerospace.

REFERENCE

1. Huang Bencheng, Tong Jingyu, “Space Environment Engineering”, Chinese Science and Technology Inc., 2010;

2. Huang Bencheng, “Space Environment Test Facility for Manned Spacecraft”, Chinese Space Science and Technology, 2002,(3) :1-6;

3. Peng Guangdong,Qi Xiaojun,Chen Li, “Development of KM5A space environmental simulation test facility”, Spacecraft Environment Engineering, 2010, 27(4):485-488.

Design and Manufacturing of Spacecraft Shell Structures Based on Automated Fiber Placement

Ke Wang, Zhihui Li, Li Chen, Tao Song

Shanghai Institute of Spacecraft Equipment

Abstract: This paper presents the design and manufacturing of the typical shell, the central tube of a satellite, based on the AFP technique. Fiber placement paths are designed and optimized aiming at the cylinder-cone shell. The finite element model of the central tube is constructed, and the FEA results demonstrate that the central tube meets the structural requirements.

Keywords: Carbon Fiber Reinforced Plastic, Automated Fiber Placement, Central Tube

1. Introduction

The shell is one of the most popular shapes of the spacecraft structure such as the central tube of satellites and the space telescope cylinder [1, 2]. Automated Fiber Placement (AFP) technique is a new advanced composites material structures manufacturing technology and is suitable for the shell [3-6]. AFP technique is widely employed in the aerospace field. For example, the central tube of Eurostar 3000 satellite platform is manufactured by using AFP [7], as well as the fuselage of Boeing 787 [8].

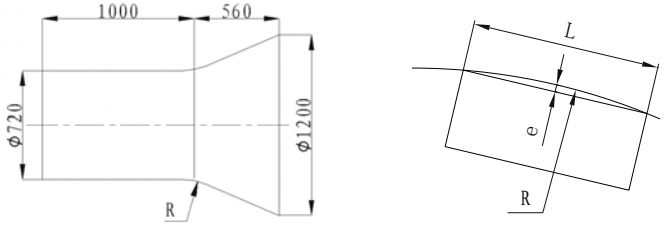
A typical AFP system consists of the machine tool and the fiber placement head, with at least 7 axis pairs in simultaneous coupled motion. The fiber placement head is crucial for fast and precise tows lay-up. It is able to manipulate each tow independently, including delivering, clamping, cutting off and restarting. Also, it is equipped with cooler and heater to adjust the viscosity of prepreg tows. Moreover, it can control the tension and pressure imposed on the tows automatically. The advantages of AFP technique are at least three-fold: 1. the prepreg tow orientation and position is controlled accurately; 2. the manufacturing process is simulated in advance; 3. the configuration of manufacture lay-up is recorded truly. Therefore, we employ the AFP technique to design and manufacture the central tube, which is described as follows in detail.

2. Fiber placement paths design

Fiber placement paths design is a key process of AFP [5, 9-10]. Generally, in accordance with the designed central tube, the 3D mandrel model is made, and then the fiber placement paths is generated and optimized with restrictions such as mechanics properties and manufacturing processes. In case of collisions between the fiber placement head and the machine tool (including the mandrel), fiber placement simulation is necessary, and helpful in locally modifying the fiber placement paths.

The designed cylinder-cone central tube is 1560mm in height. The upper part is a cylinder of 1000mm in height and 720mm in diameter, while the lower part is a cone of 560mm in height and 720mm/1200mm in diameter, as shown in Fig.1(a). The cylinder-cone shell of the central tube is designed in the form of a monolithic construction. The part of cylinder-cone transition is carefully treated in case of the body interference due to its saddle shaped curvature. Assume the width of compaction roller is L, the maximum elastic deformation of the compaction roller is e, and the

compaction factor is K. Then the minimum axial radius R of the cylinder-cone transition should be larger than K (e / 2 + L^2 / 8e), as shown in Fig.1(b).



(a) **(b)**

Fig.1 Minimum Axial Radius R of Cylinder-Cone Transition

The lay-up configuration of the shell is listed in Table 1. Hence there are three forms of the fiber placement paths, 0, 90, and ±45 (±45 can be taken as the same one). Assume the cylinder part is h_1 in height and Φ_1 in diameter, while the cone part is h_2 in height and Φ_2 in larger diameter; the number and width of prepreg tow are n and w respectively.

Table 1 Lay-up Configuration

Area	Lay-up	Number	Thickness/mm
Cylinder	[±45/0/90/0/90/0/±45]s	18	1.8
Cone	[±45/0/90/0/±45/±45/0/90/0/±45]s	28	2.8

For cylinder part, when placement angle is 0, relative to generatrix, the fiber placement path is along the generatrix, and the mandrel rotates 2nw / Φ_1 after each placement. When placement angle is 90, the true placement angle is 90 – arcsin (nw / $\pi\Phi_1$) to make sure that the fiber is continuous. When placement angle is ±45, as shown in Fig.2(a), the geometric equation is

$$\tan 45° = \frac{rd\theta}{\sqrt{(dr)^2 + (dz)^2}} = \frac{r_0 d\theta}{dz}, \tag{1}$$

where $r_0 = r = \Phi_1 / 2$. The fiber placement path is

$$\theta = \frac{z}{r_0} + \theta_0, \tag{2}$$

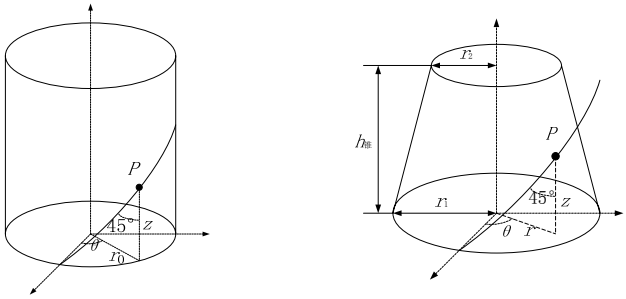
where z is the coordinate of the placement point, and θ_0 is the initial angle.

For cone part, when placement angle is 0, relative to generatrix, the fiber placement path is along the generatrix, and the mandrel rotates 2nw / Φ_2 after each placement. If tows are placed from larger end face of the cone to smaller one, the number of tows needs to be reduced, and vice versa. When placement angle is 90, the true placement angle is 90 – arcsin (nw / 2πr) to make sure that the fiber is continuous, where $r = r_1 - (r_1 - r_2) z / h_2$，$r_1 = \Phi_1 / 2$，$r_2 = \Phi_2 / 2$. When placement angle is ±45, as shown in Fig.2(b), the geometric equation is

$$\tan 45° = \frac{rd\theta}{\sqrt{(dr)^2 + (dz)^2}} = \frac{(r_1 + kz)d\theta}{\sqrt{1+k^2}dz}, \tag{3}$$

where $k = -(r_1 - r_2) / h_2$，故有 $r = r_1 + kz$，$dr = kdz$. The fiber placement path is

$$\theta = \frac{\sqrt{1+k^2}}{k} ln\left(\frac{r_1}{k} + z\right) + \theta_0. \tag{4}$$



(a) **(b)**

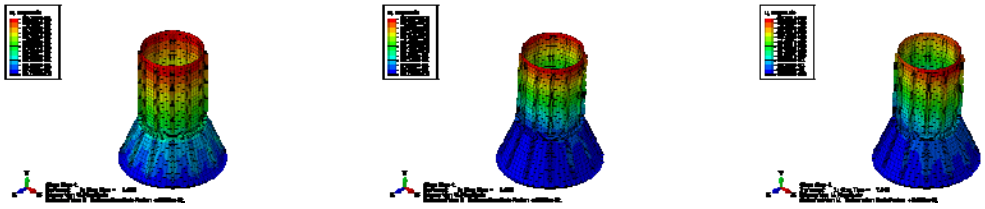
Fig.2 Fiber Placement Paths of Cylinder and Cone when Placement Angle is 45

3. FEA of central tube

The design capability of the satellite platform is 3 ton, and the axial and lateral load are 9 g and 2 g separately. The safety factor is 1.5. In order to evaluate the capability of the central tube, the finite element model is constructed and analyzed.

3.1. Statics analysis

Fig.3 shows the distortion of the central tube under several different static loads. Fig.3(a) is the condition with axial load only, and the maximum distortion is 0.915mm. Fig.3(b) is the condition with lateral load only, and the maximum distortion is 2.857mm. Fig.3(c) is the condition with axial and lateral load simultaneously, and the maximum distortion is 3.201mm.



(a) (b) (c)

Fig.3 Distortion of Central Tube

Fig.4 shows the Maximum Failure Factor (MFF) based on Tsai-Wu criteria of the central tube under several different static loads. The strength data of M55J/AG80 CFRP material are listed in Table 2. Fig.4(a-b) is the condition with axial load only, and the MFF of cylinder part is 0.274, in 3rd lay-up (0 layer), while the MFF of cone part is 0.225, in 3rd lay-up (0 layer). Fig.4(c-d) is the condition with lateral load only, and the MFF of cylinder part is 0.377, in 9th lay-up (+45 layer), while the MFF of cone part is 0.292, in 4th lay-up (90 layer). Fig.4(e-f) is the condition with axial and lateral load simultaneously, and the MFF of cylinder part is 0.533, in 7th lay-up (0 layer), while the MFF of cone part is 0.447, in 3rd lay-up (0 layer).

Table 2 Strength data of M55J/AG80 CFRP material

	Tensile strength 1 σ_{1T}/MPa	Tensile strength 2 σ_{2T}/MPa	Compressive strength 1 σ_{1C}/MPa	Compressive strength 2 σ_{2C}/MPa	In-Plane shear strength τ_B/MPa
M55J/AG80	1309	29.8	800	150	48

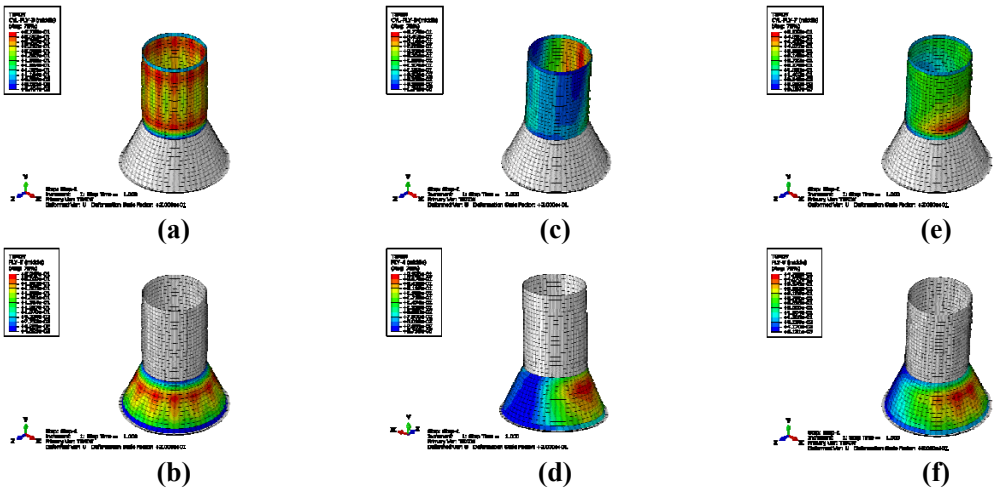


Fig.4 Maximum Failure Factor of Central Tube

3.2. Mode analysis

Fig.5 shows the main modes of the central tube. Fig.5(a) is the first order mode, and the fundamental frequency is 355.73Hz. Fig.5(b) is the first order bending mode, and the frequency is 400.53Hz.

Fig.5 Mode of Central Tube

3.3. Buckling analysis

Fig.6 shows the Stability Factor (SF) of the central tube under several different static loads. Fig.6(a) is the condition with axial load only, and the SF is 5.418mm with buckling in the cylinder part. Fig.6(b) is the condition with lateral load only, and the SF is 6.064mm with buckling in the cylinder part. Fig.6(c) is the condition with axial and lateral load simultaneously, and the SF is 3.023mm with buckling in the cylinder part.

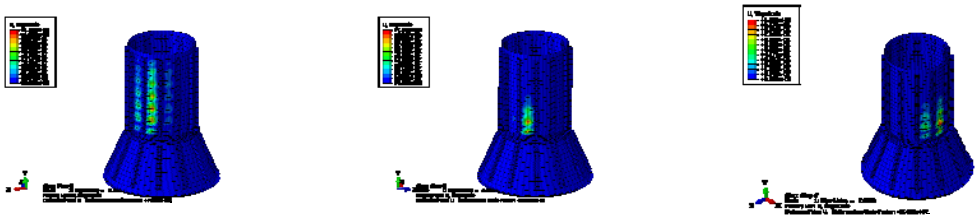


Fig.6 Stability Factor of Central Tube

3.4. Results

The results of FEA demonstrate that the designed central tube meets all the structural requirements, such as stiffness, strength, dynamic modes, and buckling, as shown in Table 3.

Table 3 FEA Results and Structural Requirements of Central Tube

	Stiffness[1]/mm	Strength[2]	Dynamic Mode[3]/Hz	Buckling[4]
FEA Results	3.201	0.533	355.73	3.023
Requirements	≤3.5	≤1	≥300	≥2

1. The maximum distortion under axial and lateral load simultaneously
2. The Maximum Failure Factor based on Tsai-Wu criteria under axial and lateral load simultaneously
3. The basic frequency
4. The stability factor under axial and lateral load simultaneously

4. Manufacturing based on AFP

To shape the cylinder-cone shell of the central tube, the mandrel is important and high-class. The surface roughness Ra should be less than 0.8μm for high quality of shell and easy demolding. The taper of the cylinder part of mandrel is 1:1000. One should notice that the size of mandrel is slightly smaller than the cylinder-cone shell because of the expansion during curing.

The manufacturing process of the cylinder-cone shell consists of three main stages. First, the fiber placement path programming and simulation are conducted and improved iteratively to optimize the placement quality and efficiency. Second, the fiber tows are placed on the clean mandrel surface or the previous layer by using the AFP machine in accordance with the NC programs to form the laminate shell, as shown in Fig.7. Third, the formed carbon fiber reinforced plastics shell is cured in the autoclave and machined to the final dimensions. The former two stages are very different from the traditional hand lay-up. The automation and computer aided technologies make the manufacturing process more precise, efficient and economical.

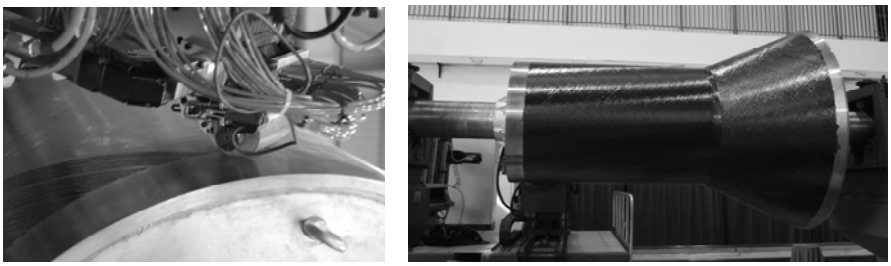


Fig.7 Manufacturing of Central Tube Based on AFP

5. Conclusion

The central tube of monolithic cylinder-cone shell has been designed based on the AFP technique. The fiber placement paths have been designed and optimized for cylinder part and cone part. The FEA results prove the central tube meets the structural requirements. By using the AFP system, the monolithic cylinder-cone shell has been manufactured successfully. The physical tests are going to be carried out in the near future.

References

[1]. http://www.esa.int/Our_Activities/Telecommunications_Integrated_Applications/First_journey_for_Alphabus.

[2]. http://www.esa.int/TEC/Structures/SEMREFWUP4F_0.html

[3]. http://www.mtorres.es/en/aeronautics/products/carbon-fiber/torresfiberlayup

[4]. Rob F, Justin N, Todd R. Production Implementation of Multiple Machine. High Speed Fiber Placement for Large Structures. 2010 SAE International, Paper No. 2010-01-1877.

[5]. Shao Z. Research on Key Technology of Fiber Placement Machine. Ph.D. dissertation, Harbin Institute of Technology, 2010.

[6]. Huan D. Fundamental Research on CAD/CAM Technology of Composite Material Automated Placement. Ph.D. dissertation, Nanjing University of Aeronautics and Astronautics, 2010.
[7]. Pascual J, Trigo J. Application of new technologies for Eurostar central tube. in European Conference on Spacecraft Structures, Materials and Mechanical Testing 2005, Noordwijk, The Netherlands, 10-12 May 2005.
[8]. Xiao J, Li Y, Li J. Application of Automated Placement Technique in Manufacturing of the Composites Structures of Large Aircraft. Aeronautical Manufacturing Technology. 2008(1): 50-53.
[9]. Shirinzadeh B, Foong C W, Tan B H. Robotic fiber placement process planning and control. Journal of Assembly Automation, 2000, 20(4): 313-320.
[10]. Schueler K, Miller J, Hale R. Approximate geometric methods in application to the modeling of fiber placed composite structures. Journal of Computing and Information Science in Engineering, 2004, 4(3): 251-256.

Author: Ke Wang, Shanghai Institute of Spacecraft Equipment, No.251, Huaning Road, Shanghai 200240, Email: chqwxylwk@163.com

Design of Unification Dataflow on Spacecraft Based on CCSDS AOS

WANG Li-sheng, GAI Jian-ning, Chen Jian-yue，Tu Zhi-jun，HAN Xiao, Lai An-xue

(Shanghai Key Laboratory of Spacecraft Mechanism, Shanghai Institute of Aerospace System Engineering, Shanghai,201109)

Abstract: With the development of spacecraft, the idea of unification dataflow based on CCSDS AOS is put forward. The framework and characteristic of the information system is introduced in this article. And the original data and virtual channel is sorted. The design and analyze of the dataflow is developed and the project can be applied to the giant spacecraft with a lot of payloads.
Keywords: CCSDS AOS; Spacecraft; Information System; Unification Dataflow

1. Introduction

With the continuous development of space technology, the task of space data processing is more and more difficult, the load type and quantity is various, and the data rate is different from platform. In addition to the conventional telemetry, trace and control data, there are along with a variety of scientific experiments data, voice communication, motion graphics and still images. Due to the higher and higher data volume and rate, a unified standard is urgently needed for dynamic management, and CCSDS AOS (advanced orbiting systems) just meets this demand.

CCSDS(Consultative Committee for Space Data System) is a international cooperation organization that was established in 1982 by NASA, ESA and other multiple national Space agency. It aims to establish the data system standard for spatial information exchange and to create a global open virtual space data system, promoting the international support, cooperation and information exchange. Its standard includes both space data link protocol, such as telemetry, control, AOS, and SCPS(space communication protocol specification) including four parts:SCPS-FP(file protocol), SCPS-TP(transmission), SCPS-SP(security protocol) and SCPS-NP(network protocol). The SCPS is equivalent to the ground network protocol FTP and TCP/IP on the function.

CCSDS is the authority of space data systems technology in the world, and its standard applicable to the whole space data link transmission types and different rate. As the development of space technology and international cooperation, the CCSDS standard is the trend of The Times in the field of space data system. Therefore, the integration and unified data flow design based on CCSDS AOS of variety of high or low speed test data has very important practical significance.

2. The AOS system

2.1. Characteristics of AOS

The main technical features of AOS is oriented of the data communication demand from small capacity and low speed to the large capacity, high rate, multiple user and payload. By using reasonable mechanism of road, AOS dynamically manages and schedules the virtual channel, which improves the utilization rate of channel. AOS standard supports multiple business, such as comprehensive information transport of images, voice, high and low speed data, transmission mode

of asynchronous, synchronous and isochronous, services of bit stream and subcontract, as well as different levels of mistake proofing, corrective measures. All of this ensure that the data, different type and rate, can share space transmission link with high efficiency and low mistake rate. Main technical characteristics of the AOS is:

a. By the idea of layered structure, the whole communication network is divided into several layers, and different layers have different functions and different data structure. Headers are set in the data unit of each layer, which is used to record some useful information, and to receive and extract user data correctly.

b. By the multiplexing mechanism of package channel and virtual channel, and channel dynamically sharing mechanism, the efficiency of spatial data channel is increased.

c. Setting up eight different business to deal with different data.

d. According to the different requirements of users for the business quality, three kinds of different business level is used, which both simplifies the system configuration, and meets the needs of different quality requirements of the user.

e. Virtual channel is set in CCSDS space link subnet(CSLS), each time-sharing taking up the physical channel for a CSLS application, so the long data source cannot monopolize channel and the channel can be orderly managed.

f. The data of different format is used by different types of users, and finally all data is put together into a data flow, which is transmitted through physical channel.

g. Identifier is set and distributed to the spacecraft, so the data from different spacecraft can be obtained and distinguished at the target end.

2.2. AOS protocol

AOS is more flexible than general system, not only compatible with the conventional system, but also providing a variety of data processing, especially meeting with the high speed data transmission such as dynamic video. For example, a AOS space system can be compatible with a free spacecraft which using conventional CCSDS standard, or in the same system, upward data without sound and image using conventional subcontract remote control standard, while down link to use AOS standard.

CCSDS Principal Network (CPN) is the most important part in space data system, and it acts as space planning and data management, providing end-to-end data stream processing for the Space mission, including satellite carrier network and space link subnet (SLS) in space orbit. CPN is a integrated network of a space network and a ground network, and it is composed of a satellite carrier network on an orbital section connecting with a ground network or a satellite carrier network on another track section by CCSDS space link subnet. The constitution of the CPN is shown in figure 1:

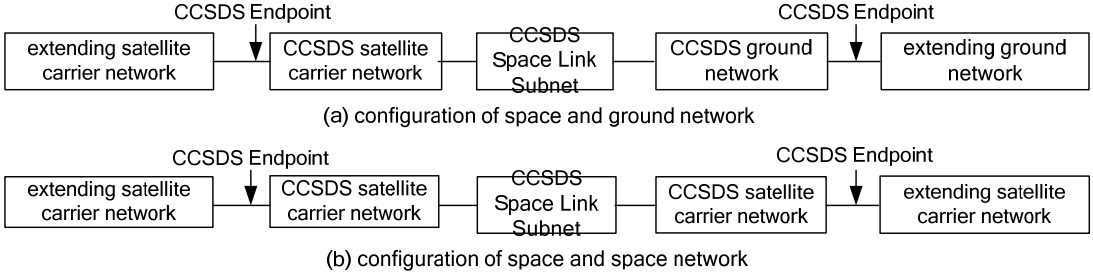


Fig.1 Framework of CCSDS Principal Network

Space link sub-layer(SL) corresponds to the data link layer of the OSI model, as shown in figure 2. CCSDS AOS provides space link sub-layer protocol and business, and does not limit the internal structure, the protocol and extending network business of the satellite carrier network and ground network, namely the AOS protocol specifies the standard behavior of the space link sub-layer.

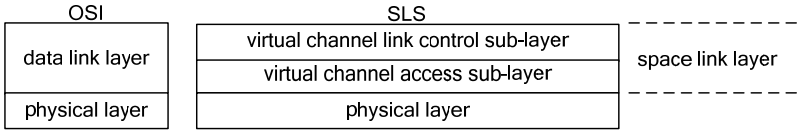


Fig.2 CCSDS Space Link Subnet

Spatial channel decline more serious due to the interference, so the signal transmission adopts the way of continuous wave, which is advantageous to the system synchronization and anti-interference, also facilitate to integrate the data transmission, speed measuring and ranging into a channel. By the basic way of time division mechanism the continuous wave data transmission realizes to share the channel among multiple users. In the traditional time division mechanism, continuous data flow is divided into fixed length of frame, starting with the frame synchronization code identification. Each frame arranges each user's data in fixed format, forming a fixed format of data blocks. The acquisition and transmission of user data is synchronous, and the advantage of this approach is simple and reliable. But when the space system exists complex source and is in a distributed environment, the way of fixed partition channel above not only reduce the operation efficiency of the system, but also is unable to support a high level of user business.

CCSDS kept fixed frame length, and at the same time the concept of virtual channel is established. Virtual channel is the way to share the channel among multiple users based on time division mechanism, and its main mechanism is to distribute channel resource dynamically. Before sent into the channel, the user data from different business types is formed to a fixed length VCDU first. VCDU is composed of leading head, user data and error correction coding data. Virtual channel data unit must also be through the store-away forward into the physical channel. Virtual channel data unit with fixed length is advantageous to synchronize and to adopt the unified error-correcting codes, and also is advantageous to prevent the large capacity data source to monopolize channel. On the other hand, due to increased user data processing, such as the long bag into data unit, short package merged into a data unit and so on, the dominant header information operation is increased, thereby increasing the workload of data processing and channel overhead.

3. The spacecraft information system

3.1. Information system architecture

The function of the spacecraft information system is to process the data of different types and different rate. It includes multiple high-speed source and source at low speed, and the rate of the high speed source vary from decades Mbps, hundreds Mbps, even to several G bps, while the rate of low-speed source vary from several kbps to hundreds of kbps. The high speed data is by high speed Ethernet, and the rate can be up to 10 G bps, meeting the needs of high-speed data transmission business. The low-speed data, such as measurement data, control data and management data, is transmitted by the mature and reliable 1553B bus. The whole information system architecture is shown in figure 3. The structure is divided into two parts, low-speed network and high speed network. Low-speed part adopts hierarchical topology structure, divided into the control layer and

user layer. Control layer is composed of management unit and multiple sets of 1553B bus, while user layer is composed of different application system of equipment and the corresponding bus. Based on Ethernet switching technology, high-speed network is divided into top level switches and access switches. As a gateway between Ethernet and bus, the control unit connect the low-speed network with the high speed network.

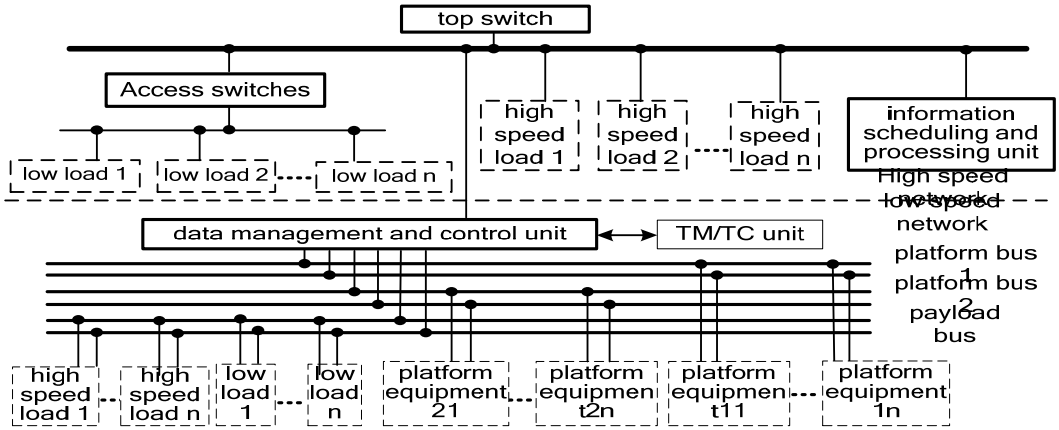


Fig.3 Framework of Information System of Spacecraft

3.2. The characteristics and requirements

Large spacecraft and load has the following characteristics and requirements:

a. platform independent management ability is enhanced, and the types and quantity of the measurement and control data increase.

b. increased load task variety, quantity, and type;

c. increased load test data , and the rate is up to several hundred Mbps or even a few G bps.

According to the different data types, it is needed to determine the reasonable space network architecture and data interface, so meeting the needs of the data transmission process and integrate design.

3.3. The data flow unified design ideas

Applying CCSDS protocol on the handling of information source is more flexible. Through AOS packing, multiplexing, and virtual channel scheduling, formatting, integration, finally the data became into virtual channel data unit. For high speed information source the corresponding proprietary virtual channel can be allocated to ensure the data of each source can be smoothly transfer. For low-speed source, to avoid reduce the usage of hardware facilities, the separate virtual channel generally don't be assigned, but holding a virtual channel after multiplexing.

The data management and control unit schedules and processes the information of multiple low-speed platform equipment through 1553B bus, then transmits it to the information scheduling and processing unit through Ethernet. The information scheduling and processing unit receives and processes the business data of the high-speed load through Ethernet. All information is scheduled and processed in accordance with the AOS protocol in the information scheduling and processing unit, then sent to the ground through the downward channel. While the platform information is transmitted to transponder after processing, then sent to the ground. The whole information system data flow is shown in figure 4.

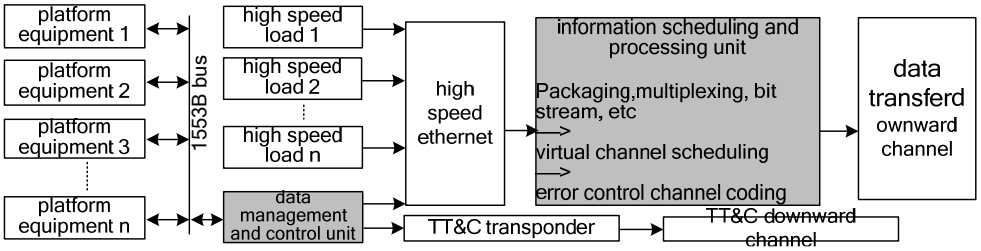


Fig.4 Unification Dataflows of Spacecraft

4. Data flow design

4.1. Data source packages and virtual channel

According to types, the data of spacecraft is divided into the following categories:

a. telemetry data: all of the real-time telemetry data collected on-orbit.

b. management data: all kinds of process data, software data and control parameters for platform autonomous management,.

c. payload data: the payload data is classified into several types of source data packets according to the type of business and data rate.

Accordingly, the virtual channel data unit is divided into telemetry data VCDU, managing data VCDU, payload data VCDU etc.. The original packages of different application process that belong to the same virtual channel are put together to form a protocol data unit, setting different priorities. And according to the actual demand, four kinds of standard AOS business, packaging, multiplexing, virtual channel access and bit stream are adopted.

4.2. The data flow design

In the whole process from generating of source data to transmission to ground, the source data packing, bus protocol, network protocol, and data transmission protocol between space and ground are used. Among them, 1553B bus interface and Ethernet interface support platform equipment adding or deleting, which can enhance the scalability and generality of the platform. The bus protocol and Ethernet protocol can be cut according to the practical application requirements. The data processing process of information system is shown in figure 5.

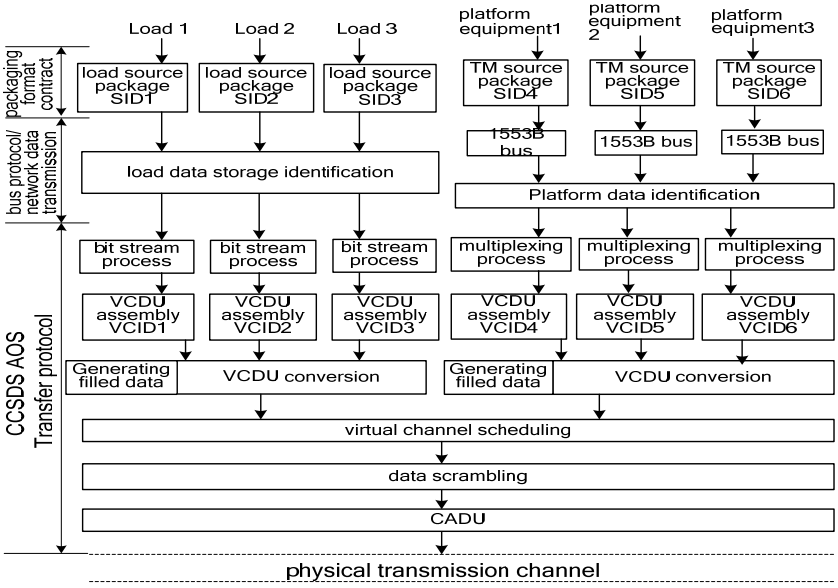


Fig.5 Data Process of Information System

5. Conclusion

According to the characteristics and requirements of the spacecraft information system, by combining with the characteristics of CCSDS AOS protocol, an design idea of unified data flow based on CCSDS AOS is put forward. And the concrete analysis of the data flow design has been carried on. It has been find that the design idea is suitable for large spacecraft with multiple payload and multiple tasks to processing and transmitting the data flow.

Reference

[1] Ma Yongkui, simulation analysis of on-board data system based on CCSDS AOS, telemetry and remote control, March 2002, 23 (2).

[2] Ji Peng, design of data flow oriented high-speed satellite payload, information technology, 2011, 11.

[3] Sun Baibo, strengthen the understanding and research of CCSDS standard, telemetry and remote control, July 1999, 20 (4).

[4]Li Shu-feng, CCSDS AOS throughput performance analysis ,avionics technology, December 2005, 36 (4).

[5] Li Bing, dacac, modeling and simulation of CCSDS AOS protocol based on opneT, Chinese high and new technology enterprise, May 2011.

[6] Song Wei-dong, research and design of data stream management system, master degree theses of Nanjing university of aeronautics and astronautics, March 2005.

Estimation of Star Sensor Alignments Using Landmarks

Yue Gu Qinghua Zhu Chensheng Cai WeiQing Wu

(Shanghai Institute of Spaceflight Control Technology, Shanghai, 201109, China)
(Shanghai Key Laboratory of Aerospace Intelligent Control Technology, 201109, China)

Abstract: Periodical changes of relative alignments between attitude sensors and payloads have adverse impacts on spacecraft directing determination accuracy. In this paper, by applying the least square batch method, star senor relative alignments in a time interval on orbit are estimated from star senors and landmark data. The time varying alignment in a whole orbital period is obtained by a nonlinear least square fitting method. Simulation results verify the effectiveness of the method. Furthermore, experiment procedures are conceived for alignment estimation from on-orbit data.
Keywords: Relative misalignment; Landmarks; Attitude estimation

1. Introduction

Alignments between star sensors and payloads may vary with launch shock and environment changes in orbit. Spacecraft attitude pointing determination accuracy and target positioning accuracy will be reduced if time varying alignments are identified inaccurately. In the past twenty years, the problem of alignment estimation has been concerned by many scholars and engineers [1-4].

The relative alignment observation equation is usually derived using sensor measurements and spacecraft attitude angles and statistical properties of installation error angles are obtained according to the maximum likelihood estimation equations. The redundancy problem in the measurement set was solved by a factorization method in [1], which makes use of sensor synchronous sampling data and does not need dynamic information and is insensitive to data loss. However this method proposed [1] does not account for the time varying alignment estimation problem. In [2], a Kalman filtering method was developed for estimating installation errors of star sensors and gyroscopes in spacecrafts. Some state variables were added to the conventional attitude filter, such as gyroscope installation errors, gyroscope ratio errors and star sensor absolute installation errors. The derived equations in [2] did not contain payload installation state variables, and the payload axis could be used as the reference axis so that the payload data were used to update the sensor alignment parameters. In order to deal with the instrument distortion error and reduce the system error associated with the star sensor data, a batch least square attitude determination method for gyros and star sensors was proposed in [3]. At the initial time, both star sensor data and gyroscope data were used for attitude quaternion and gyroscope drift estimation, while in the selected batch interval, only gyroscope data were used to correct attitude quaternion estimation. The limitation of this method is that algorithm accuracy depends on gyroscope performance and batch interval selection. In [4], according to the payload observation vector in several orbits, star sensor low frequency error parameters in the Fourier series form were obtained by the least square algorithm.

In this paper, based on observation equations of the installation error angles between the full attitude sensors and the single vector sensors [1], the installation error angles in orbit of star sensors in a period are estimated by the least square method. The basic equations of time varying installation

error angles of star sensors in one orbit period are obtained by a nonlinear fitting method. The validity of the method is verified by simulation results. In addition, applied steps will be provided for alignment estimation between star sensors and payloads.

2. Relative misalignment estimation

Owing to environment changes in orbit, relative misalignments between star sensors and payloads have long periodical time varying characteristic. Firstly, based on the relative misalignment observation equation, approximate constant misalignments of different time intervals in one orbit period are estimated by the least square method. Then the relative misalignment basic equation is obtained by the sine wave curve fitting method.

2.1 Observation equation

The observation equation of relative misalignments between star sensors and payloads[1] is shown in equation (1)

$$z_{ij,k} = \mathrm{P}(\hat{W}^o_{j,k})\hat{W}^{o*}_{ij,k} \times \hat{W}^o_{j,k} = \mathrm{P}(\hat{W}^o_{j,k})(\theta_j - \theta_i) + \Delta z_{ij,k}, \tag{12}$$

where

$$\Delta z_{ij,k} = \mathrm{P}(\hat{W}^o_{j,k})\left[\xi_{i,k} - \lfloor \hat{W}^o_{j,k} \rfloor \Delta\hat{W}^o_{j,k}\right], \tag{13}$$

$\hat{W}^o_{j,k}$ denotes the measured landmark direction in the spacecraft body frame, and $\hat{W}^{o*}_{ij,k}$ is the estimated value of $\hat{W}^o_{j,k}$. Observation matrix $\mathrm{P}(\hat{W}^o_{j,k})$ is a 2×3 matrix which projects three-dimensional vectors onto the two-dimensional space orthogonal to $\hat{W}^o_{j,k}$. $\theta_j - \theta_i$ is the relative misalignment angle vector between star sensor i and payload j. $\xi_{i,k}$ is the measurement noise of star sensor i at time instant t_k, and $\Delta\hat{W}^o_{j,k}$ is landmark vector measurement noise in the spacecraft body frame at time instant t_k. In addition, matrix $\lfloor \hat{W}^o_{j,k} \rfloor$ is defined as

$$\lfloor \hat{W}^o_{j,k} \rfloor = \begin{bmatrix} 0 & \hat{W}^o_{(j,k)3} & -\hat{W}^o_{(j,k)2} \\ -\hat{W}^o_{(j,k)3} & 0 & \hat{W}^o_{(j,k)1} \\ \hat{W}^o_{(j,k)2} & -\hat{W}^o_{(j,k)1} & 0 \end{bmatrix}. \tag{14}$$

The observation equations (1)~(2) are briefly derived as follows:

On one hand, the vector $\hat{W}^o_{j,k}$ can be directly obtained from landmark observation vector $\hat{U}_{j,k}$ in payload j at time t_k and payload initial installation matrix S^o_j. The relationship between the vector $\hat{W}^o_{j,k}$ and other variables is shown in equation (4).

$$\begin{aligned} \hat{W}^o_{j,k} &\equiv S^o_j \hat{U}_{j,k} \\ &= M_j^T \hat{W}_{j,k} = M_j^T A_k \hat{V}_{j,k} + \Delta\hat{W}^o_{j,k} = \left(I - \lfloor \theta_j \rfloor\right) A_k \hat{V}_{j,k} + \Delta\hat{W}^o_{j,k}, \\ &= A_k \hat{V}_{j,k} - \lfloor \theta_j \rfloor A_k \hat{V}_{j,k} + \Delta\hat{W}^o_{j,k} \end{aligned} \tag{15}$$

where A_k is real attitude matrix, M_j is the misalignment matrix of sensor j given by

$$M_j = e^{\lfloor \theta_j \rfloor} = I + \left(\frac{\sin|\theta_j|}{|\theta_j|}\right)\lfloor \theta_j \rfloor + \left(\frac{1-\cos|\theta_j|}{|\theta_j|^2}\right)\lfloor \theta_j \rfloor^2 = I + \lfloor \theta_j \rfloor + O(|\theta_j|^2). \tag{16}$$

On the other hand, the vector $\hat{W}_{ij,k}^{o*}$ can be obtained from the estimated matrix of the spacecraft attitude, $A_{i,k}^{o*}$, which is determined by star sensor i at time t_k and the landmark reference vector $\hat{V}_{j,k}$. The relationship between the vector $\hat{W}_{ij,k}^{o*}$ and other variables is shown in equation (6).

$$\begin{aligned}\hat{W}_{ij,k}^{o*} &\equiv A_{i,k}^{o*}\hat{V}_{j,k} \\ &= (\delta A_{i,k}^{o})A_{i,k}^{o}\hat{V}_{j,k} = (\delta A_{i,k}^{o})M_i^T A_k\hat{V}_{j,k} = \left(I + [\xi_{i,k}]\right)\left(I - [\theta_i]\right)A_k\hat{V}_{j,k} \\ &= A_k\hat{V}_{j,k} - [\theta_i]A_k\hat{V}_{j,k} + [\xi_{i,k}]A_k\hat{V}_{j,k} - [\xi_{i,k}][\theta_i]A_k\hat{V}_{j,k}\end{aligned} \quad , \quad (17)$$

where $\delta A_{i,k}^{o}$ and $\xi_{i,k}$ are the measurement noise matrix and the measurement noise of star sensor i, and equation (7) shows their relationship.

$$\delta A_{i,k}^{o} = I + [\xi_{i,k}] + O\left(\left|\xi_{i,k}\right|^2\right) \quad (18)$$

From (4) and (7), we further have

$$\begin{aligned}&\hat{W}_{ij,k}^{o*} \times \hat{W}_{j,k}^{o} \\ &= \left(A_k\hat{V}_{j,k} - [\theta_i]A_k\hat{V}_{j,k} + [\xi_{i,k}]A_k\hat{V}_{j,k} - [\xi_{i,k}][\theta_i]A_k\hat{V}_{j,k}\right) \times \left(A_k\hat{V}_{j,k} - [\theta_j]A_k\hat{V}_{j,k} + \Delta\hat{W}_{j,k}^{o}\right) \\ &= A_k\hat{V}_{j,k} \times A_k\hat{V}_{j,k} + A_k\hat{V}_{j,k} \times \left(-[\theta_j]A_k\hat{V}_{j,k}\right) + A_k\hat{V}_{j,k} \times \Delta\hat{W}_{j,k}^{o} \\ &\quad + \left(-[\theta_i]A_k\hat{V}_{j,k}\right) \times A_k\hat{V}_{j,k} + \left(-[\theta_i]A_k\hat{V}_{j,k}\right) \times \left(-[\theta_j]A_k\hat{V}_{j,k}\right) + \left(-[\theta_i]A_k\hat{V}_{j,k}\right) \times \Delta\hat{W}_{j,k}^{o} \\ &\quad + [\xi_{i,k}]A_k\hat{V}_{j,k} \times A_k\hat{V}_{j,k} + [\xi_{i,k}]A_k\hat{V}_{j,k} \times \left(-[\theta_j]A_k\hat{V}_{j,k}\right) + [\xi_{i,k}]A_k\hat{V}_{j,k} \times \Delta\hat{W}_{j,k}^{o} \\ &\quad + \left(-[\xi_{i,k}][\theta_i]A_k\hat{V}_{j,k}\right) \times A_k\hat{V}_{j,k} + \left(-[\xi_{i,k}][\theta_i]A_k\hat{V}_{j,k}\right) \times \left(-[\theta_j]A_k\hat{V}_{j,k}\right) \\ &\quad + \left(-[\xi_{i,k}][\theta_i]A_k\hat{V}_{j,k}\right) \times \Delta\hat{W}_{j,k}^{o}\end{aligned} \quad . \quad (19)$$

Ignore the infinitely small items of second order, and apply the equation $[\theta]A_kV_k = -\theta \times A_kV_k$. Then the product $\hat{W}_{ij,k}^{o*} \times \hat{W}_{j,k}^{o}$ can be simplified as

$$\begin{aligned}&\hat{W}_{ij,k}^{o*} \times \hat{W}_{j,k}^{o} \\ &\approx A_k\hat{V}_{j,k} \times \left(-[\theta_j]A_k\hat{V}_{j,k}\right) + \left(-[\theta_i]A_k\hat{V}_{j,k}\right) \times A_k\hat{V}_{j,k} + \Delta z_{ij,k}^{o} \\ &= A_k\hat{V}_{j,k} \times \left(\theta_j \times A_k\hat{V}_{j,k}\right) + \left(\theta_i \times A_k\hat{V}_{j,k}\right) \times A_k\hat{V}_{j,k} + \Delta z_{ij,k}^{o} \\ &= -\left(\theta_j \times A_k\hat{V}_{j,k}\right) \times A_k\hat{V}_{j,k} + \left(\theta_i \times A_k\hat{V}_{j,k}\right) \times A_k\hat{V}_{j,k} + \Delta z_{ij,k}^{o}\end{aligned} \quad (20)$$

where

$$\Delta z_{ij,k}^{o} \approx A_k\hat{V}_{j,k} \times \Delta\hat{W}_{j,k}^{o} + [\xi_{i,k}]A_k\hat{V}_{j,k} \times A_k\hat{V}_{j,k} \approx -[\hat{W}_{j,k}^{o}]\Delta\hat{W}_{j,k}^{o} - \left(\xi_{i,k} \times A_k\hat{V}_{j,k}\right) \times A_k\hat{V}_{j,k} \quad (21)$$

The item $\left(\theta_j \times A_k\hat{V}_{j,k}\right) \times A_k\hat{V}_{j,k}$ is the inverse projection of vector θ_j in the two-dimensional plane perpendicular to vector $A_k\hat{V}_{j,k}$. Then the equations (9) and (10) can be further transformed into (11) and (12), respectively,

$$
\begin{aligned}
&P(\hat{W}_{j,k}^{o})\hat{W}_{ij,k}^{o*}\times\hat{W}_{j,k}^{o}\\
&=P(\hat{W}_{j,k}^{o})\left[-\left(\theta_j\times A_k\hat{V}_{j,k}\right)\times A_k\hat{V}_{j,k}+\left(\theta_i\times A_k\hat{V}_{j,k}\right)\times A_k\hat{V}_{j,k}+\Delta z_{ij,k}^{o}\right] \qquad (22)\\
&=P(\hat{W}_{j,k}^{o})\left(\theta_j-\theta_i\right)+\Delta z_{ij,k}
\end{aligned}
$$

$$
\Delta z_{ij,k}=P(\hat{W}_{j,k}^{o})\left[\xi_{i,k}-\hat{W}_{j,k}^{o}\Delta\hat{W}_{j,k}^{o}\right] \qquad (23)
$$

Consequently, equations (1) and (2) are obtained.

2.2 Data Estimation

Now we apply a recursion least square method for estimating the relative misalignments in a certain period. The observation equation is given by

$$
z_k=h_kX+v_k\,. \qquad (24)
$$

Then the state estimation is obtained as follows:

$$
\hat{X}_k=\hat{X}_{k-1}+K_k\left(z_k-h_k^T\hat{X}_{k-1}\right) \qquad (25)
$$

where $K_k=P_{k-1}h_k\left(I+h_k^TP_{k-1}h_k\right)^{-1}$ and $P_k=P_{k-1}-P_{k-1}h_k(I+h_k^TP_{k-1}h_k)^{-1}h_k^TP_{k-1}$。

2.3 Data Fitting

In this paper, the Fourier series form in equation (15) is considered to represent the basic equation of misalignments between star sensors and payloads.

$$
F(t)=x_0+x_1\sin(wt)+x_2\cos(wt)+x_3\sin(2wt)+x_4\cos(2wt)+... \qquad (26)
$$

Then the next task is to search for the specific value of the coefficients x_0, x_1, x_2,… such that the obtained data can be appropriately approximated. Specifically, we use a nonlinear least square fitting method that minimizes the following objective function,

$$
J(x)=\left\|F(x,xdata)-ydata\right\|_2^2
$$

That is, the estimated value of the Fourier coefficients in equation (15) could be obtained from

$$
x^*=\arg\min_x J(x) \qquad (27)
$$

2.4 Simulation Results and Analysis

In this subsection, the effectiveness of the above nonlinear curve fitting algorithm is verified by some simulation results. Simulation data (such as attitude angles of the star sensors and the observation vector of landmark points) are randomly generated by using the following model parameters. The data output frequency is 1Hz. The measurement accuracy and the model of the misalignment angles of star sensors are set as

$$
q_{errx}=5'',\ q_{erry}=5'',\ q_{errz}=40'',
$$

$$
\theta_1=\begin{bmatrix}60''\sin(wt)+15''\sin(2wt)\\40''\sin(wt)+13.33''\sin(2wt)\\100''\sin(wt)+50''\sin(2wt)\end{bmatrix}.
$$

Orbital angular velocity is $w=\dfrac{2\pi}{T}$ with the orbital period T being 64895 seconds. The field angle of the payload is 17 degree, and the payload measurement error is represented by QUEST model:

$$E\left\{\Delta\hat{U}_k\Delta\hat{U}_k^T\right\}=\sigma_k^2\left(I-\hat{U}_k^{true}\hat{U}_k^{tureT}\right),$$

where $\sigma=1''$ and $\hat{U}_k^{true}$ is the real observation vector. The payload misalignment model is set to the following constants:

$$\theta_2=\left[20''\quad 12''\quad 30''\right]^T.$$

Then the relative misalignment model between the star sensor and the payload is given by equation (17).

$$\theta_{21}=\begin{bmatrix}\theta_{21x}\\ \theta_{21y}\\ \theta_{21z}\end{bmatrix}=\begin{bmatrix}\theta_{2x}-\theta_{1x}\\ \theta_{2y}-\theta_{1y}\\ \theta_{2z}-\theta_{1z}\end{bmatrix}=\begin{bmatrix}20''\text{-}60''\sin(wt)-15''\sin(2wt)\\ 12''\text{-}40''\sin(wt)-13.33''\sin(2wt)\\ 30''\text{-}100''\sin(wt)-50''\sin(2wt)\end{bmatrix}. \tag{28}$$

Fourteen time intervals of data are selected from the entire orbital period, and each time interval is 30 sec. Different constant misalignment angles corresponding to the time intervals are obtained by applying equations (12), (13), (24) and (25). The model relative misalignment angles and the corresponding estimation are show in Table 1.

Table 1 Model and estimated results of different time intervals

Number	Time intervals(s)	Model Relative Misalignments(″)			Estimated Relative Misalignment(″)		
		X axis	Y axis	Z axis	X axis	Y axis	Z axis
1	1~30	18.69	11.03	27.09	18.97	11.58	39.44
2	500~530	-21.29	-18.33	-59.82	-17.41	-18.65	-38.24
3	1000~1030	-43.77	-33.59	-99.36	-45.69	-33.59	-76.59
4	1500~1530	-42.77	-30.54	-79.78	-41.43	-30.59	-80.91
5	2000~2030	-25.38	-15.96	-28.41	-28.44	-15.86	-36.28
6	2500~2530	-4.14	-0.8	14.46	-2.72	-0.39	9.8
7	3000~3030	13.22	8.91	29.46	16.03	8.64	33.77
8	3500~3530	28.024	15.68	30.88	26.54	15.83	46.26
9	4000~4030	45.76	25.85	48.01	46	25.69	50.36
10	4500~4530	67.09	41.29	92.73	63.66	41.32	80.93
11	5000~5030	83.62	55.36	143.10	82.75	55.3	153.88
12	5500~5530	82.83	57.06	158.48	88.21	57.4	164.03
13	6000~6030	58.53	40.34	114.09	56.78	40.66	120.11
14	6500~6530	17.78	10.36	25.07	17.96	10.73	1.88

Based on the data in Table 1, the Fourier series coefficients in equation (15) could be obtained by the nonlinear least square fitting method in Subsection 2.3，which are shown in Table 2. Consequently, the basic equation of the relative misalignment angles can be expressed in (18).

Table 2 Estimated Fourier series coefficients of relative misalignment angles

Direction	$x0$ (″)	$x1$ (″)	$x2$ (″)	$x3$ (″)	$x4$ (″)
X axis	20.13	-59.84	0.42	-15.75	0.79
Y axis	12.08	-40.03	0.11	-13.52	0.08
Z axis	33.48	-99.09	2.03	-44.53	1.4

$$\tilde{\theta}_{21}=\begin{bmatrix}\tilde{\theta}_{21x}\\ \tilde{\theta}_{21y}\\ \tilde{\theta}_{21z}\end{bmatrix}=\begin{bmatrix}20.13''-59.84''\sin(wt)+0.42''\cos(wt)-15.75''\sin(2wt)+0.79''\cos(2wt)\\ 12.08''-40.03''\sin(wt)+0.11''\cos(wt)-13.52''\sin(2wt)+0.08''\cos(2wt)\\ 33.48''-99.09''\sin(wt)+2.03''\cos(wt)-44.53''\sin(2wt)+1.4''\cos(2wt)\end{bmatrix} \tag{29}$$

Fig. 1 depicts the misalignment angles, their estimation and the estimation error in one orbit period. The black solid lines and the red dashed lines in the three sub-graphs on the left are the model curves and the estimated curves of relative misalignment angles, respectively. In each axis, these two kinds of curves nearly overlap with each other, which indicate that the time varying misalignment angles are well estimated. The blue solid lines on the right side of Fig.1 are the estimation error curves in one orbit, which shows that the estimation errors of X axis and Y axis are smaller than 1.5" and 0.5", respectively. The estimation errors of Z axis are relatively larger (the largest value is about 12"). In our opinion, why larger misalignments estimation errors of Z axis are obtained can be explained as follows.

According to equation (1), the observation matrix $\mathrm{P}(\hat{W}_{j,k}^{o})$ denotes a two-dimensional space perpendicular to the landmark observation vector $\hat{W}_{j,k}^{o}$. When the angle of view of the TDICCD camera is 17 degree, the vector $\hat{W}_{j,k}^{o}$ only varies in a narrow range $[0,0,-1]^{T}\sim[0,\ \pm 0.149,-0.989]^{T}$. Suppose an arbitrary vector $\hat{W}_{j,k}^{o}=[0,0.011,-0.9999]^{T}$ is chosen from the above range, then the corresponding matrix $\mathrm{P}(\hat{W}_{j,k}^{o})$ is given by

$$\mathrm{P}(\hat{W}_{j,k}^{o})=\begin{bmatrix}0 & \boxed{-0.9999} & -1.098\times10^{-2}\\ \boxed{-1} & -1.37\times10^{-6} & 1.251\times10^{-4}\end{bmatrix}.$$

It is found that the proportions of X axis and Y axis are much larger than Z axis. Intuitively speaking, due to the geometric constraint of the observation space, the actual observation data provide more information for X and Y axes, but much less for Z axis. Therefore, the estimating accuracy of the misalignments of Z axis is more susceptible to observation noise or other factors.

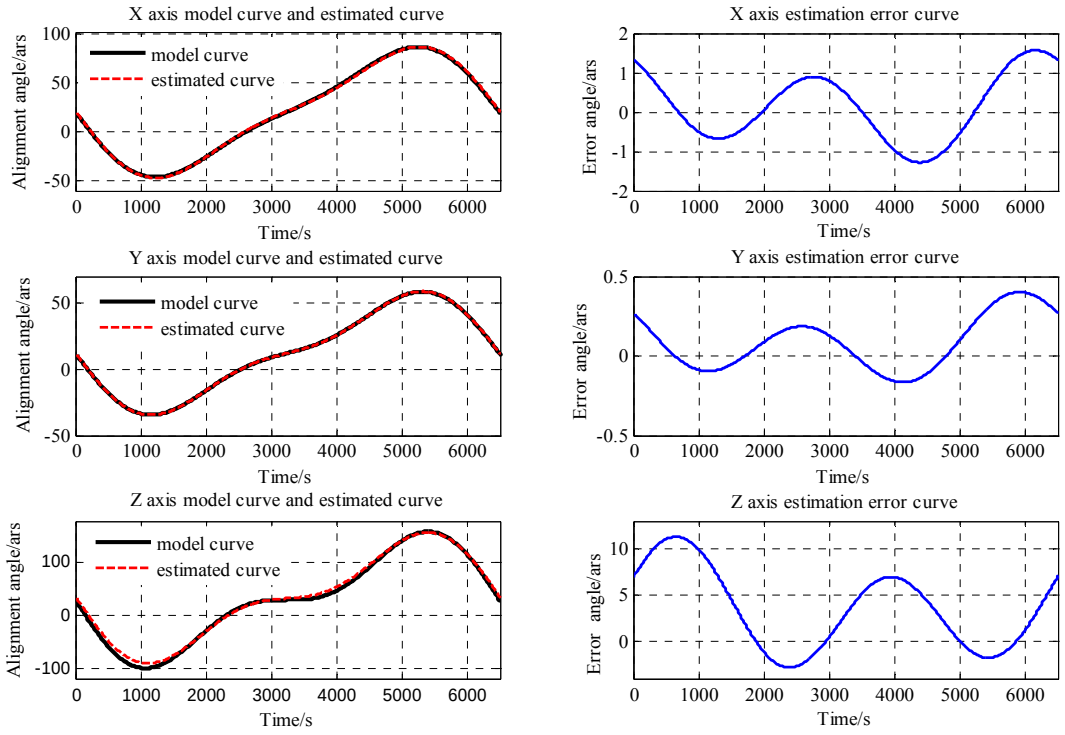


Fig. 1 Fitting curves and estimation error curves of misalignment angles in one orbit period

3. Test Verification Method

Combined with remote sensing data and image, we briefly discuss how to design an experimental verification method for the relative misalignments between star sensors and payloads in this section.

3.1 Data Processing

First we introduce the needed experimental data and the acquisition methods. The necessary data for estimating misalignment angles between star sensors and payloads are spacecraft body attitudes $A_{i,k}^{o*}$ and landmark reference vectors $\hat{V}_{j,k}$ in J2000.0 geocentric inertial coordinate system, landmark observation vectors $\hat{U}_{j,k}$ in payload body coordinate system and payload alignment matrix S_j^o determined by the prelaunch alignment calibration.

(1) Platform Attitude Data

The attitude matrix A_{bo} of spacecraft body relative to orbital coordinate and the transformation matrix A_{oi} of orbital coordinate relative to inertial coordinate system can be obtained from image assistant data. From (19), the platform attitude matrix $A_{i,k}^{o*}$ can be calculated by

$$A_{i,k}^{o*} = A_{bo}^{*} A_{oi}^{*} . \tag{30}$$

(2) Landmark Reference Vector

The landmark reference vector $\hat{V}_{j,k}$ is a unit vector $\mathbf{R}_{ts}$ from the landmark point to the spacecraft body center O_s. Denote by $\mathbf{R}_{et}$ the direction vector from the geocentric inertial coordinate system origin O_e to the landmark point. Denote by $\mathbf{R}_{es}$ the direction vector from the geocentric inertial coordinate system origin O_e to the spacecraft body center O_s. The reference vector $\hat{V}_{j,k}$ and the vectors $\mathbf{R}_{et}$, $\mathbf{R}_{es}$ are related with each other by the equations in (20) and (21).

$$\mathbf{R}_{ts} = \mathbf{R}_{es} - \mathbf{R}_{et}, \tag{31}$$

$$\hat{V}_{j,k} = \frac{\mathbf{R}_{ts}}{\|\mathbf{R}_{ts}\|}. \tag{32}$$

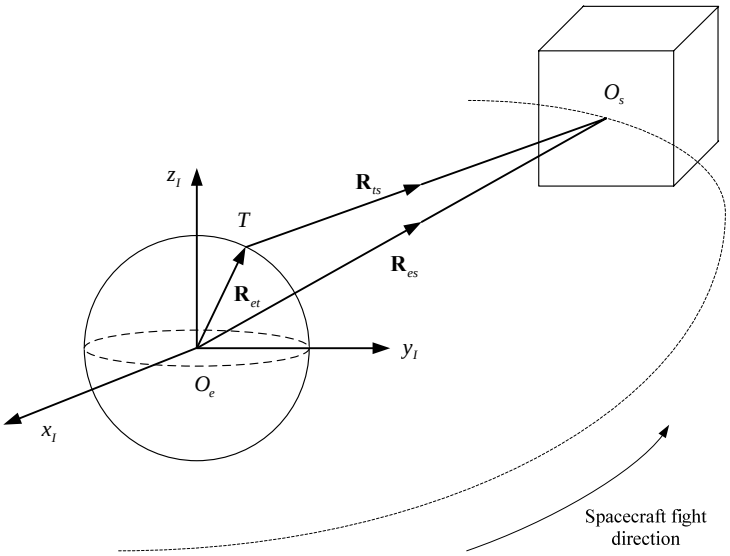


Fig. 2 Landmark reference vector relationship

①Firstly, use Google Earth to get the longitude α_e and latitude δ_e of the landmark point in WGS84. Then the direction vector $\mathbf{R}'_{et}$ in WGS84 could be obtained through equation (22). Based on time information in image assistant data, the transformation matrix $A_{WGS84\text{To}J2000}$ from WGS84 coordinate system to J2000 inertial coordinate system can be attained. Then vector $\mathbf{R}_{et}$ is calculated as follows,

$$\mathbf{R}'_{et} = \begin{bmatrix} \cos\delta_e \cos\alpha_e \\ \cos\delta_e \sin\alpha_e \\ \sin\delta_e \end{bmatrix}, \tag{33}$$

$$\mathbf{R}_{et} = r_e A_{WGS84\text{To}J2000} \mathbf{R}'_{et}, \tag{34}$$

where r_e is the radius of the earth.

② Secondly, vector $\mathbf{R}_{es}$ can be obtained according to the spacecraft position in the inertial coordinate system, which is denoted by

$$\mathbf{R}_{es} = \begin{bmatrix} x & y & z \end{bmatrix}^T. \tag{35}$$

③ Thirdly, substituting vectors $\mathbf{R}_{et}$ and $\mathbf{R}_{es}$ into equations (31)~(32), we can get the landmark reference vector $\hat{V}_{j,k}$.

(3) Landmark Observation Vector $\hat{U}_{j,k}$

The observation vector $\hat{U}_{j,k}$ could be obtained from the landmark point coordinate (x_C, y_C) in payload camera coordinate system and the focal length parameter f according to equation (25).

$$\hat{U}_{j,k} = \frac{1}{\sqrt{x_C^2 + y_C^2 + f^2}} [x_C, y_C, -f]^T . \quad (36)$$

3.2 Estimation Process

Now we need to convert original payload image and its auxiliary data into the needed form. First, one can obtain the constant misalignment angles corresponding to different time intervals using equations (12), (13), (24) and (25). Then the Fourier series coefficients $(x_0, x_1, x_2, x_3, x_4 ...)$ of the basic equation (5) of time varying misalignment angles can be computed according to (26) and (27). Now time varying misalignments fitting curves can be easily obtained.

The spacecraft may enter the region without sunlight or landmark points. Therefore it is not always possible to acquire spacecraft data through the entire orbit. However, as long as the data in some time intervals could reflect the variation of the misalignments, one can also integrate the intermittent data to obtain the basic equation of the time varying misalignments.

4 Conclusion

In this paper, the problem of estimating relative misalignments between star sensors and payloads has been studied. The proof of the misalignment angle observation equation has been presented. Based on the observation equation, a least square method and a nonlinear least square fitting method are applied to get the basic alignment equation from measurement data. The effectiveness of the proposed approach has been illustrated by a numerical example. For practical application, an experimental verification method of misalignments estimation using landmarks also has been introduced.

References

[1] Malcolm D.Shuster, Daniel S.Pitone and Gerald J.Bierman. Batch Estimation of Spacecraft Sensor Alignments I. Relative Alignment Estimation[J]. The Journal of the Astronautical Sciences, 1991, 39(4):519-548.

[2] Mark E. Pittelkau. Kalman Filtering for Spacecraft System Alignment Calibration[J]. Journal of Guidance, Control, and Dynamics, 2001, 24(6):1187-1195.

[3] Sungkoo Bae and Bob Schutz. Precision Attitude Determination Using Gyro and Star Tracker Data With A Batch Least Squares Estimator[J]. American Astronautical Society, 2006, 123(5):260-268.

[4] Xiong Kai, Tang Liang and Liu Yiwu. Calibration of Star Sensor's Low Frequency Error Based on Landmark Information[J]. Aerospace Control and Application, 2012, 38(3):11-15.

[5] Tu Shangcheng. Satellite Dynamics and Control(2)[M]. Beijing：Chinese Aerospace Press, 2009, 94~95.

Research on Scheduling Method of Dynamic Resources with Satellites Parallel Test

LIU Yifan[1] LIN Shumin[2] WANG Huamao[1] YAN Jindong[1]

1. Beijing Institute of Spacecraft System Engineering, Beijing 100094
2. Infosky Technology Co. LTD, Beijing 100022
Email：ayifan2004@163.com

Abstract：In the research and batch production of satellites, it is an urgent issue to be solved to reasonably and efficiently schedule dynamic resources. It is provided with huge spacecraft, mass testing items, great indeterminacy, different optimizing targets and etc. It's much more complex than normal resource-constrained project scheduling problem. This paper present the genetic algorithm model in parallel test resources of satellites，proposed scheduling method, and study a scheduling system for parallel test resources. First of all use genetic algorithm and experience data to solve the model, in order to get a baseline plan. When the plan is upset by actual test, schedule the rest of the testing items immediately to get the new plan. So scheduling is continued to ensure feasibility of the plan. Finally it is proved that the resource scheduling method applied well in Beidou Navigation Satellite System.

Key words：Satellites parallel test；Dynamic resources；Scheduling method

1 Introduction

In the batch production of satellites, parallel test mode has become the main form of satellites production. The resources are no longer monopolized by a single satellite, such as equipments, personnel, time, data bandwidth and so on. Through the satellites testing process analysis and optimization, satellites testing process reveals the inherent laws and issues, and to find a means of parallel satellites optimal overall efficiency of the testing process methods and approaches. Furthermore, analyzing and restructuring the logical relationships of testing process, is very necessary to break the boundaries of satellites and test satellites simultaneously.

2 Resource Scheduling Model

Satellites testing is to obtain and evaluate the qualitative and quantitative parameters, as well as to confirm the function and performance of satellites to meet the design requirements. Satellites parallel testing is generally considered as completed a number of satellites testing tasks at the same time.

satellites parallel testing contains three levels of parallelism:

(1) Parallel testing for different satellite parameters;

(2) Executed in parallel with test items of one satellite;

(3) Parallel test for test items of different satellite.

Level (1) depends mainly on testing tools, level (2) and (3) are more dependent on the rational allocation of resources, which are the focus of this paper.

Test resources are the key of satellites resource scheduling, including test equipment, human resources, site resources and funding resources. The purpose of scheduling is to achieve optimized

allocation, thereby improving the efficiency of parallel testing.

The actual scheduling problem of satellites parallel test resources is described as follows: N satellites are tested simultaneously. There are i_J test items in No. i spacecraft. These test items to meet the following property:

(1) There are timing constraints among test items of one spacecraft. Each test item can not begin until the test item just before it is completed.

(2) There is no causal timing constraint among test items of different satellite, but different satellite have different priorities, namely high-priority spacecraft test item with first priority rights of test resources.

(3) The test items which have been started cannot be interrupted.

(4) There are K kinds of renewable resources in all test items require. The measure of each resource is named R_k. Once renewable resource is freed by one test item, it can be used immediately in other items, and its usage is unchanged throughout the parallel testing process.

(5) No. j test item of No. i spacecraft is named A_{ij}. the test time of A_{ij} is a_{ij}, and the measure of No. k resources is demand as r_{ijk}.

(6) Under the premise of meeting the resource constraints, timing constraints, precedence relationships and optimize indexes, determine the start time and end time of each test item.

According to the above requirements, the resource scheduling problem can be described by the following mathematical model:

$$\min\{\max_{1\le i\le N} f_{ii_J}\} \tag{37}$$

s.t

$$S_{ij} \ge f_{uv}, i = 1,\cdots,N, j = 0,\cdots,i_J - 1, \forall A_{uv} \in P_{ij} \tag{38}$$

$$\sum_{A_{ij}\in I_t} r_{ijk} \le R_k, i = 1,\cdots,N, j = 0,\cdots,i_J - 1, k = 1,\cdots,K \tag{39}$$

$$f_{ij} = S_{ij} + d_{ij}, i = 1,\cdots,N, j = 0,\cdots,i_J - 1 \tag{40}$$

$$S_{ij} \ge 0, i = 1,\cdots,N, j = 0,\cdots,i_J - 1, k = 1,\cdots,K \tag{41}$$

In this model, S_{ij} and J_{ij} are stand for the start time and end time of test items A_{ij}. P_{ij} represents the immediately preceding set of test item A_{ij}. i_t indicates the set of test items being performed in time t.

The objective function (1) requires a minimum total duration.

Function (2) represents the timing constraints of test items, and the start time is no less than the end time of its immediately preceding item.

Function (3) indicates the resources, which are used less than the total amount of resources at any time.

Function (5) represents non-negative constraints, namely scheduled to start at time 0.

3 Scheduling Method Design

Traditional resource scheduling way for parallel test is a method for determined test environment. In other words, the testing is considered to be in the ideal case, all of the model parameters are determined and constant, then the model is solved to determine the best scheduling

plan. In order to overcome the limitations of traditional methods, based on genetic algorithm, it is designed that a scheduling method of satellites parallel test resources in uncertain environment.

The design idea is using genetic algorithms to solve the model estimates by experience or estimate data, to obtain a baseline plan. Once the baseline plan is broken by perturbations, reschedule the rest of test items immediately to obtain a new plan. So keep scheduling to ensure a feasibility plan and a shorter total duration, until the test is completed, shown in Fig. 1.

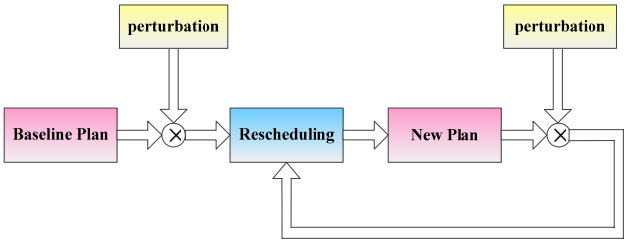


Fig.1 The flow chart of parallel resources scheduling method

Based on the ideas above, design the steps of the method as following:

(1) Genetic algorithms

The key to using genetic algorithm for solving scheduling problem of satellites is to design encoding, decoding, population initialization, genetic operators and so on.The flow chart of genetic algorithm in spacecraft as Fig. 2 shown, the computing step comprises:

a. Encoding

Choose priority-based encoding with a task list.

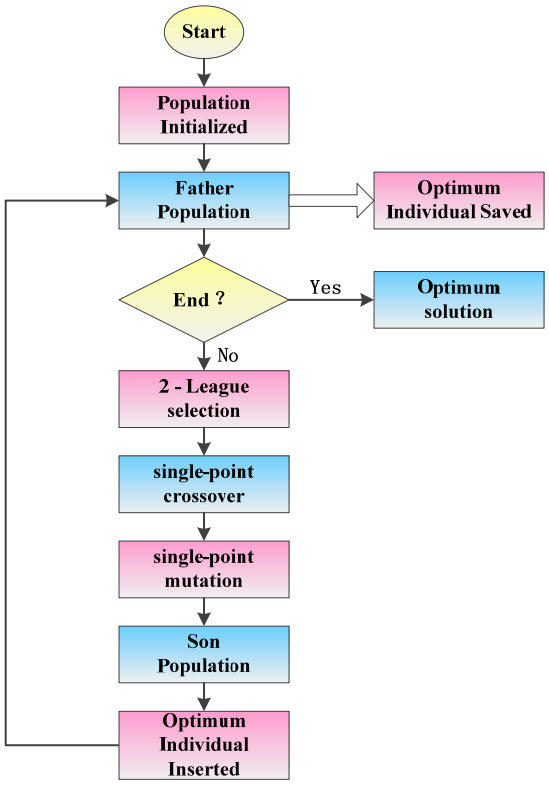


Fig.2 The flow chart of genetic algorithm in spacecraft

b. Decoding

Choose a serial decoding method. From left to right chromosomal sequence, determine the earliest start time of each test item in order, The earliest start time meets 3 conditions: (1) no less

than the end time of the item immediately before; (2) meeting the resource constrain throughout the whole process; (3) the minimum time meeting condition (1) and (2).

c. Population Initialized

The initial population can be produced by random manner, priority rules or a combination of both in order to get a more scattered individual.

d. Genetic operators

Operator Selection: 2-League selection mechanism;

Crossover operator: single-point crossover;

Mutation operator: single-point mutation.

(2) Baseline plan

Based on the Genetic algorithms above, a baseline scheduling plan is obtained before the beginning of spacecraft testing by experience or estimate data. It can be seen as a preliminary plan of the entire testing process. Obviously, the more accurate experience or estimate data, the better feasibility of baseline plan.

(3) Rescheduling strategy

There are a variety of confounding factors in the actual testing process, such as the amount of available resources change, test time change, test items addition or deletion, start time early or late. The baseline plan needs to be rescheduled, in order to remain viable and the total duration shorter. Rescheduling can be used in three ways:

a. Right shift strategy

When perturbation occurs(set the time for T1), the unfinished items are set in ascending order of their start time， and then arranged in descending order of priority. Right shift strategy is one of the simplest rescheduling strategies, but not optimized. The items closer to T1 are changed little, and so are the items execution orders of each spacecraft. It is recommended when the influence of perturbation factors are small.

b. Entirely rescheduling strategy

Once perturbation occurs (set the time for T2), the unfinished items are rescheduled by genetic algorithms. Entirely rescheduling strategy is optimized, and has the shortest total duration of three ways generally. However, it may cause some start time of the items changed greatly, especially the items near T2, and the changes of items execution order may also occur, which increase the costs of testing. When the requirement of total duration is strict, you can use this approach.

c. Partially rescheduling strategy

When perturbation occurs(set the time for T3， a period of time for ΔT), schedule items by right shift strategy if the start time is in $[T, T + \Delta T)$, and schedule items by genetic algorithms if the start time is more than $T + \Delta T$.

There is a time window ΔT plus in partially rescheduling strategy. Within the time window, use right shift strategy to change little about the start time and execution order of test items. When the time window ends, use genetic algorithms to ensure shorter total test duration.

Partially rescheduling strategy splits the difference between right shift strategy and entirely rescheduling strategy. It has the advantages of both, shorter duration and lower test costs.

(4) End Design

The first step and the second step are completed before starting test, the third step is actually a dynamic process and continued until the completion of the test.

4 Application in Beidou Navigation Satellites

Beidou Navigation Satellite System is an important spatial information infrastructure, which launched 14 satellites in three years. In the parallel testing practice of Beidou satellites, using the resource scheduling method proposed in this paper, the testing team has completed multi-satellite test items remarkably.

Take Parallel Test of three Compass Navigation Satellites for example. Each spacecraft has six testing items, and the time sequence and resources measure are shown in Fig. 3, Table 1 and Table 2. The descending order of Spacecraft importance is Spacecraft 1, Spacecraft 2 and Spacecraft 3.

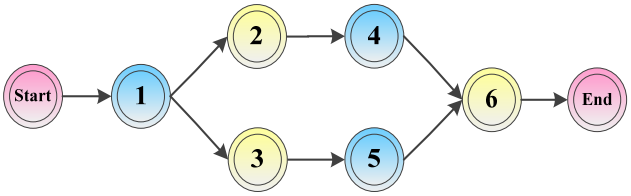


Fig.3 Time sequence of testing items

Table 1 Time and resource measure of each item

Item	Time	Resource 1	Resource 2	Resource 3	Resource 4
1	5	2	4	4	2
2	5	4	0	4	0
3	9	2	0	5	2
4	4	2	2	6	4
5	8	6	4	0	2
6	7	4	6	3	5

Table 2 Total measure of test resources

	Resource 1	Resource 2	Resource 3	Resource 4
Total	10	8	12	8

(1) Genetic algorithm calls

Based on the resource data, set parameters in the method of parallel test resources: 0.9 crossover rate, 0.1 mutation rate, the maximum generation (stop condition) is 500, then make genetic algorithm calls.

(2) A baseline plan obtained

Using the Scheduling Method on the basis of experience data, a baseline plan is obtained as Fig. 4. The total duration is 59, wherein the abscissa is the test time, and the vertical coordinate is the number of satellites, boxes represent the test task, the length of the box represents the testing time, the start time for the left margin of each block, the end time for the right margin, with Fig.5 to Fig. 7 reasonable.

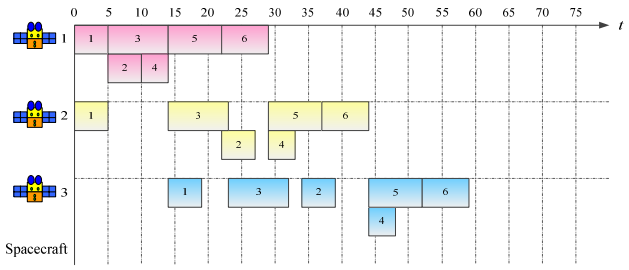


Fig.4 The baseline plan

(3) The implementation of rescheduling strategy

Assuming the baseline plan is perturbed, the 4th test item of satellite 1 is changes, the test time increased by 4, so the scheduling time is T=14. The new scheduling plans are as follows:

a. Right shift plan

As shown in Fig. 5, using the right shift strategy, the total duration is changed from 59 to 73. The increase of total duration is much more than 4. But there is little change right after the scheduling time, such as the 5th test item of satellite 1 and the 3rd test item of satellite 2. It is conducive that the original preparation is not possible to break.

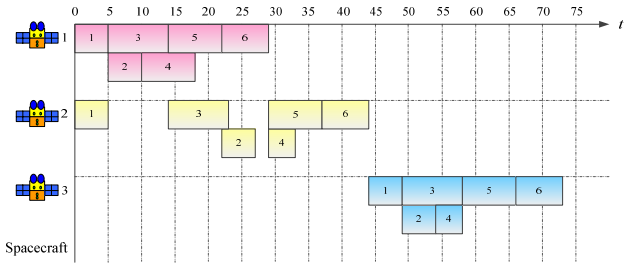


Fig.5 The right shift plan

b. Entirely rescheduling plan

The usage of entirely rescheduling strategy makes the total duration from 59 to 62, as shown in Fig. 5. The increase of total duration is less than 4, and the total duration was significantly less than that of the right shift strategies. However, there are great changes right after the scheduling time, such as the 2nd test item of satellite 2 and the 1st test item of satellite 3. The original preparation may fail, resulting in larger cost.

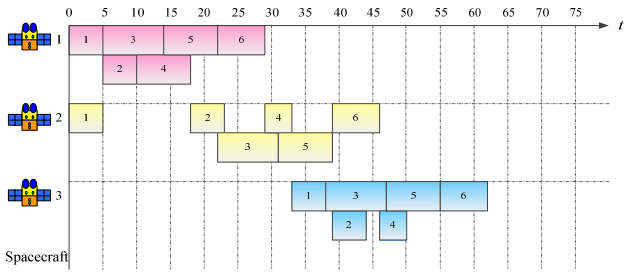


Fig.6 The entirely rescheduling plan

c. Partially rescheduling plan

Set ΔT for 10 and use the partially rescheduling strategy. As shown in Fig. 7, the total duration is changed from 59 to 63. There is little change right after the scheduling time, such as the 5th test item of satellite 1 and the 2rd test item of satellite 2. While the start time is more than T + ΔT = 24, the test items were scheduled to ensure that the total duration is not too long. Generally, the total

duration of partially rescheduling is between entirely rescheduling and the right shift strategy, so are the costs.

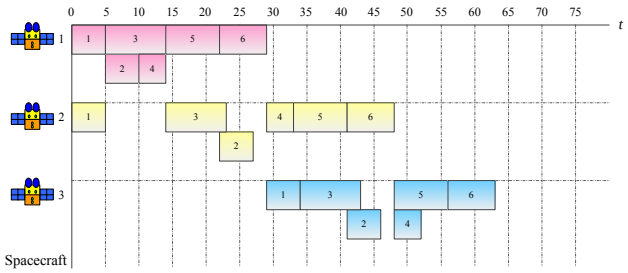


Fig.7 The partially rescheduling plan

5 Conclusions and Prospects

Parallel testing is an inevitable trend of satellites batch production. Reasonable and efficient resource scheduling is the core problem in satellites parallel testing. This paper abstracted actual scheduling problem of satellites resource as a mathematical model and present a scheduling method. It shown that the method can effectively solve the problems by the application example in Beidou navigation satellites.

The method has the following advantages: considering the complexity and uncertainty of resource constraints, the scheduling plan can always maintain viability; a shorter total duration of the test; and be able to effectively deal with the priority relationship problem in large scale. Therefore, the scheduling method of satellites parallel test resources is supposed to overcome the limitations of traditional resource scheduling methods, improve test efficiency and reduce test costs in more applications.

References

[1] Tyson R. B., Ali A. Y.. Resource-Constrained Multi-Project Scheduling: Priority Rule Performance Revisited [J]. International Journal of Production Economics, 2010, 126(2): 212-228

[2] Goncalves J. F., Mendes J. J. M., Resende M. G. C.. A Genetic Algorithm for the Resource Constrained Multi-Project Scheduling Problem[J]. European Journal of Operational Research, 2008, 189(3): 1171-1190

[3] Lin S. M., Zheng Z., Cai K. Y.. Robust Task Scheduling of Multi Satellite Parallel Test[A]. The 30th Chinese Control Conference[C]. Shanghai: Shanghai Systems Science Press, 2011: 2152-2157

[4] Stijn V. V., Ballestin F., Demeulemeester E., et al. Heuristic Procedures for Reactive Project Scheduling[J]. Computers & Industrial Engineering, 2007, 52(1): 11-28

[5] Eugene D. H.. Mixture Densities for Project Management Activity Times: A Robust Approach to PERT[J]. European Journal of Operational Research, 2008, 188(2): 450-459

[6] Liao L. M., Huang C. J.. Tabu Search for Non-Permutation Flowshop Scheduling Problem with Minimizing Total Tardiness[J]. Applied Mathematics and Computation, 2010, 217(2): 557-567

[7] Denise S. Y., Vinicius A. A., Manuel L.. Robust Optimization Models for Project Scheduling with Resource Availability Cost[J]. Journal of Scheduling Archive, 2007, 10(1): 67-76

[8] Bohle C., Maturana S., Vera J.. A Robust Optimization Approach to Wine Grape Harvesting

Scheduling[J]. European Journal of Operational Research, 2010, 200(1): 245-252
[9] Scaparra M. P., Church R. L.. A Bilevel Mixed-Integer Program for Critical Infrastructure Protection Planning[J]. Computers & Operations Research, 2008, 35(6): 1905-1923
[10] Zhao J., Zhang Y., Qu H., Qi H.. Ant Colony Optimization for Resource-Constrained Multi-Project Scheduling[A]. Intelligent Systems and Applications[C], 2009: 1-5
[11] Lova A., Tormos P.. Combining Random Sampling and Backward-Forward Heuristics for Resource-Constrained Multi-Project Scheduling [A]. Proceedings of the Eight International Workshop on Project Management and Scheduling[C], 2002: 244-248
[12] McDonell R.. Benefits of Parallel Testing[OL]. http://www.ni.com/Developerzone

Biography:
LIU Yifan Born in 1984, he received M.Sc. from Harbin Institute of Technology in 2009. His main research interest is spacecraft system-level test, EGSE system design and integration.

An Algorithm for Detecting Projection Ellipse of Docking Ring Component Based on Arc Segment Combination

Liu Fucheng, TIAN Shaoxiong, LU Shan

(Shanghai Aerospace Control Technology Institute,Shanghai,200233,China;
Shanghai Key laboratory of Aerospace Intelligent Control Technology,Shanghai 200233,China)

Abstract: Ellipse detection has important applications in spacecraft feature recognition. In order to overcome the shortcomings of the dependence on ellipse integrity in ellipse detection, this paper presents an algorithm for detecting projection Ellipse of docking ring based on arc segments combination. We detect the edges of image first. Then boundary tracing is carried on edge image. And the arc segments is classified into groups. Finally, direct least squares fitting algorithm is used to complete ellipse fitting. In this paper, we use the proposed algorithm to test on real image. The algorithm can detect the ellipse accurately and has the advantages of fast speed and avoiding noise.
Key words: Ellipse fitting, Edge tracing, Arc segments combination, Direct least squares fitting

1. Introduction

The technology of non-cooperative target capturing,aiming at maintenance of on-orbit traditional spacecraft, the clearance of obsolet satellites,has been a research focus in the field of space robot technique. The attitude measurement of non cooperative target is the key technique to achieve the target capturing .

There are no pre-set feature points for the non-cooperative targets.Its typical components must be used to accomplish pose measurement. Considering spacecrafts generally have the docking ring,it is an effective way to measure attitude by detecting the projection ellipse of the docking ring. Hough transform[1] is a common method to detect the ellipse parameters.Hough transform has 5 free parameters and it has a large memory useage. Some scholars proposed the random Hough transform[2]. It reduces the requirement for memory. But it also has a large amount of computation when detecting complex images.

In this paper,we propose an algorithm for detecting projection Ellipse of docking ring based on arc segment combination. Because arc segments contain more accurate geometry information than the discrete boundary points ,this algorithm can detect the projection ellipse accurately and has advantage of avoiding noise and fast speed.

2. Ellipse model

In the two-dimensional plane coordinate system, An ellipse can be expressed as follows:

$$ax^2 + 2bxy + cy^2 + 2dx + dey + f = 0 \tag{1}$$

Another way is to use the geometric parameters of the plane coordinate system,that is, the ellipse center position (x_c, y_c) , the semi-major axis and the semi-minor axis (a,b) , the angle between the long axis and the X axis θ . Any ellipse in a two dimensional plane can be uniquely determined by these 5 parameters. The geometric meaning of the parameters is shown in Figure 1. The parameters of two representation forms can be transformed using Eq. (2) ~Eq. (5).

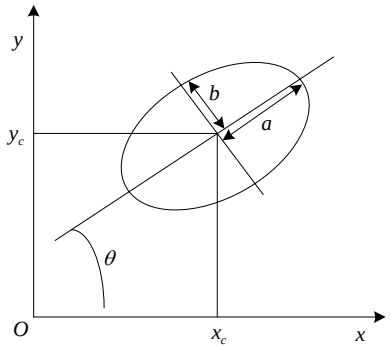


Figure 1.Ellipse models

$$X_c = \frac{be - 2cd}{4ac - b^2} \tag{1}$$

$$Y_c = \frac{bd - 2ae}{4ac - b^2} \tag{2}$$

$$a = 2\sqrt{\frac{-2f}{a + c - \sqrt{b^2 + (\frac{a-c}{f})^2}}} \tag{3}$$

$$\mathrm{b} = 2\sqrt{\frac{-2f}{a + c + \sqrt{b^2 + (\frac{a-c}{f})^2}}} \tag{4}$$

$$\theta = \frac{1}{2}\arctan\frac{b}{a-c} \tag{5}$$

3. Edge detection and tracking

Since the subsequent arc segment extraction is based on the edge, it is necessary to detect the edge of the original image.Edge detection is an important preprocessing work. LoG detector, DoG detector, Canny detector[3], etc, are commonly used in edge detection. In this paper, Canny operator is used to detect edges. The Canny detector use the first order derivative of the Gauss function as filter to smooth image and then the gradient is obtained. The local maximum is found in the vertical direction of the edge, and accurate edge is obtained by using the high and low threshold.

Edge tracking begins from a point of the edge image, searches and connecst adjacent edge points, and then detects all edge lines. In this paper, we use the Freeman chain code[4] to track the binary image in the order of eight vector directions, as shown in Figure 1. The first edge point will be encountered as the starting point of the chain code, and then the eight point around will be scaned orderly. When a point is encountered,the length of the chain code increases and the point is to avoid duplication, and then the eight point around will be scaned orderly.The tracking goes on until there is no edge points around a point.

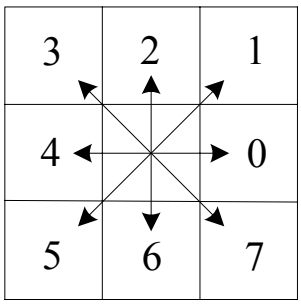


Figure.2 Eight directions of chain code

4. Direct least square method

Direct least squares is an ellipse fitting method proposed by Andrew Fitzgibbo[5]. In this method, the minimization of the distance between the data points and the distance of the ellipse is taken as the mathematical principle to seek the optimal parameters.

Planar second order curve model can be expressed as:

$$F(x,y) = ax^2 + 2bxy + cy^2 + 2dx + dey + f = 0 \tag{6}$$

The formula can be also described as implicit equations :

$$F(\alpha, X) = \alpha \cdot X \tag{7}$$

where, $\alpha = [a,b,c,d,e,f]^T$, $X = [x^2, xy, y^2, x, y, 1]$. $F(\alpha, X)$ is called the algebraic distance of the point to the second order curve Because of the error, the curve is generally second order curve, not necessarily an ellipse. In order to guarantee the fitting result is an ellipse, the following conditions should be met:

$$4ac - b^2 > 0 \tag{8}$$

The direct least square ellipse fitting method can guarantee that the fitting result of second order curve is an ellipse, and the constraint condition can be expressed as a matrix:

$$\alpha^T C \alpha = 1 \tag{9}$$

Where, C is a constraint matrix:

$$C = \begin{bmatrix} 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{10}$$

According to the least square principle, the ellipse fitting problem is converted into a minimization problem:

$$\min\left(\sum_{i=1}^{n} F(a, x_i)^2\right) \tag{11}$$

Because the ellipse parameter vector a satisfies the constraint condition, the minimization problem can be solved by using the generalized feature system:

$$\begin{cases} DD^T\alpha = \lambda C\alpha \\ \alpha^T Ca = 1 \end{cases} \tag{12}$$

Where, $D = [x_1, x_2, \cdots, x_n]^T$, The generalized feature system has only one positive eigenvalue. The eigenvector of the positive eigenvalue is the parameter vector of the ellipse.

5. Algorithm flow

The algorithm flow of this paper is as follows:

1)Detect the edge of the original image, extract the edge points.

2)Track the edge image, and remove the noise by setting a threshold.

3)Perform fitting of each arc section. Classify the arc segments by the fitting results of ellipse parameter.

4)Choose the arc segments corresponding to the largest ellipse. Perform fitting of these arc segments

6. Experimental results

To verify the effectiveness ,the algorithm is tested by using real images of size 2048×2048.

The first one (Figure.3) is the original image. There are two ellipses in the image.We need recognize the bigger one. The second one (Figure.4) is the edge image detected by Canny detector. The third one (Figure.5) is the image after edge tracking.The rough line represents edge lines after removing the noise. We can obtain five arc segments by setting threshold.Five fitting results can be obtained by performing ellipse fitting of the five arc segments .The arc segments are classified according to the fitting results. The arc segments with similar ellipse parameters are classified into one group. Because the projection ellipse of the dock ring is the larger one, we classified arc segments with larger long and short axis into one group and carried out fitting of these arc segments. The fourth one (Figure.6) is the final fitting result, and the projection ellipse of the dock ring is detected accurately. The ellipse parameters as shown in Table 1

Figure.3 Original image

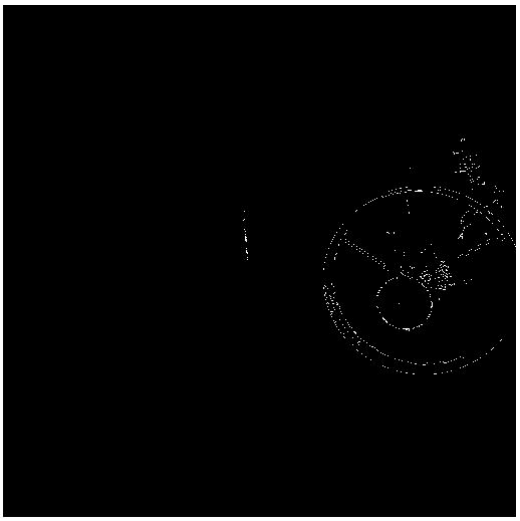


Figure.4 Edge image

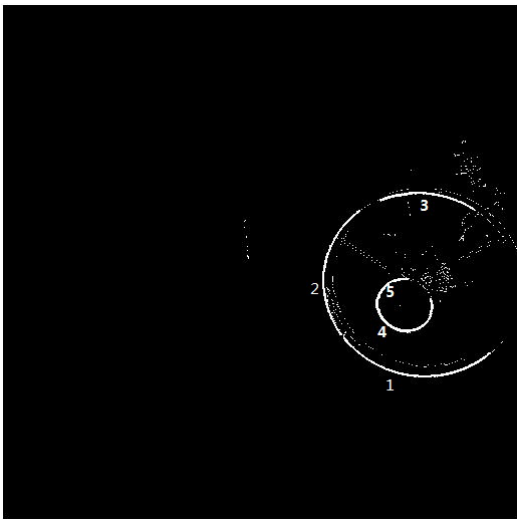


Figure.5 Arc segments after tracking

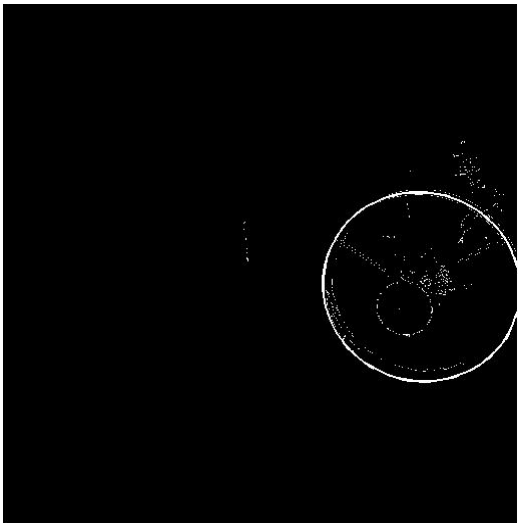


Figure.6 Extracted ellipses

Table.1 Extracted ellipse parameters

x_c	y_c	a	b	θ
1669.3	1111.9	393.6742	368.1623	-167.6°

7. Conclusion

In this paper, an algorithm for detecting projection ellipse of docking ring based on arc segment combination is proposed. Boundary tracing is firstly carried on the image, and the ellipse boundary chain code is formed. After that,the arc segments is classified and the fitting of ellipse is completed by the direct least square fitting algorithm. Finally, the feasibility of the algorithm is validated and the result shows algorithm in the paper can finish the detection and parameters fitting of the ellipse accurately.

Reference

[1] D.H.Ballard.Generalizing the Hough Transform to detect arbitrary shapes[J].Pattern Recognition,1981,13(2):111 — 122.

[2] McLaughlin R A.Randomized hough transform：Improveed ellipse detection[J].Digital Signal Processing Application,1998,19(3/4):299-305

[3] Canny J F.A ComPutational APProaeh to Edge Detection[J].IEEE Transaction On Pattern Analysis and Machine Intelligenee,1986,8(6):679 — 698.

[4] Freeman H.Computer processing of line-drawing images[J].Computing surveys，1974,6(1):57-97

[5] Andrew Fitzgibbon,Manrizio Pilu,Robert B.Fisher.Direct Least Square Fitting of Ellipses[J].IEEE Transaction on Pattern Analysis and Machine Intelligence,1999,21(5):476 — 480.

New Concept Study on Electromagnetic Force Deorbit of Waste Satellite with Active Power-on System

Yue Sun, Shan Lu, Hailei Wu, Chaozhen Liu, Chenglong Jia

（ Shanghai Institute of Spaceflight Control Technology, Shanghai 201109

Shanghai key Laboratory of Aerospace Intelligence Control Technology, Shanghai 201109）

Failure abandoned satellites will become space debris occupying precious orbit resources, and pose a threat to the security of other spacecraft on orbit. There are many ways of deorbiting waste satellites on orbit. Since electromagnetic force do not need to carry a lot of propellant, the way of making satellites away from current rail based on electromagnetic force acquire more and more attention. The research of using the electromagnetic force for deorbiting is mainly focused on electric tethered at present. The basic principle of electric tethered deorbit is that conductive tethered movement in the earth's magnetic field will cut lines of magnetic force and produces electromotive force. When free electrons and ions in the tethered and the atmosphere ionosphere constitute a closed loop, it can produce electric current in the tether, and then produce electromagnetic force due to interact with earth's magnetic field. In this paper, two new deorbit methods based on the electromagnetic force are put forward, a charged satellite and electricity metal ring. A charged satellite is using satellite carried positive (negative) ions, movement in the earth's magnetic field, and get Lorentz force. Though the Lorentz force does not work, but by changing the electric polarity can achieve the goal of changing orbit. The basic principle of electricity metal ring is that using the magnetic field generated from the metal ring interacts with the earth's magnetic field to produce magnetic attractive force, using magnetic propulsion enter into the atmosphere. Firstly this paper introduces the principle and control method of three kinds of schemes, and then analyzes the requirement of attitude control system and the electronic control system on the star, last comprehensive analysis and comparison of their respective applicable scope, time-consuming of away from the orbit and advantages and disadvantages.

Key words: electromagnetic force, deorbit, failure abandoned satellites.

1 introductions

With the development of space activities, space debris environment is getting worse. As of April 5, 2012, the number of space debris bigger than 1cm up to several hundred thousand, of which over 16,280 have been cataloged, posed a serious threat to the safe operation of spacecraft in orbit. Domestic and foreign scholars have carried out a lot of space debris protection research and experimental work; get a lot of theoretical results with empirical data. Foreign already have the related project, to be carried out in-orbit demonstration, while the domestic research is still in the theoretical stage.

The main way in the research of debris removal are: First， using a tethered capture system for controlled de-orbitation through pulling where the capture is performed using throw-nets or alternatively a har-poon. Second，using rigid capture systems with a particular emphasis on tentacles (potentially combined with a robotic arm). Here the de-orbitation is achieved through a push-maneuver. These methods are based on the output force of thruster, another method is based on

the electromagnetic force. Since electromagnetic force do not need to carry a lot of propellant, the way of making satellites away from current rail based on electromagnetic force acquire more and more attention.

The way based on the electromagnetic force of passive away from rail is mainly electric tethered. This paper proposed two kinds of new method for away from current rail based on the electromagnetic force: a charged satellite and electricity metal ring. We will discuss the principle and control method of these three methods respectively in this paper below.

2 Electric tethered

Electric tethered de-orbitation system, the conductive tether move with the orbital speed in the earth's magnetic field, in the department of rope induced electromotive force, the charged particles in the ionosphere is collected at the top of the line, then be launch out in the end of the tether, the stable current is produced. The magnetic field produces the Lorentz force on the electric tether, the electric tension. Since the direction of the Lorentz force is opposite with the speed direction, so the Lorentz force does negative work, then reduce satellite orbit energy, reduce orbital altitude. Using this system can let the satellites away from current rail more quickly than rely on atmospheric drag.

Electric tethered de-orbitation speed is faster, set the quality of the satellite is 500 kg, cross-sectional area of 1.5 m^2, lashing material is aluminum, cross sectional diameter of 0.4 mm; the quality on end of the rope is 30 kg; orbit altitude of the Satellite is 967.2 km, orbital inclination is 50 °, the length of tether is 3000 m, lashing quality is 1.0179 kg, estimated that time used for away from the rail is 81 days (to reach 120 km orbit). [1]

Electric tethered method which makes debris away from the rail is an important way to study, to implement applications on the engineering as soon as possible, we need to carry out the following research work:

1) Precise geomagnetic modeling and electric tethered system modeling

The principle of electric tethered de-orbitation is that using electric power pull dead satellites out of orbit, and eventually into the atmosphere and burned. Compared with the traditional rocket de-orbitation system, electric tethered de-orbitation can significantly reduce the quality of the de-orbitation system. Electric power pull is caused by geomagnetic, so need to establish a high precision magnetic model, and tethered system dynamics model is also established.

2) Research of electric tethered bending deformation dynamics

Electric tethered system dynamic model usually use rigid rod model, the rope is equivalent to a rigid rod, assumes that the rope will not occur bending deformation in the process of de-orbitation, ignore the elastic properties of tethered. Rigid rod model is the basic tools of tethered space plan, the model is simple, small amount of calculation, but there is error with the actual characteristics of the tethered system. When estimating and calculating the time of de-orbitation, need to use a tethered bending deformation of the elastic model to satisfy the high precision requirement. Therefore, on the basis of the rigid rod model, assuming that the electric tension of tethered is uniform distribution and equal, flexible model is established, and researching the influence of the electromotive force and Lorentz force induced by bending deformation.

3) Vibration characteristic analysis of electric tethered system

Electric tethered system is composed by satellite ontology, the mass at the end of the tether and the conductive tether have begun. The change frequency of rope tension is associated with the cycle of oscillation, and also associated with the vibration frequency of tether. Electric rope is wiggling in

the low frequency, lashing also slightly higher frequency of vibration. The motion of the tether, the motion of the end mass and the slight vibration of tether are coupled, and large wiggle will cover up the influence of high frequency vibration. When researching the vibration and the tension change of the tethered system in a short period of time, we can use the above conclusion, ignore the motion of the end mass, simplified the satellite and the end mass as fixed, separate the tether swing caused by the motion of the end mass and tether lateral vibration, focuses on the vibration of the tether.

4) Electric tethered device design

Electric tethered de-orbitation devices are usually composed by conductive tethered, tethered control units, electronic transmitters and the release recycle unit. Need the detailed design of each part.

5) Electric tethered system de-orbitation time estimate

Electric tethered de-orbitation system is a fairly complex state, involves the electromagnetic field and soft body dynamics, control, orbit, microelectronics, solar radiation and gravity gradient and other various factors, these factors have certain influence on process of de-orbitation, influence on the de-orbitation time estimate precision. Accurate estimates of de-orbitation time relies on reasonable satellite orbit kinematics model, the geomagnetic field model, electric tethered system dynamics model, the host satellite dynamics model, the disturbing force model. It is necessary to fully consider the impact of these factors on the process of de-orbitation. Using orbit elements iterations accumulative algorithm for de-orbitation time estimate, fully consider every moment change of magnetic field, current, force in the process of de-orbitation.

6) De-orbitation effect analysis of electric tethered system parameter change

Electric tension is not only related with the geomagnetic model, and also related with orbit state of the electric tethered system. On the basis of accurate geomagnetic model, according to the orbit state of electric tethered system can accurately determine the value and direction of the electric tension under geomagnetic field model, so the effects on the de-orbitation of parameter changes can be obtained.

Limitations in application of this method lie in:

1) Thousands of meters of conductive tether led to high space collision risk.

3 Charged satellites

Charged satellite flying in the earth's magnetic field cutting magnetic lines will be affected by Lorentz force, under the action of the Lorentz force, can produce satellite orbit changing, and then away from current orbit. According to the theory of field or electromagnetism and theoretical mechanics research methods are based on Lagrange function, first get motion equation of charge or charged satellite in the magnetic field through Lagrange's equation, and then analysis the force on charged satellite from the equation of motion, put it into Lagrange orbit elements variation equation, finally the impact of the magnetic field on satellite orbit can be concluded.

The motion equation of Charge Q of satellites in earth's gravitational field and magnetic field is[2]:

$$F_r = -\frac{GEm}{r^2} + \frac{\mu_0 M}{4\pi} Q \frac{\dot{\theta}}{r^2}$$

$$F_\theta = -\frac{\mu_0 M}{4\pi} Q \frac{\dot{r}}{r^3}$$

In the formula, θ is the true anomaly of the satellite; M is the satellite quality. The right end of the first formula is the space force in radial direction, the first is the earth's gravitational field force component, the second is the earth's magnetic field force component, and the right end at the second formula is the component of the magnetic field on the satellite, and the force is perpendicular to the radial direction. Here only study the impact of the earth's magnetic field on a charged satellite movement, radial and perpendicular to the radial acceleration is,

$$S = \frac{\mu_0 M}{4\pi m} Q \frac{\dot{\theta}}{r^2}$$
$$T = -\frac{\mu_0 M}{4\pi m} Q \frac{\dot{r}}{r^3}$$

Do not break general, using the characteristics of elliptical orbit:

$$\dot{\theta} = \frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{na^2\sqrt{1-e^2}}{r^2} \qquad r = \frac{a\left(1-e^2\right)}{1+e\cos\theta}$$

Can be derived,

$$\dot{r} = \frac{hae\sin\theta}{\sqrt{1-e^2}}$$

Then

$$S = \frac{\mu_0 M}{4\pi m} Q \frac{na^2\sqrt{1-e^2}}{r^4}$$
$$T = -\frac{\mu_0 M}{4\pi m} Q \frac{nae}{\sqrt{1-e^2}} \frac{\sin\theta}{r^3}$$

According to the Lagrange orbit elements variation equation:

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \frac{2}{n\sqrt{1-e^2}}\left[Se\sin\theta + \frac{P}{r}T\right]$$
$$\frac{\mathrm{d}e}{\mathrm{d}t} = \frac{\sqrt{1-e^2}}{na}\left[S\sin\theta + T\left(\cos E + \cos\theta\right)\right]$$
$$\frac{\mathrm{d}\omega}{\mathrm{d}t} = \frac{\sqrt{1-e^2}}{nae}\left[-S\cos\theta + T\left(1+\frac{r}{p}\right)\sin\theta\right]$$

After integral, the influence on the satellite orbit semimajor axis, the eccentricity and perigee amplitude angle can be obtained under the geomagnetic perturbation forces,

$$\Delta a = 0$$
$$\Delta e = -\frac{\mu_0 MQ\left(1-e^2\right)}{4\pi m\sqrt{GEp^3}}\left(\cos\theta - \cos\theta_0\right)$$
$$\Delta\omega = -\left[\frac{\mu_0 MQ}{2\pi m\sqrt{GEp^3}}\left(\theta-\theta_0\right) + \frac{\mu_0 MQ\left(1+e^2\right)}{4\pi me\sqrt{GEp^3}} \times \left(\sin\theta - \sin\theta_0\right)\right]$$

l) The magnetic field do not effect a charged satellite orbit semimajor axis, so it is no effect on orbital period, but the impact on the orbital eccentricity is periodic, the perigee precession has not only periodic of perturbation effect, but also long-term changes of perturbation.

2) The earth's magnetic do not effect a charged satellite orbit semimajor axis, neither periodic perturbation nor long-term perturbation. The results from the perturbation theory are the same with magnetic field not doing any work on satellite.

Change the orbit inclination and semimajor axis by changing polarity of charge-mass ratio. Switching polarity of charge-mass ratio when satellite pass by perigee and apogee, this kind of variable polarity solution can increase continuously orbit semimajor axis, it can even make the semimajor axis grows until the satellite into escape orbit. Sets the initial orbit of the satellite to the geostationary transfer orbit: a =24462km， e = 0.723 , i = 0°, when satellites are in their way from perigee to apogee, make charge-mass ratio (q/m) =-5C/kg, when satellites are in their way from apogee to perigee, make charge-mass ratio (q/m) =5C/kg. Through simulation, satellite run from geostationary transfer orbit to escape the earth's gravitational field need 1.84 years (670 days). [3]

Limitations in application of this method lie in:

1) Charge of natural charged state is less, it is hard to meet a series of Lorentz force orbit demand, need to adopt active charged technology solutions, increased the difficulty of technology.

2) Asked for a certain electric control ability on the satellites, able to change the charge polarity, or semimajor axis will not change.

4 Electricity metal ring

Japan plans to build a giant metal mesh, to reach several kilometers in diameter, for removal of large numbers of space junk in earth orbit. Metal net connected to a satellite, the system launch into space, and then the metal mesh unfolds, and run along the earth's orbit, cleaning the space junk encountered along the way. In the process of operation, the metal net get charged gradually, making it gradually lower orbit under the influence of the earth's magnetic field, and finally into the earth's atmosphere, metal mesh and space junk will be burned when fall in the upper atmosphere.

On the basis of Japan’s metal mesh plan, put forward the idea that satellites is installed on three orthogonal electric coil, by changing the current value and direction of each coil, can produce any value, direction of electromagnetic control force and moment. The electric coil installed on the satellites produce strong magnets, and interacts with the earth space magnetic field; the magnetic moment vector of spacecraft (value and direction) can carry on the active control, so make the spacecraft to obtain the desired effective. Thrust, after the satellite cannot serve, it can de-orbit independently.

Assume electrify coil as the magnetic dipole, the strength of the magnetic dipole μ can be expressed as, [4]

$$\mu = nI\pi R_c^2$$

In the formula, n for the coil number of turns, I as the coil current, Rc for radius of coil. The direction of the magnetic dipole and the coil current direction conform to the right-hand rule.

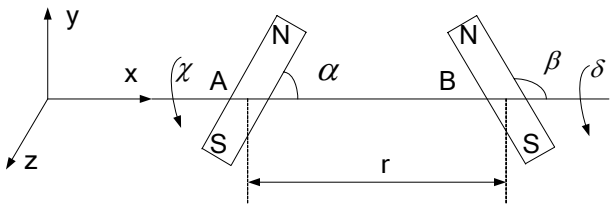


Fig.1 Three-dimensional model of the dipole

Due to the distance from the center of the earth to the satellite is far, (the separation distance is greater than 6 ~ 8 dipole radius), Taylor series expansion of the first-order approximation, can use dipole model of the far field. The forms of far field model are the most simple and convenient, and forms of midfield and near field models are too complex, so is very difficult for application. The dipole three-dimensional formula for the far field model is, [5]

$$F_x = \frac{3}{4}\frac{\mu_0\mu_A\mu_B}{\pi r^4}\left(2\cos\alpha\cos\beta - \cos(\delta-\chi)\sin\alpha\sin\beta\right)$$

$$F_y = -\frac{3}{4}\frac{\mu_0\mu_A\mu_B}{\pi r^4}\left(\cos\alpha\sin\beta\cos\delta + \cos\chi\sin\alpha\cos\beta\right)$$

$$F_z = -\frac{3}{4}\frac{\mu_0\mu_A\mu_B}{\pi r^4}\left(\sin\alpha\cos\beta\sin\chi + \cos\alpha\sin\beta\sin\delta\right)$$

$$M_x = -\frac{1}{4}\frac{\mu_0\mu_A\mu_B}{\pi r^3}\sin\alpha\sin\beta\sin(\delta-\chi)$$

$$M_y = \frac{1}{4}\frac{\mu_0\mu_A\mu_B}{\pi r^3}\left(\cos\alpha\sin\beta\sin\delta + 2\sin\alpha\cos\beta\sin\chi\right)$$

$$M_z = \frac{1}{4}\frac{\mu_0\mu_A\mu_B}{\pi r^3}\left(\cos\alpha\sin\beta\cos\delta + 2\sin\alpha\cos\beta\cos\chi\right)$$

In the formula, permeability of vacuum $\mu_0 = 4\pi\times10^{-7}H/m$, magnetic moment $\mu_A = I_A \bullet S_A$, $\mu_B = I_B \bullet S_B$.

If two magnetic dipole is parallel, for example α=β=0, the magnetic field force is largest, then the electromagnetic force can be simplified as,

$$F_x = \frac{3}{2\pi}\frac{\mu_0\mu_A\mu_B}{r^4}$$

If two magnetic dipole is vertical, for example α=π/2, β= 0, then the electromagnetic force can be simplified as,

$$M_z = \frac{1}{2\pi}\frac{\mu_0\mu_A\mu_B}{r^3}$$

For maximum eliminate the magnetic moment effect, use the magnetic force line tracking strategy [6] -- Makes the point of magnetic moment vector and the local magnetic field lines tangent parallel, thus make the aircraft produces effective propulsive force in flight direction, at the same time, avoid to produce additional magnetic moment, and by adjusting the positive and negative of the magnetic moment to make magnetic thrust always stay unified in flight direction projection.

Limitations in application of this method lie in:

1) This method does not apply to the orbit of 0 ° angle. Because when the orbit inclination is 0°,

magnetic field lines vertical pass through the closing coil, does not have a magnetic thrust. If the orbital inclination is bigger, the magnetic thrust is larger.

2) Asked for a certain attitude control ability on the satellites, otherwise under the earth's magnetic moment, the attitude will occur deflection, and the direction of the magnetic will change and without controlled, difficult to guarantee the satellite fragments can drift away from the orbit, and then into the atmosphere, likely oscillation near the current location.

The three way of de-orbitation based on electromagnetic force is compared in the following table.

	Electric tethered	Charged satellites	Electricity metal ring
de-orbitation time	fastest	fast	slower
Limitations in application	Thousands of meters of conductive tether led to high space collision risk	need to adopt active charged technology solutions, increased the difficulty of technology	does not apply to the orbit of 0 ° angle
		asked for a certain electric control ability on the satellites	asked for a certain attitude control ability on the satellites

Reference

[1] Dafu X. Electric tethered de-orbitation system modeling and dynamic analysis. Harbin institute of technology, 2010.

[2] Linsen L. Influence on a charged satellite movement of the magnetic field. Guangxi physical, 1994, 15(5): 44-46.

[3] Yang G. Charged satellite Lorentz force propulsion: orbit change and propulsion concepts. The space exploration seminar papers, 2009: 293-301.

[4] Min H, Guoqiang Z, Zhaohui D. Nonlinear feedback control of fractionated spacecraft electromagnetic formation flying in low-Earth orbit. Chin. J. Space Sci. , 2012, 32(3): 417-423.

[5] Long W, Leping Y, Juxiao X. Electromagnetic formation flight and electromagnetic docking key technology and progress. Equipment command technical college journals, 2009, 20(1): 75-78.

[6] Kailei L, Wenhao L, Heng Z. Leo spacecraft magnetic propulsion method. Journal of aerospace, 2012, 33(7): 985-990.

Binocular Vision Pose Measurement of Spacecraft in Short Range Based on Docking Ring

CAO Shu-qing[1,2] LU Shan[1,2] LIU Zong-ming[1,2] ZHANG Yu[1,2] YANG Guang[1]

（1. Shanghai Institute of Spaceflight Control Technology, Shanghai 200233, China;
2. Shanghai Key Laboratory of Aerospace Intelligent Control Technology, Shanghai 200233, China）

ABSTRACT: According to the docking ring which is a typical component on space vehicle can provide circle feature，a binocular vision-based pose measurement method of spacecraft in short range based on docking ring was proposed to solve the technology of relative measurement in the final approaching phase operations of rendezvous and docking. Firstly, a median filter was used to smooth the original images and the edges in the smoothed images were processed by using the canny edge-detection algorithm， then removed the noise and isolated point and acquired the sectional continuous ellipse's edge-chain. By taking the boundaries of the image edge extraction to piecewise fit ellipse and adopting the least square fitting algorithm based on locally optimal iterative weighted least-square to achieve the entire ellipse fitting. The fundamental matrix was calculated by using the left and right camera calibration intrinsic and extrinsic parameters and at least six matching points in the left and right images were derived based on epipolar constraints and space 3-dimensional circle parameters were acquired by iteratively calculating. finally centroid method was adopted to calculate space 3-dimensional circle center then the position of the target satellite were acquired, singular value decomposition was adopted to calculate the normal vector of the circular planar, then the pitch angle and yaw angle of the target satellite were acquired According to the triangular relationship of the geometry. Simulation results show that this algorithm has strong stability and can satisfy the measuring requirement of relative pose for RVD spacecrafts.
Keywords: Edge-detection; Ellipse Fitting; Iterative Weighted Least-square; Locally Optimal; singular value decomposition; Pose Measurement

Fund Project: Program for Excellent Technical Leaders of Shanghai Science and Technology Committee(14XD1421500)
Auther Profile: CAO Shuqing(1990-),female(Han),Assistant Engineer, Master, Research Area of Image Processing and Pattern Recognition.(shuqing070432@163.com)

INTRODUCTION

In consideration of circular docking rings are the basic configuration of the majority of non-cooperative target satellites for rendezvous and docking in space environment, so we can detect the target satellite's circular features to achieve ultra-close non-cooperative target satellites high-precision relative attitude estimation. In this paper, we firstly introduce the basic idea of edge processing, ellipse fitting and least square method. Then based on the characteristics of the typical least square method[1], the weighted least square ellipse fitting method with boundary arcs is proposed. Finally, using local optimization iterative method to achieve final ellipse features accurate extraction and estimate the space round parameter and then by decomposing the space

round parameter to acquire the relative attitude of the spacecrafts and verify the effectiveness of the algorithm with experiment. Algorithm flow chart shown in Figure 1:

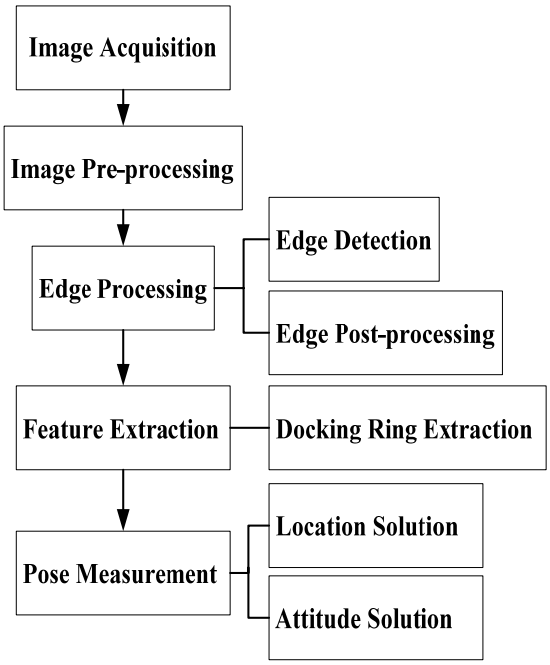


Fig.1 Flowchart of pose measurement algorithm

1 IMAGE PRE-PROCESSING

Image pre-processing includes image logarithmic transformation and lens distortion rectification. An assumption that the gray level of a certain pixel of the original image is B, then the gray level of the pixel with gray logarithm transformation is converted to:

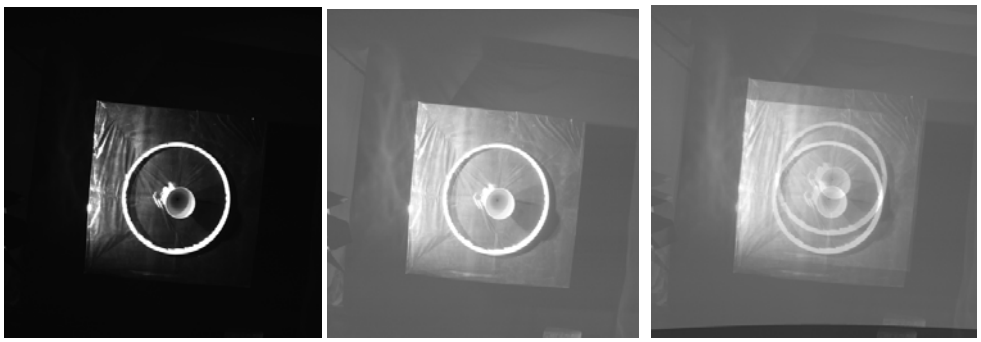
$$B'=k\times\log(1+B) \tag{1}$$

K is a constant k which is used to guarantee image gray scale in 0~255.

Assuming the image measurement coordinate is (x, y), logarithmic transform coordinate is $(x-x_0-\Delta x, y-y0-\Delta y)$, (x_0, y_0) is the main point, Δx and Δy are the image distortion, then:

$$\begin{aligned}\Delta x&=(x-x_0)[k_1(r^2-r_0^2)+k_2(r^4-r_0^4)]+p_1[r^2+2(x-x_0)^2]+2p_2(x-x_0)(y-y_0)+c_1(x-x_0)+c_2(y-y_0)\\ \Delta y&=(y-y_0)[k_1(r^2-r_0^2)+k_2(r^4-r_0^4)]+p_2[r^2+2(y-y_0)^2]+2p_1(x-x_0)(y-y_0)\end{aligned} \tag{2}$$

The inverse solution method is used to solve the image distortion rectification parameter position coordinate. Then the gray values of the interpolation image are interpolated by bilinear interpolation method to complete the image distortion rectification. Image pre-processing effect as shown in Figure 2:



（a）original image （b）logarithm transform （a）distortion rectification

Fig.2 Image pre-processing effect charts

2 EDGE PROCESSING

2.1 Canny Edge-detection

Edge-detection is the precondition of feature extraction and pose measurement. This paper uses canny algorithm to detect the edge of the image. Figure 3 and 4 are the steps of the canny algorithm and the pre and post comparison effect charts of the edge-detection:

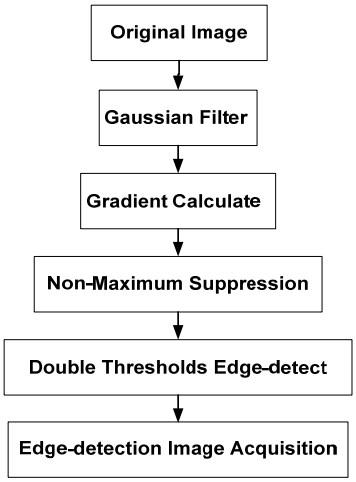
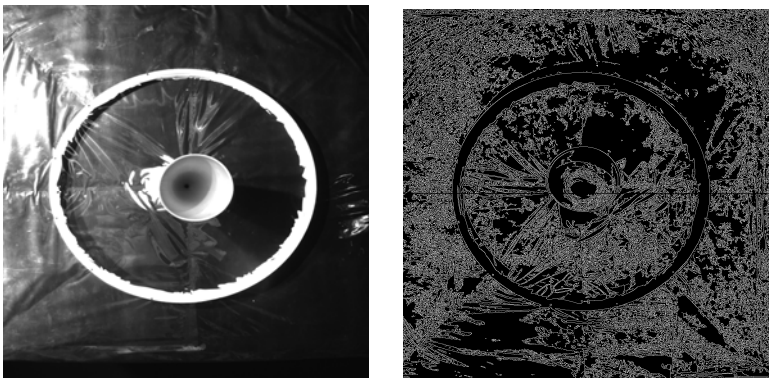


Fig.3 Flowchart of canny edge-detection



（a）original image （b）canny edge-detection

Fig4 Pre and post effect charts of the edge-detection

2.2 Edge Post-processing

As shown in figure 4 (b), there are many chaos edge and noise, which is not conducive to the extraction of the ellipse characteristics. In this paper, edge tracking method is used to process he

canny edge-detection image which can remove the noise and isolated point and acquire the sectional continuous ellipse's edge-chain and then take the boundaries of the image edge extraction piecewise fit ellipse[2][3]. The flowchart and effect chart of the edge post processing shown in Figure 5 and 6:

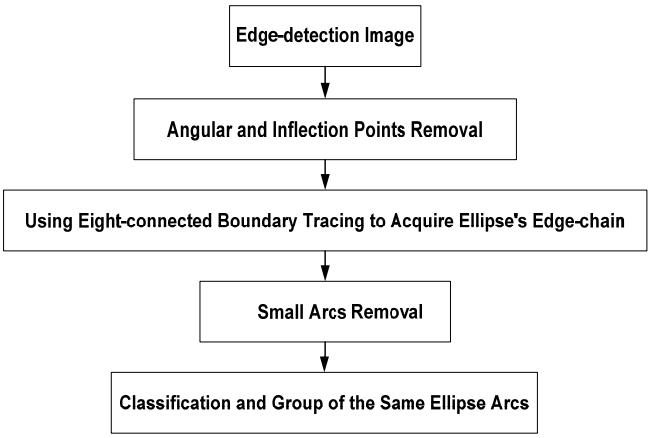


Fig.5 Flowchart of edge post processing

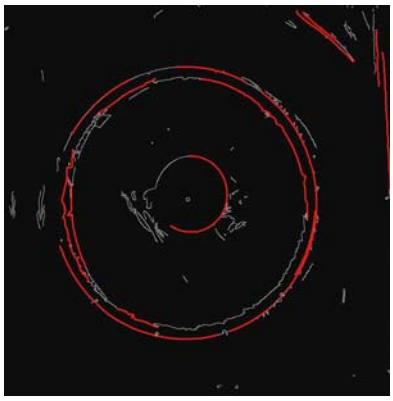


Fig.6 Effect chart of the edge post processing

3 FEATURE EXTRACTIONS

Based on the round outer edge of docking ring can provide a single round feature, the single space circular feature of the docking ring parts is extracted. By circle fitting way item by item to detect possible arc segment boundaries in the boundary tracking image, through merging arcs to retain the most possible ellipse edges and then piecewise fitting ellipse to complete the final ellipse detection [4].

3.1 Arc Fitting

The equation of standard circle is:

$$(x-x_c)^2+(y-y_c)^2=R^2 \tag{3}$$

Expand the standard circle:

$$-2xx_c-2yy_c+x_c^{\,2}+y_c^{\,2}-R^2=-x^2-y^2 \tag{4}$$

Equation (4) can be written as:

$$xA+yB+C=L \tag{5}$$

where：

$$L = -x^2 - y^2 \tag{6}$$

The center（ x_c， y_c ）and the radius parameters R as follows:：

$$x_c = -\frac{A}{2} \tag{7}$$

$$y_c = -\frac{B}{2} \tag{8}$$

$$R = \sqrt{\frac{A^2 + B^2}{4} - C} \tag{9}$$

Substitute each edge points in figure 6 into Equation（5）and list the error equation, the solution to the equation is the circle equation. Circle fitting is very important in the extraction of ellipse. By considering the edge as a circle and preserving only the edge of the circular arcs can filter out a lot of unnecessary ellipse fitting dates.

3.2 Ellipse Fitting

3.2.1 Ellipse Function Model Representation

Elliptic curve is a standard quadratic equation, it can be expressed as :

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \tag{10}$$

where：

$$m = [a\ b\ c\ d\ e] \tag{11}$$

$$n = [n_1\ n_2\ n_3\ \cdots\ n_i] \tag{12}$$

$$n_i = [x_i^2\ x_i y_i\ y_i^2\ x_i\ y_i] \tag{13}$$

m is the unknown ellipse parameters， n_i is the coefficient matrix of the observation point i (x_i, y_i) coordinate， N is a total observation points numbers. For each point on the edge (x_i, y_i)， the constant $a = f$ =1， you can get the error equation:

$$v_i = x_i^2 + bx_i y_i + cy_i^2 + dx_i + ey_i + 1 \tag{14}$$

Written in the form of a matrix:

$$V = l + n \cdot m \tag{15}$$

l are column vector which all the elements are constant 1:

$$l = [1\ 1 \cdots 1]^T_{1\times N} \tag{16}$$

Using the least square method to solve the least squares solution of the coefficient m of the elliptic equation:

$$\hat{m} = -(n^T n)^{-1} n^T l \tag{17}$$

The equation (17) can obtain the elliptic equation of the least square sense [5], in order to further enhance the ability to resist the noise of the least squares, the use of weighted processing for further optimization.

For the error equation (14), the weight function is defined:

$$p_i = \frac{1}{\sqrt{F_x^2 + F_y^2}} \tag{18}$$

Where:

$$F_x = 2x_i + by_i + d \tag{19}$$

$$F_y = bx_i + 2cy_i + e \tag{20}$$

For each edge point, the weight function based on the formula (19) can lists error equations in the form of a matrix:

$$V = l + n \cdot m \quad \text{weight matrix P} \tag{21}$$

The corresponding least squares solution is:

$$\hat{m} = -(n^T Pn)^{-1} n^T Pl \tag{22}$$

Using formula (18) to update the weight of the given error equation continuously, and until the error results are stable and the final elliptic algebraic equation can be obtained.

3.2.2 Ellipse Geometric Parameters Solution

In order to facilitate the subsequent solution of the target pose, the geometric parameters of the ellipse is necessary [5]. The first three in the elliptic equation (10) is actually the standard quadratic form. The geometric parameters of the ellipse can be solved by using orthogonal congruent transformation of the linear algebra theory including the center position (x_o, y_o), the long half axis a and the short half axis b, and the rotation angle θ, as shown in Figure 7:

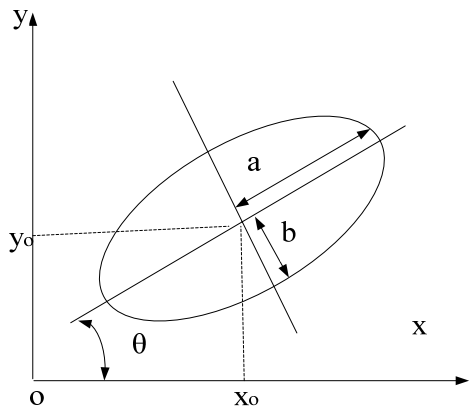


Fig.7 Ellipse two-dimensional plane schematic diagram

The ellipse general coefficient of the formula (10) represented by $m = (a \;\; b \;\; c \;\; d \;\; e)$, according to the configuration of the quadratic term matrix B, such that:

$$ax^2 + bxy + cy^2 = [x \;\; y]\, B\, [x \;\; y]^T \tag{23}$$

$$B=\begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix} \tag{24}$$

By using eigenvalue decomposition to obtain the feature vector and feature value of the formula (24)：

$$[M \quad N]=eig(B) \tag{25}$$

Contract quadratic form into standard form with orthogonal congruent transformation:

$$m_1=\left[N(1,1) \quad N(2,2) \quad [m(4) \quad m(5)]\times M\right] \tag{26}$$

Where the rotation matrix:

$$M=\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{27}$$

$$\theta=\arcsin(M(2,1)) \quad \theta\in\left(-\frac{\pi}{2},\frac{\pi}{2}\right] \tag{28}$$

while $N(1,1)>0$ ， $N(2,2)>0$,complete analytical calculation of the ellipse parameters. According to resolution algorithm, oval standard equation can be obtained:

$$\frac{(x^{'}-x_o)^2}{p^2}+\frac{(y^{'}-y_o)^2}{q^2}=1 \tag{29}$$

The long half axis of the ellipse:

$$p=\sqrt{\frac{1+\frac{c_o^2}{4a_o}+\frac{d_o^2}{4b_o}}{a_0}} \tag{30}$$

The short half axis:

$$q=\sqrt{\frac{1+\frac{c_o^2}{4a_o}+\frac{d_o^2}{4b_o}}{b_0}} \tag{31}$$

The center position:

$$\begin{bmatrix} x \\ y \end{bmatrix}=M\begin{bmatrix} x^{'} \\ y^{'} \end{bmatrix}-M\begin{bmatrix} x_o \\ y_o \end{bmatrix} \tag{32}$$

where : $a_0=N(1,1)$ ， $b_o=N(2,2)$ ， $c_0=m(4)\times M$ ， $d_0=m(5)\times M$ 。

$$\begin{bmatrix} x_o \\ y_o \end{bmatrix} = \begin{bmatrix} \frac{c_o}{2a_o} \\ \frac{d_o}{2b_o} \end{bmatrix} \tag{33}$$

4 POSE MEASUREMENT

Visual pose measurement using a single circle feature which is offered by the outer edge of the satellite can obtain the pose parameters included of the three-dimensional coordinates of the space circle and the normal vector of spatial circle support plane. Due to the normal vector around the center of space circle have the characteristic of rotation symmetry. When the support plane normal vector is determined, spatial circle image will not be changed no matter how much angle obtained by circle around its rotational axis is rotated. So the attitude angle of the space circle is determined by only two degrees of freedom, namely the pitch angle and the yaw angle.

4.1 Location Solution

In the image processing technology, after extracting the ellipse around the image, we can use the camera parameters of the left and right images and the initial several homonymy points obtained by epipolar constraints to estimate the space circle parameters. Though adjusting the parameters of the space circle continuously to obtain the optimum space circle parameters and then get the 3D reconstruction results.

The nuclear line equations are as follows. If the point coordinate of the left image is $u = (x, y, -f)^T$, the corresponding point coordinate on the right image is $u' = (x', y', -f')^T$. Then according to the computer vision or photogrammetry collinearity equation, the epipolar constraints can be obtained as follows:

$$u^T R_1^T \left[B\right]_\times R_2 u' = 0 \tag{34}$$

R_1, B and R_2 are given by the external orientation elements of the image. Actually, the expressions of the above formula（34）is the nuclear line equation. Through a point on the left image, we can get the epipolar equation of the right image, on the contrary can also be easily.

In order to obtain the initial parameters values of the space circle, the six points are selected on the left image and the nuclear line equations are calculated with the above method for the right image and the six homonymy points can be obtained from the corresponding ellipse .Finally the 3D coordinates of the six points can be obtained by using the method of photogrammetry. If the 3D coordinates of these six points are $U_i = (X_i, Y_i, Z_i), i = 1,2,3,4,5,6$.

Calculating the gravity center of the above six points, if $\bar{U} = (\bar{X}, \bar{Y}, \bar{Z})$， then through the following methods we can calculate the normal direction and circle center of the circle:

(1) Matrix:

$$N = \sum_{i=1}^{6} (U_i - \bar{U})(U_i - \bar{U})^T \tag{35}$$

（2）Make singular value decomposition of matrix N, the feature vector corresponding to the minimum singular value of the decomposition is the normal vector of the space circle, which is assumed to be:

$$\vec{n} = (n_x, n_y, n_z)^T \tag{36}$$

where: $n_z > 0$.

（3）Projecting the centers of all points into the plane of the normal vector, then the 3D coordinates of these points are: $u_i = (x_i, y_i, z_i), i = 1,2,3,4,5,6$.

（4）Assuming that the centre of spatial circle is positioned $\bar{U} = (\bar{X}, \bar{Y}, \bar{Z})$ and the normal vector is $\vec{n}$， then we can list the following equations:

$$\begin{cases} n_x(x_c - \bar{X}) + n_y(y_c - \bar{Y}) + n_z(z_c - \bar{Z}) = 0 \\ (x_i - x_c)^2 + (y_i - y_c)^2 + (z_i - z_c) = R^2, i = 1,2,3,4,5,6 \end{cases} \tag{37}$$

We can obtain the initial space circle center (x_c, y_c, z_c) by calculating the above equation.

4.2 Attitude Solution

As shown in Figure 8, the attitude angle of the space circle plane is defined, and the n is the normal vector of the circle in 3D space. In the coordinate $O_0 - X_0Y_0Z_0$, the pitch angle is defined as the angle between the normal vector n and the plane X_0O_0Y and the range is $[-\frac{\pi}{2}, \frac{\pi}{2}]$, the yaw angle is defined as the forward angle between projection vector n_p of the normal vector n in the plane X_0O_0Y and the O_0X_0 axis. The yaw angle range is $[0, 2\pi]$. When the spatial circular plane normal vector (n_x, n_y, n_z) is known, the pitch and yaw angle can be calculated according to the geometrical triangle [7][8].

$$\phi = \arctan \frac{n_z}{\sqrt{n_x^2 + n_y^2}} \tag{38}$$

$$\theta = \arctan \frac{n_y}{n_x} \tag{39}$$

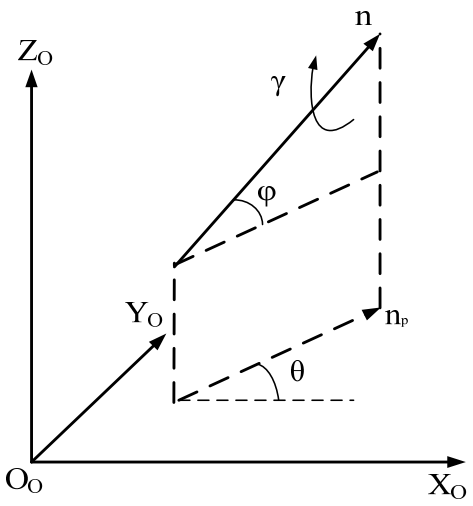


Fig .8 Definition of pose parameters of space circle

4.3 Experimentation Results

In order to verify the validity and feasibility of the proposed method, we build a close rendezvous and docking ground to test the relative position and attitude of the target spacecraft and the track spacecraft in the near field. The measuring platform is shown in Figure 9 below. Track satellite with binocular camera fixed on the target bracket. Six degrees of freedom coordinate translation stage can achieve the continuous movement of three position degrees of freedom and three attitude degrees of freedom. Longitudinal rails can be spliced to realize the approximation process of the target track.

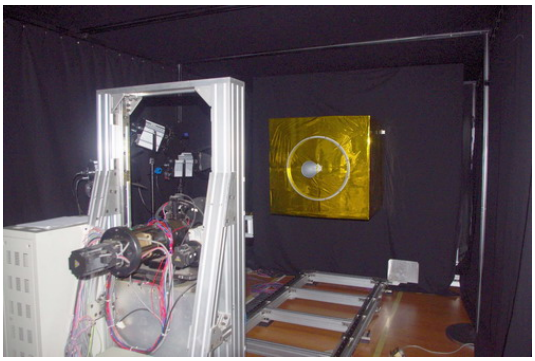


Fig .9 Rendezvous and docking ground testing system

Pose error results are shown in Figure 10 - Figure 14:

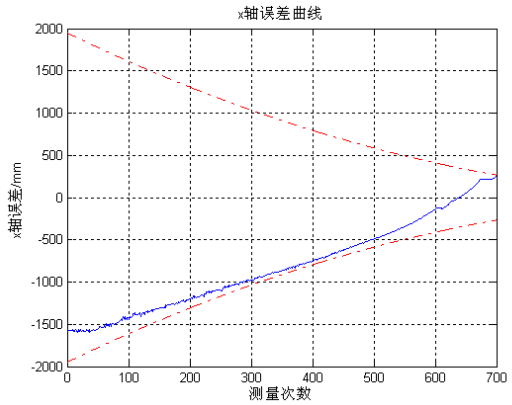


Fig .10 X axis position error

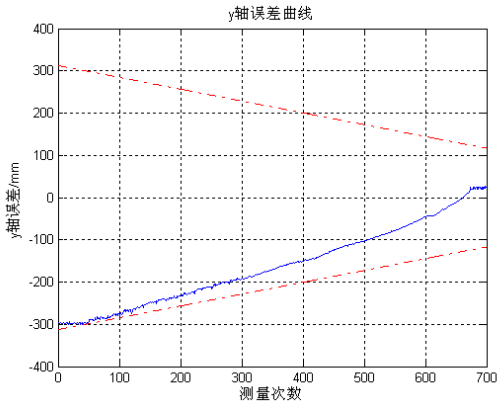


Fig .11 Y axis position error

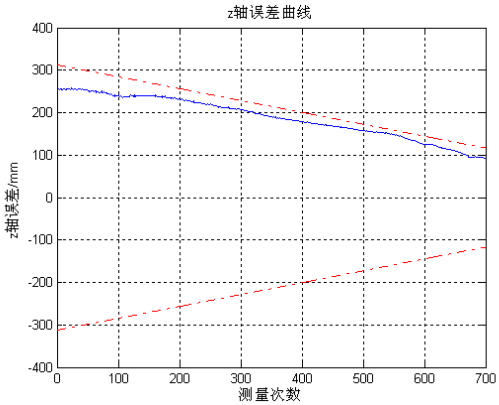


Fig .12 Z axis position error

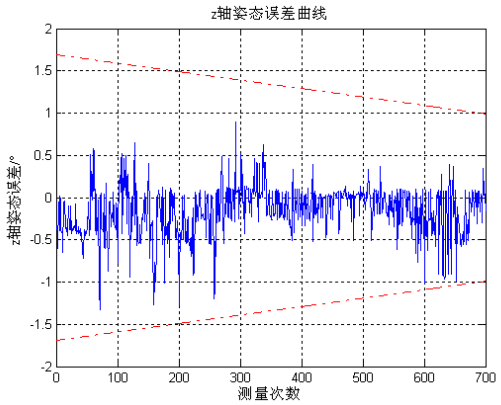


Fig .13 Yaw angle error

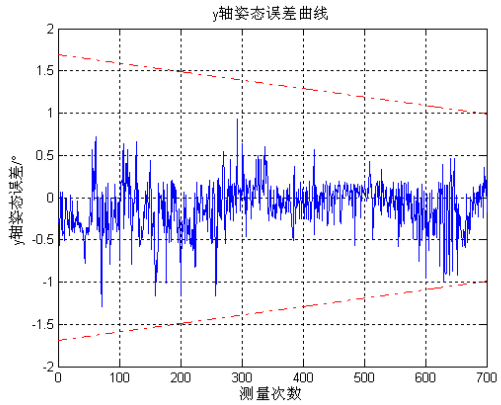


Fig .14 Pitch angle error

5 CONCLUSION

This paper presents a method binocular vision pose measurement method of spacecraft in short range based on docking ring. Using the least square method to fit the elliptic provided with the satellite docking ring and using gravity center method and singular value decomposition method to solve the solutions of the space circle center and circle plane normal vector respectively and then get the target satellite positions and pitch angle and yaw angle. Experimental results show that the algorithm have a highly stability, within the measurement range R is from 0.2 to 10m, attitude angles error is better than 0.6 + 0.1 (R-0.1) °, the position errors is better than 0.01 + 0.028 (R-0.2) m. So this algorithm can meet the requirements of super close rendezvous and docking in orbit relative pose measurement requirements effectively.

REFERENCES

[1] Zou Yimin, Wang Bo. Fragmental ellipse fitting based on least square algorithm [J].Chinese Journal of Scientific Instrument.2006, 27(7), and 808.

[2] Hu Haiou,Zhu Jianzhong. Direct least squares ellipse fitting based on arc group [J].Journal of Hangzhou Normal University (Natural Science Edition).2011,10(6),556-560.

[3] Research on relative position measurement of non-cooperative spacecraft [D].Harbin: Harbin Institute of Technology, 2012:19-23.

[4] WANG Wan-guo,WANG Shi-rong,XU Zheng-fei,YANG Wen-bo,WANG Zhen –li etc.Optimal ellipse fitting algorithm of square principle based on boundary[J].Computer Technology and Development.2013,23(4),68.

[5] ZHANG Yuan-yuan,YANG Guo-dong,WANG Feng-yan.Application of Danish robust esrimation in tunnel ellipse fitting[J].Global Geology,2012,31(1),200-201.

[6] DAI Tian-shi,CHEN Dian-you.Linear algebra[M].Beijing:Help Education Press,2004:169-193.

[7] Monocular vision pose measurement based on docking ring component [J].Acta Optica Sinica, 2013, 33(4), 0412006-1~0412006-9.

[8] Improvement of position and orientation measurement algorithm of monocular vision based on circle features [J].Journal of Hefei University of Technology, 2009, 32(11), 1659~1673.

Development of an Equivalent Equipment for Satellite Power Supply and Distribution System based on LabVIEW

Xiaolong Liu[1], Yixin Geng[2], Shuhai Chen[1]

[1] Shanghai Institute of Satellite Engineering， 251 Huaning Rd., Minghang, Shanghai, China

[2] Shanghai Institute of Aerospace Electronic Technology， 1777 Zhongchun Rd., Minghang, Shanghai, China

xiaolong9215@sina.com

ABSTRACT： Based on tests and the evaluation of power supply and distribution systems for several satellites, a compatible and configurable equivalent equipment has been designed to generally cover all needs of the power supply and distribution system ground testing, using LabVIEW to develop the applicable software. The universal equivalent equipment allows all kinds of signals as input and gives the outputs respectively according to the user-configured excitation-response mode. Real-time communication with the server was guaranteed to monitor states of a power supply and distribution system and other loads online as well as the equivalent equipment's own. The equivalent equipment underwent usability and robustness assessment with all aspects of power supply and distribution system ground testing in practical satellite development. Overall, all functional requirements were satisfied and the design showed an advantage in generality, maintenance, and applicability. Besides, areas for improvement based on feedback were identified.

1 INTRODUCTION

Power supply and distribution system[1](PSDS), as an important part of a satellite, administrates the production, storage, transformation, regulation and distribution of the electrical energy of the satellite. Its reliability is critical to the functional and operational life of a satellite. An equivalent equipment for a PSDS is used to detect the satellite's power supply, wire control, telemetry acquisition, as well as to correct pathways, control unit, actuator, solar array simulator, ground power cables, etc. In a word, it works as a testing payload to evaluate if the PSDS is reliable and it is essential.

Earlier equivalent equipments utilize as simple excitation-response models, where drive signals are directly used as the locus of control or a signal source, thus output certain preset response. Such an equivalent equipment is particularly designed for a specific power supply distribution system, and won't be used again in other situations.

As the power supply and distribution system becomes more complicated, e.g. equipped with more test units, with the original equivalent equipment design, the efficiency of the joint debugging has dropped dramatically, as well as the accuracy of fault location and diagnosis. Besides, all equivalent equipment designs are specific for one model of satellites, even specific for only one grade of a model.

To conclude above, it is quite necessary to improve the automation, intelligence and generality level of systematic joint testing by promote the design of the equivalent equipment, an important part, to reach an ultimate sizable goal of a shortening in PSDS development cycle, an impressive

enhancement in quality of PSDS systematic testing, a considerable reduce in resource wasting and a smart increase in multiplexing and reusing.

Our objective is to design an intelligent and automated equivalent equipment that can run all the measurements and process data without manual intervention at the same time. In the meanwhile, the design can be transplanted to models of PSDS easily by fitting and improved interface box and writing the software configuration.

2 METHODS

The universal PSDS equivalent equipment design is based on the concept of Auto Test and Virtual Instrument[2]. It runs into the preset testing mode[2] automatically following the configuration, receiving remote control commands from the comprehensive test server, and packing and broadcasting the monitored data out to the network[3]. Following the structure of the Virtual Instrument, we introduced the hardware design scheme, chose appropriate industrial computer and PCI function cards, and thus composed an intelligent measure and control system. The software, based on the language of LabVIEW, allows working on different excitation-response work modes and switching between each two of them, guaranteeing smooth communication with server and self-supervision. The hardware is designed to fulfill multi-channel environment and to meet various kinds of hardware signals, by transforming adapter box and modifying incentive-response configuration in format of Excel table, with the help of the application software.

2.1 System Design

At the end of the requirements analysis phase, Equivalent Equipment's functional requirements were briefed as the following: a) functional inspection of Actuator & Controller wire control and wire monitoring; b) access check and validation of Solar Simulator Array pathway; c) receive of Test LAN remote control commands; d) regular transmission of telemetry to Test LAN; e) closed-cycle self-checking of the PSDS test system; f) auto-test of power supply to Actuator & Controller, Solar Simulator Array and Ground Electric Cable.

The structure of the equivalent equipment is designed as drawn in Figure 1, from left to right, composed of Ground Electric Cable, Test LAN, Interface Box, Equivalent Payload, Conditioning Adapter, Industrial Control Computer Inspector, and somatological devices for visualization and user input. Equivalent Equipment is connected to Actuator & Controller through the pull-off plugs in Interface Box, which does the signal transformation. The central detection and controlling unit consists of an Industrial Control Computer and an Acquisition Card based standard PCI protocols. Generally speaking, the Acquisition Card accomplishes data transfer and collection and command control, while leaves computing and analysis to the Application Software running on the Industrial Control Computer. Test results are analyzed and displayed at real-time.

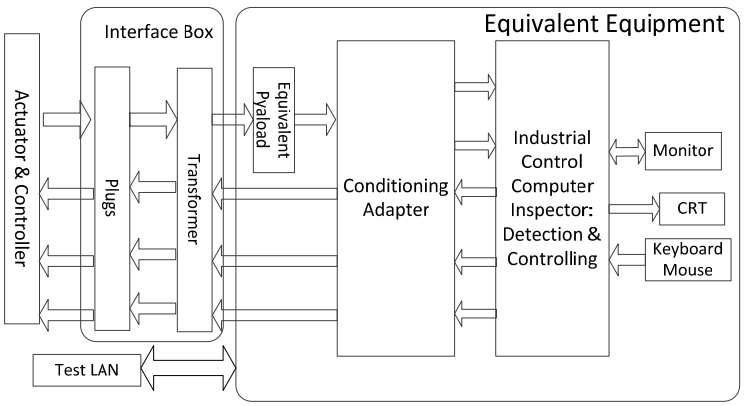


Figure 1: System Composition

2.2 Software Design

The Application Software is architected in way of an Virtual Instrument[4], on the basis of LabVIEW platform. The basic framework is demonstrated in Figure 2. Take in account of network commands, input signals from Actuator & Controller, and preset "Excitation-Response" relationship, Application Software outputs directly or controls Acquisition Card to output responses in types of electrical level, contactor, 24V-fixed analog value, tunable analog value and so forth. "Excitation-Response" relationship is defined in an Excel table and can be modified offline; by doing so, PSDS testing is generalized and extended to different models of PSDS.

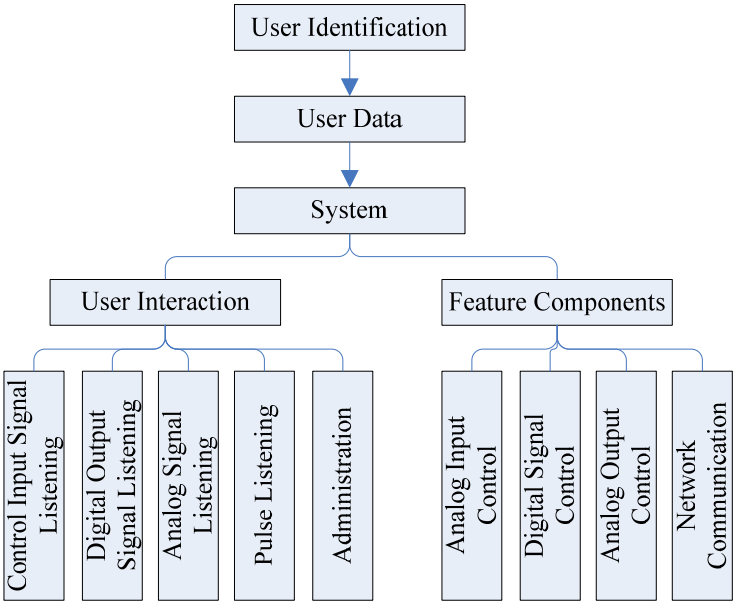


Figure 2: Software Framework

2.3 Software Modules

Application Software is divided to modules of Drivers and Functions, as shown in Figure 2. Driver modules deal with input and output of digitals, like Short-Circuit Plug signals, and analog quantities, like Solar Simulator Array voltages, Separation Plug disconnect control command pulse, etc. Functional modules, for instance, data collection module, pulse width test module, and network connection module, are describes as flows in Figure 3, 4, and 5.

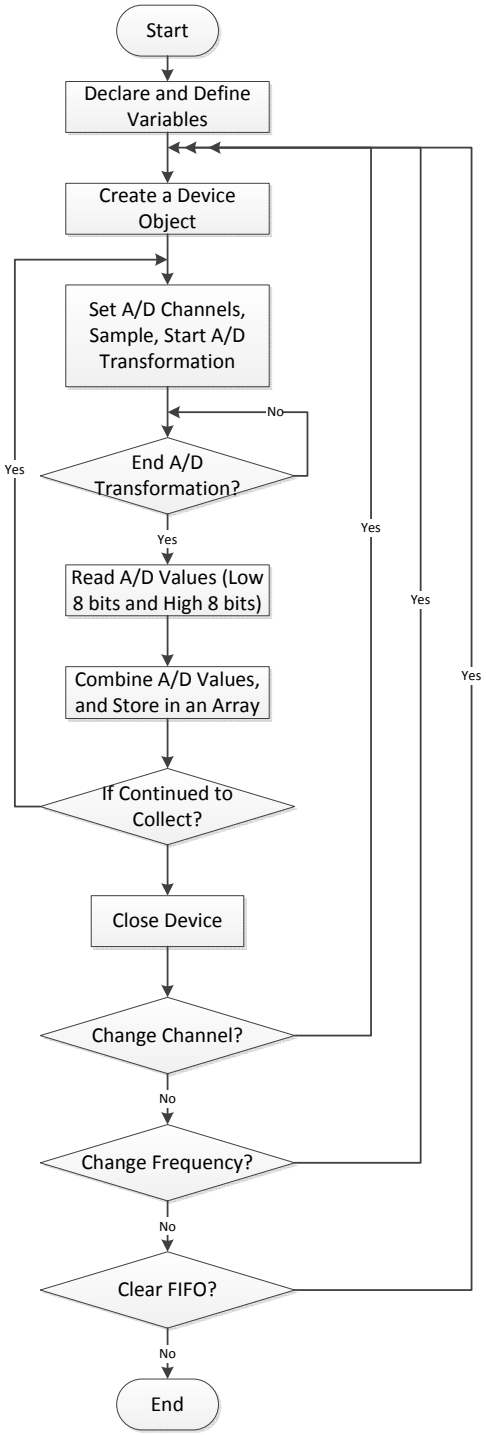


Figure 3: Process of Data Collection

Before signal sampling, some parameters, ID of the channel that is wanted the data to come from, Frequency of sampling, for example, need to be configured. A FIFO buffer is used to realize continuous data collection.

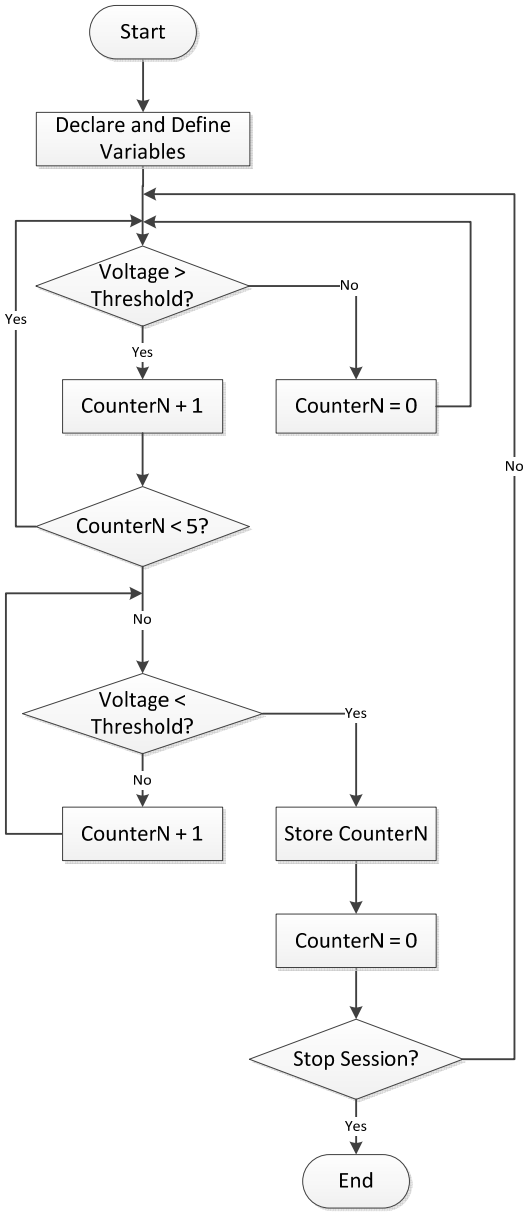


Figure 4: Process of Pulse Width Test

Pulse width is computed from every gauge point, meaning that every voltage level captured will be judged. If a voltage was greater than the maximum, make the counter increment by one; if five voltages in a row were greater than the maximum threshold, a rising edge is considered arrived, otherwise, drop the voltages; perform the same procedure to detect a falling edge. That a falling edge arrives late enough but also soon enough after a rising edge, means a complete pulse is captured, and thus the module should respond correctly according to the preset Excitation-Response relationship that can be found in the configuration table using ID of the channel as the search criteria.

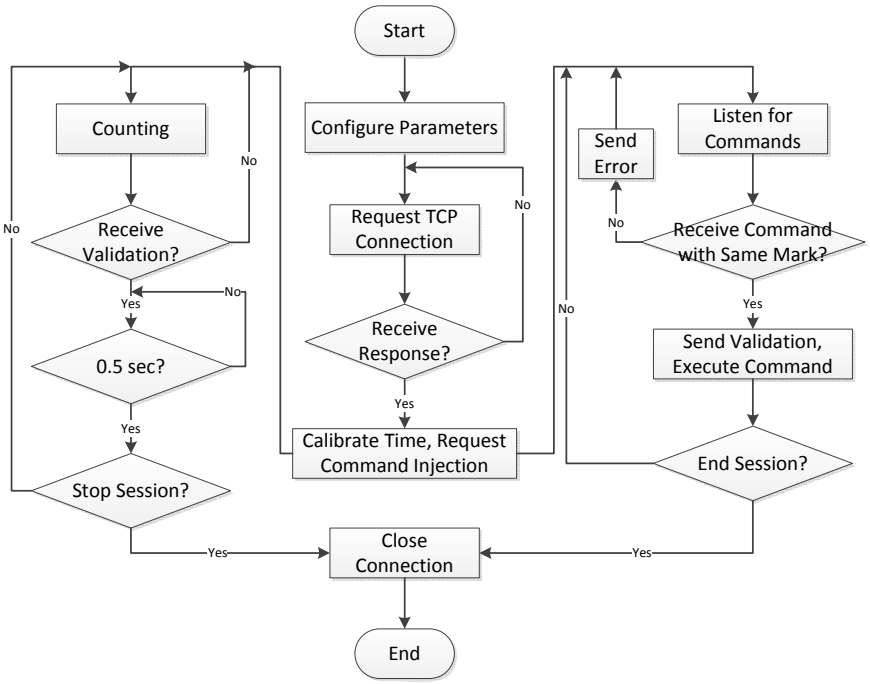


Figure 5: Process of Network Communication

Network communication module creates and closes a connection between Application Software and the server at the end of the Test LAN. The module basically accomplishes the time calibration and call corresponding modules to execute commands from the server.

2.4 Assessment of the System

We conducted an assessment on the design based on two separate aspects: function and performance. The functional assessment includes features of wire control, wire monitoring, correctness check of simulated power supply pathways, fetch of the test LAN remote control commands, transmission of regulated telemetry to the comprehensive test LAN, and closed-cycle ground test. The performance assessment includes voltage check and pulse width check through the detection of the satellite standalone control commands, monitor of the solar array simulator, detection of the separation jack-plug disconnection control, and internet communication test.

Given a thorough investigation and survey, the development of the equivalent equipment was based on the Advantech Industrial Control Computer with an Advantech IO Acquisition Card and self-designed Transforming & Conditioning Circuits.

A sample report of the assessment is attached in Chapter 3.

3 RESULTS

With our design, the equivalent equipment is reliable, highly applicable, easy for maintenance, and transplantable. The software interface is easy to configure with specific modes, and is equivalent to some PSDS payloads in a satellite: receiving power from the solar array simulator; testing the voltage of solar array simulator and separation plug; detecting pulse control instructions from control unit and actuator; achieving the level of the output, output contacts and the distribution output; receiving the remote control instructions from the test server; sending the monitoring data to the server through the test LAN. It also supports local control operations.

3.1 Sample

A sample report of assessments is presented as follows.

Functional Assessment

Actuator & Controller sends a pulse control command to Equivalent Equipment thus to form an instantaneous pathway, which is once detected, an analog feedback will be directed back to Actuator & Controller. For example, initially, a discharge switch, say A, outputs voltage at 3.6V; when a "Switch On A" command is received, A outputs 0.0V, and Switch-On Indicator Light A1 and A2 are enlightened; when a "Switch Off A1" command arrives, A outputs 1.5V, and Switch-Off Indicator Light A'1 is enlightened Switch-On Indicator Light A1 is dimed; when a message is resolved as "Switch Off A2" command, A outputs 2.7V, and Switch-Off Indicator Light A'2 is enlightened Switch-On Indicator Light A2 is dimed; "Excitation-Response" table is given below.

Table 1: Discharge Switch A's State Described as "Excitation-Response" Pairs

Input Commands	Discharge Switch A's State	Response as an Indicator Light
Initial (Power-Off)	3.6V	Dim A1, A2, A'1, and A'2
Switch-On A (Both A1 and A2)	0.0V	Enlightened A1 and A2
Switch-Off A1	1.5V	Enlightened A'1, Dim A1
Switch-Off A2	2.7V	Enlightened A'2, Dim A2

Performance Assessment

Actuator & Controller sends out separation plug disconnect control signals with a pulse width between 0.1 seconds and 1.0 seconds, with a voltage around 40.0V. Equivalent Equipment collects the signals abiding by the "Main Power Supply Detection Consistency" principle (discordant voltage division), and processes the data using LS-Fit (Least Square Fit) as shown in Table 2.

Table 2: Channel Test Data and Voltage Reduction Formula of a Separation Plug (Unit: V)

		Gauge Point 1	Gauge Point 2	Gauge Point 3	Gauge Point 4
Channel 1	Collected Value U_T	6.38	6.65	7.6	8.61
	Voltage Reduction Formula	$U_Y=5.454*U_T+0.832$			
	Computed Value U_Y	35.63	37.1	42.29	47.8
	Input Voltage	35.68	37.04	42.3	47.8

Figure 6 is the after-fit curve of Channel 1 of the same Separation Plug as in Table 2.

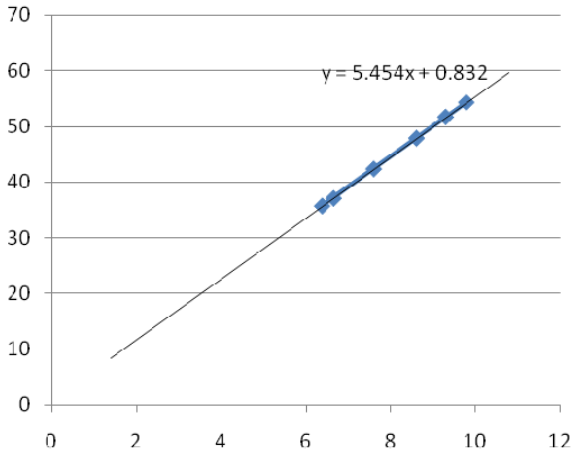


Figure 6: After-Fit Curve of Channel 1 of a Separation Plug

Result: Has met the test requirement of "Separation Plug Voltage Error in Less than 1V".

Once Acquisition Card captures a separation plug disconnect control command pulse, Equivalent Equipment Application Software starts the counter. When the counter stops, that means no pulse comes in for a test cycle, the software computes pulse width using the formula below:

$$T = N*2*\pi / f_x$$

Thereunto, T is the pulse width, N represents how many pulses are collected, f_x is the sampling frequency, which is the key parameter and in this test, is set to 1000KS/s.

The precision of the pulse width is plus-minus one padding pulse. Thus, Table 3 is constructed based on 1000KS/s test.

Table 3: Disconnect Control Command Pulse of a Separation Plug (Unit: s)

		Gauge Point 1	Gauge Point 2	Gauge Point 3	Gauge Point 4
Channel 1	Input Pulse	0.2085	0.4208	0.6883	0.9874
	Test Pulse	0.208	0.420	0.688	0.987

Result: Has met the test requirement of "Separation Plug Control Command Pulse Width Error in Less than 5ms".

4 CONCLUSION

Given that systematic and software assessments of several aspects, including functional testing, robustness confirmation, generality estimates and user-friendliness evaluation, have done in several models of satellites, we can make a conclusion that the equivalent equipment design satisfies criteria and requirements of an automated equivalent that can run measurement and process data. Beyond that, an application of such a design can help improve automation level and efficiency of the testing as well as accuracy fault location and diagnosis of a satellite power supply and distribution system.

REFERENCES

[1] Fei Gao, He-Xin Zhang, Kai Wang. Design of Missile Simulator based on ARM. Computer Measurement & Control, 2009.17:77-83.

[2] Wei Xie. Development of the Automated Test System for Astronautic Monitoring and Communication Subsystems. Huazhong University of Science and Technology, 2002:4-11.

[3] Jin-Song Tian. Development of an Environmental Online Monitoring System. Wuhan University of Technology, 2004:3-6.

[4] Jing-Ling Zhou, Wen Cai. Teaching Experiments on LabVIEW. Journal of Nantong Institute of Technology: Natural Science, 2004.3:59-62.

[5] Jin-Guang Zhao. Research and Application of Remote Measurement and Control Technology based on LabVIEW. Beijing Jiaotong University, 2008:41-47.

Parameter Influence Analysis for Doppler Velocity-Measurement Target Source on Asymmetric Spatial Heterodyne Technology

Zhang Wei[1,3], Cheng Xin[2], Du Yang[1,3]

1. Shanghai Institute of Satellite Engineering, Shanghai 200240, China;
2. Institute of Optics and Electronics Chinese Academy of Science, Chengdu 610209, China;
3. Shanghai Key Laboratory of Deep Space Exploration Technology, Shanghai 200240, China

Abstract: After introducing the principle of Doppler velocity-measurement on asymmetric spatial heterodyne technology, an analysis model is established by treating Gaussian emission line as research object and the influences of velocity-measurement precision are made from the way of wavelength bandwidth, frequency point and so on. Finally, as an example, a validation with solar actual lines is implemented by contrast for the above-mentioned ways. The research results of this paper would give some theoretical and methodological supports for the establishment of Doppler velocity-measurement target source select rules on asymmetric spatial heterodyne technology.
Keywords: asymmetric spatial heterodyne; Doppler velocity-measurement; target source; Gaussian line; influence analysis

0 Introduction

As one application branch on optics Fourier transform theory and white-light interference theory, Doppler velocity-measurement method on asymmetric spatial heterodyne technology is provided with important application value on wind field measurement, object surface shape measurement, velocity-measurement navigation for spacecraft and so on [1,2,3]. As a result of adopting heterodyne base frequency selecting, it can be acquired extremely high measurement accuracy in the case of finite sampling number[4,5,6].

The application direction of the paper to discuss is velocity-measurement navigation field for spacecraft, the state of relative motion between target source and observing point need to be inspected carefully. Because the measurement signal comes from target source, the specific influence relationship between parameter of target source and measurement outcome should be inspected, so as to clear the fundamental norm of selection for target source parameter, and to guide design and application.

1 Principle of Doppler velocity-measurement on asymmetric spatial heterodyne technology

Principle of Doppler velocity-measurement on asymmetric spatial heterodyne technology is to utilize the theorem of Fourier transform, that is to say for the reason of the target source spectral line shift in frequency domain due to relative radial movement between target source and observing point, will lead to the phase shift of the interference image in spatial domain.

The principle formula of velocity-measurement on asymmetric spatial heterodyne technology will be derived next. The essence of velocity-measurement on asymmetric spatial heterodyne technology is to utilize interference of light and Fourier transform method, record the radiation

signal of spectral line which carries the velocity information in the incidence pass-band, and accordingly carries out analysis to help the calculation for velocity information.

By inspecting a certain specified reference spectra, the principle formula of Doppler frequency shift can be written below:

$$\Delta\lambda/\lambda_0 = v_r/c \tag{1}$$

Where $\Delta\lambda$ is the wavelength shift due to Doppler effect, λ_0 is the central wavelength of reference spectra, v_r is the relative radial movement velocity, c is the light velocity in vacuum.

According to the condition of bright stripe of interference image, the bright stripe should be appeared at the position where optical path difference (OPD) equals integer multiple of the wavelength, as shown below:

$$\delta = m\lambda_0 \tag{2}$$

Where δ is the OPD corresponding the interference of a certain point in spatial domain, m is an integer, indicating the order of interference.

If a reason of relative radial movement causes wavelength shift, under the condition of invariance on reference position in spatial domain, i.e. the OPD, formula (2) can be revised as below:

$$\left(\lambda_0 + \Delta\lambda\right)\left(m + \Delta m\right) = \delta \tag{3}$$

Where Δm is the order offset due to relative radial movement. Combining formula (1), and considering the condition of relative radial movement velocity being a little approximation, we can get formula (4) as below:

$$v_r = \frac{c\lambda_0}{\delta}\Delta m \tag{4}$$

Near the close bright stripe interference order, we can get $\Delta m = \delta\varphi/2\pi$, put it into formula (4),then we can get:

$$v_r = \frac{c\delta\varphi}{2\pi\delta\sigma_0} \tag{5}$$

Where $\delta\varphi$ is the relative shifting volume of interference stripe which appeared by relative radial movements near the close bright stripe, σ_0 is the wave number of central wavelength.

Formula (5) indicates that it can be calculated the relative radial movement velocity between target source and observing point by means of design for δ and σ_0 , along with measurement for $\delta\varphi$. From the aspects of difficulty level and sensitivity about measurement, at the same time,

under the condition of σ_0 being confirmed, selection for longer δ will be more convenient to enhance the accuracy of measurements.

According to formula (5), and combining interference theory, the relationship between phase shift and OPD can be described as Fig (1) below:

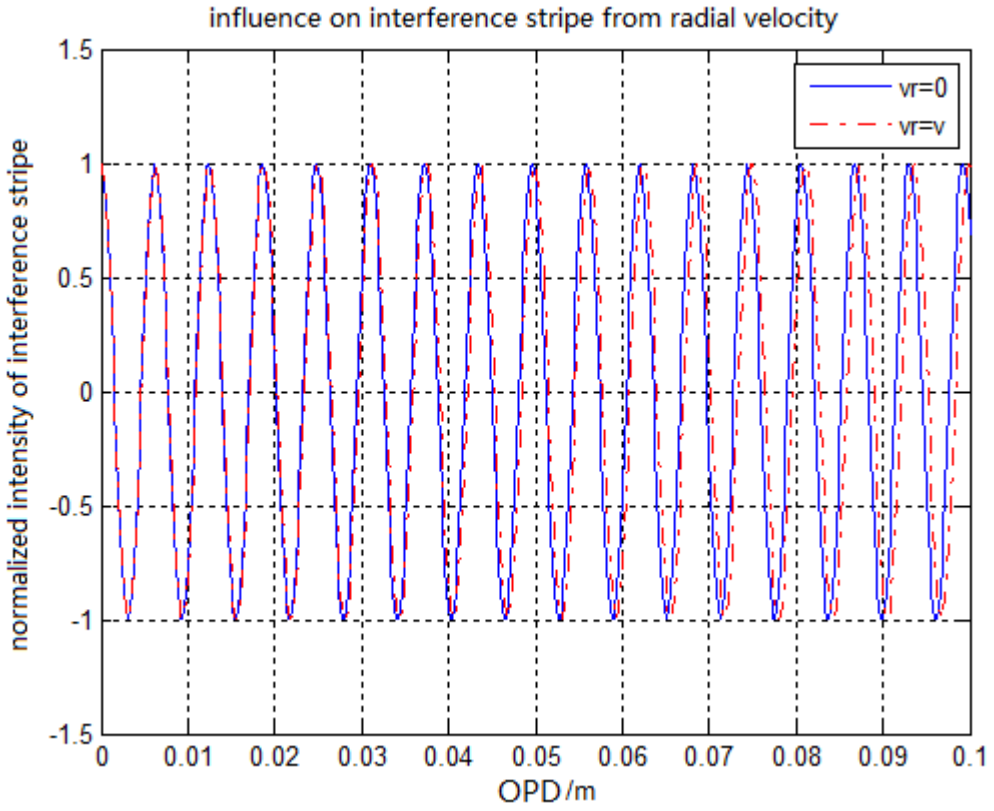


Fig. 1 Relationship between radial velocity and phase shift

From Fig (1), we can get that spectra line frequency shift caused by relative radial velocity will lead to phase shift of interference stripe, and the volume of shift will be magnified along with the OPD being magnified.

2 Influence of model from velocity-measurement target source

In last stage, the influence of model from velocity-measurement target source is not discussed. In fact, because of the influence mechanism in macroscopical and microcosmic aspects, actual spectral line of target source will be broadened, even to laser which possesses high degeneracy[7]. Spectral line broadening will directly affect the monochromaticity of interference light, aliasing of interference image will lead to degradation on degree of coherence, along with later stage analysis and identification.

Now we will establish the influence analytic model of velocity-measurement on heterodyne technology based on spectral line style of velocity-measurement target source. After survey and analysis we can get that common spectral style include Gaussian, Lorentz and Foketuo styles. Combining the research work of this paper, as study object, spectral line style of laser or star usually acted as Gaussian. We will develop on this below.

Normalized distribution function of Gaussian line can be expressed as formula (6) [2]:

$$g(\nu)=\frac{2\sqrt{\ln 2}}{\sqrt{\pi}\Delta\nu_s}\exp\left[-\left(2\sqrt{\ln 2}\frac{\nu-\nu_0}{\Delta\nu_s}\right)^2\right] \quad (6)$$

Where $g(\nu)$ is the normalized distribution function of Gaussian line, $\Delta\nu_s$ is the FWHM, and ν is the frequency of line, and ν_0 is the frequency of central line.

According to Fourier transform relationship[8] between normalized complex degree of coherence and spectral line normalized distribution function about interference theory, we can get formula (7) as below:

$$\gamma(\tau)=\int_0^{\infty} g(\nu)\exp(j2\pi\nu\tau)d\nu \quad (7)$$

Where $\gamma(\tau)$ is the normalized complex degree of coherence, τ is the transferring time difference of interference light, which corresponding relative OPD.

Put formula (6) into formula (7), we can get:

$$\gamma(\tau)=\exp\left[-\left(\frac{\pi\Delta\nu_s\tau}{2\sqrt{\ln 2}}\right)^2\right]\exp(j2\pi\nu_0\tau) \quad (8)$$

After taking modulus of formula (8), we can get the characteristics that normalized complex degree of coherence changing with time difference (OPD). For two light with same intensity, that is the contrast ratio of interference stripe, which can be shown as formula (9):

$$\left|\gamma(\tau)\right|=\exp\left[-\left(\frac{\pi\Delta\nu_s\tau}{2\sqrt{\ln 2}}\right)^2\right] \quad (9)$$

From formula (9) we can get that the contrast ratio of interference stripe will be decreased along with FWHM and OPD being magnified. In order to ensure that the contrast ratio can be surveyed during actual research, the edge of the contrast ratio of interference stripe should be set. It must be indicated that the volume of edge setting is independent of the principle, and is confirmed by actual requirement. Here we set $1/e$ as the edge, combining formula (9), so we can get the coherent length of interference stripe as below:

$$l_c=\gamma_c c=\frac{2\sqrt{\ln 2}c}{\pi\Delta\nu_s}=\frac{2\sqrt{\ln 2}\lambda_0^{\,2}}{\pi\Delta\lambda_s} \quad (10)$$

Where l_c is the coherent length of interference stripe, γ_c is the coherent time, $\Delta\lambda_s$ is the wavelength bandwidth of spectral line.

In order to acquire higher accuracy of velocity-measurement, OPD is set with coherent length. Put formula (10) into formula (5), we can get as below:

$$v_r = \frac{c\delta\varphi\Delta\lambda_s}{4\sqrt{\ln 2}\lambda_0} \quad (11)$$

By now, we have inferred the relationship among relative radial movement of target source and observing point, wavelength bandwidth of spectral line, and central wavelength. The latter two physical quantities both belong to target source parameters, and from which the influence of velocity-measurement accuracy from the selection of target source can be shown.

3 Simulation analysis

3.1 Influence analysis of wavelength bandwidth

Now the simulation analysis of velocity-measurement accuracy will be implemented by treating formula (11) as the start point. Here we set λ_0, the central wavelength, as 750nm (visible range), and simulate the relationship between Doppler relative movement radial velocity and phase shift of interference stripe under the condition of different wavelength bandwidth.

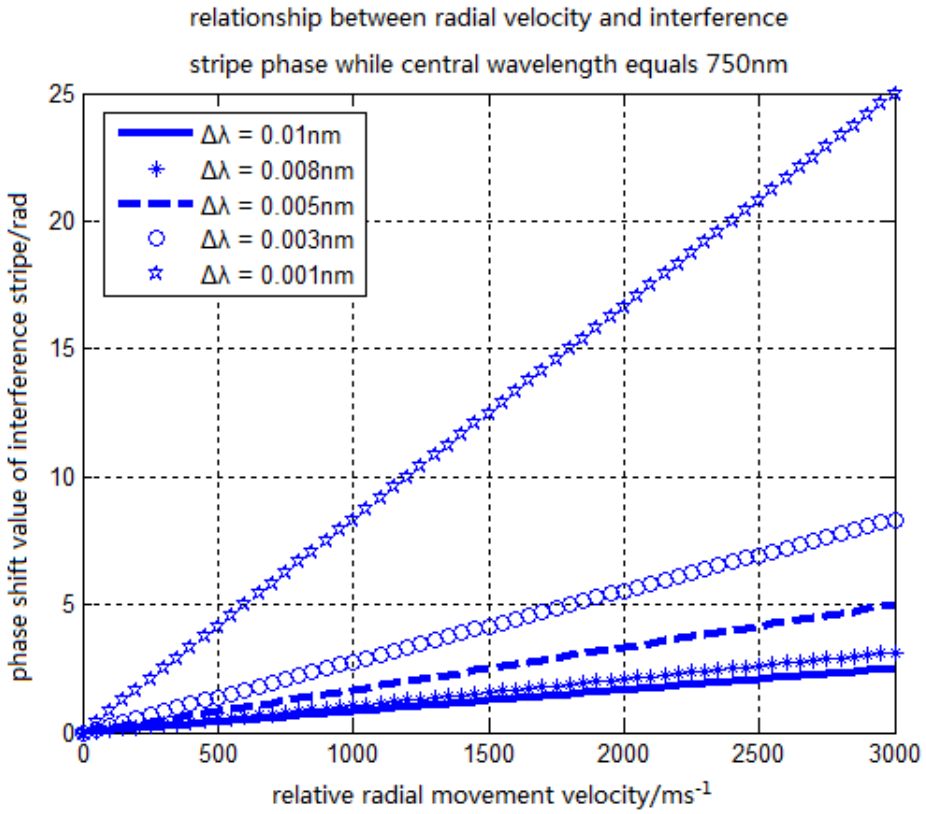


Fig. 2 Relationship between radial velocity and phase shift at 750nm for central wavelength

From Fig 2 we can get that the relationship between Doppler relative radial movement velocity and phase shift is linear when central wavelength is certain. The distribution of different wavelength bandwidth in the figure show us: the narrower wavelength bandwidth is and the larger gradient is, the higher accuracy measurement volume of phase shift can be acquired for one definite measurement accuracy.

3.2 Influence analysis of central wavelength

Here we set $\Delta\lambda_s$, the wavelength bandwidth, as 0.01nm (classic wavelength bandwidth of

star), and simulate the relationship between Doppler relative movement radial velocity and phase shift of interference stripe under the condition of different central wavelength.

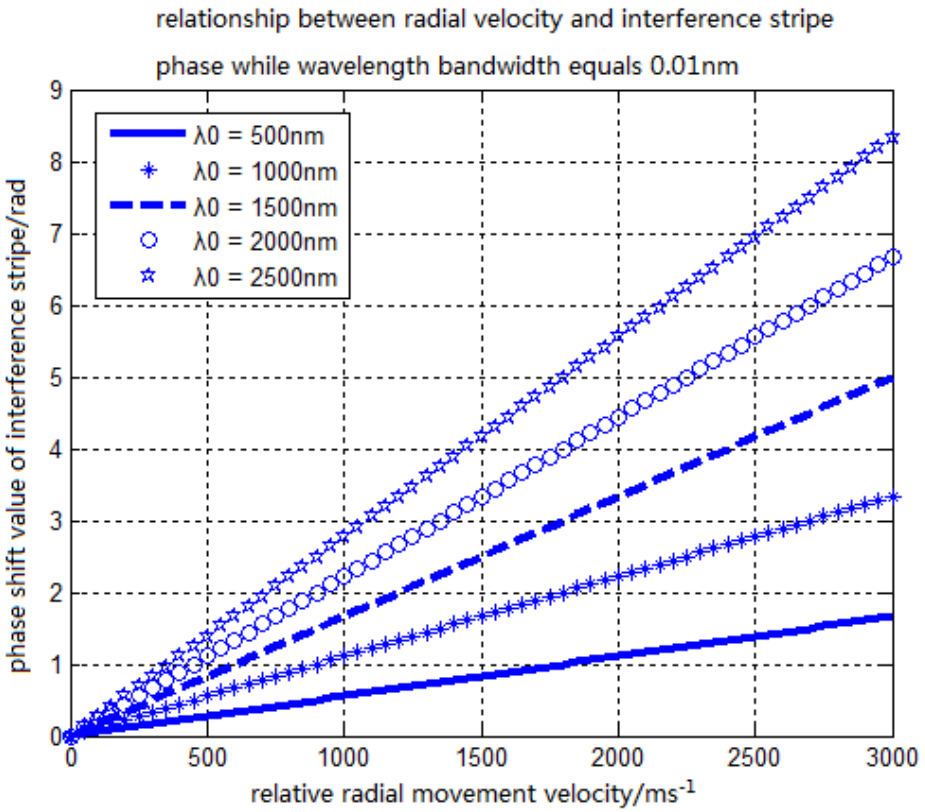


Fig. 3 Relationship between radial velocity and phase shift at 0.01nm for wavelength bandwidth

From Fig 3 we can get that the relationship between Doppler relative radial movement velocity and phase shift is still linear when wavelength bandwidth is certain. The distribution of different wavelength bandwidth in the figure show us: the longer central wavelength is and the larger gradient is, the higher accuracy measurement volume of phase shift can be acquired for one definite measurement accuracy.

4 Actual example analysis

In order to explain the influence situation of target source parameter selection furthermore, here we pick two spectral lines of actual measurement from solar ultraviolet band to compare and analyze. Parameters of this two lines can be shown as table 1 below.

Table 1 Main parameters of solar ultraviolet line (He II 304 and Fe IX 171)

Order number	Name	Central wavelength/nm	Wavelength bandwidth/nm
1	He II 304	30.3781	0.068
2	Fe IX 171	17.10729	0.071

The analysis outcome can be shown as Fig (4) below.

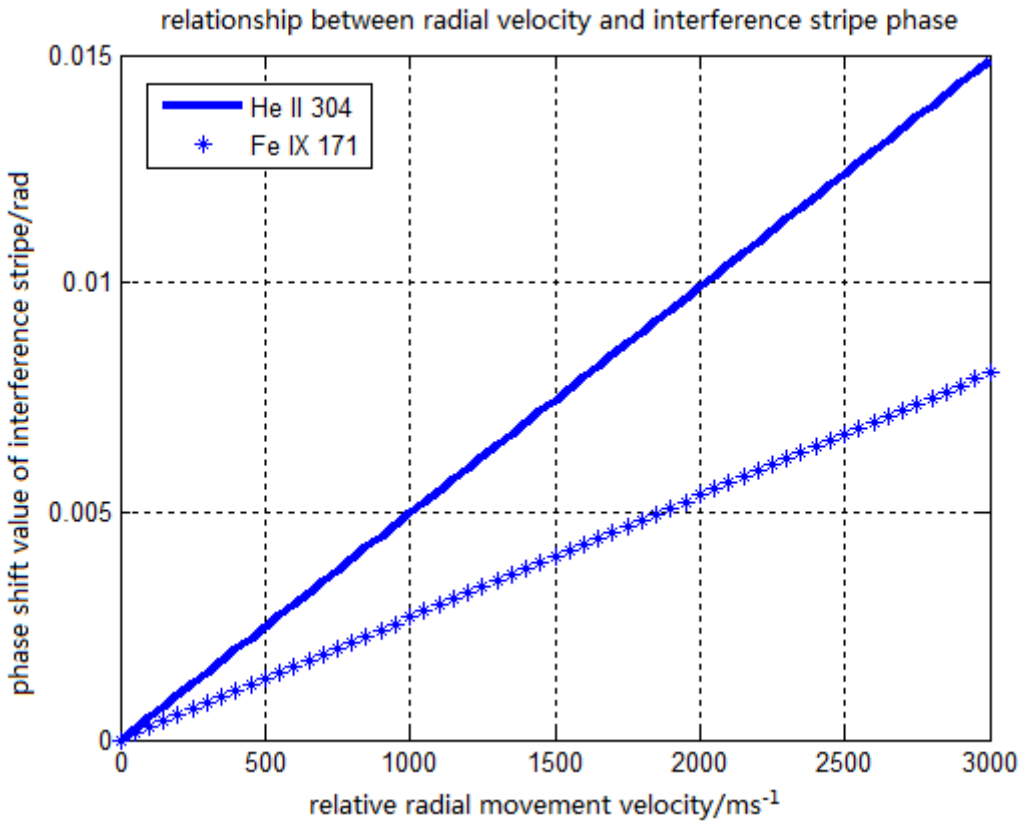


Fig. 4 Comparison on radial velocity measurement accuracy of solar ultraviolet line (He II 304 and Fe IX 171)

From Fig (4) we can get that the gradient of He II 304 is larger than that of Fe IX 171. The reason is that its central wavelength is longer and its wavelength bandwidth is narrower, so it is more suitable to be treated as observing object to implement measurement at the time of target source selecting.

5 Conclusion

It is vitally important for target source parameter selection from which Doppler velocity-measurement scheme design on asymmetric spatial heterodyne technology can be done. Under the condition of Gaussian line style and coherent length are certain, the narrower wavelength bandwidth is and the longer central wavelength is, the higher sensitivity and accuracy for measurement can be acquired. To those target sources which emit radial line mainly, it can be acted as one of the norm for selection by this.

Except the wavelength bandwidth and central wavelength mentioned in this paper, time stability and signal intensity of the target source physical parameters including these two items are subjects that required to research. At the same time, the style that spectral line object as absorbed line should be considered carefully. It will be concerned to the feasibility of actual engineering filed application, so continuously and deeply performance in future research work is necessary.

Reference

[1] Wang Li. Study on Wind Measurement of Atmosphere by Interferometry Technology [D]. Xi'an: Xi'an Institute of Optics and Precision Mechanics Chinese Academy of Science, 2007.

[2] Research on Contact and Noncontact Hybrid Measurement System of Profile Based on White

Light Interferometry [D]. Wuhan: Huazhong University of Science and Technology, 2008.
[3] Zhang Wei, Chen Xiao, You Wei, et al. New autonomous navigation method based on redshift [J], Aerospace Shanghai, 2013(2): 32-33.
[4] J.Harlander, R.J.Reynolds, and F.L.Roesler. "Spatial Heterodyne Spectroscopy for the Exploration of Diffuse Interstellar Emission Lines at Far-Ultraviolet Wavelengths," The Astrophysical Journal, 396: 730-740 (1992).
[5] Christoph R. Englert, David D. Babcock and John M. Harlander. "Doppler asymmetric spatial heterodyne spectroscopy(DASH): concept and experimental demonstration," Applied Optics, Vol. 46, No. 29, 7297-7307 (2007).
[6] J.M.Harlander, F. L. Roesler, C. R. Englert, J. G. Cardon, R. R. Conway, C. M. Brown, and J. Wimperis, "Robust monolithic ultraviolet interferometry for the SHIMMER instrument on STPSat-1," Appl. Opt. 42, 2829-2834 (2003).

Brief introduction of author: Zhang Wei (1976-), male, master, senior engineer, main direction of research is payload system, sensor for navigation in deep space exploration mission.
Address: No.251, Huaning road, Minhang district, Shanghai, China
Cell phone: (021) 34054518-6404
E-mail: zw1229@126.com

Error Analysis of the Autonomous Celestial Navigation Based on the Spectrum Velocity Measurement

Huang Qinglong, Chen Xiao, You Wei

1. Shanghai Institute of Satellite Engineering, Shanghai, 200240
2. Shanghai Key Laboratory of Deep Space Exploration Technology, Shanghai, 200240

Abstract: Research on the autonomous celestial navigation is the hotspot of the deep space exploration missions in these years. The autonomous celestial navigation based on the spectrum velocity measurement is a new navigation method that can obtain the relative velocity between the spacecraft and the stars with the spectrometer. Principle of the navigation based on the spectrum velocity measurement has been described briefly in this article. Based on the measurements of the line of sight vector of the Sun and radial velocity to the Sun and other two Stars, the observation strategy is derived as well. According to the characteristics of the navigation system, the impacts of the celestial navigation error on the navigation system are studied, such as the ephemeris errors of the Planets, the error of the dynamic model and the attitude error of the spectrometer. Then the error propagation equations of these error sources are derived. Finally, the navigation model based on the cruise phase of the Mars exploration mission is established by utilizing the UKF, which is suitable for the nonlinear system. The results show that the accuracy of positioning is better than 120 km and the error of velocity estimation can be better than 1m/s, the navigation system meets the need of the deep space exploration. The result of the navigation system is sensitive to these errors, particularly to the attitude error of the spectrometer. And once the error exceeds a certain range, the result will be easy to be divergent. Results of the study can be used as the theoretical references of the filtering algorithm and engineering design of the spectrum velocity measurement system.

Key word: Error Analysis; Spectrum Velocity Measurement; Autonomous Navigation; UKF

0 Introduction

In the deep space exploration mission implemented abroad currently, most of the deep space probes possess partial autonomous navigation ability. The autonomous navigation method has become an effective supplementary means of ground measure. There are many limitations in ground radio navigation such as real-time performance, operation cost and satellite resources, etc. On the one hand, autonomous navigation in deep space exploration can overcome these limitations, the independent survival ability of deep space probe can be improved; on the other hand, the accuracy and continuity of deep space probe navigation can be enhanced, the reliability of project success can be enhanced as well. In some special flight phase, such as closing, fly-around, landing, adhesion and increased rendezvous, precise position and velocity information relative to the target objects are required, and the autonomous navigation and control method can perform better than ground measure in deep space exploration mission.

The autonomous navigation technology depends mostly on angle measurement or ranging information to obtain the real-time estimation state of probe currently. But it is difficult to obtain the long-term real-time and continuous navigation information because of the limitation of the target object conditions. The astronomical optical information contains amount data of spectral

characteristics and frequency shift, and the velocity of the probe can be resolved by these information. If we can make full use of the space and natural resources, such as astronomical spectral information, it is advantageous to realize the high precision estimation of instantaneous velocity in deep space exploration, and the autonomous navigation accuracy of deep space exploration can be improved further [1-3].

The autonomous celestial navigation based on the spectrum velocity measurement is given on the basis of the ZHANG's literature [1] in this paper. Principle of the autonomous navigation based on the spectrum velocity measurement has been described briefly in this article. Based on the measurements of the line of sight vector of the Sun and radial velocity to the Sun and other two Stars, the observation strategy is derived as well. According to the characteristics of the navigation system, the impacts of the celestial navigation error on the navigation system are studied, such as the ephemeris errors of the Planets, the error of the dynamic model and the attitude error of the spectrometer. Then the error propagation equations of these error sources are derived. Finally, the navigation model based on the cruise phase of the Mars exploration mission is established by utilizing the UKF, which is suitable for the nonlinear system. The research in this paper is a practical guide to the development and realization of the autonomous celestial navigation based on the spectrum velocity measurement technology in the deep space exploration.

1 Principle of the autonomous navigation

Supposing the probe position relative to the stars is static when the probe is receiving a spectrum of the star, then the astronomical optical information is constant, and the wavelength of the spectrum of the star is denoted by λ_{rest}. When relative position of deep space probe to star has been changed, the wavelengths drift can be detected by the space borne spectrometer, which is denoted by $\Delta\lambda$. The wavelength drift performance on the spectrum is the movement of the spectral line [4-6]. According to the principle of Doppler, the ratio of the wavelength drift value and the wavelength value is equal to the ratio of the line of sight velocity v_r and the speed of light c, namely:

$$\frac{\Delta\lambda}{\lambda_{rest}} = \frac{v_r}{c}$$

Star spectrometer can obtain the information of the spectral frequency shift phase $\delta\varphi$, so the relative velocity of the probe and star is resolved indirectly by:

$$v_{Spe} = \frac{c \cdot \delta\varphi}{4\pi \cdot \Delta d \cdot \sigma}$$

In the equation above, Δd is the optical path difference, which is constant variable; σ is the light wave number, which is the reciprocal of wavelength of light.

The relative velocity derived from the astronomical spectra Doppler information is scalar value. Supposing the velocity of probe and stars in the heliocentric ecliptic inertial coordinate system as $\boldsymbol{v}$ and $\boldsymbol{v}_{Star}$, and the relative line of sight vector of probe and star is $\boldsymbol{l}_{Star}$. So the relative velocity value of probe and star is:

$$(\boldsymbol{v}_{Star} - \boldsymbol{v})^T \boldsymbol{l}_{Star} = v_{Spe}$$

In the equation above, $\boldsymbol{v}_{Star}$ and $\boldsymbol{l}_{Star}$ is the velocity and line of sight vector of the star in the inertial coordinate system. Because the probe position value is much smaller than the star position value, the influence of the probe position is neglected. And the value of $\boldsymbol{v}_{Star}$ and $\boldsymbol{l}_{Star}$ can be derived by the star catalogue, such as the Hipparcos and Tycho Star catalogues.

Based on the above analysis, the measurement strategy including the relative velocity measurement of the probe and the sun, and the relative velocity measurement of the probe and the other two stars can be established, which is shown in figure 1. The equation of the velocity measurement can be written as:

$$\begin{bmatrix} \boldsymbol{l}_{Star1}^T \\ \boldsymbol{l}_{Star2}^T \\ \boldsymbol{l}_{Star3}^T \end{bmatrix} \boldsymbol{v} = \begin{bmatrix} \boldsymbol{l}_{Star1}^T \boldsymbol{v}_{Star1} \\ \boldsymbol{l}_{Star2}^T \boldsymbol{v}_{Star2} \\ \boldsymbol{l}_{Star3}^T \boldsymbol{v}_{Star3} \end{bmatrix} - \begin{bmatrix} v_{Spe1} \\ v_{Spe2} \\ v_{Spe3} \end{bmatrix}$$

The equation above is linear equations about $\boldsymbol{v}$, so when the coefficient matrix $A_{l_{Star}} = [\boldsymbol{l}_{Star1} \quad \boldsymbol{l}_{Star2} \quad \boldsymbol{l}_{Star3}]$ is invertible, which means the line of sight vector of probe and the three stars is non-coplanar, the velocity of the probe in the inertial reference system can be derived directly by the measurement strategy.

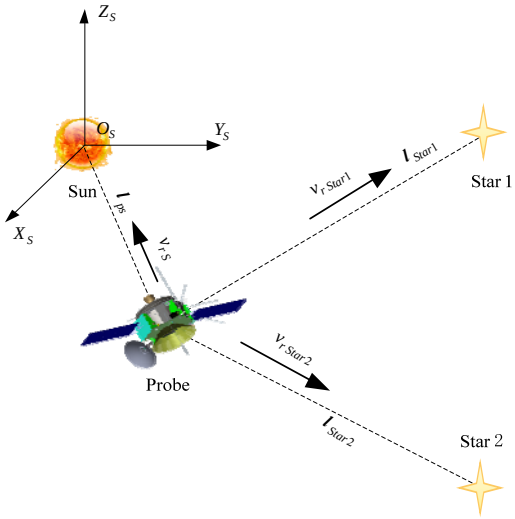


Figure1. Principle of velocity measurement navigation

Because the information of the line of sight vector of sun is included in the above scheme, and the whole velocity measurement model can be obtained by combining the measurement of the line of sight vector of sun with the velocity equation solution above.

2 Error analyses

2.1 Errors of the state model

Error of the navigation system state model means the dynamic model error, which refers to the difference between the real deep space dynamic environment and the mathematical model. The

errors mainly include the planetary perturbation acceleration error, the planetary ephemeris error, the planetary gravitational constant error, sunlight pressure perturbation acceleration error, etc.

In order to analyze the change of different planetary perturbation acceleration and its influence on navigation precision, the three body trajectory dynamics model perturbation error term $\Delta \boldsymbol{a}_P$ is introduced, namely:

$$\dot{\boldsymbol{r}} = \boldsymbol{v}$$
$$\ddot{\boldsymbol{r}} = \boldsymbol{a}_S + \boldsymbol{a}_E + \boldsymbol{a}_M + \Delta \boldsymbol{a}_P + \boldsymbol{a}_{PSR} + \boldsymbol{a}_T$$

The planetary position information in the deep space exploration orbit dynamics model can be obtained by the ephemeris. DE/LE ephemeris of JPL has been widely used in the deep space navigation, interplanetary exploration, data analysis and reduction of high precision of astronomical observation, and all kinds of the theory of gravity test as well [7-8]. In this paper, DE421 ephemeris is adopted.

In the case of earth, for example, supposing the position ephemeris error is $\Delta \boldsymbol{r}_{DEE}$, then:

$$\boldsymbol{a}_{DEE} = -\mu_E \left(\frac{\left(\boldsymbol{r}_E + \Delta \boldsymbol{r}_{DEE}\right)}{\left|\boldsymbol{r}_E + \Delta \boldsymbol{r}_{DEE}\right|^3} + \frac{\boldsymbol{r}_e - \left(\boldsymbol{r}_E + \Delta \boldsymbol{r}_{DEE}\right)}{\left|\boldsymbol{r}_e - \left(\boldsymbol{r}_E + \Delta \boldsymbol{r}_{DEE}\right)\right|^3} \right)$$

Considering $\Delta \boldsymbol{r}_{DEE}$ □ $\boldsymbol{r}_E$, the approximation of the equation above is:

$$\boldsymbol{a}_{DEE} = \boldsymbol{a}_E + \Delta \boldsymbol{a}_{DEE} \approx -\mu_E \left(\frac{\boldsymbol{r}_E}{\left|\boldsymbol{r}_E\right|^3} + \frac{\boldsymbol{r}_e - \boldsymbol{r}_E}{\left|\boldsymbol{r}_e - \boldsymbol{r}_E\right|^3} \right) - \mu_E \left(\frac{\Delta \boldsymbol{r}_{DEE}}{\left|\boldsymbol{r}_E\right|^3} - \frac{\Delta \boldsymbol{r}_{DEE}}{\left|\boldsymbol{r}_e - \boldsymbol{r}_E\right|^3} \right)$$

And the error caused by the planetary ephemeris error approximate:

$$\Delta \boldsymbol{a}_{DEE} \approx -\mu_E \left(\frac{1}{\left|\boldsymbol{r}_E\right|^3} - \frac{1}{\left|\boldsymbol{r}_e - \boldsymbol{r}_E\right|^3} \right) \cdot \Delta \boldsymbol{r}_{DEE}$$

The equation above contains two parts: the state model error changed with the planetary position and other term changed with relative position of probe and planet. The case of other planets can be resolved in the same way.

The sunlight pressure perturbation acceleration model is closely relative to satellite parameters, such as the satellite attitude control law, area-mass ratio, surface physical properties. It can be treated as a random error briefly.

$$\Delta \boldsymbol{a}_{PSR} = \eta P_{SR} A U^2 C_R \left(\frac{A_R}{m} \right) \frac{\boldsymbol{r}}{r^3} \cdot \boldsymbol{\varepsilon}_{PSR}$$

2.2 Errors of the measurement model

2.2.1 CCD sun sensor

Linear array CCD sun sensor measurement model is shown as figure 2. The reference coordinate system of the sun sensors is $O_{Sen1} X_l Y_l Z_l$, sensor aim axis coincide with the axis Z_l. Measuring axis is in the plane $X_l O_{Sen1} Y_l$ and respectively parallel to the axis X_l and the axis Y_l.

Assuming the incidence light beam is parallel to the line of sight $O_{Sen1}S$, two light lines can be derived in sensitive objects bottom by the sunlight through the optical refraction. The value of the distance deviated from the baseline can be obtained by the encoder at the bottom, respectively named as d_x and d_y, h refer to the distance between the radiation plane to the bottom of the sun sensors.

The direct measuring value by the sun sensors is $m_x = n\frac{d_x}{h}$ and $m_y = n\frac{d_y}{h}$. Considering the equation $l_x^{\ 2} + l_y^{\ 2} + l_z^{\ 2} = 1$, then:

$$l_{Sen1} = \begin{bmatrix} l_x \\ l_y \\ l_z \end{bmatrix} = \frac{1}{\sqrt{m_x^2 + m_y^2 + 1}} \begin{bmatrix} m_x \\ m_y \\ 1 \end{bmatrix}$$

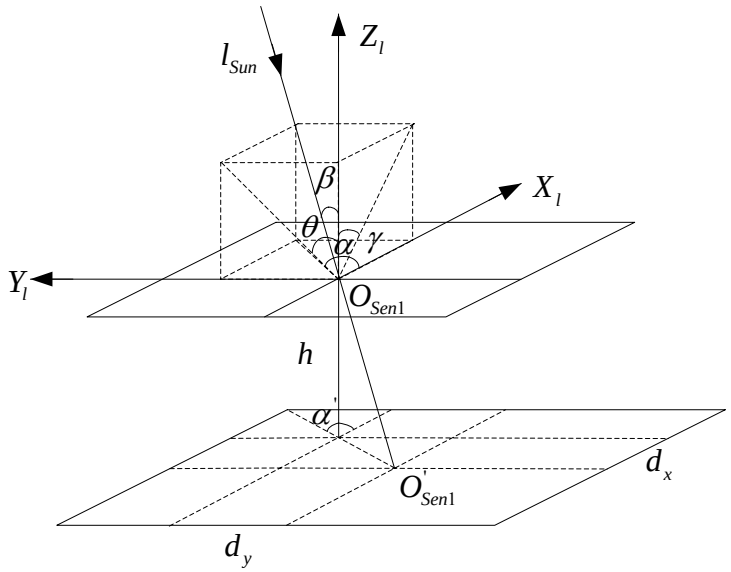


Figure2. Principle of the CCD sun sensor observation model

In the equation above, the measurement errors of sun sensors are introduced, the observation equation of the line of sight vector of Sun in the sensor fixed coordinate system is:

$$l_{Sen1}^{\varepsilon} = \frac{1}{\sqrt{m_x^2 + m_y^2 + 1}} \begin{bmatrix} m_x + \varepsilon_x \\ m_y + \varepsilon_y \\ 1 \end{bmatrix} = l_{Sen1} + \frac{1}{\sqrt{m_x^2 + m_y^2 + 1}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \end{bmatrix}$$

The error propagation matrix is:

$$A_{Sen1} = \frac{1}{\sqrt{m_x^2 + m_y^2 + 1}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

In the equation above, $v_{Sen1} = \begin{bmatrix} \varepsilon_x & \varepsilon_y \end{bmatrix}^{\mathrm{T}}$ is the observation noise vector of the line of sight vector of Sun; ε_x、ε_y are two axes measuring error, generally regarded as the gauss normal distribution.

The observation equation of the line of sight of Sun in the heliocentric ecliptic inertial coordinate system can be derived by introducing the attitude transformation matrix, namely:

$$z_1 = C_{O_S O_{Sen1}} l_{sen1}^{\varepsilon} = \left(C_{O_S O_b} \cdot C_{O_b O_{Sen1}} \right) l_{Sen1}^{\varepsilon}$$

In the equation above, $C_{O_S O_{Sen1}}$ refers to the transformation matrix from sun sensors installed coordinate system to the heliocentric ecliptic inertial coordinate system, which can be resolved by the transformation matrix $C_{O_b O_{Sen1}}$ and $C_{O_S O_b}$. And $C_{O_b O_{Sen1}}$ refers to the transformation matrix from the sun sensor installed coordinate system to the probe body coordinate system, $C_{O_S O_b}$ refers to the transformation matrix from the probe body coordinate system to the heliocentric ecliptic inertial coordinate system. The parameters of $C_{O_S O_b}$ can be obtained by the attitude control system.

In conclusion, the observation equation of the line of sight vector of sun is:

$$z_1 = h_1(X) + v_{\varepsilon 1} = C_{O_S O_{Sen1}} l_{Sen1} + C_{O_S O_{Sen1}} A_{Sen1} v_{Sen1}$$

In the equation above, $v_{\varepsilon 1} = C_{O_S O_{Sen1}} A_{Sen1} v_{Sen1}$ is the observation error of the line of sight vector of sun, which can be treated as random error.

2.2.2 The velocity measurement spectrometer

Based on the analysis above, velocity measurement equation by the spectrometer can be written as follows:

$$z_2 = h(X) + v_v = A_{l_{Star}}^{-1} \begin{bmatrix} l_{Star1}^T v_{Star1} \\ l_{Star2}^T v_{Star2} \\ l_{Star3}^T v_{Star3} \end{bmatrix} - A_{l_{Star}}^{-1} \begin{bmatrix} v_{Spe1} \\ v_{Spe2} \\ v_{Spe3} \end{bmatrix}$$

The main error factors of the velocity measurement spectrometer include star catalog ephemeris error, the identification error of the observation and the attitude error of the spectrometer, etc.

Identification error of the observation is mainly refers to the process error when the spectral data transfer into the probe velocity information. The imaging performance of navigation sensors is mainly influenced by the characteristic of the observation target star, such as the albedo, the absolute magnitude, the error of geometry，etc.

The measurement error of the spectrometer ε_{Spe} is introduced, and the equation of velocity measurement by spectrometer can be written as:

$$v_{Spe}^{\varepsilon} = \frac{c \cdot \left(\delta\varphi + \varepsilon_{Spe} \right)}{4\pi \cdot \Delta d \cdot \sigma} = v_{Spe} + \frac{c \cdot \varepsilon_{Spe}}{4\pi \cdot \Delta d \cdot \sigma}$$

So the equation of the velocity measurement model is:

$$z_2 = h(X) + v_v = A_{l_{Star}}^{-1} \begin{bmatrix} l_{Star1}^T v_{Star1} \\ l_{Star2}^T v_{Star2} \\ l_{Star3}^T v_{Star3} \end{bmatrix} - A_{l_{Star}}^{-1} \begin{bmatrix} v_{Spe1}^{\varepsilon} \\ v_{Spe2}^{\varepsilon} \\ v_{Spe3}^{\varepsilon} \end{bmatrix}$$

The error equation of the velocity measurement by spectrometer can be written as:

$$v_{\varepsilon_{Spe}} = A_{Spe_{\varphi}}^{\varepsilon} \cdot \begin{bmatrix} \varepsilon_{Spe1} \\ \varepsilon_{Spe2} \\ \varepsilon_{Spe3} \end{bmatrix}$$

In the equation above, the error transfer matrix of the velocity measurement is:

$$A_{Spe_{\varphi}}^{\varepsilon} = -\frac{c}{4\pi \cdot \Delta d \cdot \sigma} A_{l_{Star}}^{-1}$$

Based on the analysis above, the data of the position and velocity of the star is obtained by the star catalog. The Hipparcos and Tycho Star catalogues record the data of the stars by observing each star more than 100 times after three and a half years. The Hipparcos and Tycho Star catalogues contain 117982 stars information, such as star position, parallax and proper motion, etc [9-10]. The statistical specific accuracy results are shown in table 1. In this paper, the Hipparcos and Tycho Star catalogues is adopted to obtain the data of the stars.

Table1 The statistical specific accuracy results of the Hipparcos and Tycho Star catalogues

Star magnitude	Number	Position (0.001")		Parallax (0.001")	Proper motion (0.001")	
		RA	DE		RA	DE
< 6.0	5041	±0.74	±0.61	±0.80	±0.80	±0.65
6.0 ~ 6.99	10356	±0.77	±0.63	±0.84	±0.84	±0.68
7.0 ~ 7.99	25661	±0.84	±0.70	±0.96	±0.94	±0.77
8.0 ~ 8.99	41935	±1.03	±0.86	±1.17	±1.14	±0.94
9.0 ~ 9.99	24765	±1.42	±1.19	±1.53	±1.53	±1.26
> 10	10224	±2.88	±2.4	±2.59	±2.69	±2.21
> 11	2716	±7.19	±6.04	±4.98	±5.37	±4.39
Total	117982	±1.32	±1.10	±1.36	±1.36	±1.12

Assuming the position error of the stars in the star catalogue is $\left(\varepsilon_{Spe_{\alpha}}, \varepsilon_{Spe_{\beta}}\right)$, the equation of the line of sight vector of the star is:

$$l_{Star}^{\varepsilon} = C_{O_S O_{ICRS}} \begin{bmatrix} \cos\left(\alpha_{Star} + \varepsilon_{Spe_{\alpha}}\right)\cos\left(\beta_{Star} + \varepsilon_{Spe_{\beta}}\right) \\ \cos\left(\alpha_{Star} + \varepsilon_{Spe_{\alpha}}\right)\sin\left(\beta_{Star} + \varepsilon_{Spe_{\beta}}\right) \\ \sin\left(\alpha_{Star} + \varepsilon_{Spe_{\alpha}}\right) \end{bmatrix}$$

In the equation above, $C_{O_S O_{ICRS}}$ is the transformation matrix form the star catalogue reference coordinate system ICRS to the heliocentric ecliptic inertial coordinate system.

Because of the equation $\varepsilon_{Spe_{\alpha}} \ll 1, \varepsilon_{Spe_{\beta}} \ll 1$, the small second order and above error is ignored, the equation above can be rewritten as:

$$l_{Star}^{\varepsilon} = C_{O_S O_{ICRS}} l_{Star} + C_{O_S O_{ICRS}} A_{Spe_{(\alpha,\beta)}} \begin{bmatrix} \varepsilon_{Spe_\alpha} \\ \varepsilon_{Spe_\beta} \end{bmatrix}$$

In the equation above, the error transfer matrix of the line of sight vector of the star is:

$$A_{Spe_{(\alpha,\beta)}}^{\varepsilon} = \begin{bmatrix} -\sin\alpha\cos\beta & -\cos\alpha\sin\beta \\ -\sin\alpha\sin\beta & \cos\alpha\cos\beta \\ \cos\alpha & 0 \end{bmatrix}$$

And the velocity error of the star in the star catalogue ε_{Spe_v} is introduced, then:

$$v_{Star}^{\varepsilon} = v_{Star} + \varepsilon_{Spe_v}$$

3 Autonomic navigation filtering algorithm

Unscented Kalman Filter (UKF) is suitable for the nonlinear system, the unscented transformation (UT) is the core of the UKF filtering algorithm [9]. UT method select a set of sample points nearby $\hat{\boldsymbol{x}}(k \mid k)$, the mean and the covariance of these sample points respectively are $\hat{\boldsymbol{x}}(k \mid k)$ and $\boldsymbol{P}(k \mid k)$. Assuming the state variables is $n \times 1$, then the $2n+1$ sample points and their weights are:

$$\boldsymbol{\chi}_{0,k} = \hat{\boldsymbol{x}}_k, \quad \boldsymbol{w}_0 = \lambda / (n+\lambda)$$

$$\boldsymbol{\chi}_{i,k} = \hat{\boldsymbol{x}}_k + \sqrt{n+\lambda}\left(\sqrt{\boldsymbol{P}(k \mid k)}\right)_i, \quad \boldsymbol{w}_i = 1/\left[2(n+\lambda)\right] \quad i = 1, 2, ..., n$$

$$\boldsymbol{\chi}_{i+n,k} = \hat{\boldsymbol{x}}_k - \sqrt{n+\lambda}\left(\sqrt{\boldsymbol{P}(k \mid k)}\right)_i, \quad \boldsymbol{w}_{i+n} = 1/\left[2(n+\lambda)\right] \quad i = n+1, n+2, ..., 2n$$

In the equation above, $\lambda \in R$.

The standard UKF algorithm is as follows:

(1) Initialization:

$$\hat{\boldsymbol{x}}_0 = E[\boldsymbol{x}_0], \quad \boldsymbol{P}_0 = E[(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)^T]$$

(2) Computation sample points:

$$\boldsymbol{x}_{k-1} = \left[\hat{\boldsymbol{x}}_{k-1} \quad \hat{\boldsymbol{x}}_{k-1} + \sqrt{(n+\tau)P_{k-1}} \quad \hat{\boldsymbol{x}}_{k-1} - \sqrt{(n+\tau)P_{k-1}}\right]$$

(3) Time to update:

$$\boldsymbol{\chi}_{k|k-1} = f\left(\boldsymbol{\chi}_{k-1}, k-1\right)$$

$$\hat{\boldsymbol{x}}_k^- = \sum_{i=0}^{2n} w_i \boldsymbol{\chi}_{i,k|k-1}$$

$$\boldsymbol{P}_k^- = \sum_{i=0}^{2n} w_i \left[\boldsymbol{\chi}_{i,k|k-1} - \hat{\boldsymbol{x}}_k^-\right]\left[\boldsymbol{\chi}_{i,k|k-1} - \hat{\boldsymbol{x}}_k^-\right]^{\mathrm{T}} + \boldsymbol{Q}$$

$$\boldsymbol{Z}_{k|k-1} = h\left(\boldsymbol{\chi}_{k|k-1}, k\right)$$

$$\hat{\boldsymbol{z}}_k^- = \sum_{i=0}^{2n} w_i \boldsymbol{Z}_{i,\,k|k-1}$$

(4) Measurement update:

$$\boldsymbol{P}_{\hat{z}_k \hat{z}_k} = \sum_{i=0}^{2n} w_i \left[\boldsymbol{Z}_{i,k|k-1} - \hat{\boldsymbol{z}}_k^-\right]\left[\boldsymbol{Z}_{i,k|k-1} - \hat{\boldsymbol{z}}_k^-\right]^{\mathrm{T}} + \boldsymbol{R}$$

$$\boldsymbol{P}_{\hat{x}_k \hat{z}_k} = \sum_{i=0}^{2n} w_i \left[\boldsymbol{\chi}_{i,k|k-1} - \hat{\boldsymbol{x}}_k^-\right]\left[\boldsymbol{Z}_{i,k|k-1} - \hat{\boldsymbol{z}}_k^-\right]^{\mathrm{T}}$$

$$\boldsymbol{K}_k = \boldsymbol{P}_{\hat{x}_k \hat{z}_k} \boldsymbol{P}_{\hat{z}_k \hat{z}_k}^{-1}$$

$$\bar{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_k^- + \boldsymbol{K}_k\left(\boldsymbol{Z}_k - \hat{\boldsymbol{z}}_k^-\right)$$

$$\boldsymbol{P}_k = \boldsymbol{P}_k^- - \boldsymbol{K}_k \boldsymbol{P}_{\hat{z}_k \hat{z}_k} \boldsymbol{K}_k^T$$

In the equation above, $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$ is respectively refer to the system and measurement noise covariance matrix.

4 Simulation analyses

4.1 Simulation parameters

Table2 The basic orbit parameters of the Mars probe

Initial state	r=[60480784 216398917 6349369] km v=[-20.2006 10.0324 -0.4970] km/s
Initial error	Δ r=[1000 1000 1000] km Δ v=[100 100 100] m/s
Start time	2021-1-15 05:46:07 UTC
End time	2021-1-25 18:40:00 UTC

Table3 The parameters of the UKF

Step time	600 s	UKF parameters	$P = diag\left(\left[10^6, 10^6, 10^6, 10^{-2}, 10^{-2}, 10^{-2}\right]\right)$ $Q = 10^{-20}\boldsymbol{I}_{6\times 6}$, $R = 0.235\times 10^{-10}\boldsymbol{I}_{6\times 6}$ $\alpha = 1$, $\beta = 0$, $\kappa = -3$
Angle measurement precision	1 arc sec		
Velocity measurement precision	1 m/s		

The state parameters of real orbit produced by STK, and the basic orbit parameters of the Mars probe are shown in Table2. The parameters of UKF algorithm are shown in Table3. The simulation dynamical model consider the influences of the gravitational perturbations of the Sun, Mars and earth, and the sunlight pressure perturbation. The planetary information is derived from DE421. The

Hipparcos and Tycho Star catalogues are adopted to obtain the data of the stars, and the catalog number of the two stars in the Hipparcos and Tycho Star catalogues are 45348 and 172167.

4.2 Simulation results

The navigation simulation results are shown as Figure3 and Figure4. The simulation results show that the accuracy of positioning is 118.3341 km, it is better than 120 km (3σ); and the error of velocity estimation is 0.2690 m/s (3σ), it is better than 1m/s. In conclusion, the autonomous celestial navigation based on the spectrum velocity measurement meets the need of the deep space exploration.

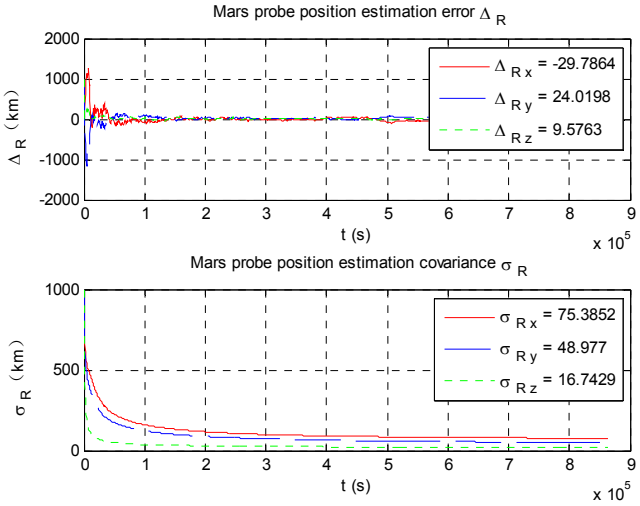


Figure3. Mars probe position estimation result

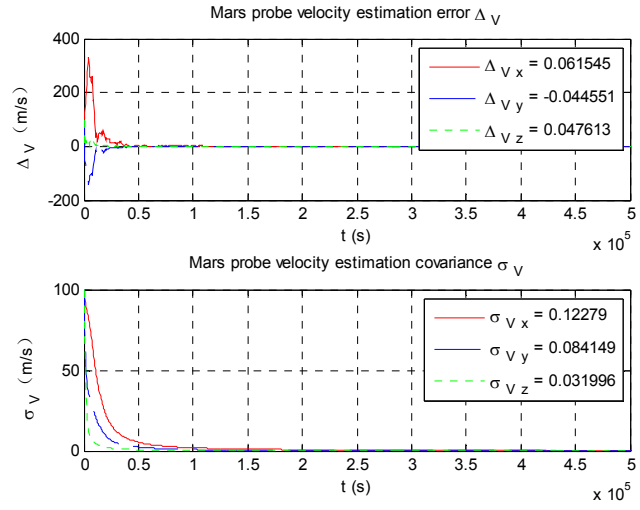


Figure4. Mars probe velocity estimation result

In order to analyze the influence of different error to navigation results, the situation above is regarded as case1; the angle measurement precision is increased to 5 arc sec, which is regarded as case 2; and the velocity measurement precision increased to 5 m/s, which is regarded as case3. The Simulation results of the three cases are shown as below.

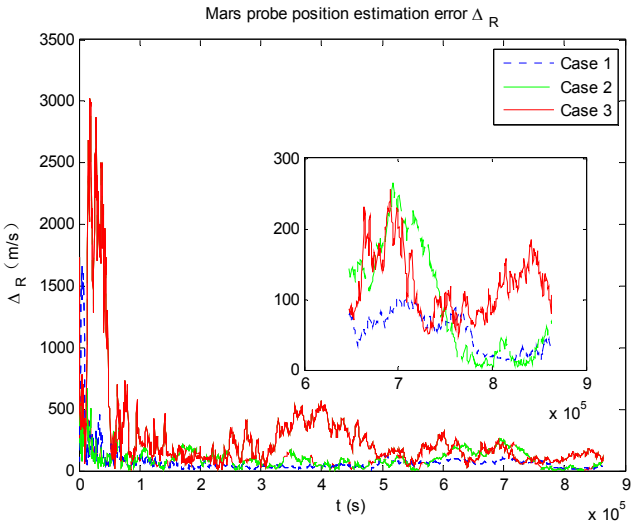


Figure5. Mars probe velocity estimation result

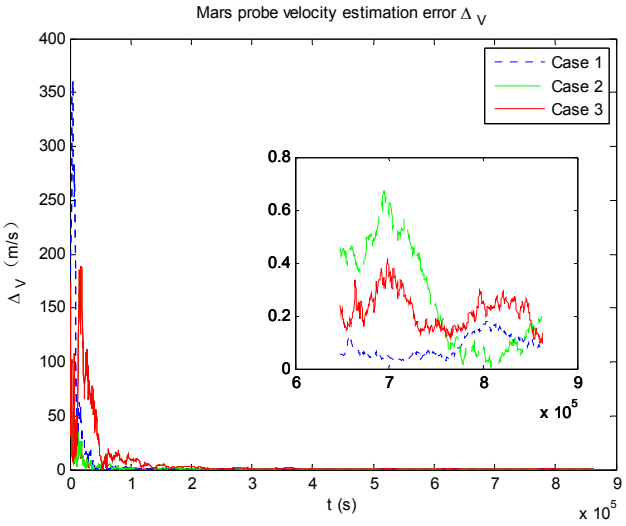


Figure6. Mars probe velocity estimation result

The navigation simulation results are shown as Figure5 and Figure6. The simulation results show that the navigation system is sensitive to these errors, particularly to the attitude error of the spectrometer. The result of case3 is worst in the three cases. And once the error exceeds a certain range, the result will be easy to be divergent.

5 Conclusions

The autonomous celestial navigation based on the spectrum velocity measurement method is derived, and the principle of the spectrum velocity measurement has been described briefly in this article. Analysis of the error propagation is respectively studied as the error of the state model and the error of measuring model in the navigation system. Then the error propagation equations of these error sources are derived. Finally, the navigation model based on the cruise phase of the Mars exploration mission is established by utilizing the UKF. The simulation results show that the accuracy of positioning is better than 120 km and the error of velocity estimation can be better than 1m/s, the navigation system meets the need of the deep space exploration. The result of the navigation system is sensitive to these errors, particularly to the attitude error of the spectrometer. And once the error exceeds a certain range, the result will be easy to be divergent. Results of the

study can be used as the theoretical references of the filtering algorithm and engineering design of the spectrum velocity measurement system.

Reference

[1] W. Wang, B. D. Fang, W. Zhang, Deceleration options for a robotic interstellar spacecraft[C]. 64th International Astronautical Congress. Beijing, 2013.

[2] X. Chen, W. Zhang, W. Wang. Preliminary Research of Mars Local Navigation Constellation[C]. 64th International Astronautical Congress. Beijing, 2013.

[3] Lightsey E G, Mogensen A, Burkhart P D, et al. Real-Time Navigation for Mars Missions Using the Mars Network [J]. Journal of Spacecraft and Rockets, 2008, 45(3):519-533.

[4] J.Harlander, R.J.Reynolds, and F.L.Roesler. Spatial Heterodyne Spectroscopy for the Exploration of Diffuse Interstellar Emission Lines at Far-Ultraviolet Wavelengths[J], The Astrophysical Journal, 1992: 730-740.

[5] Christoph R. Englert, David D. Babcock and John M. Harlander. Doppler asymmetric spatial heterodyne spectroscopy(DASH): concept and experimental demonstration[J], Applied Optics, 2007, Vol. 46, No. 29, 7297-7307.

[6] Deng X M. RAA, 2012, 12:703.

Brief introduction of author: Huang Qinglong (1989-), male, postgraduate studying in Shanghai Institute of Satellite Engineering, main direction of research is celestial navigation.

Address: No.251, Huaning road, Minhang district, Shanghai, China

Cell phone: (021) 34054518-6406

All-electric Propulsion Satellite Leading Revolution of Telecommunication Satellite Development

HU Zhao ZHOU ZhiCheng WANG Min

China Academy of Space Technology, Beijing 100081, China,
huzhao1984@aliyun.com

Abstract: The technical characteristics, market positioning and development strategies of main international all-electric propulsion satellites are introduced and analyzed, also including the technological advantages，market opportunities and the impact on communication satellites development of communication satellite industry.

1 Intruduction

All-electric propulsion satellites accomplish orbital raising, station keeping and attitude control with high specific impulse electric propulsion systems tasks [1].The biggest advantage of All-electric propulsion satellite platform is that can significantly reduce the amount of propellant carried, reduce launch weight by half[2], break the limit of chemical propulsion system, largely enhance carrying capacity of the platform and extends service life to more than 20 years. It also makes the price of telecommunication satellite more competitive.

In March 2012, Boeing 702SP pioneered all-electric propulsion satellite platform developing. In March 2015, the first two satellites (ABS-3A, Eutelsat-115WestB) were dual-launched by Falcon-9 and successfully entered the preset orbit, which have been delivered to the users in September 2015.Following Boeing, the main international satellite manufacturers joined in. Until now, there have been 15 all-electric propulsion satellites ordered.

The development momentum of All-electric propulsion satellites is very ferocious. The revolution of commercial telecommunication satellites market is coming.

2 All-electric propulsion satellite platform development process

By analyzing the velocity increment demand of GEO satellite flight missions, the Maximum demand is in orbit raising, following by the north-south station-keeping. Therefore, the high specific impulse electric propulsion systems for orbital transfer can significantly reduce the carrying amount of satellite propellant, reduce the satellite launch weight. It enhances the payload carrying capacity of the satellite platform.

The application of electric propulsion system on the satellite has gone through three major stages: the first stage, used in satellite orbital position keeping; the second stage, partly used in satellite orbit transfer, orbital position keeping and momentum wheel unloading; the third stage, used in satellite whole orbit transfer, orbital position keeping and momentum wheel unloading, etc. [3]

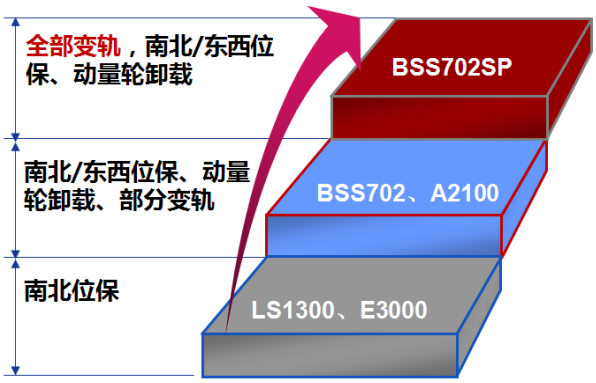


Fig. 1 Developing course of electric propulsion system application

1) The first stage

Since 1995 Russia equipped with Hall electric propulsion on the MSS-2500-GSO satellite platform, the international main telecommunications satellite platform are equipped with electric propulsion systems for orbit station keeping, such as the Boeing 601 satellite platform and 702SP platform, Lockheed Martin's A2100 platform, Loral Spaces LS3000 platform, Alcatel's Spacebus-4000 platform, Airbus' EUROSTAR-3000 platform, ESA's "Artemis" Artemis satellite, Russia's Express Communication satellites[4]. Among them, the Russian Express satellites install a set of electric propulsion system including eight SPT-100 Hall electric thrusters, and a set of anhydrous hydrazine monopropellant propulsion system composed of 12 thrusters for attitude control.

2) The second stage

The application of electric propulsion system is further expanded by applied in orbit transfer task. But limited by small thrust of electric thrusters, the application brings substantial increase the time in satellite orbit transfer. Therefore, only part of the task is completed by electric propulsion system in this stage, most of the task still need to rely on chemical propulsion system[5]. For large GEO satellite launch weight exceeding 5000kg, raising to intermediate orbit by chemical thruster, and then accomplishing the rest of orbit transfer by electric propulsion, which can effectively reduce the amount of propellant carried, thereby enhance the carrying capacity of the payload [6] . For example, a satellite based on AEHF satellite A2100M platform, using the electric propulsion system to part of the orbit transfer and position keeping, the orbit transfer time increased to 100 days. Although the orbit time increases, the benefits of increased payload carrying capacity is 700 kg.

3) The third stage

With efficient power technology and long-life high-thrust multi-mode electric propulsion technology development, electric propulsion systems can be used to perform the entire process of orbit transfer, orbital position keeping and momentum wheel unloading, thus achieving all-electric propulsion platform. All-electric propulsion platform can greatly improve payload carrying capacity, or reduce the satellite launch weight. For GEO telecommunication satellites, the launch weight can reduce by half , but increase the payload carrying efficiency from about 10% to about 25%. [7]

3 The advantages of all-electric propulsion platform

3.1 Substantial increase in the efficiency of the satellite carrying

Taking a medium-capacity GEO communication satellite as an example, assume payload weight of 500 kg, power 8 kW, the satellite life of 15 years, respectively, FIG.2 contrast full

chemical propulsion, hybrid propulsion and all-electric propulsion 3 kinds of satellite by launch weight and carrying efficiency (Payload weight divided by the weight of the satellite launch). In the case of identical weight and payload in-orbit life, the deeper the electric propulsion system applies, the less weight and higher carrying efficiency the satellite can get [8].

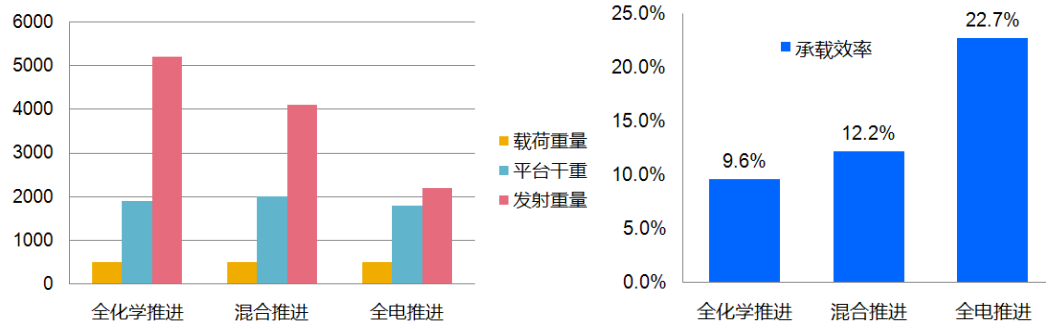


Fig. 2 Effect of different propulsion modes on launch weight and carrying efficiency

3.2 Reduce overall satellite costs, improve market competitiveness

The payload carrying efficiency of satellite platform is improved by all-electric propulsion, which means the development costs of the satellite is cut back. On the other hand, due to the small launch weight of all-electric propulsion satellite, it makes dual-launch for telecommunication satellites a reality, which can reduce the cost of launch vehicles.

Take ABS-3A satellite based on 702SP platform for example, dual-launched by "Falcon -9" , the cost $ 65 million is less than " Ariane-5" rocket (about 100 million US dollars) and the "Proton "rocket (about 80 million US dollars). ABS-3A satellite loaded with 51 C-band, Ku-band transponders, a single satellite cost about $ 100 million, the cost of each rocket $ 32.5 million. Considering the development cost and launch costs, etc.),each transponder price is only $3 million, far below the current $ 5 million , the average price per transponder on the international market. Therefore, the competitive advantage is significant.

3.3 Breakthrough the limit of chemical propulsion and increase satellites service life in orbit

For example, a all-electric propulsion satellite platform dry weight is 1300 kg, the average annual consumption of xenon for orbital position keeping and attitude control to maintain is 7 kg. A satellite with chemical propulsion propellant consumption per year is about 55kg correspondingly. In addition, all-electric propulsion satellites spend less propellant to push it away from preset orbit at the end of life.

According to the analysis, all-electric propulsion system can apparently extend the service life of the satellite in orbit, even extends service life to more than 20 years. Because of the expansion of service life, the economic benefits for satellites operators increase significantly.

For all-electric propulsion satellites, the operators would worry about the long orbit transfer period which delays in the delivery to users. One solution is to optimize the platform to develop processes, compress satellite development cycle as a remedy; the second is to take full advantage of carrying excess capacity of launch vehicles and high specific impulse electric propulsion to increase satellite service life.

4 Leading all-electric propulsion satellite communications satellite industry changes

With a new force suddenly rises of all-electric propulsion satellite industry chain recently, all-electric propulsion satellite has begun to change the market structure of telecommunication satellites.

First, the world's leading satellite manufacturers and operators have joined the all-electric satellite camp. In addition to Boeing, American Orbital Sciences Corporation, Loral, Lockheed Martin, Airbus, Germany OHB company, Russian Satellite Communications Company have been carrying out all-electric propulsion research; For satellite operators, only two medium-sized satellite operators ABS and Mexico satellite company dare to try initially. To present, neither is the top operator SES and Eutelsat, or new operators India Anne Yala operators joined the market competition quickly.

Secondly, a new satellite development model spawned by all-electric satellite is changing the traditional relationship between the satellite manufacturers, operators, satellite parts manufacturers. October 2013, ESA and OHB German company and the world's second-largest satellite Luxembourg operator SES company signed an agreement to jointly develop all-electric propulsion satellite platform -Electra platform, the company's first SES satellites planned for launch in 2018. In addition, most of all-electric propulsion satellite platform make global tender for their components. It not only guarantees the quality and development progress of products but also further reduce the cost of satellite development by choosing the best satellite components provider in the world wide.

Thirdly, all-electric propulsion technology significantly reduce the satellite launch weight,and demand for low-cost satellite services will profoundly affect the market of low-cost commercial satellite launch vehicles. Cutting the launch weight of current mainstream telecommunication satellites from 5~ 7 tons down to 2~ 4 tons, which makes medium-weight satellites be able to carry the same amount of payload to large satellites. In some way, all-electric propulsion technology gives medium-weight satellites a new life. The continuous application of 2 to 4 tons communications satellites in geostationary orbit will affect commercial high orbit launch vehicle capability and market positioning. Satellites of about 2 tons can choose low-cost Falcon 9 rocket to dual-launch, and satellites of about 4 tons can choose Falcon 9 to launch alone or select other large launch vehicles to launch as load of another large satellite like Ariane 5 rocket.

Satellite operators, satellite manufacturers, satellite component suppliers, launch service providers are faced with the challenges of all-electric push satellite development, it is also faced with the development opportunities, challenges and opportunities.

5 All-electric propulsion satellite platforms market opportunities

In June 2013, the United States space transportation management associations and the Advisory Committee on Commercial Space Transportation jointly issued the "2013 Commercial Space Transportation Forecast Report". The statistics of the international telecommunications satellite market was showed, and the number of telecommunications satellites to be launched in the coming years was predicted. According to statistics, from 2015 to 2022, the small-capacity and medium-capacity telecommunication satellites weighing less than 4200kg would occupy 41% international market share. Obviously, low-cost medium-sized communications satellite will occupy an important position in the international market.

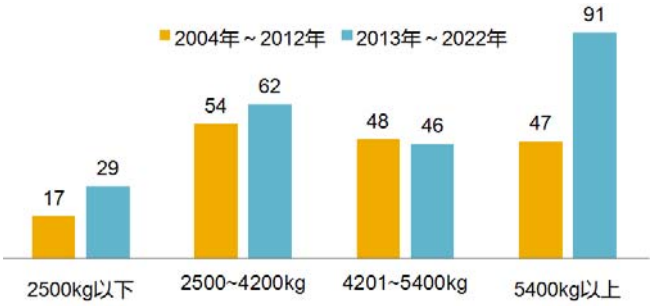


Fig. 3 Market demand of telecommunication satellites in the future

According to statistics, by 2026 there will be 310 satellites are expected to retire in the world, the average annual number is 21 (Figure 4). For china, a lot of medium-capacity telecommunication satellites launched in early 2000 are also expected to reach the service life.

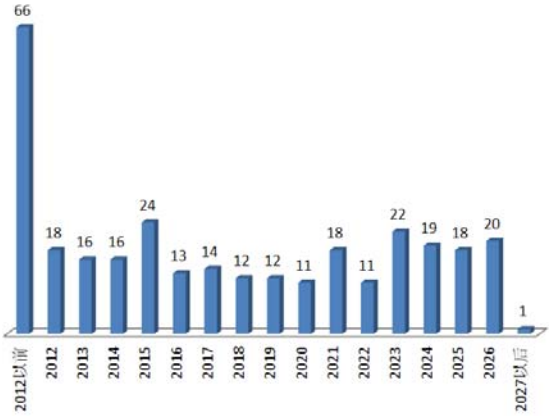


Fig. 4 Sum of civil and commercial telecommunication satellites in orbit next retired

From the development of future communications satellite market, small-capacity and medium-capacity telecommunication satellites will gradually replace large satellite and become the dominant of the market. In 2011, the global primary communications satellite operators had accomplished replacement for their large and very large communications satellites. After that, the global demand for large, very large communications satellite has stabilized, and there is a decreasing trend year by year [9]. At the same time, medium and small capacity telecommunications satellites with its lower cost, short development cycle and high flexibility become even more popular. All-electric propulsion satellite is very fit future market demand, both for emerging market development and supplement for retirement satellite launch market.

Airbus satellite department CEO Beranger in September 3, 2014 of "Space News" pointed out: In the past year, nearly a quarter of operators issued a full electric propulsion satellites tender requirements. French Space Agency (CNES) in March 9, 2015 the "Space News" on predicted: In 2020, 50% of commercial communications satellites will be all-electric propulsion satellites.

6 Conclusion

At present, all-electric propulsion satellite has become a hot area in telecommunication satellites development. The developing process of all-electric propulsion satellites have the following characteristics:

1)The conversion from new technology to practical results is very fast. From the conception of all-electric propulsion satellites coming out to the flight test in orbit, it took only 3 years.

2)The response of commercial satellite market to all-electric propulsion satellite is very fast.

From 2012 to the present, there have been 15 all-electric propulsion satellites ordered, and the number is still growing.

3)Satellite manufacturers join the all-electric propulsion satellite camp very fast. Until now, at least 10 satellite manufacturers have proposed their own all-electric propulsion satellite development plan.

4)Satellite manufacturers join the all-electric propulsion satellite market very fast. International major communications satellite operators have joined one after another.

With two satellites based on Boeing 702SP platform successfully delivered to the user in orbit, all-electric propulsion satellite has entered the stage of practical application. All-electric propulsion satellites with high payload carrying capacity and low price are occupying the international commercial telecommunications satellite market quickly. We can foresee that in the near future, all-electric propulsion satellite will become dominant of the market.

References:

[1] Zhou Zhicheng, Gao Jun. Development approach to all-electric propulsion GEO satellite platform[J]. Spacecraft Engineering, 2015, 24(2): 1-6

[2] Konstantinov M. The analysis of influence of electrical propulsion characteristics on efficiency of transport maneuvers[C]//Proceedings of the 30th International Electric Propulsion Conference, IEPC-2007-212. Florence, Italy, 2007: 1-18

[3] Wen Zheng, Wang Min, Zhong Xiaoqing. Multi-task mode electric propulsion technologies[J]. Spacecraft Engineering, 2014, 23(2): 118-123

[4]Bian Bingxiu, Wei Yanming. Key techniques for the application of electric propulsion system in the geostationary satellite platform[J]. Aerospace Control and Application, 2008, 34(1): 20-24

[5] Notarantonio A. Final stage of ARTEMIS orbit raising[C]//2nd AIAA "Unmanned Unlimited" Systems, Technologies, and Operations. San Diego, California, 2003

[6] Liu Yue. Analysis of all-electric propulsion satellite platform for future growth prospects [J]. Space International, 2014(7): 11-15

[7] Duan Chuanhui, Chen Liying. Research and inspiration of all-electric propulsion technology for GEO satellite[J]. Spacecraft Engineering, 2013, 23(6): 99-104

[8] Rathsman P. ELECTRA: The implementation of all-electric propulsion on a geostationary satellite[C]// 64th International Astronautical Congress. Beijing, 2013-09

[9] Monika A K. Optimization of electric thrusters for primary propulsion, AIAA-2001-3347[R] [9] Ocampo C. Geostationary orbit transfer using electric propulsion for the Hughes HS-702 Satellite[C]//49th International Astronautical Congress. Melbourne, Australia , 1998-09-28

Orbit Control Manoeuvre Module for Thaichote Satellite

Manop Aorpimai*, Pornthep Navakitkanok** and Veerachai Malayavej*

*Thai-Paht Satellite Research Centre, Mahanakorn University of Technology, Nongchok, Bangkok, Thailand (phone: +66 2 9883655, fax: +66 2 9884040, email: manop@mut.ac.th)

**Orbit and Space Object Analysis Division, Geo-Informatics and Space Technology Development Agency (GISTDA), Ministry of Science and Technology, Thailand (phone: +66 38 490858, fax: +66 38 490858, email: pornthep@.gistda.or.th)

Abstract: In this paper, we present an orbit control manoeuvre module. Its aim is for controlling the Earth-observation satellites in low-Earth orbit, where the groundtrack pattern and local solar time of the orbital line-of-nodes need to be maintained. The delta-V strategy for achieving and keeping the satellite's orbit at desired conditions, in spite of the astrodynamics disturbances, especially from the non-spherical Earth and atmospheric drag, will be proposed. Some practical issues relating the conversion from theoretical delta-V into practical implementation plan that is ready for the thrusters' execution will be described in details. Some experimental results obtained during the orbit control manoeuvre operations of the Thaichote satellite, which confirm the maturity of the proposed module for practical use, will be presented and discussed.

Keywords: Orbit Control Manoeuvre, Earth-Observation Satellite, Thaichote Satellite

1. Introduction

Thaichote, also known as THEOS, is Thailand's first commercial Earth-observation satellite [1]. It has been operating since March 2008, in a Sun-synchronous orbit at an altitude of about 822 km. The Local Solar Time (LST) of the orbital descending node is assigned at 10:00 AM, and the satellite's ground-track repeats every 369 revolutions or about 26 days. The overall operational orbit configuration of the mission, as well as its control criteria are summarized in Table 1.

Table 1. Operational orbit configuration of Thaichote-I satellite

Orbital Parameters	Value
Altitude	822 km
Inclination	98.76 deg.
Frozen eccentricity	0.001146
Mean Motion	14+5/26 revolutions/day
LST of Descending Node	10.00 A.M. +/- 2 min.
Groundtrack revisit	26 days (369 revolutions)
Groundtrack Maintenance	+/- 40 km w.r.t. the reference longitude

The Orbit Control Manoeuvre (OCM) Module is a part of the spacecraft's Flight Dynamics System (FDS) [2]. It takes responsibilities on achieving and keeping the above mentioned orbital conditions, in order to serve the mission operations throughout its life time. In this paper, we describe an in-house development of the Thaichote's OCM Module. Its aim is to be used as a backup of the operating commercial software, and the module is expected to be extended for the next generation of Thaichote satellites. The proposed module comprises of 2 main parts, namely the Orbit Control Computation and the Orbit Control Implementation, which will be described in details in the next sections. Some practical OCM scenarios will be given and discussed in the experimental results section, and the conclusions will be drawn in the final section.

2. Orbit Control Computation

Naturally, the astrodynamics disturbances, especially those from the Earth's oblateness and atmospheric drag, trend to cause the satellite orbit deviated from its desired initial conditions. Consequently, in order to keep the profile of the controlled parameters, i.e. the groundtrack pattern, frozen eccentricity and the LST of the line of node, within a tolerance gap, delta-V executions are required periodically. The Orbit Control Computation (OCC) sub-module takes responsibility on calculating the required delta-V magnitude and direction, as well as the optimal execution time, while taking into account the minimization of the onboard propellant expenditure as a primary constraint.

2.1 Groundtrack and Frozen Orbit Maintenance

The main cause of groundtrack drift from its reference pattern is the disturbance from atmospheric drag, which trends to reduce the orbital energy, hence nodal period. Consequently, the satellite's groundtrack will drift eastward relative to its reference track. In groundtrack maintenance, a control window is typically assigned, so that the groundtrack drift is allowed only within a pre-defined tolerance. If a constant atmospheric drag is assumed during a control period, a parabolic profile of ground-track drift and drift rate is expected, as depicted in Figure 1. Starting from an initial condition with a positive drift rate, the groundtrack will naturally drift westward. Such condition can be achieved by putting the spacecraft into an orbit with higher altitude than the reference value. Under influence from atmospheric drag, the drift rate reduces and the groundtrack profile is brought eastward. An OCM is required at the positive boundary of the control window to increase the semimajor axis and resume the control profile.

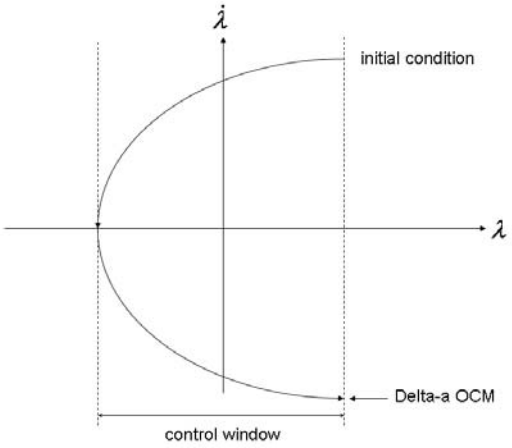


Figure 1. A typical groundtrack profile drifting inside the control window

Only the along-track component of delta-V is required for resuming the orbital energy. The required change in semimajor axis (Δa) is a function of the estimated orbital decay rate caused by atmospheric drag, D, and the control cycle period, T_c, as

$$\Delta a = (a - a_0) + \frac{T_c}{2} D \tag{1}$$

which, to first order approximation, relates to the required delta-V through

$$\Delta V = \frac{\Delta a \cdot V}{2a_0(1 - e\cos n)} \tag{2}$$

where V is orbital velocity, a_0 is the reference semimajor axis, e is eccentricity and n is true anomaly.

As the frozen orbit is also preferable for general Earth-observation missions for altitude variation minimization, each burn of the energy-resuming delta-V can also be utilized for optimally adjusting the eccentricity vector towards frozen [3]. In order to preserve the eccentricity vector in the case that the orbit is already in the frozen condition, the required delta-V can be divided into 2 equivalent parts and applied at the opposite orbital phases.

2.2 Line-of-Node's LST Maintenance

The desired relative angle between the orbital descending nodes and the Sun's direction for Thaichote mission is at 10:00 AM with a tolerance gap of ± 12 min. The third-body attraction, especially the Sun, trends to lean the orbit inclination down at a rate of about 0.04 deg/year, as shown in Figure 2. This causes the Line-of-Node's LST (± 1) drifts in a quasi-parabolic profile as shown in Figure 3, and a similar delta-V strategy as the groundtrack maintenance can again be applied.

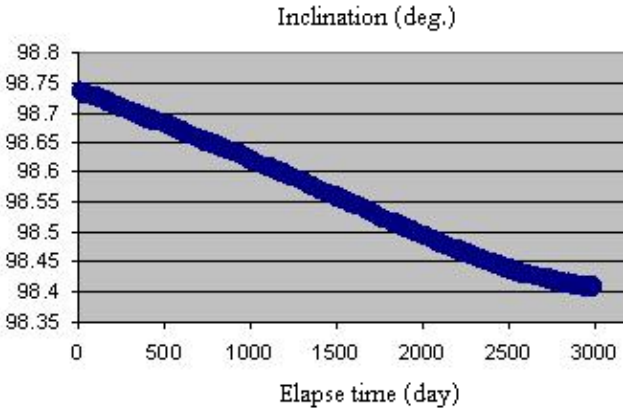


Figure 2. Long-term inclination drift caused by the third-body attraction

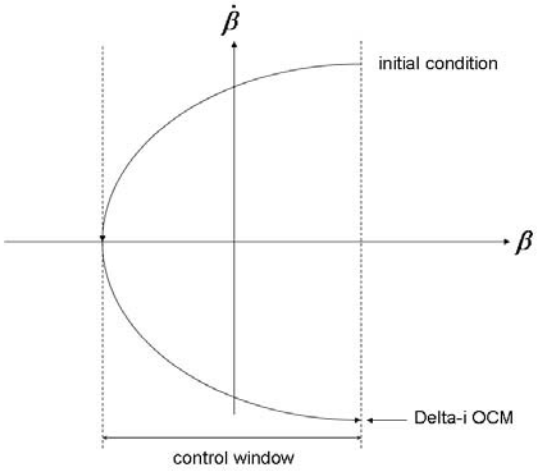


Figure 3. A typical LST profile drifting inside the control window

The initial condition can be set accordingly to the disturbance effect, and a delta-V along the cross-track direction can be applied periodically to resume the inclination:

$$\Delta V = \frac{\Delta i \cdot V}{(\cos l)(1 - e\cos n)} \qquad (3)$$

where l is the argument of latitude. Note that the most effective firing position is at the equator crossing.

3. Orbit Control Implementation

The Thichote's propulsion system available for OCM comprises 4 1-N Monopropellant Hydrazine thrusters [4]. The Orbit Control Implementation (OCI) sub-module calculates the relevant parameters obtained from the conversion of theoretical delta-V into practical values where real data of the onboard propulsion system are taken into account. Various constraints that can affect the OCM are also examined in this sub-module.

3.1 Manoeuvre Duration

With its limited thrust level, the onboard propulsion system can approximately implement the required impulsive delta-V by execution through a short period of time:

$$\Delta V = \int_{t_1}^{t_2} \frac{F(t)}{m(t)} dt \qquad (4)$$

Note that both thrust level, F, and the spacecraft's mass, m, are time-varying variables. They can be formulated as a function of pressure, P, as

$$\Delta V = \int_{P_1}^{P_2} \left(\frac{F}{m}\right)(P) j'(P) dP \qquad (5)$$

where

$$j'(P) = \frac{dt}{dP} = -\frac{r P_1 V_1}{P^2} \frac{1}{(1 - omc_1)q_1 + \ldots (1 - omc_n)q_n}$$

and r is the propellant density, q_i and omc_i are the propellant flow rate and off-modulation coefficient of each thruster. Finally, firing duration can be found from

$$\Delta t = \int_{P_1}^{P_2} j'(P) dP \qquad (6)$$

3.2 Manoeuvre Centroid

As thrust level varies with time, the centroid of accumulated delta-V, t_c, trends to move forward because of continuously reduction of pressure, and it can be calculated from

$$t_c = \frac{1}{\Delta V} \int_{t_i}^{t_f} t \frac{F}{m}(t) dt \qquad (7)$$

3.3 Thrust Level

As the thrust can be different for each of the 4 thrusters, the thrust level shall represents the mean value over the number of thrusters used (2 or 4 thrusters according the configuration) at manoeuvre start epoch, i.e.

$$F = \frac{1}{P_f - P_i}\left(\int_{P_i}^{P_f} F_1(P)dP + \ldots \int_{P_i}^{P_f} F_n(P)dP\right) \tag{8}$$

3.4 Number of Pulses

The number of pulses, NoP, is the global number of pulses assuming that a pulse corresponds to the continuous activation of one thruster during a Flight Software (FSW) Cycle. Therefore, the number of pulses is dependent on thruster configuration. Once the expected duration is known, having considered the right number of thrusters according to the thruster configuration and the off-modulation coefficients, the value of NoP can be computed as

$$NoP = Nearest\ Integer\left(\frac{Theoretical\ Duration}{FSW\ Cycle\ Duration} \times \left(1 - omc_1\right) + \ldots \left(1 - omc_n\right)\right) \tag{9}$$

3.5 Manoeuvre Quaternion

The quaternion represents the rotation that transforms the local orbital frame, +Z towards the earth, +Y opposite from the orbital angular momentum, +X completes the tridron, roughly along the delta-V direction, into OCM attitude. The OCM quaternion is the resultant quaternion of 7 successive rotations:

$$\begin{aligned} Q\_R1\_to\_R7 = {} & [Q\_R7\_to\_R6] \otimes [Q\_R6\_to\_R5]\ [Q\_R5\_to\_R4] \\ & [Q\_R4\_to\_R3] \otimes [Q\_R3\_to\_R2]\ [Q\_R2\_to\_R1] \end{aligned} \tag{10}$$

where $[Q\_R2\_to\_R1]$ is the rotation that transforms the local orbital frame (LVLH) into the TNW frame where T is along the velocity direction, W is along the kinetic momentum and N completes the tridron. $[Q\_R3\_to\_R2]$ is the rotation matrix around W axis to align with the in-plane delta-V, $[Q\_R4\_to\_R3]$ rotates the spacecraft around its own X axis to point the +Z direction towards nadir, $[Q\_R5\_to\_R4]$ aligns the spacecraft along the cross-track delta-V direction. Finally, $[Q\_R6\_to\_R5]$ and $[Q\_R7\_to\_R6]$ are the rotations around Z and Y axis, respectively, to correct the thruster real canting direction.

3.6 OCM Constraints

It is vital to check OCM constraints before generating the OCM plan, in order to guarantee that each OCM execution is safe and efficient. The following constraints are to be checked:

a. Manoeuvre Maximum Duration: This constraint affects both the safety of the propulsion system and the effectiveness of delta-V execution, because the theoretical OCM calculation assumes impulsive delta-V.

b. Maximum M_{sat} 碱 V : This constraint is to limit the magnitude of delta-V that can be safety executed under the operational onboard propulsion, as well as the attitude control configuration.

c. Minimum duration between two OCMs: This constraint is to guarantee a safety margin of the duration between the end epoch of an OCM and the start epoch of the next one.

d. Minimum duration before/after automatic transition: As the attitude manoeuvre to go into OCM attitude shall not occur during an automatic transition, it is therefore necessary to check the duration before/after automatic transition. This value does not depend on the wheel configuration.

f. Minimum duration in a mode other than Sun pointing: This value corresponds to the maximum duration during which the spacecraft can be in a mode other than Sun pointing mode.

g. Minimum angle between Sun direction and +Z axis: This value corresponds to the payload enlighten constraint. The payload shall not be enlighten by the sun during the attitude manoeuvre to go to OCM, the OCM and the attitude manoeuvre to go back to default pointing.

4. Experimental Results

The proposed OCM module has been successfully tested during groundtrack maintenance of the Thaichote satellite. With the practical constraints given in Figure 4 have been verified, the OCM plan was generated and tele-commanded for the practical execution of the delta-V.

Figure 4. Thaichote mission's OCM Constraints

Figure 5 shows the satellite's groundtrack profile before and after the OCM execution, and Figure 6 shows the mean eccentricity is preserved around the frozen value after the manoevre.

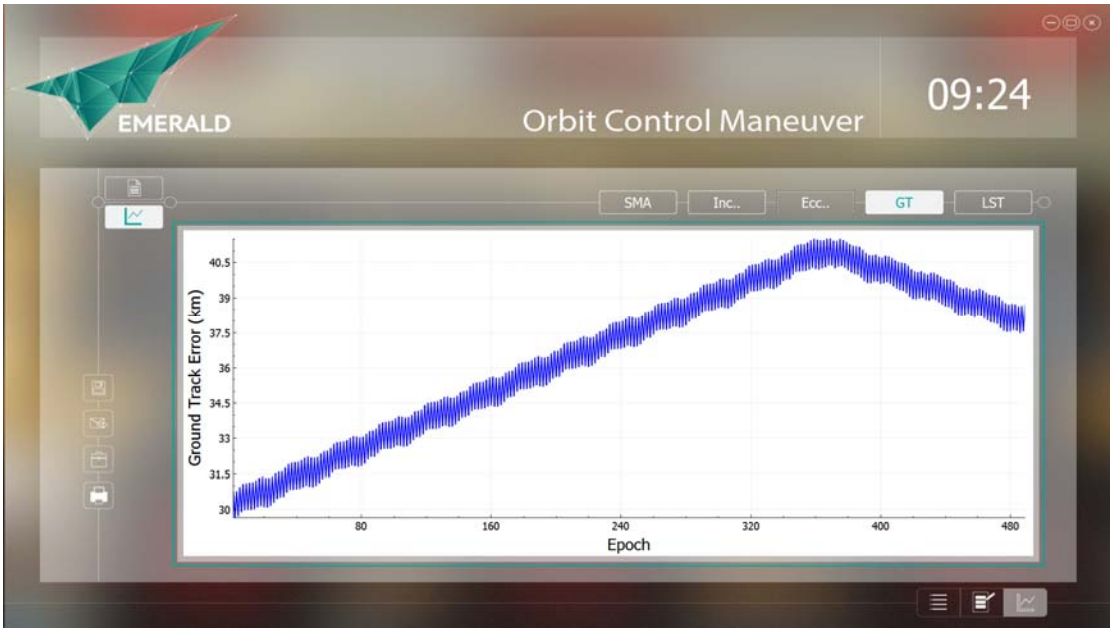


Figure 5**.** Groundtrack profile before and after OCM

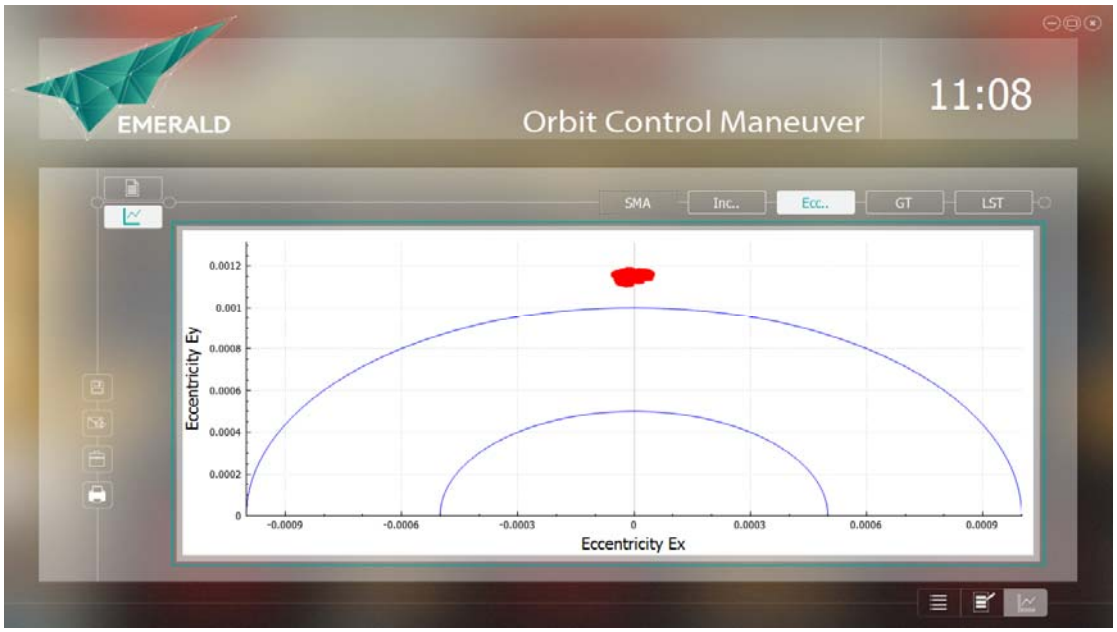


Figure 6. Eccentricity vector motion around the frozen condition during OCM

We also conducted a node's LST maintenance OCM. As the required change in inclination is 0.103 deg, the corresponding delta-V of 13.3896 m/s needed to be divided in to a series of 6 burns. Figure 7 shows the increasing of inclination during OCM, and Figure 8 shows the profile of LST after the OCM which was verified by the onboard GPS orbit determination system.

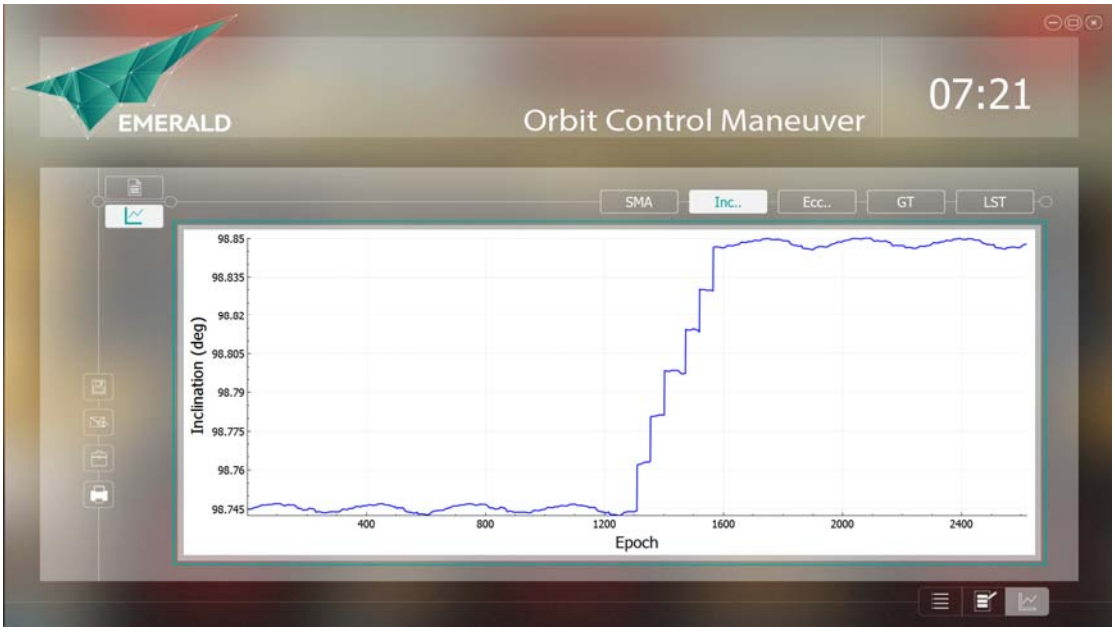


Figure 7. Change in orbital inclination during OCM

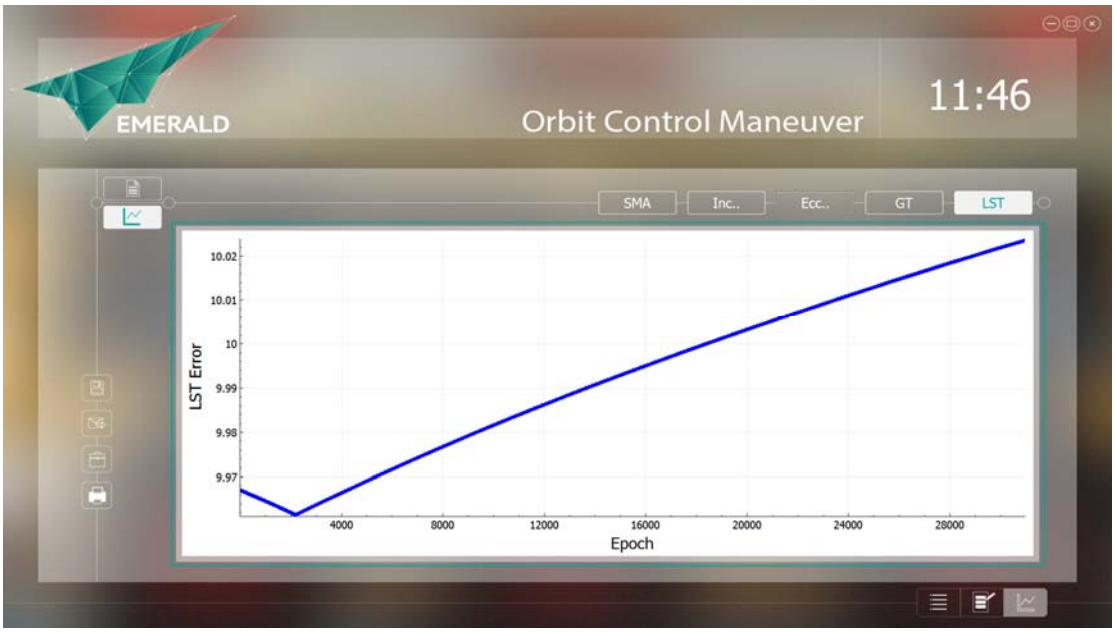


Figure 8. Profile of local solar time of the orbital node

5. Conclusions

In this paper, we have presented an in-house development of the orbit control manoeuvre module for the Thaichote mission. The module has been designed for friendly user interface operation. The simplicity of the computation and the reliability of the verification and plan generation is the key part of the OCM module. A successful demonstration of both groundtrack and LST maintenance were achieved using the proposed system, which guarantees its maturity for deployment on the current mission and further extension for our future Earth-observation missions.

References

[1] P. Navakitkanok, “THEOS Orbit Maintenance: Assessment of 2 Years Operations”, The 22nd International Symposium on Space Flight Dynamics, Sao Jose dos Campos, Brazil, 28th February – 4th March, 2011.

[2] M. Aorpimai and P. Navakitkanok, “Automatic Flight Dynamics System for Thaichote Satellite”, 2nd Conference on Dynamics and Control of Space Systems, Rome, Italy, March 24-25, 2014.

[3] M. Aorpimai, P. Palmer, “Analysis of Frozen Conditions and Optimal Frozen Orbit Insertion”, Journal of Guidance, Control, and Dynamics, Volume 26, Number 5, 2003, pp. 786-793.

[4] T.B.V. HA, QUARTZ++ LEO FDS USER GUIDE, 2007

Satellite Plate Structure Selection and Design Analysis

Wan Shuang Rong Hua Liu Yebao

（China Academy of Launch Vehicle Technology, Beijing 100076,China)

Abstract The strength and stiffness of parts to meet the needs of the satellite to work is the basic structural design requirements. Satellite box plate structure has a simple structure, large interior space, low cost and many other advantages, it has been widely used. The panel is the main component tank bearing plate structure, which can provide an external contour that satellite, but also can be used as a variety of instruments and equipment installation interface. Due to the lower box plate structure bearing efficiency, we design foam sandwich board structure, with the traditional veneer comparative analysis showed a substantial increase in this program box plate structure bearing efficiency. Using advanced composite materials replace aluminum panels as a sandwich structure, so the structure of bearing efficiency has been further improved. In this paper, the use of common business software MSC.Patran / Nastran finished create three panel structural finite element model. Three models for analysis and calculation to obtain the displacement and stress distribution of different configurations. While the failure mode structure was predicted, it provides an important reference for the subsequent design and improvement.

Keywords Satellite, Panel, Structural Design, Simulation Analysis

Introduction

Select the structure of the program needs from the performance, cost and other aspects into account, while demand strength and stiffness of parts to meet the normal operation of the satellite is the basic requirement satellite structural design[1]. Currently there are satellite common configuration box plate, truss, center bearing cartridge[2], multi-box combination of several stacked structure[3]. Wall plate structure in case of these types of programs are the main load-bearing member, which can provide an external contour that satellite, but also can be used as a variety of instruments and equipment installation interface. Box plate structure has a simple structure, large interior space, low cost and many other advantages, it has been widely used, typical satellite box board modular structure shown in Figure 1. But there are also lower box plate structure bearing efficiency, the design was not prone to insufficient strength drawbacks. In order to improve the efficiency of container carrier plate structure, design foam sandwich plate structure, and with the traditional veneer were analyzed. Foam core with high toughness, good impact resistance, energy absorption and fatigue resistance, low moisture absorption, chemical resistance, heat insulation, sound insulation effect, is an ideal structural material[4]. Application of advanced composite materials, lightweight aircraft structures, Integration and high performance plays a vital role[5], which has many excellent characteristics specific strength and specific stiffness, performance can be design, can significantly lose weight and significantly improve the structure of anti-fatigue properties. In this paper, high-performance composite materials replace aluminum panels as a sandwich structure, access to the structure carrying efficiency improvement again.

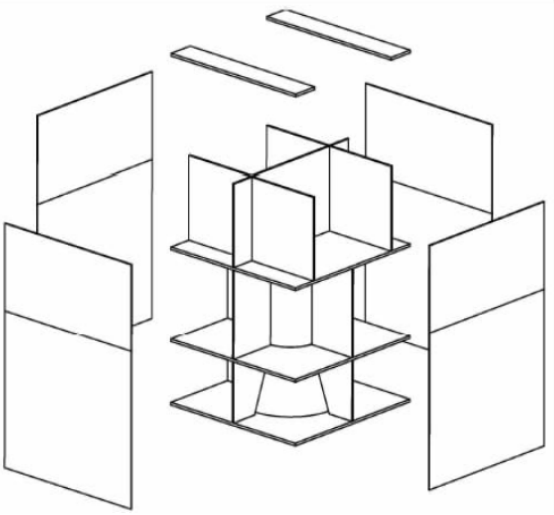


Fig.1 A typical combination of satellite box board structure

This paper designed aluminum single plate, aluminum foam sandwich panels, composite foam sandwich plate structure of three panels, the use of common business software MSC.Patran / Nastran are finished creating three finite element models. Of the three models analyzed and calculated to obtain the displacement and stress distribution in different configurations, the predicted structural damage model for the subsequent design and improvement provide a reference.

1 Aluminum single plate analysis

In this paper, the structure of the selected panel length 600mm, width of 400mm, flat panel dimensions are shown in 2 (b) below. Support clamped bottom panel, two side simply supported at the top end load applied compressive load.

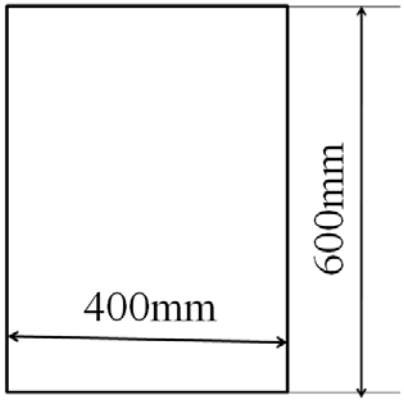
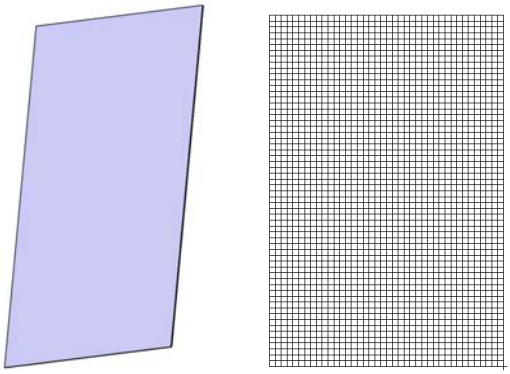


Fig.2 panel schematic plane size

Aluminum panel single structure made of aluminum alloy 2A14 manufacturing, plate thickness of 2mm, as shown in Figure 3 (a) below. Aluminum single-panel structure of the total mass of 1.344kg. Use MSC.Patran / Nastran for single panel numerical simulation analysis. Since the panel of shell structure, so as to Quad4 mesh, mesh side length of 10mm, including a total of 2501 nodes, 2400 elements, finite element mesh shown in Figure 3 (b) below.



（a） （b）

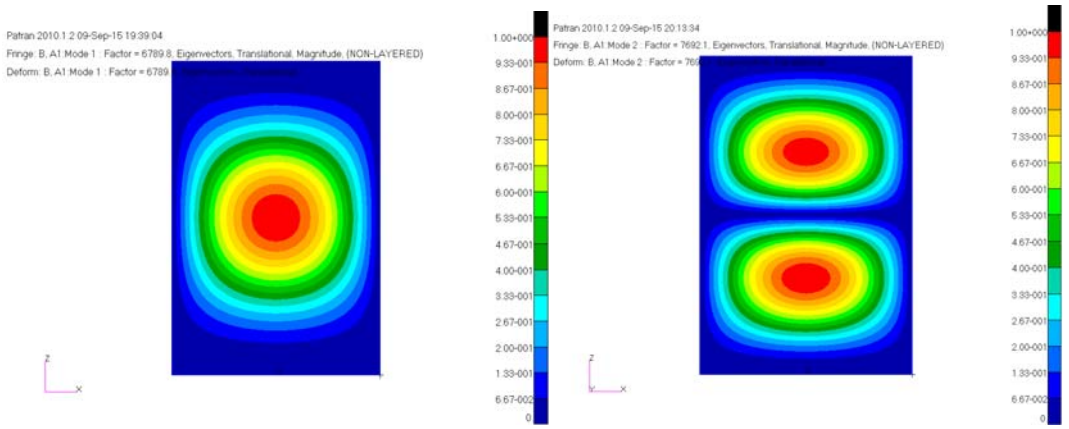
Fig.3 Aluminum single panel schematic

The mechanical properties of aluminum alloy 2A14 parameters shown in Table 1。

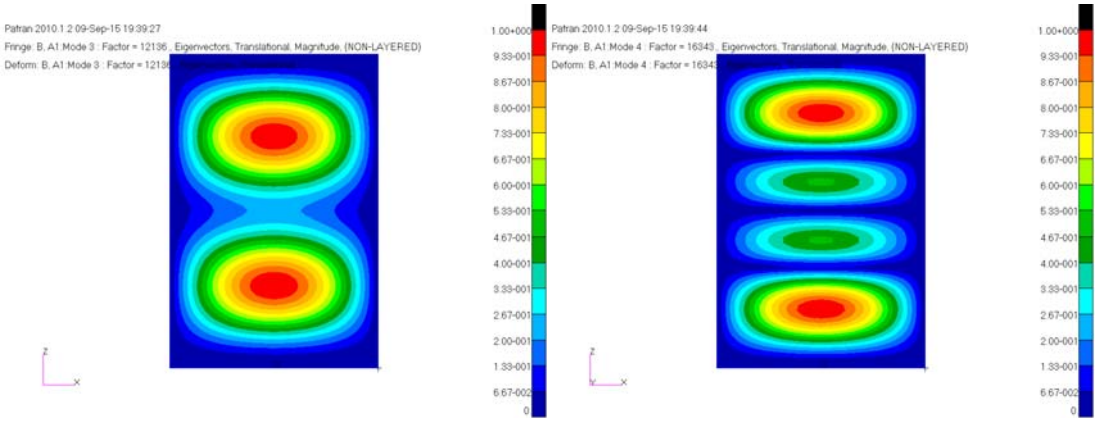
Tab.1 2A14 mechanical performance parameter information

Material name	E (GPa)	μ	Density （kg/m3）	Yield Strength （MPa）	Ultimate strength （MPa）
2A14	70	0.3	2800	375	440

Using MSC.Nastran finite element analysis software Buckling solver (SOL105) linear buckling analysis, when applied to the total load of 1N aluminum single panel structure, the first order buckling factor 6789.8N, buckling load F = 1 × 6789.8N = 6.79kN. Buckling is located in the middle of the rectangular panel, as shown in Figure 4.



（a）The first order buckling （b）Second Order buckling



（c）Third-order buckling （d）Fourth Order buckling

Fig.4 Buckling schematic

When the compression load of 298.97kN, aluminum single board structure to achieve 2A14 yield stress 375MPa, shown in Figure 5. In this case the maximum displacement of the structure of 3.18mm, as shown in Figure 6. The load is much larger than the first order buckling load analysis, the aluminum single panel structure will be the first occurrence of buckling failure.

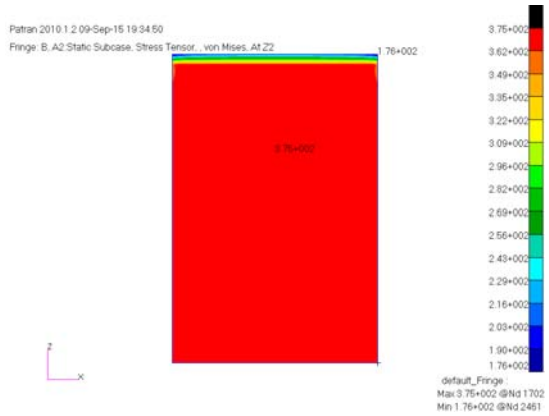


Fig.5 Aluminum single plate stress cloud

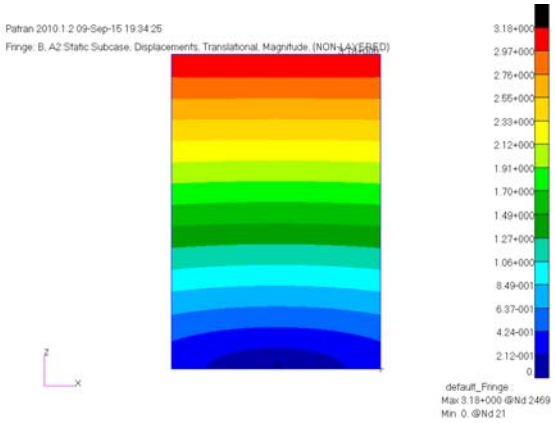
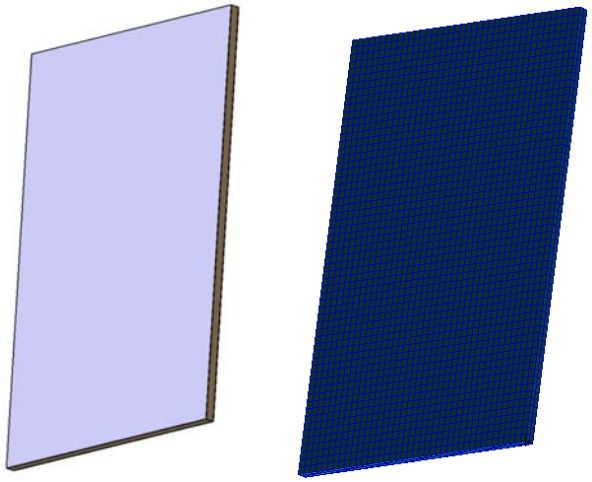


Fig.6 Aluminum single plate displacement contours

2 Aluminum foam sandwich panel analysis

Aluminum foam sandwich panel made of aluminum alloy 2A14 manufacturing structure on both sides of the panel, the panel thickness are 0.5mm, middle foam core thickness of 13.5mm, shown in Figure 7 (a). Aluminum foam sandwich panel construction with a total mass 1.3362kg. MSC.Patran with MSC.Nastran use of aluminum foam sandwich panel for numerical simulation. Since the aluminum panel shell structure, so as to Quad4 mesh, mesh side length of 10mm, foam core entity structure, using Hex8 creation unit, all models includes a total of 7503 nodes, 4800 housing units, 4800 hexahedral element, finite element mesh shown in Figure 7 (b) below.



（a） （b）

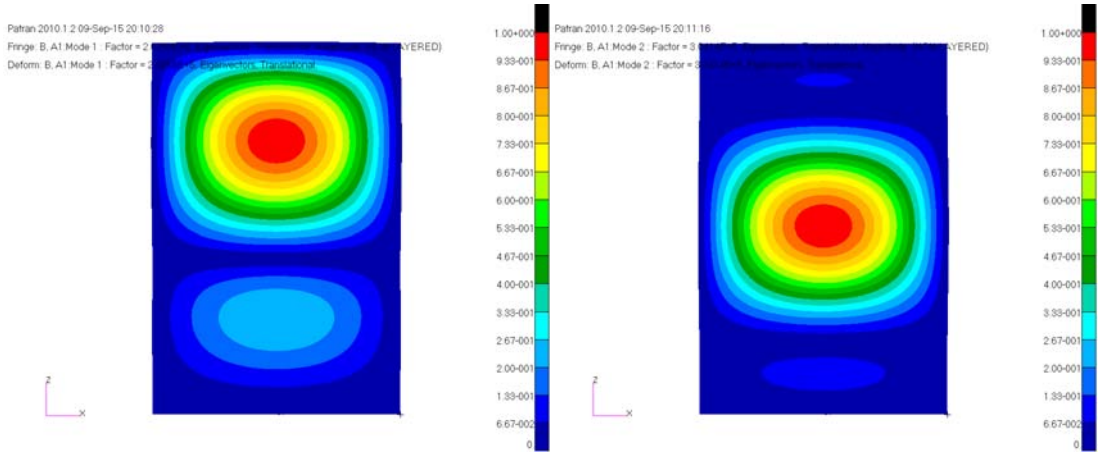
Fig.7 Aluminum foam sandwich panel schematic

Mechanical properties of the foam material parameters shown in Table 2.

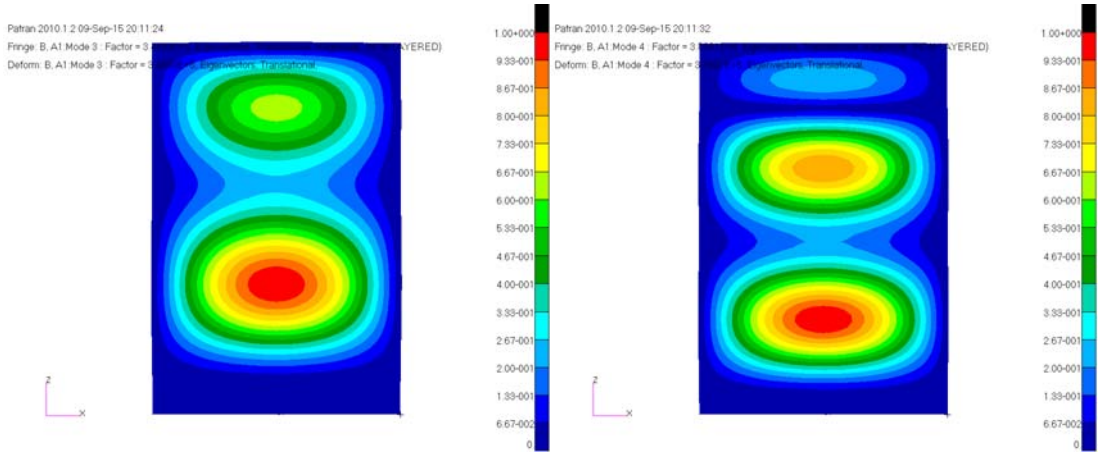
Tab.2 Mechanical properties of the foam parameter information

Material name	E (MPa)	μ	Density （kg/m3）	Compressive strength （MPa）	Shear strength （MPa）
foam	270	0.35	205 ± 35	6.4	3.6

Using MSC.Nastran finite element analysis software Buckling solver (SOL105) linear buckling analysis, when applied to the total load of 1N aluminum foam sandwich panel structure, the first order buckling factor 262580N, buckling load F = 1 × 262580N = 262.58kN. Buckling is located in the middle of the rectangular panel, as shown in Figure 8.



（a）The first order buckling　　（b）Second Order buckling



（c）Third-order buckling　　（d）Fourth Order buckling

Fig.8 Buckling schematic

When the compression load of 156.80kN, the aluminum panels reach 2A14 yield stress 375MPa, shown in Figure 9. In this case the maximum displacement of the structure of 3.17mm, as shown in Figure 10. The load is far less than the first order buckling load, and aluminum single plate different circumstances, aluminum foam sandwich panel structure buckling does not occur, it will be the first damage occurred aluminum panels.

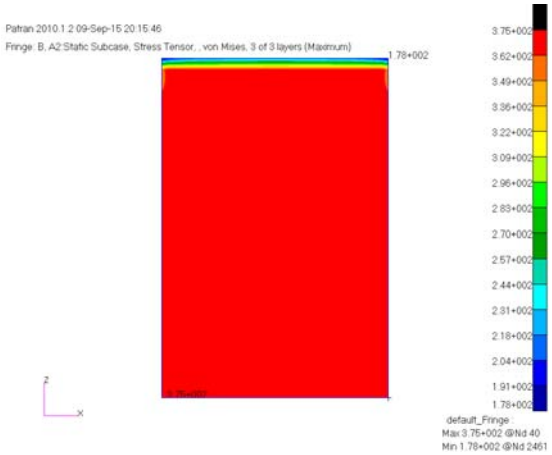


Fig.9 Aluminum foam sandwich panel stress

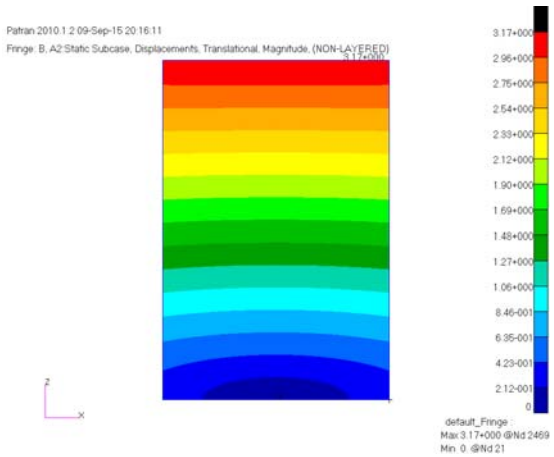


Fig.10 Aluminum foam sandwich panel displacement contours

3 Composite foam sandwich plate analysis

Composite foam sandwich panel structure using carbon fiber resin matrix composite material on both sides of the panel, the panel thickness are 0.8mm, middle foam core thickness of 13.5mm, its shape schematic see Figure 7 (a). Composite foam sandwich panel construction with a total mass 1.2978kg. MSC.Patran with MSC.Nastran use of composite foam sandwich panel numerical simulation analysis. Since the composite foam sandwich panel with aluminum sandwich panel the same structure, so share a finite element mesh, given the different property values, finite element mesh schematic see Figure 7 (b). Composite material used is carbon fiber woven fabric, which overlay information as shown in Table 3, the mechanical properties of woven material parameters shown in Table 4.

Tab.3 Composite panel overlay detailed information

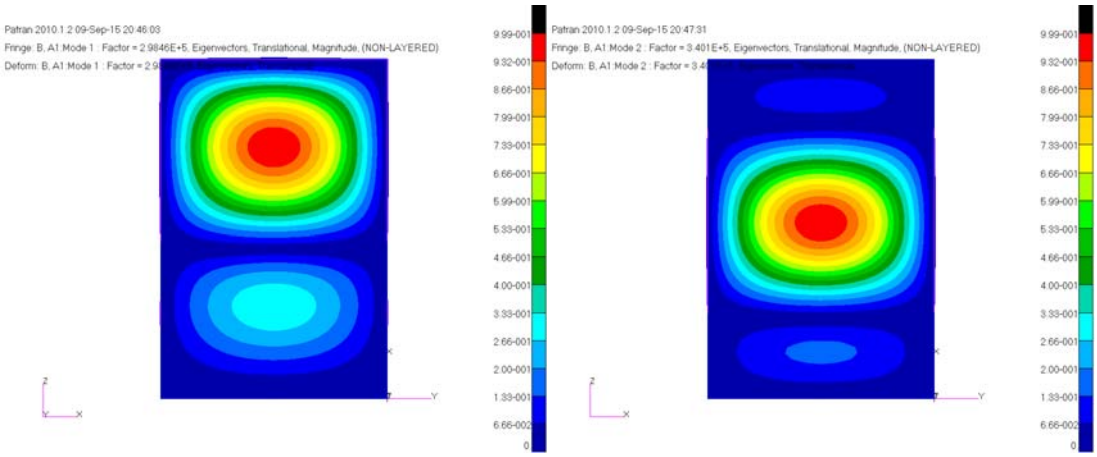
Location	Overlay direction	Number of plies	Layer thickness/mm	Total thickness/mm
panel	[(+45)/(0)/(+45)/(0)]	8	0.2	0.8

Tab.4 Woven mechanical performance parameters

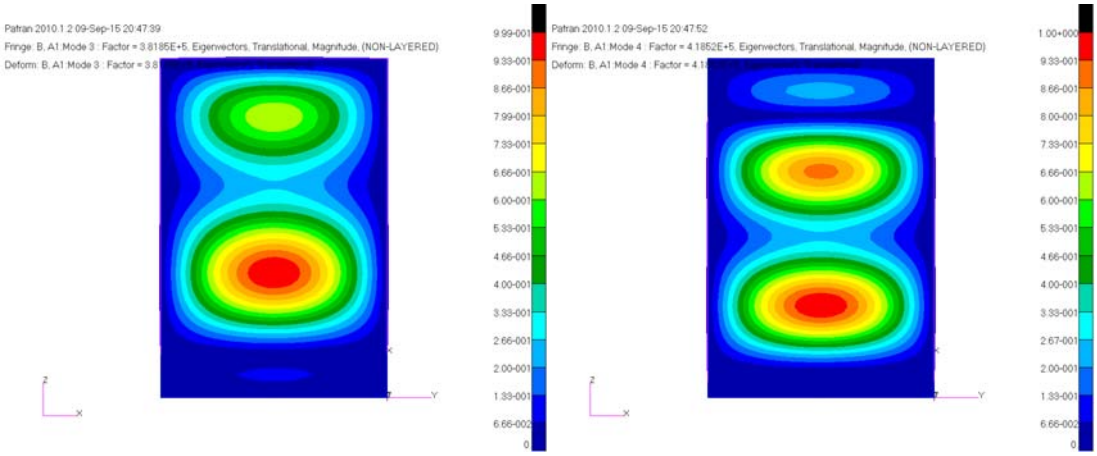
serial number	item	Numerical
1	0°Tensile modulus /GPa	73.6
2	0°Tensile Strength /MPa	983
3	Main Poisson's ratio	0.06
4	90°Tensile modulus /Gpa	71.6

serial number	item	Numerical
5	90°Tensile Strength /MPa	930
6	Poisson ratio	0.06
7	0°Compressive strength /MPa	590
8	0°Compression modulus /GPa	79.8
9	90°Compressive strength /MPa	543
10	90°Compression modulus /GPa	81
11	0°Bending modulus /GPa	61
12	0°Bending strength /MPa	993
13	90°Bending strength /MPa	922
14	90°Bending modulus /GPa	55.8
15	shear modulus /GPa	4.64
16	Shear strength/MPa	123
17	0°Interlaminar shear strength /MPa	70.4

Using MSC.Nastran finite element analysis software Buckling solver (SOL105) linear buckling analysis, when applied to the total load of 1N aluminum foam sandwich panel structure, the first order buckling factor 298460N, buckling load F = 1 × 298460N = 298.46kN. Buckling, also located in the middle of the rectangular panel, as shown in Figure 11.



（a）The first order buckling　　（b）Second Order buckling



（c）Third-order buckling　　（d）Fourth Order buckling

Fig.11 Buckling schematic

Aluminum alloy material with isotropic different anisotropic composite woven, the strength in different directions are not identical. As used herein, Tsai-Wu criterion woven laminated composite panel failure prediction, which was expressed as follows[6]:

$$F\ I = F_1\sigma_x + F_2\sigma_y + F_{11}\sigma_x^2 + F_{22}\sigma_y^2 + 2F_{12}\sigma_x\sigma_y + F_{66}\tau_{xy}^2 \qquad (1)$$

among them:

$$F_1 = \frac{1}{TX} - \frac{1}{CX}; F_2 = \frac{1}{TY} - \frac{1}{CX}; F_{11} = \frac{1}{TX \times CX}; F_{22} = \frac{1}{TY \times CY;}$$

$$F_{12} = \frac{IXY}{\sqrt{TX \times CX \times TY \times CY}}; F_{66} = \frac{1}{SXY^2}$$

When $F\ I \geq 1$, composites failure occurs.

When the compression load of 292.19kN, the composite panel Tsai-Wu criterion maximum coefficient of 1, as shown in Figure 9, the composite material of panel damage. In this case the maximum displacement of the structure of 5.01mm, as shown in Fig. The load is slightly smaller than the first order buckling load, in this case with the same aluminum foam sandwich panels, composite foam sandwich panel structures are not buckling occurs, it will be the first carbon fiber woven panel damage occurred.

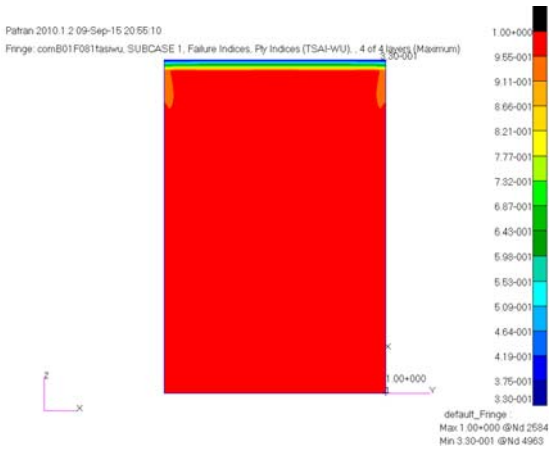


Fig.9 Composite foam sandwich panel Tsai-Wu criterion coefficient cloud

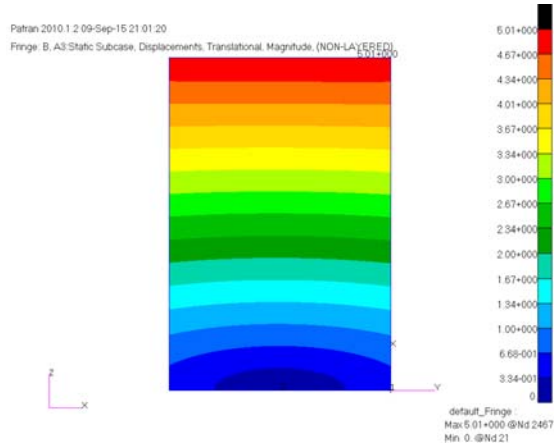


Fig.10 Composite foam sandwich panel displacement contours

4 Comparative analysis results

Comprehensive structural analysis of the results of the three panel, as shown in Table 5, three configurations in a single layer aluminum panel maximum weight, minimum failure load, which is due to a single-layer panel prone to buckling under compression loads destruction, failed to play the carrying capacity of the material; composite foam sandwich panel construction minimum weight, maximum failure load, this is due to a sandwich structure with good resistance to buckling failure, the other is due to the carbon fiber resin-based composite material itself has a specific strength and specific stiffness characteristics, carrying capacity; aluminum foam sandwich structure and properties of the three living in the middle. Can be seen from the table in the last column, failure load weight ratio composite foam sandwich panel reached 225.14kN / kg, which has a very high bearing efficiency.

Tab.5 Three configurations comparison

Configuration	Weight kg	Failure load kN	Failure load/ Weight kN/kg
Aluminum single plate	1.344	6.79	5.05
Aluminum foam sandwich panels	1.3362	156.8	117.35
Composite foam sandwich plate	1.2978	292.19	225.14

5 in Conclusion

(1) For boxboard combined satellite structure, the paper its main load-bearing member panel conducted several configuration analysis, using MSC.Patran / Nastran finite element model is created, to the selection panel analysis provides a feasible engineering approach;

(2) The application of finite element model analysis, various configurations under compression loads and failure modes and failure load;

(3) the failure load and the weight ratio of composite foam sandwich panel reached 225.14kN / kg, which has a very high bearing efficiency, the application of this configuration is designed to significantly reduce the weight of the satellite structure;

(4) It should be noted that this article only carried out static analysis, in practical engineering application process, still need to structure modal, dynamic response, thermal deformation / heat stress all aspects of the assessment to determine the final structure of the program .

Reference

[1] Fang Bo. Micro-satellite structure design method [J], Shenyang Aerospace University learned journal，2013,30(5):37-40.

[2] Yao Jun，Zhang Hongying，Cui Wei，etc. Cylinder for modal analysis and testing [J]. Spacecraft Environment Engineering，2010,27（1）：83-86

[3] Chin Feng，Zu Chang Hong，Jeng Shing Chern，etal。structure design for TUUSAT-1A Micosatellite{J}。Wseas Transactions On Applied And Theoretical Mechanies，2010,1（5）：45-58

[4] Ma Li，Zhu Dalei. Satellite structure compared with PVC foam core aluminum honeycomb core sandwich panels [J], Composite manufacturing，2013,2（1）:18-21

[5] Li Zhengneng. Learn aircraft structure [M],Beijing：Beijing University of Aeronautics and Astronautics Press，2010，P382.
[6] Yang Jian,Zhang Pu,Chen Huohong. New MD Nastran Finite Element tutorial examples [M], Bei Jing：Machinery Industry Press，2008，P224.

Numerical Modeling for the Dynamic Tracking of Dynamic-hydrostatic Hybrid Gas Seals for liquid oxygen pump

ZHANG Shu-qiang, WANG Liang, CHEN Jie, ZHAO Wei-gang

（Xi'an Aerospace Propulsion Institute, Xi'an 710100, China）

Abstract: One of important components for liquid oxygen pump is mechanical seal, which prevents leakage of propellant. At present the contacting mechanical seal is widely adopted in liquid oxygen/kerosene propellant rocket engine. Usually, the speed of liquid oxygen pump is high, so the seal would easily lead to face wear, and ultimately premature seal failure. In view of this, the dynamic-hydrostatic hybrid gas seals (DHHGS) are designed for liquid oxygen pump to prevent the seal friction failure and improve reliability. The gas film stiffness and damping coefficients for DHHGS are obtained from the unsteady nonlinear Reynolds equation using the perturbation method. The dynamic tracking of stator is analyzed under the simple harmonic excitation caused by the rotor motion due to its runout and misalignment. Stability of axial free vibration of stator is investigated for the active regulation, and critical instability criterion is given. Results showed that stator still has the perfect tracking motion even if the amplitudes of simple harmonic excitation are large enough. The damping ratio of free vibration of stator approximately equals to 1.0 for the active regulation, the motion of stator is non-reciprocating attenuation vibration.

Keywords: liquid oxygen pump, dynamic-hydrostatic hybrid gas seals (DHHGS), dynamic characteristic coefficient, tracking characteristics, active regulation

1 Introduction

Liquid oxygen (LOX) turbopumps are high-speed, high-pressure, and high-power density machines that require effective sealing for the impeller entrance of the machine [1]. At present the contacting mechanical seals are widely adopted in LOX turbopumps, because of the vaporization in the seal face could cause seal failure by overheating, face temperature is high and graphite gets very heavy wear of the contacting mechanical seals. Although the face temperature and graphite wear rate could be decreased by optimizing the seal structure parameter and choosing high performance frictional pair materials, the improvements are not significant for the seals which are operated at high parameters.

Dynamic-hydrostatic hybrid gas seal (DHHGS) based on the principles of dynamic gas seal and hydrostatic gas seal is known for its extremely low leakage, low wear or non-wear, and low heat generation [2-3]. It has been used in almost every sector of industrial pumping applications and has been expanding its application in Aero-engines [4-5]. However, DHHGS has not been applied to LOX turbopumps of liquid oxygen/kerosene propellant rocket engine, only theoretical research and experimental study are taken by researchers [6]. Because of the DHHGS is operated with no face contact, the vaporization of LOX in the seal face is basically avoided. According to different quench gas pressure which create static pressure forces, DHHGS can be divided into self-pressurized

dynamic-hydrostatic hybrid gas seals (SDHHGS) and external-pressurized dynamic-hydrostatic hybrid gas seals (EDHHGS).

The aim of this paper is to study the dynamic tracking of EDHHGS which are designed for LOX turbopumps of liquid oxygen/kerosene propellant rocket engine. A two-dimensional numerical model is presented and the stiffness and damping coefficient of the gas film are calculated for EDHHGS using the perturbation method. The dynamic tracking of stator is analyzed under the simple harmonic excitation caused by the rotor motion due to its runout and misalignment, and the influence of quench gas pressure on the amplitude of stator is investigated. Meanwhile, stability of axial free vibration of stator is investigated for the active regulation of EDHHGS, and critical instability criterion is given.

2 The Working Principle of EDHHGS

A schematic of the EDHHGS model is shown in Fig. 1. One seal ring (stator) is flexibly mounted to the housing by an elastomeric o-ring, acting as the secondary seal, and a spring, providing closing force. The other seal ring (rotor) is rigidly mounted to the rotating shaft. The double-divert spiral grooves are etched into one of the faces and are highlighted here for contrast. The restrictive orifice feed holes and static pressure equalizing groove are connected directly, and they are machined into the stator surface. The purpose of these restrictive orifice feed holes is to permit quench gas (the pressure is p_s) to flow into the stator and rotor interface area (the pressure drop to p_d). A properly designed restrictive orifice is used for increasing the coning stiffness as well as being very effective at improving the axial stiffness. During static conditions the closing force is balanced by the static pressure forces created by p_d such that there exists a clearance, thus, creating an initially open condition. Once the rotor starts rotating and the system pressure increases, the static pressure forces and the balancing hydrodynamic film force between the stator and rotor become the dominant forces. In order to obtain different sealing effectiveness, the structural form and geometric parameters of double-divert grooves and restrictive orifice feed holes can be changed. In general, by changing the value of the quench gas pressure, the sealing clearance can be regulated to control leakage.

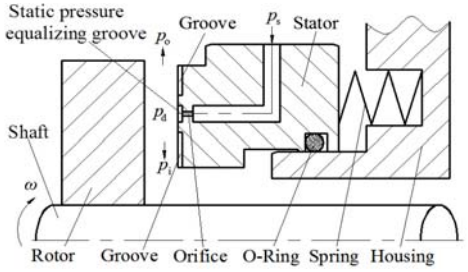
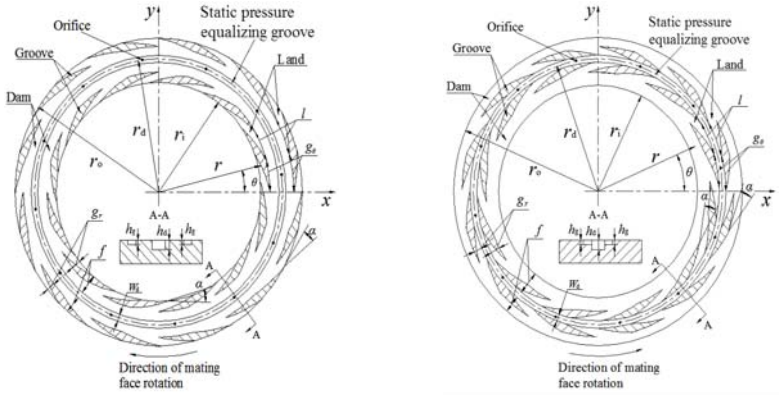


Fig. 1 Schematic of EDHHGS

Due to the excellent performance of spiral groove, this paper takes it as the research object. Two different end face structures of EDHHGS are shown in Fig. 2. The “pump in” EDHHGS (pump in) is shown in Fig. 2(a), the double-divert spiral grooves are located at the inner and outer place of the seal face. The quench gas flows into the stator and rotor interface area through the restrictive orifice feed holes. Once the rotor starts rotating, the double-divert spiral grooves pump the gas into the mid place of seal face. The “pump out” EDHHGS (pump out) is shown in Fig. 2(b), the double-divert spiral grooves are located at the inner and outer place of static pressure equalizing groove, the double-divert spiral grooves and static pressure equalizing groove are connected directly.

Similarly, the quench gas flows into the stator and rotor interface area through the restrictive orifice feed holes. On the contrary, once the rotor starts rotating, the double-divert spiral grooves pump the gas into the inner and outer place of seal face.



(a) pump in (b) pump out

Fig. 2 End face structure of EDHHGS

The following analysis is based on the parameters listed in Table 1, which includes seal geometry, material properties, operating conditions, and rotor runout and misalignments.

Table 1 Parameters for the Steady-state Tracking Analysis of EDHHGS

Structural Parameters		**Operating Conditions**	
Seal ring inner radius r_i/mm	47	Pressure at the inner radius p_i/MPa	0.1
Seal ring outer radius r_o/mm	66	Pressure at the outer radius p_o/MPa	0.6
Radial position of orifice and static pressure equalizing groove r_d/mm	57	Pressure of quench gas p_s/MPa	0.7
Radius on which the O-ring is installed r_s/mm	54	Viscosity μ/Pa·s	1.79×10^{-5}
Static pressure equalizing groove width W_d/mm	0.4	Shaft speed n/rpm	12000
Static pressure equalizing groove depth h_d/mm	0.2	Steady-state equilibrium film thickness of dam area h_l /mm	0.004
Double-divert spiral grooves depth h_g/mm	0.006	**Physical Properties**	
Land-to-groove ratio $\gamma=l/g_\theta$	1.0	Stator mass m/kg	0.2
Sealing groove ratio $\delta=g_r/f$,	0.7	Stator transeverse moment of inertia I_x , I_y/kg·m^2	3.5×10^{-4}
Spiral groove angle α/°	20	Spring stiffness k_s/N·m^{-1}	5×10^{4}
Number of double-divert spiral grooves N_g	12	O-ring damping c_s/N·s·m^{-1}	1×10^{3}
Number of restrictive orifice feed hole N_o	12	**Face Runout and Misalignment**	
Diameter of restrictive orifice feed hole d/mm	0.2	Rorot axial runout A_{rz}/μm	5
Balance ratio B	0.67	Rotor misalignment A_{rx} , A_{ry}/rad	5×10^{-4}

3 Numerical Solution of the Lubrication Equations

3.1 Lubrication Equations

The gas flow is assumed to be isothermal and isoviscous, so the gas film pressure in the seal is governed by the unsteady compressible Reynolds equation, expressed in the polar coordinates is given by [7]

$$\vec{\nabla}\cdot\left[ph^3\vec{\nabla}p-6\mu\omega rph\,\vec{e}_\theta\right]+12\,\mu RT\dot{m}\delta_j=12\,\mu\frac{\partial(ph)}{\partial t} \tag{1}$$

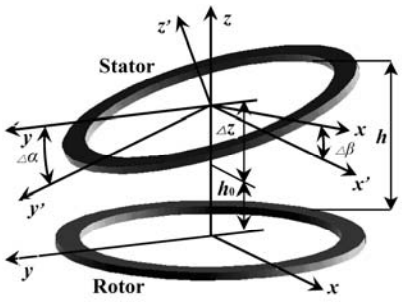


Fig. 3 Seal face kinematic model

As shown in Fig. 3, when the seal is perturbed by a small oscillatory motion (Δz, Δα, Δβ) with frequency υ, the film thickness can be expressed as the following:

$$h=h_0+\Delta z(t)+r\sin\theta\Delta\alpha(t)-r\cos\theta\Delta\beta(t) \tag{2}$$

The following motion is assumed for the stator:

$$\Delta z(t)=|\Delta z|e^{i\upsilon t},\ \Delta\alpha(t)=|\Delta\alpha|e^{i\upsilon t},\ \Delta\beta(t)=|\Delta\beta|e^{i\upsilon t} \tag{3}$$

which leads to

$$\Delta\dot{z}(t)=i\upsilon\Delta z(t),\ \Delta\dot{\alpha}(t)=i\upsilon\Delta\alpha(t),\ \Delta\dot{\beta}(t)=i\upsilon\Delta\beta(t) \tag{4}$$

These small motions cause a small pressure perturbation Δp about the equilibrium pressure p_0. Using a first order Taylor series expansion, the total pressure (including the exit pressure of the restrictive orifice feed holes) can be expressed as

$$p=p_0+\Delta p=p_0+p_z\Delta z(t)+p_x\Delta\alpha(t)+p_y\Delta\beta(t)++p_{\dot{z}}\Delta\dot{z}(t)+p_{\dot{x}}\Delta\dot{\alpha}(t)+p_{\dot{y}}\Delta\dot{\beta}(t) \tag{5}$$

The small pressure perturbation Δp is expressed as follows:

$$\Delta p=p_z\Delta z(t)+p_x\Delta\alpha(t)+p_y\Delta\beta(t)+p_{\dot{z}}\Delta\dot{z}(t)+p_{\dot{x}}\Delta\dot{\alpha}(t)+p_{\dot{y}}\Delta\dot{\beta}(t) \tag{6}$$

where p_z, p_x, p_y, $p_{\dot{z}}$, $p_{\dot{x}}$ and $p_{\dot{y}}$ are perturbation pressure caused by unit perturbation displacement and unit perturbation velocity, respectively. Substituting equations (3) and (4) into equation (5), the total pressure can be written as

$$p=p_0+p_z^*\Delta z(t)+p_x^*\Delta\alpha(t)+p_y^*\Delta\beta(t) \tag{7}$$

where

$$\begin{cases} p_k^*=p_{kr}+ip_{ki} \\ p_{kr}=p_k \qquad (k=z,x,y) \\ p_{ki}=\upsilon p_{\dot{k}} \end{cases} \tag{8}$$

where r and i represent real part and imaginary part.

In Eq. (1), the gas mass flow of $\dot{m}$ is given by

$$\dot{m}=\dot{m}_0+\left.\frac{\partial\dot{m}}{\partial p_d}\right|_0\Delta p_d=\dot{m}_0+k_m\Delta p_d \tag{9}$$

The $\dot{m}_0$ and k_m can be evaluated by the Eqs. (10) and (12):

$$\dot{m}_0=p_s\phi\sqrt{\frac{2}{RT}}\psi_0 \tag{10}$$

where ψ_0 is the flow function, is given by

$$\psi_0=\begin{cases}\left[\dfrac{k}{2}\left(\dfrac{2}{k+1}\right)^{\frac{k+1}{k-1}}\right]^{1/2} & \dfrac{p_{d0}}{p_s}\le\left(\dfrac{2}{k+1}\right)^{\frac{k}{k-1}}\\ \left\{\dfrac{k}{k-1}\left[\left(\dfrac{p_{d0}}{p_s}\right)^{\frac{2}{k}}-\left(\dfrac{p_{d0}}{p_s}\right)^{\frac{k+1}{k}}\right]\right\}^{1/2} & \dfrac{p_{d0}}{p_s}>\left(\dfrac{2}{k+1}\right)^{\frac{k}{k-1}}\end{cases} \tag{11}$$

$$k_m=\left.\frac{\partial\dot{m}}{\partial p_d}\right|_0=p_s\phi\sqrt{\frac{2}{RT}}\left.\frac{\partial\psi}{\partial p_d}\right|_0 \tag{12}$$

where

$$\left.\frac{\partial\psi}{\partial p_d}\right|_0=\begin{cases}0 & \dfrac{p_{d0}}{p_s}\le\left(\dfrac{2}{k+1}\right)^{\frac{k}{k-1}}\\ \dfrac{1}{2}\left\{\dfrac{k}{k-1}\left[\left(\dfrac{p_{d0}}{p_s}\right)^{\frac{2}{k}}-\left(\dfrac{p_{d0}}{p_s}\right)^{\frac{k+1}{k}}\right]\right\}^{1/2} & \dfrac{p_{d0}}{p_s}>\left(\dfrac{2}{k+1}\right)^{\frac{k}{k-1}}\\ \times\dfrac{k}{k-1}\left[\dfrac{2}{k}\left(\dfrac{p_{d0}}{p_s}\right)^{\frac{2-k}{k}}\dfrac{1}{p_s}-\dfrac{k+1}{k}\left(\dfrac{p_{d0}}{p_s}\right)^{\frac{1}{k}}\dfrac{1}{p_s}\right] & \end{cases} \tag{13}$$

Introducing the dimensionless quantities, as shown in Eq. (14), including dimensionless excitation frequency Γ and compressibility number Λ.

$$H_0=\frac{h_0}{h_l},R=\frac{r}{r_i},P_0=\frac{p_0}{p_a},\Gamma=\frac{\upsilon}{\omega},\Lambda=\frac{6\mu\omega r_i^2}{p_a h_l^2},P_{kj}=\frac{p_{kj}h_l}{p_a}\text{(k=z)},P_{kj}=\frac{p_{kj}h_l}{p_a r_i}\text{(k=x，y)}$$

$$E_0=\frac{12\mu RTr_i^2}{p_a^2h_l^3},E_1=\frac{12\mu RTr_i^2}{p_a h_l^3} \tag{14}$$

where j=r, i.

Substituting Eqs. (2), (7) and (9) into Eq. (1), neglecting higher order terms, and introducing the dimensionless quantities yields the following seven equations which govern the dimensionless steady-state pressure distribution P_0 at the equilibrium position h_0 and three separate pairs of dimensionless variables (P_{zr}, P_{zi}), (P_{xr}, P_{xi}), and (P_{yr}, P_{yi}):

$$\frac{\partial}{R^2\partial\theta}\left(H_0^3\frac{\partial P_0^2}{\partial\theta}\right)+\frac{\partial}{R\partial R}\left(RH_0^3\frac{\partial P_0^2}{\partial R}\right)=2\Lambda\frac{\partial(P_0H_0)}{\partial\theta}-2E_0\dot{m}_0\delta_j \tag{15}$$

$$\frac{\partial}{R^2\partial\theta}\left[H_0^3\frac{\partial(P_0P_{zr})}{\partial\theta}+\frac{3}{2}H_0^2\frac{\partial P_0^2}{\partial\theta}\right]+\frac{\partial}{R\partial R}\left[R\left(H_0^3\frac{\partial(P_0P_{zr})}{\partial R}+\frac{3}{2}H_0^2\frac{\partial P_0^2}{\partial R}\right)\right]=\Lambda\left[\frac{\partial(P_0+P_{zr}H_0)}{\partial\theta}-2\Gamma P_{zi}H_0\right]-E_1k_mP_{zr}\delta_j \tag{16a}$$

$$\frac{\partial}{R^2\partial\theta}\left[H_0^3\frac{\partial(P_0P_{zi})}{\partial\theta}\right]+\frac{\partial}{R\partial R}\left[RH_0^3\frac{\partial(P_0P_{zi})}{\partial R}\right]=\Lambda\left[\frac{\partial(P_{zi}H_0)}{\partial\theta}+2\Gamma(P_0+P_{zr}H_0)\right]-E_1k_mP_{zi}\delta_j \tag{16b}$$

$$\frac{\partial}{R^2\partial\theta}\left[H_0^3\frac{\partial(P_0P_{xr})}{\partial\theta}+\frac{3}{2}H_0^2R\sin\theta\frac{\partial P_0^2}{\partial\theta}\right]+\frac{\partial}{R\partial R}\left[RH_0^3\frac{\partial(P_0P_{xr})}{\partial R}+\frac{3}{2}H_0^2R^2\sin\theta\frac{\partial P_0^2}{\partial R}\right]=\Lambda\left[\frac{\partial(P_0R\sin\theta+P_{xr}H_0)}{\partial\theta}-2\Gamma P_{xi}H_0\right]-E_1k_mP_{xr}\delta_j \tag{17a}$$

$$\frac{\partial}{R^2\partial\theta}\left[H_0^3\frac{\partial(P_0P_{xi})}{\partial\theta}\right]+\frac{\partial}{R\partial R}\left[RH_0^3\frac{\partial(P_0P_{xi})}{\partial R}\right]=\Lambda\left[\frac{\partial(P_{xi}H_0)}{\partial\theta}+2\Gamma(P_0R\sin\theta+P_{xr}H_0)\right]-E_1k_mP_{xi}\delta_j \tag{17b}$$

$$\frac{\partial}{R^2\partial\theta}[H_0^3\frac{\partial(P_0P_{yr})}{\partial\theta}-\frac{3}{2}H_0^2R\cos\theta\frac{\partial P_0^2}{\partial\theta}]+\frac{\partial}{R\partial R}[RH_0^3\frac{\partial(P_0P_{yr})}{\partial R}-\frac{3}{2}H_0^2R^2\cos\theta\frac{\partial P_0^2}{\partial R}]=\Lambda[\frac{\partial(-P_0R\cos\theta+P_{yr}H_0)}{\partial\theta}-2\Gamma P_{yi}H_0]-E_1k_mP_{yr}\delta_j \tag{18a}$$

$$\frac{\partial}{R^2\partial\theta}[H_0^3\frac{\partial(P_0P_{yi})}{\partial\theta}]+\frac{\partial}{R\partial R}[RH_0^3\frac{\partial(P_0P_{yi})}{\partial R}]=\Lambda[\frac{\partial(P_{yi}H_0)}{\partial\theta}+2\Gamma(-P_0R\cos\theta+P_{yr}H_0)]-E_1k_mP_{yi}\delta_j \tag{18b}$$

3.2 Boundary Conditions

The boundary conditions are

$$P_0=p_i/p_a\ ,\ (R=1);\ P_0=p_o/p_a\ ,(R=r_o/r_i)$$

for Eq. (15) and

$$P_{kj}=0,\ (R=1,\ R=r_o/r_i)$$

for Eqs. (16)-(18).

where k=z, x, y and j=r, i.

3.3 Numerical Method

Eq. (15) is a nonlinear equation that governs the dimensionless steady-state pressure distribution P_0. In order to solve the equation, the finite element method is used [3,8]. The solution P_0 will be used as the input to solve the remaining equations. Eqs. (16), (17), and (18) form three sets of simultaneous equations governing three separate pairs of dimensionless variables (P_{zr}, P_{zi}), (P_{xr}, P_{xi}), and (P_{yr}, P_{yi}) respectively. Luckily, the Eqs. (16), (17), and (18) can be solved according to literature [9].

Solving the model by MATLAB. As shown in Fig. 4, the four-noded isoparametric element with a total of 2231 nodes is used for the entire computational domain. Such an element mesh provides a good balance between numerical accuracy and computational efficiency. Special note: The boundaries of grooves must coincide with the boundaries of elements. The centre of restrictive orifice feed holes are must located at the nodes.

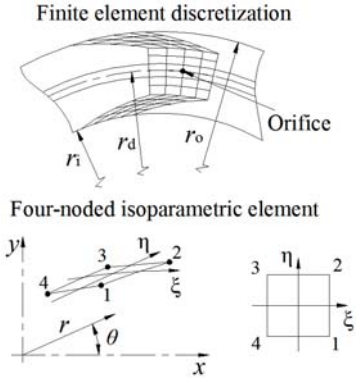


Fig. 4 Finite element discretization model

4 Gas Film Stiffness and Damping Coefficients

Now the axial force (ΔF_z) and angular moments (ΔM_x and ΔM_y) in response to the perturbation Δh can be expressed in the following form:

$$\begin{pmatrix}\Delta F_z\\ \Delta M_x\\ \Delta M_y\end{pmatrix}=\int_0^{2\pi}\int_{r_i}^{r_o}\begin{pmatrix}1\\ r\sin(\theta)\\ -r\cos(\theta)\end{pmatrix}\Delta p r\,dr\,d\theta \tag{19}$$

Substituting Eq. (6) into Eq. (19)

$$\begin{pmatrix}\Delta F_z\\ \Delta M_x\\ \Delta M_y\end{pmatrix}=\int_0^{2\pi}\int_{r_i}^{r_o}\begin{bmatrix}p_z & p_x & p_y\\ r\sin(\theta)p_z & r\sin(\theta)p_x & r\sin(\theta)p_y\\ -r\cos(\theta)p_z & -r\cos(\theta)p_x & -r\cos(\theta)p_y\end{bmatrix}r\mathrm{d}r\mathrm{d}\theta\begin{bmatrix}\Delta z(t)\\ \Delta\beta(t)\\ \Delta\beta(t)\end{bmatrix}+\int_0^{2\pi}\int_{r_i}^{r_o}\begin{bmatrix}p_{\dot z} & p_{\dot x} & p_{\dot y}\\ r\sin(\theta)p_{\dot z} & r\sin(\theta)p_{\dot x} & r\sin(\theta)p_{\dot y}\\ -r\cos(\theta)p_{\dot z} & -r\cos(\theta)p_{\dot x} & -r\cos(\theta)p_{\dot y}\end{bmatrix}r\mathrm{d}r\mathrm{d}\theta\begin{bmatrix}\Delta\dot z(t)\\ \Delta\dot\beta(t)\\ \Delta\dot\beta(t)\end{bmatrix} \tag{20}$$

Because the gas film possesses certain stiffness and damping properties, so the dynamic forces are given by

$$\begin{pmatrix}\Delta F_z\\ \Delta M_x\\ \Delta M_y\end{pmatrix}=-\begin{bmatrix}k_{zz} & k_{zx} & k_{zy}\\ k_{xz} & k_{xx} & k_{xy}\\ k_{yz} & k_{yx} & k_{yy}\end{bmatrix}\begin{bmatrix}\Delta z(t)\\ \Delta\alpha(t)\\ \Delta\beta(t)\end{bmatrix}-\begin{bmatrix}c_{zz} & c_{zx} & c_{zy}\\ c_{xz} & c_{xx} & c_{xy}\\ c_{yz} & c_{yx} & c_{yy}\end{bmatrix}\begin{bmatrix}\Delta\dot z(t)\\ \Delta\dot\alpha(t)\\ \Delta\dot\beta(t)\end{bmatrix} \tag{21}$$

where k_{ij}, c_{ij} (i, j=z, x, y) are stiffness and damping coefficients, the negative sign reflects blocking perturbation motion.

Comparing Eqs. (20) and (21), the stiffness and damping coefficients are easily identified. The dimensionless quantities can be expressed in the following form:

$$\begin{bmatrix}K_{zz} & K_{zx} & K_{zy}\\ K_{xz} & K_{xx} & K_{xy}\\ K_{yz} & K_{yx} & K_{yy}\end{bmatrix}=-\int_0^{2\pi}\int_1^{R_o}\begin{bmatrix}P_{zr} & P_{xr} & P_{yr}\\ R\sin(\theta)P_{zr} & R\sin(\theta)P_{xr} & R\sin(\theta)P_{yr}\\ -R\cos(\theta)P_{zr} & -R\cos(\theta)P_{xr} & -R\cos(\theta)P_{yr}\end{bmatrix}R\mathrm{d}R\mathrm{d}\theta$$
$$\begin{bmatrix}C_{zz} & C_{zx} & C_{zy}\\ C_{xz} & C_{xx} & C_{xy}\\ C_{yz} & C_{yx} & C_{yy}\end{bmatrix}=-\frac{1}{\Gamma}\int_0^{2\pi}\int_1^{R_o}\begin{bmatrix}P_{zi} & P_{xi} & P_{yi}\\ R\sin(\theta)P_{zi} & R\sin(\theta)P_{xi} & R\sin(\theta)P_{yi}\\ -R\cos(\theta)P_{zi} & -R\cos(\theta)P_{xi} & -R\cos(\theta)P_{yi}\end{bmatrix}R\mathrm{d}R\mathrm{d}\theta \tag{22}$$

The dimensional quantities of the stiffness and damping coefficients are given by

$$\left.\begin{aligned}k_{zi}&=\frac{K_{zi}p_a r_i^2}{h_l},\ k_{xi}=\frac{K_{xi}p_a r_i^4}{h_l},\ k_{yi}=\frac{K_{yi}p_a r_i^4}{h_l}\\ c_{zi}&=\frac{C_{zi}p_a r_i^2}{h_l\omega},\ c_{xi}=\frac{C_{xi}p_a r_i^4}{h_l\omega},\ c_{yi}=\frac{C_{yi}p_a r_i^4}{h_l\omega}\end{aligned}\right\}(i=z,x,y) \tag{23}$$

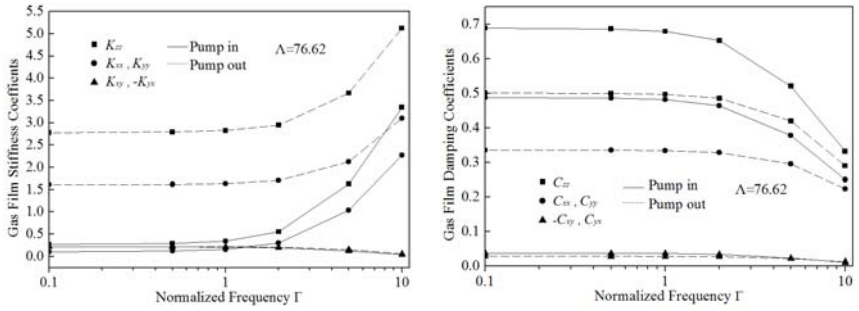
This paper and literature [9] have the same conclusion about the dynamic force coefficients: (1) All cross coupling force and moment coefficients between the axial mode and angular modes are zero, demonstrating decoupling of the axial mode and angular modes. (2) Due to the symmetric nature of the seal and the fact that it undergoes rigid body motions, the diagonal elements of the moment coefficients are identical and the off-diagonal elements are equal in magnitude but opposite in signs. With these features, the dynamic coefficients can be written in a simplified form

$$\begin{bmatrix}K_{zz} & K_{zx} & K_{zy}\\ K_{xz} & K_{xx} & K_{xy}\\ K_{yz} & K_{yx} & K_{yy}\end{bmatrix}=\begin{bmatrix}K_{zz} & 0 & 0\\ 0 & K_{xx} & K_{xy}\\ 0 & K_{yx} & K_{yy}\end{bmatrix}\quad\begin{bmatrix}C_{zz} & C_{zx} & C_{zy}\\ C_{xz} & C_{xx} & C_{xy}\\ C_{yz} & C_{yx} & C_{yy}\end{bmatrix}=\begin{bmatrix}C_{zz} & 0 & 0\\ 0 & C_{xx} & C_{xy}\\ 0 & C_{yx} & C_{yy}\end{bmatrix} \tag{24}$$

The dynamic force coefficients of EDHHGS for a wide frequency range (from Γ=0.1 to 10) are computed and plotted in Fig. 5. One of the prominent characteristics of these coefficients is that they are all frequency dependent. It is seen that the stiffness and damping coefficients are less sensitive to the frequency for Γ up to 1. For Γ>1, the diagonal stiffness coefficient increases with frequency, and the off-diagonal moment coefficients are much smaller and decreases with frequency, while the diagonal damping coefficients drop quickly with increasing frequency, and the off-diagonal damping coefficients are basically unchanged with frequency.

Perturbation frequency has the same influences on the gas film dynamic force coefficients for pump in and pump out. The off-diagonal stiffness and damping coefficients of two seals are almost identical at the same perturbation frequency. For two seals, the diagonal stiffness coefficients of pump in are smaller, but the diagonal damping coefficients are larger.

In view of the theoretical research for mechanical gas face seals focus on spiral groove face seals, we make a comparison with the results for a spiral groove face seals by Ruan [9] to verify the current model. The correlation between the two models is fairly good as shown in Fig. 6.



(a) Gas film stiffness coefficients (b) Gas film damping coefficients

Fig. 5 Influence of perturbation frequency on gas film dynamic force coefficients

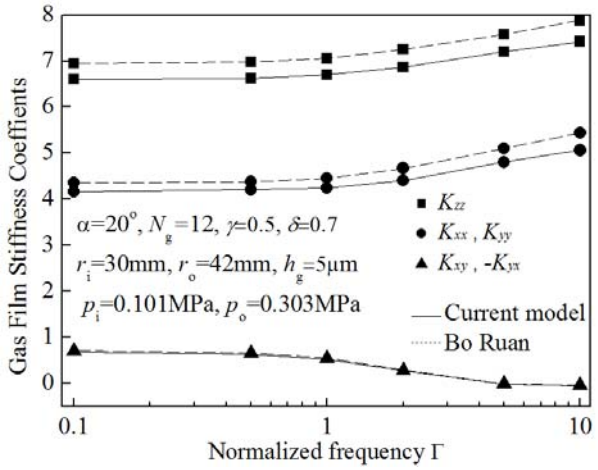


Fig. 6 Literature validation of numerical calculation

5 Tracking Motion of EDHHGS

The tracking motion analysis for EDHHGS belongs to forced vibration category. That is to say, the aim of tracking motion analysis is to solve stator dynamic response due to a prescribed rotor motion. Owing to the response frequency of stator is same as the forcing frequency, we can use the forcing frequency to get the dynamic force coefficients of gas film in the Fig. 5. The stiffness and damping coefficients are less sensitive to the frequency for Γ up to 1 on the one hand, on the other hand, the rotor is rigidly mounted to the rotating shaft, which easily leads to the case that the forcing frequency is equal to the shaft rotational speed. So for the steady-state synchronous tracking motion, only one frequency (Γ=1) is considered here. The axial mode and angular modes are decoupling as discussed above. The gas film of the seal can now be represented by a spring-damper system, and the corresponding equations of motion for the seal can be expressed in terms of the gas film stiffness and damping as well as that of the secondary O-ring and spring, that is,

$$m\ddot{z} + \left(c_{zz} + c_s\right)\dot{z} + \left(k_{zz} + k_s\right)z = c_{zz}\dot{z}_r + k_{zz}z_r \tag{25}$$

for the axial motion and

$$\begin{bmatrix} I_x & 0 \\ 0 & I_y \end{bmatrix}\begin{bmatrix} \ddot{\alpha} \\ \ddot{\beta} \end{bmatrix}+\begin{bmatrix} c_{xx}+c_{sx} & c_{xy} \\ c_{yx} & c_{yy}+c_{sy} \end{bmatrix}\begin{bmatrix} \dot{\alpha} \\ \dot{\beta} \end{bmatrix}+\begin{bmatrix} k_{xx}+k_{sx} & k_{xy} \\ k_{yx} & k_{yy}+k_{sy} \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \end{bmatrix}=\begin{bmatrix} c_{xx} & c_{xy} \\ c_{yx} & c_{yy} \end{bmatrix}\begin{bmatrix} \dot{\alpha}_r \\ \dot{\beta}_r \end{bmatrix}+\begin{bmatrix} k_{xx} & k_{xy} \\ k_{yx} & k_{yy} \end{bmatrix}\begin{bmatrix} \alpha_r \\ \beta_r \end{bmatrix} \quad (26)$$

for the angular motion about x and y axes. In Eqs. (25) and (26), the rotor axial motion z_r is caused by the axial runout and the rotor angular motion α_r, β_r are caused by the angular misalignment, m is stator mass, I_x and I_y are transverse moments of inertia of the stator (as shown in Eq. (27)), k_s is the axial stiffness of the spring, c_s is axial damping of the O-ring, k_{sx} and k_{sy} are angular stiffness of the spring, c_{sx} and c_{sy} are angular damping of the O-ring. The angular stiffness and damping are given in Eq. (28) [10].

$$I_x = I_y = \oint r^2 \mathrm{d}m \quad (27)$$

$$\begin{aligned} k_{sx} &= k_{sy} = 1/2 k_s r_s^2 \\ c_{sx} &= c_{sy} = 1/2 c_s r_s^2 \end{aligned} \quad (28)$$

Without the loss of generality, the following motion is assumed for the rotor:

$$z_r = A_{rz}\sin(\omega t), \alpha_r = A_{rx}\cos(\omega t), \beta_r = A_{ry}\sin(\omega t) \quad (29)$$

The initial conditions for the stator that define the quasi-equilibrium configuration are

$$z(0) = h_0, \dot{z}(0) = \dot{\alpha}(0) = \beta(0) = \dot{\beta}(0) = 0, \alpha(0) = A_{rx} \quad (30)$$

To solve Eq. (25), a Laplace transform is performed as the following:

$$Z(s) = \frac{(c_{zz}s + k_{zz})Z_r(s)}{[ms^2 + (c_{zz} + c_s)s + (k_{zz} + k_s)]} \quad (31)$$

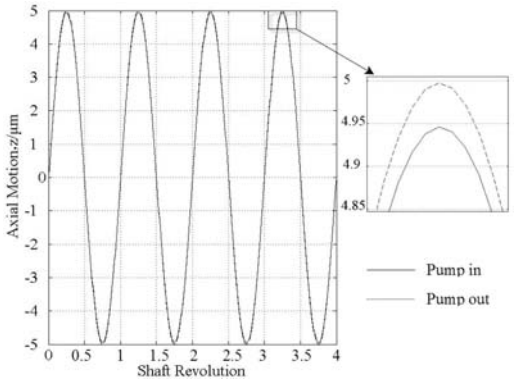
Similar derivation can be made for the angular modes of Eq. (26) with the following Laplace transform equations

$$\begin{Bmatrix} A(s) \\ B(s) \end{Bmatrix} = D^{-1}\left\{\begin{bmatrix} c_{xx}s + k_{xx} & c_{xy}s + k_{xy} \\ c_{yx}s + k_{yx} & c_{yy}s + k_{yy} \end{bmatrix}\begin{Bmatrix} A_r(s) \\ B_r(s) \end{Bmatrix} - \begin{bmatrix} c_{xx} & c_{xy} \\ c_{yx} & c_{yy} \end{bmatrix}\begin{Bmatrix} \alpha_r(0) \\ \beta_r(0) \end{Bmatrix} + \begin{bmatrix} I_x s + c_{xx} + c_{sx} & c_{xy} \\ c_{yx} & I_y s + c_{yy} + c_{sy} \end{bmatrix}\begin{Bmatrix} \alpha(0) \\ \beta(0) \end{Bmatrix}\right\} \quad (32)$$

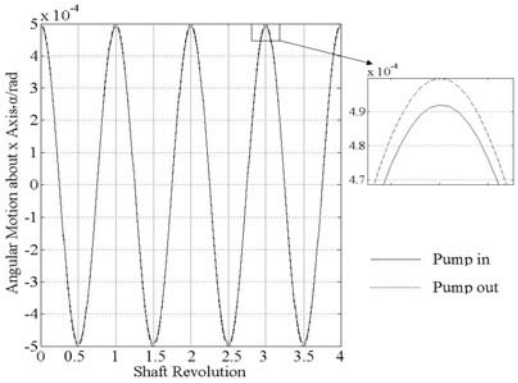
where

$$D = \begin{bmatrix} I_x s^2 + (c_{xx} + c_{sx})s + k_{xx} + k_{sx} & c_{xy}s + k_{xy} \\ c_{yx}s + k_{yx} & I_y s^2 + (c_{yy} + c_{sy})s + k_{yy} + k_{sy} \end{bmatrix}$$

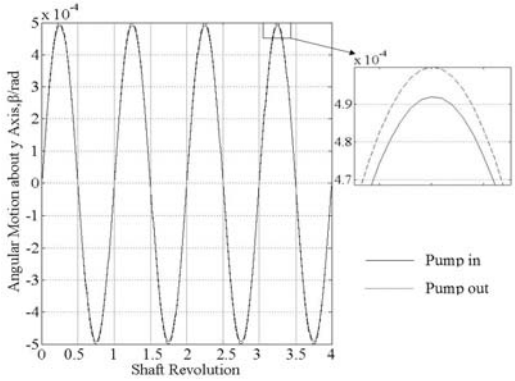
The time responses of these motions can now be recovered by an inverse Laplace transform for both Eqs. (31) and (32) numerically. The axial and angular motions are plotted as a function of shaft revolution in Fig. 7. Both the axial and angular responses are in synchronization with the rotor motions. The amplitude of the stator motion is only slightly reduced, and will not cause direct face contact. The results also show that the effect of the initial conditions on the tracking motion is quickly damped out and hardly noticeable. From the Fig. 7 we can see that the tracking motion of pump out is better than pump in.



(a) Axial motion



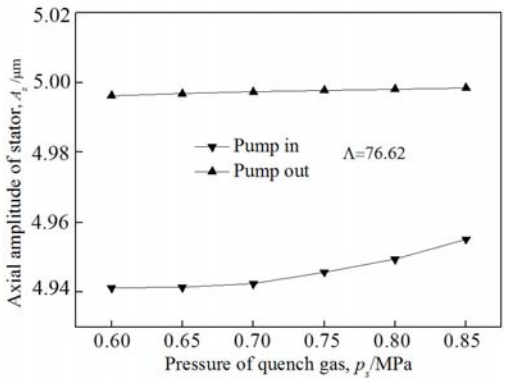
(b) Angular motion about x axis



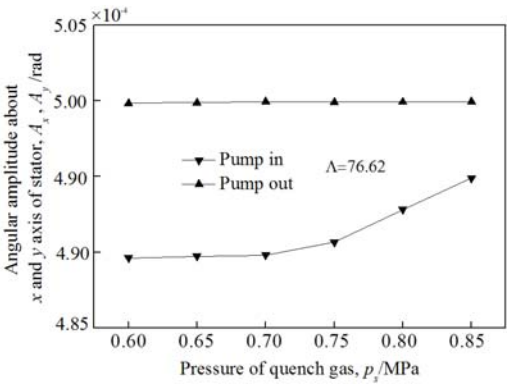
(c) Angular motion about y axis

Fig. 7 Axial and angular responses of stator

The value of quench gas pressure is restricted by many factors which should be neither too high (seal instability or leakage rate is too large) nor too low (seal face is hard to open). Generally, the value of quench gas pressure is 0.05~0.25MPa higher than seal cavity pressure. Therefor, in the current study, the influence of quench gas pressure (from 0.60MPa to 0.85MPa) on the amplitude of stator is investigated and plotted in Fig. 8. It is seen that the axial and angular amplitudes increase with quench gas pressure for pump in, while the axial and angular amplitudes are less sensitive to quench gas pressure for pump out. From the Fig. 8 we can also see that the amplitude of pump out is higher than pump in at the same quench gas pressure.



(a) Axial amplitude

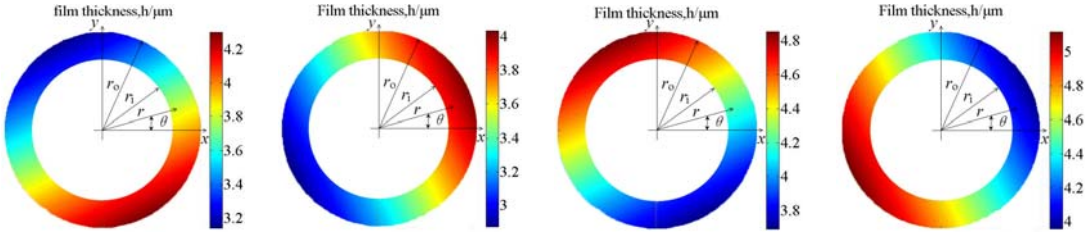


(b) Angular amplitude

Fig.8 Influence of quench gas pressure on amplitudes of stator

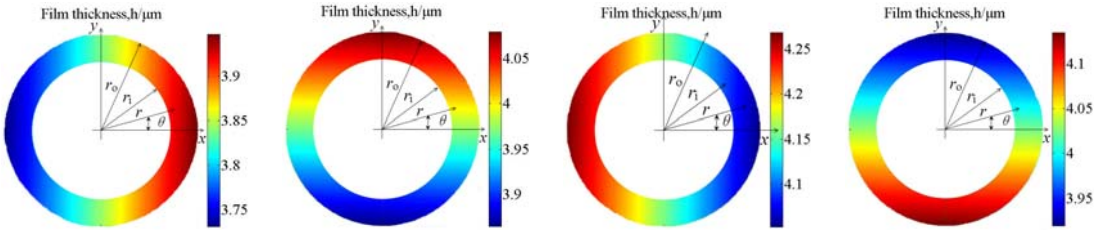
The gas film thickness (ignoring groove depth) of four times are shown in Fig. 9. For pump in, the maximum gas film thickness is 5.13 μm and appears at t=(n+0.75)T moment, and the minimum gas film thickness is 2.87 μm and appears at t=(n+0.25)T moment. For pump out, the maximum gas film thickness is 4.27 μm and appears at t=(n+0.5)T moment, and the minimum gas film thickness is

3.73 μm and appears at t=nT moment. Therefore, the gas film of EDHHGS is stable even though the perturbation amplitudes of rotor are large enough. The EDHHGS will not cause direct face contact.



(a1) t=nT (a2) t=(n+0.25)T (a3) t=(n+0.5)T (a4) t=(n+0.75)T

(a) Pump in



(a1) t=nT (a2) t=(n+0.25)T (a3) t=(n+0.5)T (a4) t=(n+0.75)T

(b) Pump out

Fig. 9 Film thickness (n is natural number, T=2π/ω)

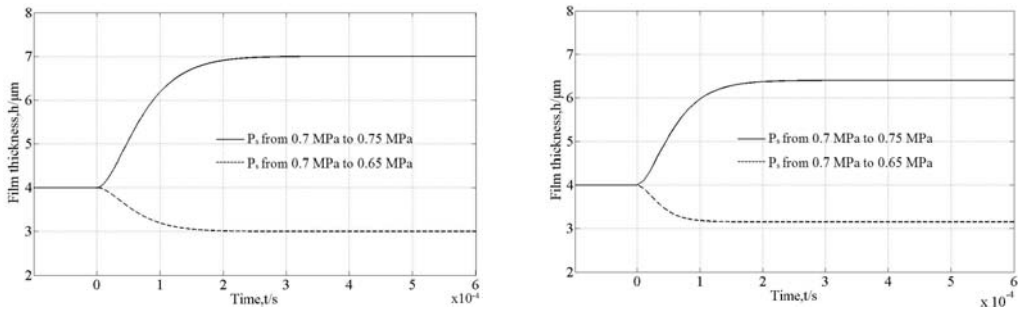
6 Stability of axial free vibration for the active regulation

The sealing performance parameters of EDHHGS could be changed through quench gas pressure adjustment online (active regulation). That is to say, the sealing clearance can be regulated by changing the quench gas pressure p_s. One of the advantages of active regulation is that the tracking motion can be enhanced to avoid excessive leakage or face wear. In addition, another advantage of active regulation is that the gas film stiffness and damping coefficients can be changed to optimize dynamic property of EDHHGS.

The quench gas pressure adjustment online (the pressure from p_{s1} to p_{s2}) can lead to axial free vibration of stator. The equilibrium position of axial free vibration is the new equilibrium position corresponds to the quench gas pressure p_{s2}. Because the closing force is constant, the new equilibrium position can be solved. Based on the mentioned method, the variation of the gas film stiffness and damping coefficients with frequency at new equilibrium position can be obtained. The equation of axial free vibration for the stator can be expressed by

$$m\ddot{z}+\left(c_{zz}+c_s\right)\dot{z}+\left(k_{zz}+k_s\right)z=0 \tag{33}$$

Due to the frequency-dependency of the gas film stiffness and damping coefficients, solution to the Eq. (33) in general requires time integration technique. That is to say, the Eq. (33) and lubrication equation should be solved simultaneously. As shown in Fig. 10, the quench gas pressure is changed at zero moment. The damping ratio of axial free vibration of stator approximately equals to 1.0 for the active regulation, so the motion of stator is non-reciprocating attenuation vibration. The two seals can quickly reach the new equilibrium position, and will not be unstable in the axial free vibration. Compared with the pump in, the pump out spends less time to reach the new equilibrium position.



(a) Pump in (b) Pump out

Fig. 10 Active regulation lead to film thickness changes

In order for an undamped axial free vibration to occur, the parts of the damping coefficients (in Eq. (33)) must vanish at the free frequency. Thus, the criteria for seal stability can be stated as the following:

$$c_{zz}\big|_{\upsilon=\upsilon_0} = -c_s \tag{34}$$

where υ_0 is the critical frequency. The critical mass is given by

$$m_0 = \frac{k_{zz}\big|_{\upsilon=\upsilon_0} + k_s}{\upsilon_0^2} \tag{35}$$

It is readily clear that the computed axial damping coefficients c_{zz} (Fig. 5(b)) in the specific example provided never become negative in the frequency range examined (up to Γ=10). Therefore, the seal will not be unstable in the axial free vibration of active regulation according to the stability criterion.

7 Conclusions

The gas film stiffness and damping coefficients for external-pressurized dynamic-hydrostatic hybrid gas seals are obtained from the unsteady nonlinear Reynolds equation using the perturbation method. The influence of the perturbation frequency on dynamic characteristics coefficients is analyzed.

The three-dimensional dynamic tracking motion of dynamic-hydrostatic hybrid gas seals is analyzed under the simple harmonic excitation caused by the rotor motion due to its runout and misalignment. Results showed that the gas film of dynamic-hydrostatic hybrid gas seals is stable even though the perturbation amplitudes of rotor are large enough, and the dynamic-hydrostatic hybrid gas seals will not cause direct face contact.

The quench gas pressure adjustment online (the pressure from p_{s1} to p_{s2}) can lead to axial free vibration of stator. The damping ratio of axial free vibration of stator approximately equals to 1.0, so the motion of stator is non-reciprocating attenuation vibration. Further analysis showed that the dynamic-hydrostatic hybrid gas seals will not be unstable in the axial free vibration of active regulation.

It should be cautioned that although this method is effective in solving the tracking motion, it is only applicable in the case of a quasi steady-state tracking at a given frequency. For transient motions or motions involving different frequencies, this method is inapplicable due to the frequency dependency of the gas film stiffness and damping.

Nomenclature

A_{rz}	=	rotor axial runout
A_{rx} , A_{ry}	=	rotor misalignment
B	=	balance ratio
c_s	=	axial damping of O-ring
c_{sx} , c_{sy}	=	angular damping of O-ring, $1/2\ c_s(r_s)^2$
c_{ij}	=	gas film damping (i, j=z, x, y)
C_{ij}	=	dimensionless gas film damping (i, j=z, x, y)
C_p , C_v	=	specific heat at constant pressure and specific heat at constant volume, respectively
d	=	diameter of restrictive orifice feed hole
$\vec{e}_\theta$	=	unit vector in the angular direction
f, g_r	=	radial width as shown in Fig. 2
ΔF_z	=	dynamic force
h	=	film thickness
h_d	=	static pressure equalizing groove depth
h_g	=	double-divert spiral grooves depth
h_l	=	steady-state equilibrium film thickness of dam area
h_0	=	steady-state equilibrium film thickness
H_0	=	dimensionless steady-state equilibrium film thickness, h_0/h_l
I_x , I_y	=	stator transeverse moment of inertia
k	=	gas specific heat ratio, C_p/ C_v
k_m	=	perturbation gas mass flow through unit section of the restrictive orifice feed holes, and caused by unit perturbation pressure
k_s	=	axial stiffness of spring
k_{sx} , k_{sy}	=	angular stiffness of spring, $1/2\ c_s(r_s)^2$
k_{ij}	=	gas film stiffness (i, j=z, x, y)
K_{ij}	=	dimensionless gas film stiffness (i, j=z, x, y)
l, g_θ	=	circumferential width of the land and the groove
m	=	stator mass
$\dot{m}$	=	gas mass flow through unit section of the restrictive orifice feed holes
$\dot{m}_0$	=	steady-state gas mass flow through unit section of the restrictive orifice feed holes
ΔM_x , ΔM_x	=	dynamic moments about x and y axes, respectively
n	=	shaft speed
N_g	=	number of double-divert spiral grooves
N_o	=	number of restrictive orifice feed hole
p	=	gas pressure

p_a = atmospheric pressure
p_d = exit total pressure of the restrictive orifice feed holes
p_{d0} = steady-state exit total pressure of the restrictive orifice feed holes
p_i = pressure at ID boundary
p_o = pressure at OD boundary
p_s = pressure of quench gas
p_0 = pressure at equilibrium piont h_0
P_0 = dimensionless pressure at equilibrium piont h_0, p_0/p_a
r_d = radial position of orifice and pressure equalizing groove
r_i = seal ring inner radius
r_o = seal ring outer radius
r_{gyr} = radius of gyration
r_s = radius on which O-ring is installed
R = gas constant, dimensionless radial coordinates, r/r_i
s = Laplace variable
t = time
T = gas temperature
W_d = static pressure equalizing groove width
z = stator axial displacement
z_r = rotor axial displacement
$Z(s)$, $Z_r(s)$ = Laplace transform of $z(t)$, $z_r(t)$
$A(s)$, $B(s)$ = Laplace transform of $\alpha(t)$, $\beta(t)$
$A_r(s)$, $B_r(s)$ = Laplace transform of $\alpha_r(t)$, $\beta_r(t)$
(x, y) = Cartesian coordinates
(r, θ) = polar coordinates
α = spiral groove angle, stator angular motion about x axes
α_r = rotor angular motion about x axes
β = stator angular motion about y axes
β_r = rotor angular motion about y axes
γ = land-to-groove ratio, l/g_θ
δ = sealing groove ratio, g_r/f
δ_j = Kronecker number, at the restrictive orifice feed holes is one, otherwise is zero,
μ = gas viscosity
υ = excitation frequency
ϕ = flow coefficient, generally takes 0.8
ψ_0 = flow function
ω = shaft rotational speed
Λ = compressibility number, $6\mu\omega(r_i)^2/(p_a(h_l)^2)$
Γ = dimensionless excitation frequency, υ/ω
$\bar{\nabla}$ = polar coordinate gradient operator

References

[1] Zhang S Q, Wang L, Zhao W G. Research on temperature field and heat deformation of mechanical seal in liquid rocket engine turbopump. Journal of Rocket Propulsion, 2014, 40(5): 92-98. (in Chinese)

[2] Zhang S Q, Li S X, Cai J N, et al. Numerical solution for the tracking property and active regulation vibration characteristic of dynamic-hydrostatic hybrid gas seals. Acta Aeronautica et Astronautica Sinica, 2012, 33(7): 1336-1346. (in Chinese)

[3] Zhang S Q. Dynamic behaviors analysis of dynamic-hydrostatic hybrid lubrication gas face seal. Beijing: Beijing University of Chemical Technology, 2012. (in Chinese)

[4] Berard G, Zheng X Q. Analysis and design of a double-divert spiral groove seal. AIAA-2006-4753, 42nd AIAA/ASME/SAE/ASEE Joint Propulsion Conference. 2006.

[5] Zheng X Q, Berard G. Development of non-contacting, low-leakage, large-diameter air sea. AIAA-2008-4507, 44nd AIAA/ASME/SAE/ASEE Joint Propulsion Conference. 2008.

[6] Wilbur S, Robert H. Seal technology for liquid oxygen turbopumps. NASA CR-174866, 1985.

[7] Liu T, Liu Y H, Chen S I. Hydrostatic gas lubrication. Harbin: Harbin Institute of Technology Press, 1990: 122-126. (in Chinese)

[8] Ruan, B. Finite element analysis of the spiral groove gas face seal at the slow speed and the low pressure conditions-slip flow consideration. Tribology Transactions, 2000, 43(3): 411-418.

[9] Ruan, B. A semi-analytical solution to the dynamic tracking of non-contacting gas face seals. Journal of Tribology, 2002, 124(1): 196-202.

[10] Green I, Etsion I. Stability threshold and steady-state response of noncontacing coned-face Seals. ASLE Transactions, 1985, 28(4): 449-460.

Research on Dynamic Simulation of an Aerospace pyrotechnic device

YE Yao-Kun[1], LI Xiao-gang [2] ,Ding Feng[1] ,Man Jian-feng[1]

[1]Beijing Institute of Spacecraft System Engineering, Beijing 100094； [2] Beijing Institute of Technology, Beijing 100081 [a]yeyaokun@163.com

Abstract: Based on 3D-LS DYNA finite element model, it makes a explosion dynamics simulation analysis of a typical explosive bolts to study the inside product structural response during the process of doing work. To explore the weaknesses of product structure in the design process through analyzing the stress distribution of the bolt body in product processes, it cannot only provide effective support for the theory of products optimal design, but also offer a useful reference for the simulation and design optimization of similar aerospace pyrotechnic device.

Key Words: aerospace pyrotecshnic device dynamics numerical simulation 3D-LSDYNA

CLC number: TJ450 WM: A Article ID

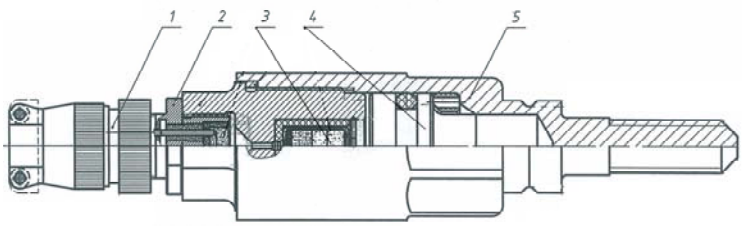
0 Introduction

As a typical aerospace pyrotechnic device, explosive bolt is commonly used in link and unlock between the satellite and the rocket, as well as spacecraft cabins, which initially plays a major role in mechanical connection, and then unlocks satellite and the rocket or spacecraft cabin after receiving the predetermined separation command, according to the detonation role of inside explosives. Therefore, the function of explosive bolts determines the success of spacecraft systems tasks. Generally, the unlock function of explosive bolts are mainly used the detonation wave generated by internal detonation explosive to push the piston move, resulting in broken of bolt body along its preset grooves place, leading to separation of two connected objects, so as to achieve the separation purposes. In the unlocked separation process, it is impossible to observe the internal organization movement and the forces to bolt body during exercise in the actual test, thus it is necessary to analyze the mechanism of unlock action and key dynamics parameters by means of numerical simulations.

Based on the 3D-LSDYNA numerical simulation of finite element model, the study replays the transient phenomenon detonation under closed state, and gives stress distribution curves of explosive bolts or internal structures during the explosive process. It provides technical reference for the optimal design of the product by researching the internal system structural response during work done process and analyzing the differences in morphological changes after the product doing work

1 Work principle analysis

It shows in figure 1 that the explosive bolt is composed by electrical connector, the igniter, the main charge, the piston, main body and other components. The work principle is that electric powered firearms produce flame by warming up, and ignites flame detonator through the fire-hole to output detonation wave accelerating the piston to shock the bolt body, breaking the shell at a preset channel slot, to achieve separation.



1. electrical connector； 2. igniter； 3. main charge； 4. piston； 5. main body

Fig.1 Configuration of explosive bolt

2 Finite Element Simulation

2.1 Finite Element Model

The work process of explosive bolts is a transient impact process, so the simulating calculation needs to describe the whole work process of charge system. Because the bolt body and its related components are axis-symmetric structure, it builds the finite element model in a quarter of the bolt in order to reduce the computational model and the computing cycle. The 1/4 sectional view of the bolt shows in Figure 2.

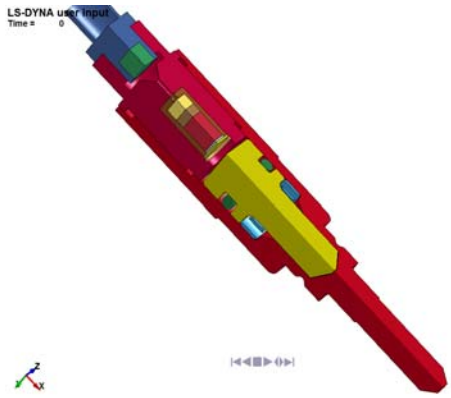
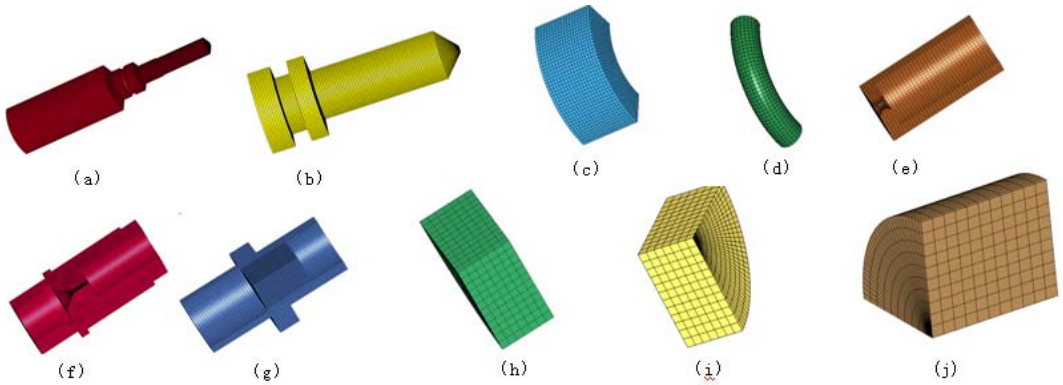


Fig. 2 the 1/4 section view of explosive bolts

This study makes a related calculating research on detonation transfer process of non-electric detonation transfer system by using numerical simulation software ANSYS/LS-DYNA. It establishes model by ANSYS, does explicit solver by LS-DYNA, and then deals with the simulation result using the standard LS-PREPOST. Explicit finite element code LS-DYNA and ANSYS program powerful combination after treatment obtains effective combination in the software. It can quickly solve large deformation dynamics, large deformation and non-linear quasi-static problem of multiple and complex exposure collision with LS-DYNA explicit transient algorithm. It can also pass geometric information and results information between ANSYS and ANSYS-LS-DYNA to perform continuous implicit - explicit / explicit - implicit analysis to apply in numerical simulation of detonation problems research.

2.2 Mesh Generation

This research built the basic finite element computation model by splits the parts at first, and then combined with face cutting method, meshing each part based on hexahedral and tetrahedral supplemented. The finite element model of each part can be shown from figure 3.



（a）the shell of explosive bolts（b）piston （c）washer（d）seal ring（e）detonator set （f）tube socket （g）igniter shell（h）detonator charge（i）detonator charge (j) igniter charge

Fig. 3The finite element model of explosive bolt's parts

Based on the research method of summarizing, decomposing and summarizing, it can obtain the explosive bolts FEM overall model of 3D-LSDYNA identification after integrating the components finite element model. The size of finite element mesh is about 0.02 ~ 0.04cm, and the total number of grid is more than 100,000.

2.3 Material characteristics

Finite element calculation uses the cm-g-μs as unit, and the derived units are: pressure 10^7N, stress/pressure Mbar, speed 10km / s, density 10^3kg/m^3. During the calculation process, it chooses blank material model for the air field, for explosive charge, selects *MAT_HIGH_EXPLOSIVE_BURN model and JWL state equation for lord sent charge material model, applies the Lee-Tarver ignition and growth reaction rate model of LS-DYNA program for reaction rate for excited drugs, and the material model compatible with the ignition and growth model is blank material or elastic-plastic fluid dynamics material model. Other key material parameters are shown in Table 1.

Table 1 characteristics of main material

element	material	constitutive	Densityρ /(10^3kg/m^3)	elasticity modulus E/GPa	Poisson's ratioλ	yield strength Es/MPa
bolt	30CrMnSiNi2A	bilinear kinematic harden	7.85	207	0.27	1390
piston	30CrMnSiNi2A	bilinear kinematic harden	7.85	207	0.27	1400
socket shell	30CrMnSiNi2A	bilinear kinematic harden	7.85	207	0.27	1000
socket tube	30CrMnSiNi2A	bilinear kinematic harden	7.85	207	0.27	1000

2.4 Charging parameters testing

In this study, there are three drug boxes in the detonator element, which are passivation Lead Styphnate, carboxymethyl cellulose lead azide, and PETN. Although the charge of explosive bolt is difference under different conditions, the press-fit process and press-fit pressure of detonator elements are same, so the charge density of detonator can be no difference. There is impact detonator during the initiation process, therefore, it needs multilayer ignition growth model when doing the numerical calculation by numerical simulation software, to collect and test the explosive speed of related charge, detonation pressure and other parameters.

According to the document data, the detonation velocity of Lead Styphnate is about 3.8 km/s when the density is $2.08\times10^3 kg/m^3$. For lead azide and PETN, the study obtains the result by manganin piezoresistive sensor, which shows in table 2.

Table 2 Detonation velocity result of lead azide and PETN

drug	density/($10^3 kg/m^3$)	detonation velocity /(km/s)	detonation velocity average /(km/s)
lead azide	3.92	4.70	4.696
	3.90	4.66	
	3.91	4.73	
PETN	1.59	7.2	7.116
	1.585	7.01	
	1.60	7.14	

3 The simulation calculation and result analysis

Considered the complex parts and large model features of explosive bolt, it takes the calculation pattern of local first and total later in simulation process. By calculating local part to adjust the indicators of grid model, eliminating the grid distortion loop, endless loop, and other troubles, then it does the overall computing and acting process simulation of system model. It is divided into the following three parts based on the major acting process.

（1）Simulation research on detonation process of ignition to primer

This detonation process is caused by the flame of electrical ignition tube through the igniting hole to fire the detonator which outputs a strong detonation. Relatively speaking, the impact strength of the flame is small and much smaller to detonator chamber after decayed by igniting hole, therefore, the function of electric ignition tube can be weaken by simplify its structure, replaced by high explosive with same energy, simulating this process by ignition - growth model. The pressure nephogram of igniter tube after igniting shows in figure 4. After igniting the igniter tube, the chemical reactions output energy to the detonator through igniting hole which can blast the Lead Styphnate. This process can be expressed in figure 5, where the red part is representing the high-pressure wave front passing the hole. After that, the detonation wave gradually grows by firing explosives and contains stability that leads to a detonating impact to the following charge, and the pressure nephogram of this process can be seen in figure 6. The entire ignition and shock initiation process can be described by depicting pressure curve of detonator charge, like it shows in figure 7, selecting several finite elements along the axle wire of detonator’s three charge from top to bottom, drawing the pressure curve, from which can be observed that the pressure of detonator near the end face of the igniter tube begins to appear after 2μs, which indicates that the igniter tube acts successfully to start the detonation. As the time past, pressure peak of the charge is increasing, and reaches maximum at the time of 3.3μs, then the pressure dropped suddenly, which means the depletion of the first stage charge. Followed that, a escalating pressure wave shows up, which declares that the next stage charge is detonated and the energy transferred successfully. The pressure on each finite element shows a similar variation which is rapidly raised first and then dropped down, with a concussion secondary peak following, where the first peak is the pressure caused by shock wave passing, the next peak is the impact and pressure generated by the protons in the cavity after wave shock.

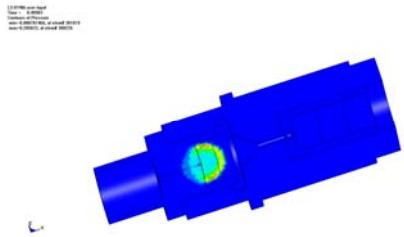


Fig. 4 the pressure nephogram after igniting

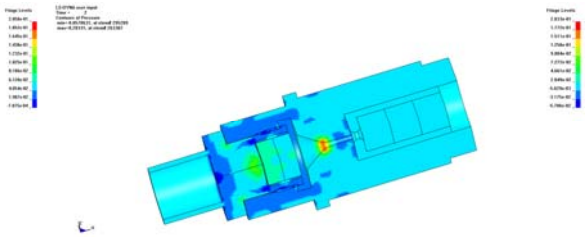


Fig. 5 the pressure nephogram of fire and initiation process

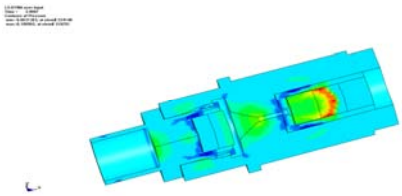


Fig. 6 Ignition growth simulation picture

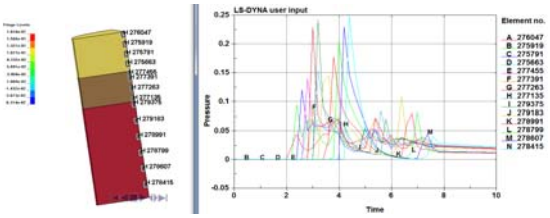


Fig. 7 pressure course curve

（2）Simulation research on the process of detonator push piston movement

The lead azide of detonator is fired by Lead Styphnate and output energy forward to PETN. The output energy falloff after pressing on gaskets, air and other media then loaded on the piston, whose delivery and pressure loading process is shown in figure 8.

（3）Simulation research on the process of bolt fracture caused by piston movement

The piston speeds up along the direction of force vector after being loaded by shock wave pressure, while the front of the piston presses the bolt body resulting in stretching and shrinking at the presetting V-shaped groove, with ultimately rupture, completing the separation function of explosive bolt. The piston forward movement and rupture process of the explosive bolt is shown in figure 9.

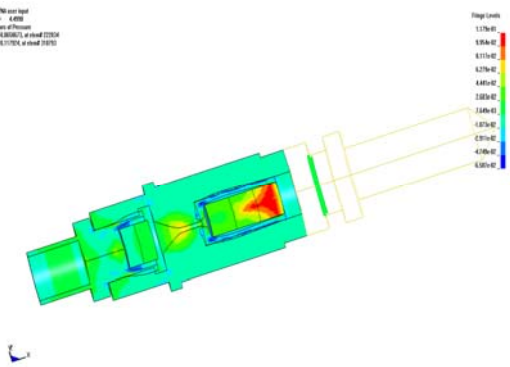


Fig. 8 simulation picture of the pressure loading

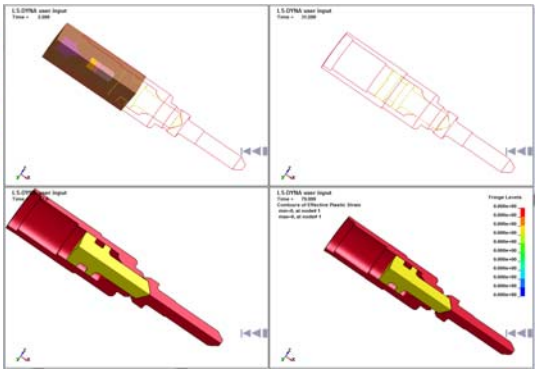


Fig. 9 the simulation model of bolt rupture process

（4）The simulation research on overall effect process

Based on the first three step simulation models, the study establishes an overall explosive bolts model, and sets an air field in the area of drug installed so as to form a complete EULER field. The calculation model with an air field is like the expression of figure 10.

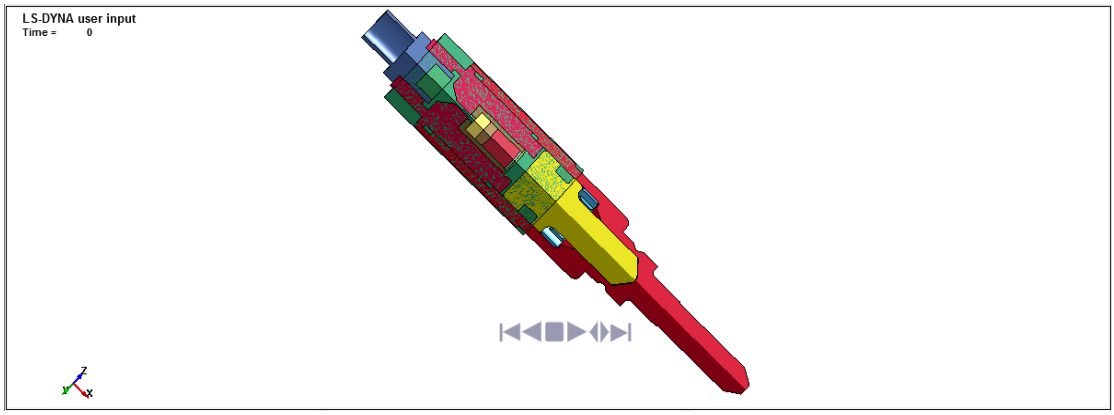


Fig. 10 the overall simulation model of explosive bolt

a. Stress analysis of bolt cavity

The charge of the detonator follows a step of 150mg asphalt passivation lead styphnate, 220mg carboxymethylcellulose lead azide, and 310mg PETN. Because of the impact of explosive device, the explosive bolt generates different pressure in different time from 2.599μs to 5.599μs to 15.99μs and other period during the process of explosive bolt doing work. The different pressure can be seen in figure 11.

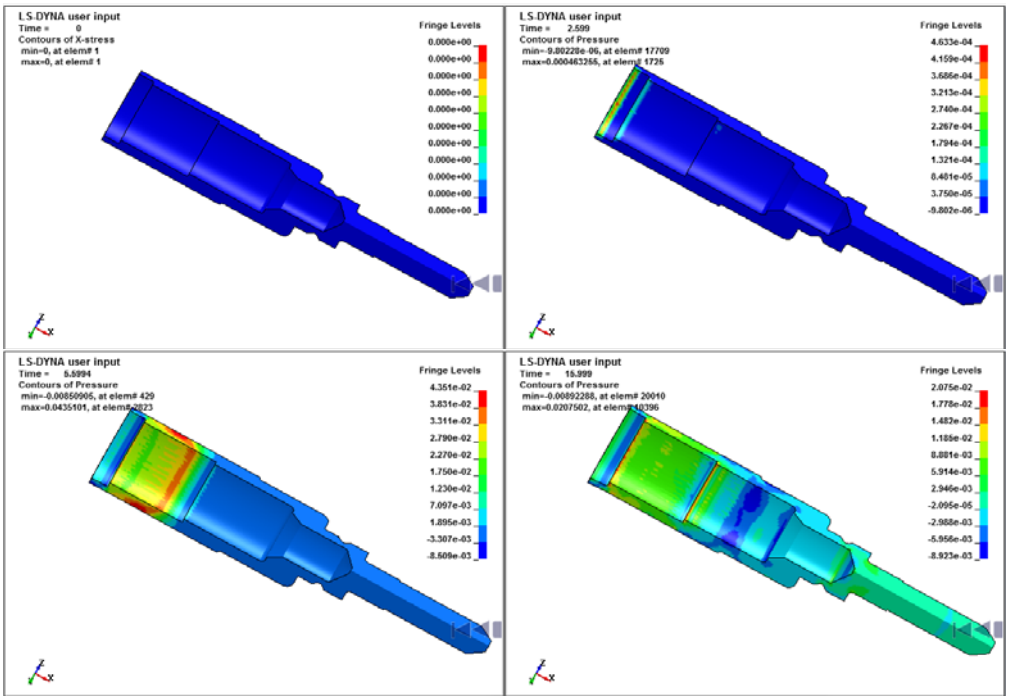


Fig. 11 Stress diagram of bolt body

Select the finite elements from the inner wall of the bolt body to prearranged V-groove like the figure 12, and the stress on each finite element is shown in Figure 13.

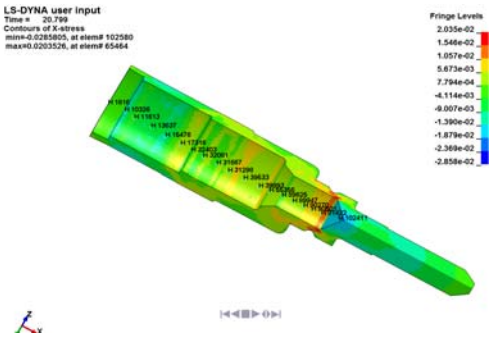


Fig. 12 the way of selecting finite element

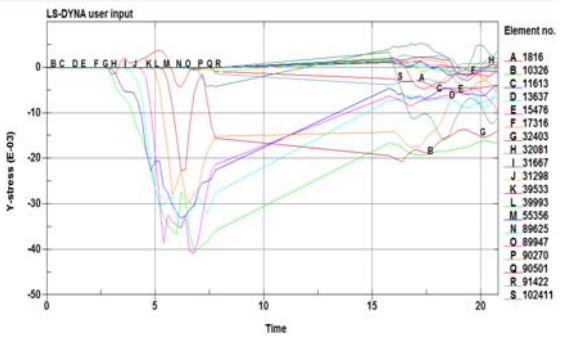


Fig. 13 the stress on each finite element

It can be seen from figure 13 that the stress of finite element A, B, C, D, E, F, and G are relatively close, the stress peak is around 300MPa to 400MPa, but the stress of finite element H and the following points are all bellow 200MPa. It is easy to conclude that the strong stress field overlaps the area of explosive drug installation. The detonation wave of drug explosive directly affects on the bolt wall, resulting in swell and strong impact, so the stress is stronger. Referring the stress calculating method of thin wall cylinder, its stress can be estimated according to the following formula.

$$\sigma=(\Pi D^2/4-\Pi d^2/4)\sigma_b \quad (1)$$

The D and d in the formula respectively represent the inner and outer diameter of cylinder, while the σ_b means the stress of bolt body after heat treatment choosing the number of 1390MPa. For different degrees of impact against the bolt body, as shown in figure 14it divides the main pressed area into a, b, and c three sections according to different thickness, which are 2.75mm, 4.75mm, 4mm in turn.

Fig.14 the wall thickness of bolt body

It can calculate the tensile stress of each section according to formula (1), where the result is shown in table 3.

Table 3 the tensile stress calculation result of bolt body

Sections	Outer diameter (cm)	Inner diameter (cm)	Stress (MPa)
a	2.75	2.2	2972
b	2.75	1.8	4719
c	2.6	1.8	3843

From the data in the table 3, the tensile stress of both section b and c are greater than the level of 3000MPa except section a, which is 2972MPa, less than 3000MPa, so it means that the section b and c usually will not break or lose efficacy before section a. From the single calculation result, section a would have been crack or lose efficacy after strong impact, however, it is the periphery of bolt socket with the wall thickness of 5mm which can bear the tensile or shear stress about 4200MPa, which is much higher than the simulating stress, that is why section a contains available and efficacy. Therefore, the structural strength of the explosive bolts is strong enough to resist the impact of detonation to keep crack and lose efficacy away. However, considered the outer diameter and

thickness of the bolt, it can be compared to the thick-walled cylinder. Under a internal pressure, the plastic zone expands outward from inner wall, forming a body with an inner plastic zone and outer elastic zone. Under the restriction of outer elastic zone, the inner plastic deformation and elastic deformation remains the same standard, the unrestricted plastic flow becomes possible as long as the total cross-section becomes a plastic state. Thus, it is possible that the bolt body turns out crack or loses efficacy if the total cross-section does not become a plastic state during the detonation process, or the plastic zone increased because of the change of pressure and temperature during the storage process.

b. Analysis of piston acting process

The piston is installed in the bolt cavity between the detonator and V-shaped pre-groove, whose front connect with the bolt closely. From the simulation analysis result, it obtains that the explosive wave continues to spread forward after detonation, and loads on the surface of piston after 5.99μs, leading an acceleration of the piston which slides forward along the internal cavity of the bolt body, and after 18μs, the piston reaches a speed of 36m/s. The front of the piston is pressed firmly on the bolt body, so it slows down under a force of reaction when it moves forward, and the speed reduces to 19m/s at the time of 33μs, where formed a concentration stress at the V-shaped pre-groove. With continued advance of the piston, the V-shaped groove begins to stretch and shrink, bringing the lose efficacy of cell at 48μs, while the speed of the piston keeps around 19m/s. After 54μs, the bolt body appears obvious crack zone and the speed of the piston increases slightly, but there is no obvious decrease of stress on the V-shaped groove. With the continuous movement, the bolt breaks intensified, and half of the V-shaped groove breaks at 63μs, with stress value decreased an order of magnitude, the impact of shock wave on piston is stronger than the restriction force of the bolt, so the piston speed increases rapidly. As the piston continues pushing, the last cell of the V-shaped groove fractured after 68.4μs, the piston carries the head of the separate bolt to accelerate together. At the time of 158μs, the doing work process is completed when piston shaft shoulder against the bolt body, with the residual velocity of 94.5m/s. For the deformation of the piston has a friction with the inner wall of the bolt body forming a concentration area with part of the stress, a severe lose efficacy is generated on the cell because of the directly hit of shaft shoulder against bolt body. The process is shown in figure 15.

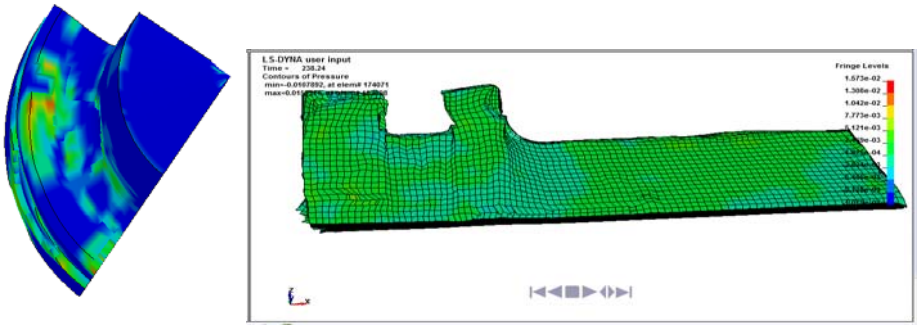


Fig. 15 the effect picture of shaft shoulder damage process after piston work

4 conclusion

In this paper, it can conclude the following results by finite element simulation research on explosive bolt.

(1) After the ignition of explosive bolts can achieve the desired effect of normal separation function, and access to the explosive bolts role during peak stress of about 4000MPa, and explosive

bolts can withstand considerable pressure limit, the function may occur after the explosive bolts to complete the shell body delayed cracking phenomenon;

(2) The bolt-body's separating speed is 94.5m/s after piston completing specified function, and then the remaining energy will cause the piston shaft shoulder crack, it is clear that this phenomenon occurs after the completion of pre-separation function explosive bolts;

(3) The research results provide a useful reference for similar devices' optimization.

Reference

[1]. YANG Jian-zhong, QI Yu-feng, LOU Han-wen. Releasable Unlock Connection and Separation Device for Spacecraft. [J].Spacecraft Engineering, 2003, 12(1)，pp. 47-49.

[2]. HE Chun-quan, YAN Nan, YE Yao-kun. Study of Stage Separation Device for Missile [J].Spacecraft Recovery & Remote Sensing, 2009, 30(3)，pp. 70-77.

[3]. LIU Ming-fang, LIU Sheng-xin. Manual of Mechanical Properties of Metallic Materials [M]. Beijing: Machinery Industry Press, 2011，pp. 525-600.

[4]. JIN Zhi-ming. Firearm Interior Ballistics [M]. Beijing: Beijing Institute of Technology Press, 2004:11-16.

[5]. Hibbitt, Karlsson, Sorensen. Getting Started with ABAQUS/Standard [M]. Pawtucket: Hibbitt, Karlsson & Sorensen, Inc., 2000，pp. 1-50.

[6]. SHA Ding-guo. Error Analysis and Evaluation of Uncertainty [M]. Xi'an: China Metrology Press, 2003，pp. 33-89.

Application of GPS in TT&C Time Synchronization

ZHANG Yanping, SHANG Jianzhong, BAI Yungao, LI Haichao

（Xi'an Satellite Control Center·Xi'an·710043）

Abstract: Sometimes there aren't common view GPS satellites for different TT&C stations in the three-way measurement, especially in the deep-space TT&C. In order to achieve TT&C systems time synchronization under the condition, a method is posted by using the time difference between the satellite clock and GPS system time which is included in the down-way data of GPS satellite. It's proved to be effective by comparing and analyzing the data of A and B stations' common view receivers.

KeyWord: Time Synchronization, Non Common View, Common View

1 Introduction

In the three-way measurement[1] one TT&C station transmits radio signal, the responder beacon (Abbr. RSP) receives the signal and transmits the coherent signal, and another TT&C station receives the signal which is transmitted by the RSP. The station which transmits radio signal is called main station, while the station which receives radio signal is called subsidiary station. The three-way measurement includes coherent and non-coherent types. In the paper only the coherent three-way measurement is talked about. The error of the time synchronization influences the precision of the ranging value apparently, so the time synchronization is very important in the three-way measurement.

2 Time Synchronization between Main and Subsidiary Stations

2.1 Time Synchronizationwith Common View GPS Satellites

Now the time synchronization between the main station and the subsidiary station is mainly achieved by GPS common view method[2, 3, and 4] in three way measurement.A GPS common view receiver is installed in each station for observing the same one satellite to measure the time difference between two stations, thus high accuracy time synchronization is achieved.

The common view receivers receive satellite time, ephemeris, tropospheric parameter,ionospheric parameter and so on from the same one satellite to recover the satellite clock precisely. Each common view receiver can measure the time difference between the station clock and the satellite clock by comparing 1pps signal of the station with the satellite clock. Then the time difference between the main station and the satelliteminus the time difference between the subsidiary station and the satellite will be the difference between the two stations.

The main station clock is marked as T_M， The subsidiary station clock is marked as T_B， The satellite clock is marked as T_S， Sothe time difference between the main station and the satellite is,

(1)

T_B

The time difference between the subsidiary station and the satellite is,

(2)

T_{MB}

Then the time difference between the main and subsidiary station will be calculated by formula（1）and（2）,

(3)

From formula (3), it can be concluded that the accuracy of the time difference between the main and subsidiary station is determined by the accuracy of the time difference between the main station and the satellite, and that between the subsidiary station and the satellite, which both are mainly influenced by satellite orbit error, the station position error, atmosphere refraction correction error and ionosphere correction error[2].

The data of common view receivers in station A and station B is as follows.

Chart 1, Data of Common View Receiver in Station A

PRN	CL	MJD	STTIME	REFSV	SRSV	REFGPS	SRGPS	DSG	IOE	MDTR	SMDT	MDIO	SMDI
			hhmmss	.1ns	.1ps/s	.1ns	.1ps/s	.1ns		.1ns	.1ps/s	.1ns	.1ps/s
15	18	56627	90200	1498862	83	-498	64	14	72	78	0	130	-7
18	18	56627	90200	-2880522	179	-692	201	17	63	139	-23	221	-38
21	18	56627	90200	3309657	-15	-606	-32	13	52	104	-1	176	-8
24	18	56627	90200	183012	9	-604	2	14	18	133	-25	223	-47

Chart 2, Data of Common View Receiver in Station B

PRN	CL	MJD	STTIME	REFSV	SRSV	REFGPS	SRGPS	DSG	IOE	MDTR	SMDT	MDIO	SMDI
			hhmmss	.1ns	.1ps/s	.1ns	.1ps/s	.1ns		.1ns	.1ps/s	.1ns	.1ps/s
15	18	56627	90200	1480763	-29	-18596	-48	9	72	88	9	373	31
18	18	56627	90200	-2898185	-3	-18356	18	7	63	69	0	298	-3
21	18	56627	90200	3291879	31	-18384	14	7	52	66	0	289	0
24	18	56627	90200	164853	-39	-18764	-46	11	18	191	-43	636	-71

PRN - the satellite vehicle PRN number; CL- the hexadecimal number that corresponds to the Common View class of the track;MJD- the five digit Modified Julian Date of the start of the track; STTIME- hour, minute, and second (in UTC) of the start of the track; REFSV- the time difference (measured in .1 nanosecond) between the laboratory reference clock and satellite time, referred to the midpoint of the pass via a linear fit; SRSV- the slope determined via the linear fit to produce REFSV (measured in .1 picosecond per second); REFGPS- the time difference (measured in .1 nanosecond) between the laboratory reference clock and GPS system time, referred to the midpoint of the pass via a linear fit; SRGPS- the slope determined via the linear fit to produce REFGPS (measured in .1 picosecond per second); DSG- the root mean square of the residuals of the linear fit used to produce REFGPS; IOE- the index of ephemeris, a three digit decimal code indicating the ephemeris used for computations; MDTR- the modeled tropospheric delay (measured in .1 nanosecond) referred to the midpoint of the pass via a linear fit; SMDT- the slope determined via the linear fit to produce MDTR (measured in .1 picosecond per second); MDIO- the modeled ionospheric delay (measured in .1 nanosecond) referred to the midpoint of the pass via a linear fit; SMDI- the slope determined via the linear fit to produce MDIO (measured in .1 picosecond per second).

The time differences between the station and the satellite15, 18, 21, 24 are got by computing REFSV and SRSV, which started at 9:02 and lasted for 780 seconds, and the atmosphere and ionosphere delay is corrected. Then the time differences between two stations are got by the time differences between station Aand the satellite minus that between station B and the same one satellite, and the average value of time differences is also calculated. The results are shown in Fig.1.

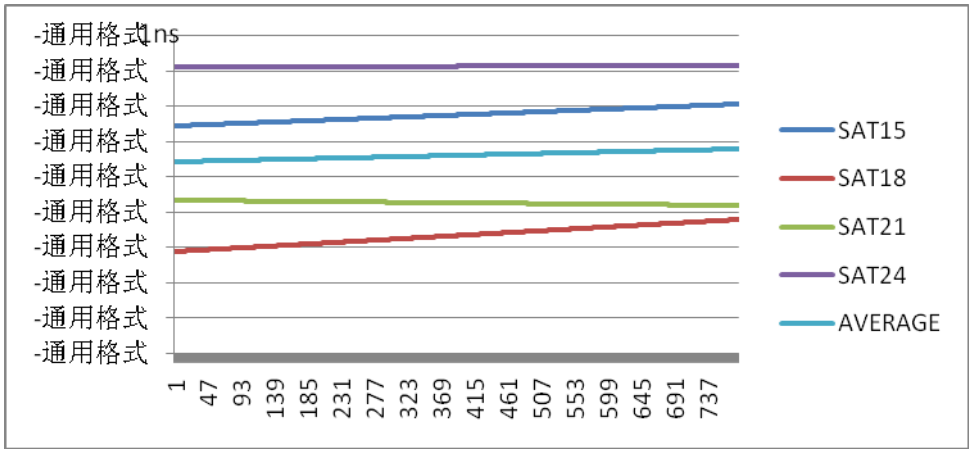


Fig.1,Time Difference between Two Stations by Common View Method

2.2 Time Synchronization without common view GPS Satellites

But the main station and the subsidiary station don't always have common view GPS satellites because of their positions in three way measurement. To solve the problem, a non-common view method is posted by using the time difference between the GPS satellite clock and GPS system time, which is included in the data transmitted from GPS satellites. In the data transmitted from GPS satellites there is the value of the time difference between the satellite clock and GPS system time. Firstly the common view receiver gets the time difference between the station clock and the satellite clock. Then it gets the time difference between the station clock and GPS system time. Once we get the time difference between the main station clock and GPS system time, and the time difference

between the subsidiary station clock and GPS system time, we can calculate the difference between the two stations clock. As a result the time synchronization is achieved.

The main station clock is marked as T_M, the subsidiary station clock is marked as T_B, the clock of the satellite which is tracked by the main station , is marked as T_S, the time difference between the ΔT_{SG}

main station tracked satellite clock and the GPS clock is marked as , the clock of the satellite which is tracked by the subsidiary station , is marked as T'_s, the time difference between $\Delta T'$

the subsidiary station tracked satellite clock and the GPS clock is marked as , thus the time difference between the main station and the GPS clock is,

(4)

The time difference between the subsidiary station and the GPS clock is,

(5)

ΔT_{MB}

Then, formula (4) minus formula (5) is the time difference between two stations ,

(6)

From formula (6), it can be concluded that the accuracy of the time difference between the main and subsidiary station is determined not only by the accuracy of the time difference between the station and the satellite, but also by the accuracy of the time difference between the satellite and GPS system clock.

The data of common view receivers in station A and station B, which is got by observing non-common view satellites is as follows.

Chart 3 Data of Common View Receiver in Station A

PRN	CL	MJD	STTIME	REFSV	SRSV	REFGPS	SRGPS	DSG	IOE	MDTR	SMDT	MDIO	SMDI
			hhmmss	.1ns	.1ps/s	.1ns	.1ps/s	.1ns		.1ns	.1ps/s	.1ns	.1ps/s
8	18	56627	90200	-103628	-193	-1026	-191	25	67	306	126	242	-2
9	18	56627	90200	-2872584	-147	-1055	-119	28	79	300	130	240	-1
26	18	56627	90200	-2345334	57	-712	-47	16	18	99	11	152	5
28	18	56627	90200	-3132027	32	-1075	62	26	29	305	-41	242	-22

Chart 4 Data of Common View Receiver in Station B

PRN	CL	MJD	STTIME	REFSV	SRSV	REFGPS	SRGPS	DSG	IOE	MDTR	SMDT	MDIO	SMDI
			hhmmss	.1ns	.1ps/s	.1ns	.1ps/s	.1ns		.1ns	.1ps/s	.1ns	.1ps/s
3	18	56627	90200	-2953662	-32	-18669	17	8	148	243	-65	667	-54
6	18	56627	90200	-1757680	-123	-18558	43	7	40	122	-9	482	-26
22	18	56627	90200	-2000441	-36	-18429	-14	7	76	100	-10	417	-37
27	18	56627	90200	26333	86	-18649	54	7	75	149	-23	543	-46

The time differences between the station A and the satellite (8, 3, 9, 6) and the time differences between the station B and the satellite (26,22,28,27) are got by computing REFGPS and SRGPS, which started at 9:02 and lasted for 780 seconds, and the atmosphere and ionosphere delay is corrected. Then the time differences between two stations are got by the time differences between station A and the satellite minus that between station B and the satellite, and the average value of time differences is also calculated. The results are shown in Fig. 2.

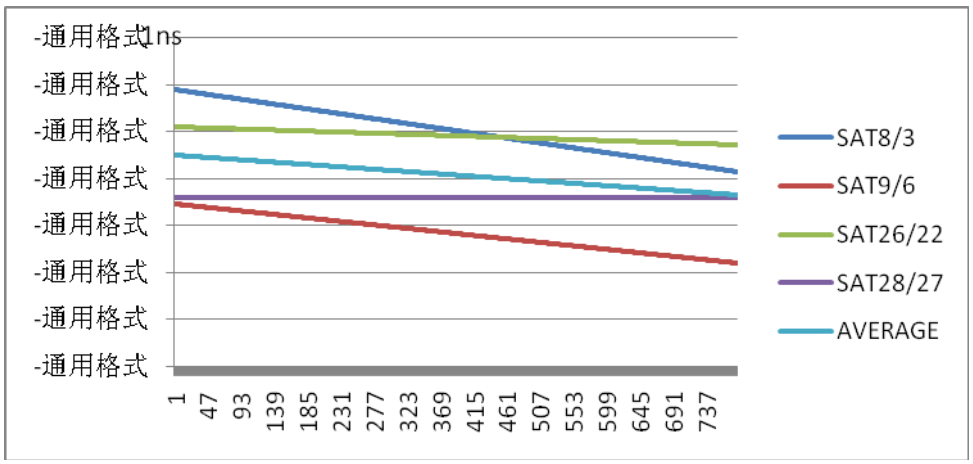


Fig. 2,Time Difference between Two Stations by Non Common View Method

2.3 Data Comparison in the Two Cases

At the same time two stations' common view receivers measured the time difference between the station clock and GPS timeby monitoring different GPS satellites, and the time difference between the station clock and the satellite time by monitoring same GPS satellites. And we calculated the time difference between the two stations separately by the common view method and by the non-common view method. The RMS of the time difference by common view method is around 40ns, while the RMS of the time difference by non-common view method is around 20ns. And the D-value of the average value of the time difference by non-common view method and the average value of the time difference by common view method is less than 45ns, as shown in Fig 3.

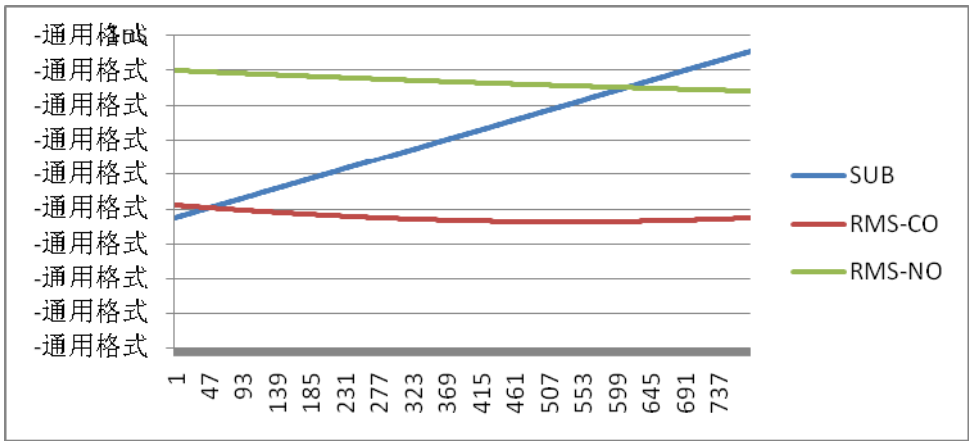


Fig.3,Comparison by Results of two different methods

3 Conclusion

The time differences between station A and station B are measured separately by common view method and non-common view method. And the precisions are compared by each other. It's proved that theprecision of the non-common view method is about same to the precision of the common view method.

Citations

[1] Huang Lei, Wang Hong, Fan Min.Study on Three Way Measurement in Deep Space [J]. Journal of Spacecraft TT&C Technology,2012,31(3), 6-10.

[2]Liu Jiyu.GPS Timing and Application[J]. Digital Communication World, 2012.8,53-58.

[3] Peng Fang, Zuo Jizhang, Chen Yufeng. Time and Frequency Synchronization Method based on Compass[J]. Journal of Projectiles,Rockets,Missiles and Guidance, 2007,27(1), 18-20.

[4]Gao Xiaoxun,Gao Yuan,Zhang Yue. Study on Time and Frequency Delivery in Remote Distance by Common View Method[J]. Acta Metrologica Sinica,2009,29(1), 80-83.

About the Authors：

ZHANG Yanping (1980-), Male, Master, Engineer, Major research direction is TT&C, fei2000a@sohu.com.

SHANG Jianzhong (1975-), Male, Bachelor, Senior Engineer,Major research direction is TT&C.

BAI Yungao (1988-), Male, Bachelor, Assistant Engineer, Major research direction is TT&C.

LI Haichao (1975-), Male, Bachelor, Senior Engineer, Major research direction is TT&C.

Thermal Expansion Coefficients Calculate and Analysis for Unidirectional Fiber-reinforced Composite of Space Truss Structure

Jian Ma

Beijing Institute of Spacecraft System Engineering, Beijing 100094, China
*Corresponding author: Tel: +86 10 18601059109
E-mail: majian0508@163.com

Abstract: Space engineering structures experience large temperature changes which lead to an extreme thermal stress and activate structural failure. Structural design with high stiffness and low thermal expansion in the specified direction is interesting and challenging. Carbon fiber reinforced composite is one of the most important structure and functional materials in space industry because of their excellent thermal-mechanical performance. Design of composite materials with low or zero isotropic thermal expansion coefficients were investigated intensively in the field of structural optimization and solid mechanics. In this paper, thermal-mechanical calculate method is developed to design the space deployable truss structure, which combines high stiffness with low thermal expansion in a predefined direction. The coefficient of thermal expansion(CTE) of composite material is investigated. To study the thermal expansion property of long fibre reinforced unidirectional (UD) composite, a new method was developed to create fibre random distribution models. In all the models, the fibre distribution state and the real periodic boundary condition have been taken into account. For models with high fibre volume fraction, a method called RDM (random disturbing method) was developed, which can generate models with fibre volume fraction not less than 65%. Random models generated by RDM were used to forecast the thermal expansion coefficient of M40J/TDE-85, and the thermal expansion coefficients of this composite were tested with high precision. The predicted results agree well with the experimental data, which shows that this random model can be used to predict the thermal expansion coefficients of UD composite correctly.

Key words: space deployable truss; composite; thermal deformation analysis ; coefficient of thermal expansion ; random distribution model

1. Introduction

As we know, composite materials have been undergoing extraordinary technological advances and enjoying widespread applications in different fields. However, as a result of their complex properties, such as wettability, chemical compatibility, anisotropic mechanics, heat absorption and conductivity abilities, their complete characterization has not been achieved so far.

Coefficient of thermal expansion (CTE) is defined as the fractional change in length of a body under heating or cooling through a given temperature range[1] and it is usually given as a coefficient per unit temperature interval at a given temperature. It is a key material property especially when a composite structure works in a temperature-changing environment. Here, the focus was placed upon

on studying the longitudinal and transverse CTEs of continuous fiber-reinforced unidirectional (UD) composites.

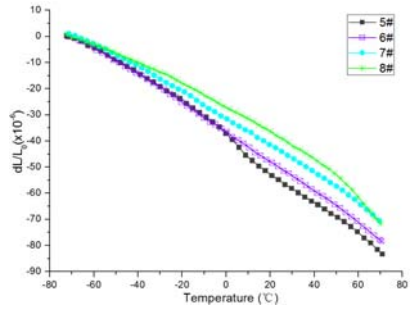
The problem of relating effective properties of a fiber-reinforced material to its constituent properties has drawn great attention. As a result, many analytical solutions have been made to predict the upper and lower bounds of CTEs of UD composites, which are composed of isotropic or anisotropic fibers and matrices[2-12]. In a series of studies[3-5] by Van Fo Fy, analytical solutions were presented to predict both axial and transverse CTEs of a UD composite through its constituent properties. However, the results were very sensitive to the elastic modulus and Poisson's ratio of the UD material. Levin6 expanded Hill's method and gave the upper bounds of a certain glass fiber-reinforced composite's CTEs, and the results were in much better agreement with the data in Van Fo Fy's study[3] than other predictions in that paper. Schapery[2] has derived expressions for longitudinal and transverse CTEs of composites with isotropic fibers embedded in isotropic matrices by adopting extreme energy principles. Chamberlain[8] Schneider[12] and Chamis[9] applied different methods and obtained the same longitudinal CTE expression, while the transverse expressions were quite different. In general, the predictions of longitudinal CTEs were always in good agreement with experimental data, while those of transverse CTEs failed to agree. An exception was Rosen and Hashin's[10] prediction as an extension of the work of Levin[6] . However, it is inconvenient to obtain results by Rosen and Hashin's solution, because to solve the CTEs of a composite, the mechanical properties of both the composite and its constituents must be determined first.

At the same time, as computation capability has grown dramatically over the last three decades, numerical solutions such as the finite element method (FEM) are being extensively applied to determine the CTEs of composite materials. Islamet et al[13] and Rupnowskiet et al[14] investigated the linear CTEs of UD composites systematically by the FEM. Karadeniz et al[15] explored the CTEs of different material systems by micromechanical modeling using the FEM, and comparisons were carried out among their results, analytical solutions, and experimental data. However, discrepancies still exist between FEM results and experimental data.

2. Test study on the thermal expansion properties of the uniaxial composites

2.1 The manufacture of the test pieces and the test method

In order to find out the thermal expansion properties of the M40/TDE85 composites comprehensively, the thermal expansion properties testing will be done on the test pieces of the M40/TDE85 composites correspondingly in this paper. The M40/TDE85 composite panels are produced in a special mould under the thermal shock before laying unidirectional material by hand and are developed in the Harbin FRP research institute. The nominal thickness of the unidirectional tape is 0.15mm. The volume content of the fibers is 60%. Each test piece of the composite contains 20 unidirectional tapes of M40/TDE85. Apply the pressure load and the heat load to the mould which has been laid the unidirectional tape. First, heat the mould to 120℃ and keep for 3 hours. Then keep heating up to 170℃ for 4 hours. Finally, allow the mould and the solidified composites come to room temperature slowly. The test method of the thermal expansion properties of the M40/TDE85 composites is continuous differential method using the thermal expansion instrument Netzsch DIL402C. Test temperature ranges for -50 ℃ to 70 ℃. All of the test processes will take place under the environment of He protection.



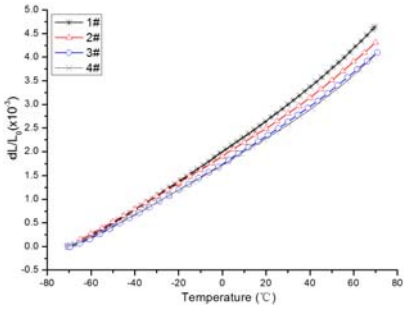


Figure 1 Vertical test results Figure 2 Transverse test results

2.2 Analysis of test results.

The longitudinal DL/L curve is less than zero and decreases with the increase of the temperature from Fig.1 and Fig.2. There is little difference between the 4 curves at low temperature. But with the increase of the temperature, the difference becomes bigger, especially at the room temperature. A similar phenomenon occurs to the transverse DL/L curve, too. The test results differ from the different test pieces. The reason is that the end faces of the test piece can't be parallel and perpendicular to the fiber direction strictly during polishing. The length direction of the test pieces should be parallel to the push rod during test process. While there is no reference substance when putting the test pieces on the support so that the angle deviation can make differences between the results.)

According to the definition of the coefficient of thermal expansion, the average thermal expansion coefficient under each temperature can be found by the above two figures. There is only the data at room temperature because of taking the numerical method to predict thermal expansion properties. In order to be convenient for comparison between the results, the thermal expansion properties at room temperature would be a key consideration. According to the measured results, the longitudinal and transverse thermal expansion properties at room temperature is:

Table 1 M40/TDE85 at room temperature and thermal expansion properties

CTE	Mean value /1.0E-6/℃	Standard deviation /1.0E-6/℃	CV %
Vertical CTE	-0.5437	0.0457	8.4095
Transverse CTE	31.9559	2.6193	8.1967

3 Micro- scale numerical investigations

3.1 Fibers randomly distributed algorithms

In unidirectional fiber-reinforced composite materials, the fibers in cross section show the characteristics of random distribution. Thus considering the fiber random distribution properties is important to accurately predict the macroscopic properties by micromechanical model. The generation methods in composite micromechanics finite element model to enhance non-uniform spatial distribution include: molecular dynamics method, Voronoi method, Monte Carlo method, random sequence adsorption (RSA) method and graphical reconstruction technology so on. RSA method is the most widely applied, but using RSA method to generate the random distribution of the long fibers, fiber volume fraction is difficult to exceed 50%. The RDM implementation process is as follows:

(1) Enter the basic parameters (ν_f , ϕ_f , k) of the composite RVE model to generate an initial model with fibers uniformly distributing, and the fiber in the initial model may be assumed as hexagonal or quadrangular. ν_f is the target fiber volume content of the RVE model; ϕ_f is the fiber diameter; k is a control parameters of the size of model, and it can determine the total number of fibers N_f ; The initial model is shown in Figure 3 (a).

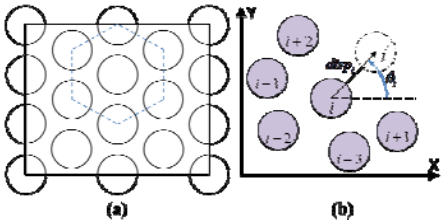


Fig. 3 schematic diagram of the initial model and fiberdisturbance

(2)In turn generate three random numbers: i_f 、 θ_i 、 $disp_i$, and the three random numbers are subject to different Poisson distribution. i_f is an integer between $1 - Num_f$, and represents the number of fibers at the first X times "disturbance". Num_f is the number of fiber has not yet been "perturbed"; θ_i is in the range of $0 - 2\pi$, and indicates the angle of the current fiber "disturbance"; $disp_i$ represents fibers "disturbance" distance in the range between $0 - dist_f$ .The parameters are shown in Figure 3 (b). And $disp_i$ is defined as:

$$dist_f = \left(\sqrt{\frac{\pi}{\sqrt{12}\nu_f}} - 1 \right) \phi_f \tag{1}$$

(3)Disturb the fiber specified by random integer i_f as the disturbance distance $disp_i$ and the disturbance angle θ_i , until all the fibers are successfully disturbed.

(4)Check the fibers position after disturbance, requiring no overlap between the current fibers disturbed with any other fibers, i.e. the distance between the axes of any two fibers ($dist_{ij}, i \neq j$) is greater than the fiber diameter ϕ_f .

(5)Check the current fiber disturbed intersects the RVE's boundary or not. If they intersect, generate another portion of the fiber cut at the opposite of the boundary, in order to make RVE satisfy periodic boundary conditions and to ensure the fiber volume fraction unchanged.

(6)Test intersection between the fiber i_f before disturbed and the RVE boundary. If they intersect, the fiber on the corresponding boundary is disturbed as the same disturbance parameters of fiber i_f .

(7)Sometimes the new position of fibers cannot meet the minimum distance condition, after they are disturbed as the random disturbance distance $disp_i$ and random perturbation angle θ_i generated by RDM. If multiple attempts are still unable to meet the conditions, the program will enter an endless loop. In order to reduce the possibility of the endless loop and accelerate the speed of model generation, RDM introduces cycles control parameter $count\_limit$. If it still cannot meet the corresponding conditions after random "disturbance" attempts reaching $count\_limit$, the program will return to the first step and regenerate new initial uniform model for follow-up process.

(8)The program automatically outputs all the coordinates of the fiber and exit, after all the fibers are successfully disturbed and the new location correctness of the fibers are verified. The process of RDM is shown in Figure 4, the description of relevant parameters are seen above. RDM can solve the unidirectional composite models of fiber volume content up to 90% theoretically.

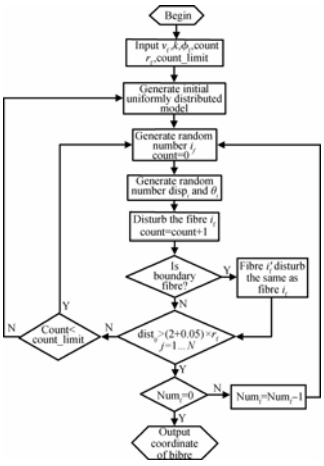


Fig. 4 RDM flow chart

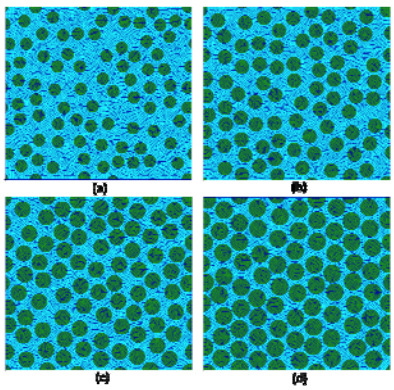


Fig. 5 Finite element model of RVE with different fiber volume content(a)30%，(b)40%，(c)50%，(d)60%

In RDM method, the fibers on the boundary require special treatment. When the fiber on each boundary is disturbed, all the corresponding fibers are disturbed as the same parameters at the same time. The purpose of this method is to make the fiber volume fraction of the whole model and target volume content strictly equal, and make the model satisfies the periodic boundary conditions. Computer programs are written with Matlab to achieve the random disturbance method. The fiber random distribution model is generated quickly using RDM by entering the appropriate parameters. And the location of the center of the fiber disturbed is output as standard text file, in order to make finite element pre-processing software generate the representative volume element mesoscopic model conveniently. Four representative mesoscopic models with different fiber volume content are shown in Figure 5, and the fiber volume content are 30%, 40%, 50%, 60% respectively, covering the composites commonly used in engineering.

3.2 FEM based on RDM

Since the coefficient of unidirectional composites' vertical and horizontal thermal expansion should be forecasted together, representative volume element requests the usage of 3D model; As

the calculation of amount and representative volume element has certain requests to scale, representative volume element is believed to have 40 fiber. Except using the method of RDM to find out the distribution of fiber, MSC.Patran is introduced to have mesh elements separated and 3D model generated. The created model of representative volume element, which is used for predicting the property of thermal expansion, is shown in Fig. 46. In that picture, the blue parts represent fiber and the rest means matrix. If the flaws exist in composites are considered, such as pores and so on, the method of random pore generation is adopted to create relative model of representative volume element that has pores. Due to the introduction of pores, the prediction results would have a larger variance. Moreover, the model of representative volume element would even be expanded to have 94 fiber instead of 40 fiber. Fig. 47 describes the representative volume element 3D model whose porosity is 2.71%. What's more, the black parts show imperfect regulated pores.

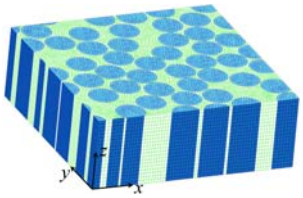


Fig. 6 Representative volume element 3D model

For all representative volume element models, the conditions of all models' boundary are the same regardless of the existence or the absence of pores. To start with, periodicity boundary conditions are applied to the counterparts and corresponding points of models to restrain the equation. After that, as long as the models do not face rigid body displacements, representative volume element models would have the least constrains in order to prevent the extra stresses and strains of displacement boundary conditions on models. Last but not least, according to the definition of the coefficient of thermal expansion, the strains of per temperature difference are material's coefficient of thermal expansion. If the models' original temperature is T_0， the temperature difference is $\Delta T = 1℃$.

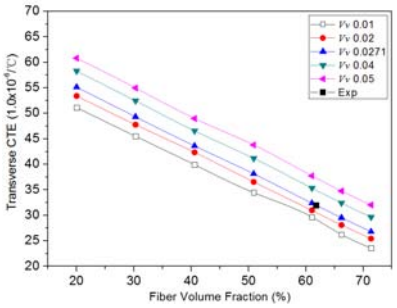
For every composite that has certain fiber and porosity, 5 independent representative volume element models should be built. Apart from that, the average prediction results from those 5 models should be regarded as the final forecasted outcome.

3.3 Analysis and comparison of prediction results

To thoroughly find out the influences of different fiber volume and pore volume on M40/TDE85 composites' vertical and horizontal thermal expansion, this chapter uses the approach of RDM to set up a series of models that apply diverse fiber volumes. Specifically, the fiber volumes in those models are separately 20.0%, 30.3%, 40.6%, 50.8%, 60.8%, 61.7%, 66.2% and 71.3%, which cover common composites' fiber volumes used by engineering. The predicted results of the transverse thermal expansion coefficient of M40/TDE85 under indoor temperature are shown in Fig. 7, and Fig. 8 shows the changing trend of the longitudinal thermal expansion coefficient as the fiber volume ratio and void ratio changing. It is can be found that the longitudinal and transverse thermal expansion properties decrease with the increase of fiber volume ratio. In Fig. 61, the five curves represent the variation trend of the transverse thermal expansion coefficient along with the fiber volume ratio, respectively. Although the void ratio of each model differs, the transverse thermal expansion properties decreasing trends have good consistence with each other, and the transverse thermal expansion properties have approximate linear relationship with fiber volume ratio.

According to Schapery's transverse prediction formula $\alpha_2 = (1+v_f)\alpha_f V_f + (1+v_m)\alpha_m V_m - \alpha_1(v_f V_f + v_m V_m)$, the transverse thermal expansion is a linear function of fiber volume ratio V_f. And it is proved that RDM method is in agreement with Schapery method on transverse prediction results with the error 0.2579×10^{-6} /℃, and the relative error is only 0.8072%.

In Fig. 62, the descend range of the longitudinal thermal expansion properties of composite materials at low fiber volume ratio is larger than that at a higher fiber volume ratio, the longitudinal thermal expansion properties have approximate hyperbola relationship with the fiber volume ratio. The error 0.0342×10^{-6} /℃, and the relative error is only 6.286%.



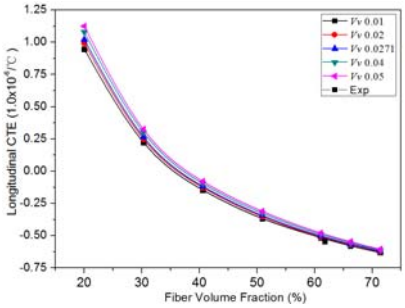


Fig.7 Transverse CTE verse fiber volume Fig.8 Vertical CTE verse fiber volume

In Fig. 7 and 8, the results of composites with same fiber volume ratio and different void ratio are different, especially when the fiber volume ratio increases. The reason is that the existence of the pore directly affects the properties of the matrix. Pore and resin blend to form a new matrix material. When the fiber volume ratio is smaller, the matrix volume ratio becomes more, finally affecting the macro-properties.

4. Conclusion

To study the thermal expansion property of long fibre reinforced unidirectional (UD) composite, a new method was developed to create fibre random distribution models. In all the models, the fibre distribution state and the real periodic boundary condition have been taken into account. For models with high fibre volume fraction, a method called RDM (random disturbing method) was developed, which can generate models with fibre volume fraction not less than 65%. Random models generated by RDM were used to forecast the thermal expansion coefficient of M40J/TDE-85, and the thermal expansion coefficients of this composite were tested with high precision. The predicted results agree well with the experimental data, which shows that this random model can be used to predict the thermal expansion coefficients of UD composite correctly. It is rapid and precise to adopt this random model to predict the thermal expansion property of composite, and it would be convenient for materials study and engineering application.

References

[1]. Z.Haktan Karadeniz, Dilek Kumlutas. A numerical study on the coefficients of thermal expansion of fiber reinforced composite materials. J Composite Structures 2007;78:1-10.

[2]. Schapery RA. Thermal expansion coefficients of composite materials based on energy principles. J Composite Materials 1968;2:380-404.

[3]. Van Fo Fy GA. Thermal strains and stresses in glass fiber-reinforced media. Prikl Mekh Teor Fiz 1965;101[inRussian].

[4]. Van Fo Fy GA. Elastic constants and thermal expansion of certain bodies with inhomogeneous regular structure. Soviet Physics, Doklady 1966;11:176.[in Russian].
[5]. Van Fo Fy GA. Basic relations of the theory of oriented glass reinforced plastics with hollow fibers. Mekhanika Polimerov 1966; 2:763 [in Russian].
[6]. Levin VM. Thermal expansion coefficients of heterogeneous materials. Mechanics of Solids 1967;2:58-61.
[7]. Schapery RA. On application of a thermodynamic constitutive equation to various nonlinear materials. Purdue Univ. Report 1968;6:No.68-4.
[8]. Rogers KF, Phillips LN, Kingston-Lee DM, Yates B, Overy MJ,Sargent JP, et al. The thermal expansion of carbon fibre-reinforced plastics. Part 1: The influence of fiber type and orientation. J Mater Sci 1977;12:718–34.
[9]. Chamis C.C. simplified composite micromechanics equations for hygrolthermal and mechanical properties. SAMPE quarterly 1984;15(3):14-23.
[10].Rosen B.W and Hashin Z. Effective thermal expansion coefficients and specific heats of composite materials. Int J Eng Sci 1970;8:157-173.
[11].Hashin Z. Analysis of properties of fiber composites with anisotropic constituents. J Appl Mech 1979;46:543-550.
[12].Sideridis E. Thermal expansion coefficient of fiber composites defined by the concept of the interphase. Compos Sci Tech 1994;51:301–17.
[13].Islam MDR, Sjolind SG, Pramila A. Finite element analysis of linear thermal expansion coefficients of unidirectional cracked composites. J Compos Mater 2001;35:1762–76.
[14].Rupnowski P, Gentza M, Sutterb JK, Kumosa M. An evaluation on the elastic properties and thermal expansion coefficients of medium and high modulus graphite fibers. Composites: Part A 2005;36(3):327–38.
[15].Z. Haktan Karadeniz, Dilek Kumlutas. A numerical study on the coefficients of thermal expansion of fiber reinforced composite materials. Composite Structures 2007;78:1-10.

Experimental Investigation for Single-phase Heat Transfer of Confined Jet Array Impingement on a Round Heated Surface in a Narrow Space

Zhongxuan Du, Xiang Zhang, Zaiteng Zhai, Han Chen, Shichen Jiang

Shanghai Institute of Satellite Engineering, 251 Huaning Rd, Minhang District, Shanghai, 200240

Abstract: Jet impingement fluid loop is considered as a promising technology for high heat flux removal in aerospace application. In the present study, a closed loop experimental setup was built for the study of single-phase heat transfer characteristics of confined jet array impingement on a round heated surface in a narrow space by using deionized water as coolant. The test section was designed that the round heated surface area was 7.1 cm^2, the orifice diameter d=0.5mm, the orifice number N=150, the ratio of jet-to-target spacing to orifice diameter H/d =4, and the ratio of jet-to-jet spacing to orifice diameter S/d=4.5. The water was averagely jetted at each orifice corresponding to the round heated surface in the test section, which made the round heated surface uniformly cooled. The experimental heat transfer results showed that the wall temperature decreased and the heat transfer coefficient increased with the increase of volume flow rate, and the pressure drop increased with the increase of mass flow rate, but was unaffected by the heat flux. The heat reject of jet impingement reached as high as 498.5W/cm^2 and the pressure drop was 0.4 bar or so under the conditions that the mass flow rate was 2.0L/min and the system pressure was 2.0 bar.

Keywords: high heat flux, confined jet array impingement, round narrow space, uniform jet, single-phase

1. Introduction

During the design of thermal control system in modern spacecraft, the high temperature and heat flux appear in some local area because of the highly integration of electronics and the increase ratio of heat consumption to surface area. The traditional passive thermal control design can only remove heat flux as high as 10 W/cm^2, far lower than the heat rejection demand.

Jet impingement is an efficient enhanced heat transfer technology, especially in the application of local high heat flux. Compared with free jet and submerge jet, the confined jet is better suitable for the micro-gravity conditions for two advantages: one is that the fluid is confined to a closed room to form a closed loop, the other is that the jet orifice is close to the heated wall, thus the impact of body forces on heat transfer can be ignored. Therefore, confined jet impingement has been considered as one of the promising high heat removal technologies for spacecraft thermal control system.

The single-phase liquid jet cooling has three overweighed virtues in comparison with two-phase jet cooling in aerospace application. First, the single-phase jet cooling is easily implemented under the ground, and these experiment results resemble those in space. Next, the heat transfer mechanism of single-phase jet is simple, so it is well predicted in space application. While the two-phase jet cooling is complicated because the vapor-liquid interface under microgravity conditions is unknown and uncontrollable. Finally, the single-phase jet cooling has sufficient heat

transfer margin far away from critical heat flux (CHF), meeting the design demands of redundancy. Therefore, the single-phase jet cooling is more reliable than the two-phase jet cooling for spacecraft thermal control system.

The confined jet cooling of single orifice has been investigated in recent decades. These researches was reviewed and summarized in [1]. Garimella [1] concluded that the local heat transfer coefficient on the target surface had a bell-shaped distribution with respect to radial distance from the stagnation point, and the maximum value occurred at the stagnation point and decreases symmetrically with radial distance.

Jet array cooling is more complicated than the single jet cooling due to the interference of neighboring jets. Most of previous works mainly focused on gas jet array cooling or free and submerged liquid jet impingement. Few were reported on the liquid confined jet array impingement.

Chang et al. [2] conducted an experimental study to compare the heat transfer for confined single- and multiple-jet impingement by using R113. The results showed that the heat transfer of multiple jets was weaker than that of single jet at the stagnation point, while it was the opposite for local average heat transfer coefficient. Correlations applicable to actual multi-jet impingement were also developed.

Sung and Mudawar [3] studied the single-phase heat transfer of a hybrid device consisting of a slot jet impinging into a micro-channel experimentally and numerically. They found that higher jet Reynolds number yielded stronger fluid attachment to the heated surface and lower surface temperature, and decreasing jet width and micro-channel height also could lower the surface temperature.

Meyer et al. [4] performed experiments to explore the effects of jet width, impingement velocity, and inlet subcooling on the cooling performance of an array of three confined rectangular FC-72 and ethanol jets impacting a 3.0 cm × 3.0 cm heated surface. The results showed that the single-phase heat transfer coefficient increased with increasing jet velocity and/or jet width. Increases in jet velocity, jet width, and subcooling broaden the single-phase region preceding the commencement of boiling.

Hong et al. [5-6] studied the single-phase and two-phase confined Jet array impingement cooling of aqueous ethylene glycol solutions with d=1mm, H/d=1, 1.5 or 3, S/d= 4 or 5, and N=32 or 50. They concluded for single-phase jet cooling that the orifice number has little effect on the average heat transfer coefficient under the same jet velocity; that the back pressure slightly affected the heat transfer characteristics; that the single-phase jet heat transfer reached the best at H/d=1.5, which indicated there was an optimal H/d to achieve the best heat transfer performance.

Rau and Garimella [7] study the heat transfer characteristics and pressure drop of dielectric liquid HFE7100 with three round orifice geometries with the same total orifice open area, a single orifice of 3.75 mm diameter, a 3×3 array of 1.25 mm diameter orifice and a 5×5 array of 0.75 mm diameter orifices. They concluded that the arrays of jets resulted in higher area-averaged heat transfer than a single jet at a fixed flow rate, however, displayed larger relative non-uniformity in local two-phase heat transfer coefficient and surface temperature. The pressure drop in each orifice configuration was independent of the applied heat flux and vapor generation.

The existing studies demonstrated that single-phase jet array impingement was affected by the structural parameters, such as d, N, H/d and S/d, and fluid flow parameters, such as mass flow rate, back pressure and liquid sub-cooling. However, the previous studies mainly focused on the plane

heated surface, no open literature was found on the study of jet impingement over the round heated surface.

2. Experimental setup

A closed loop experimental setup was built for the study of the confined jet array impingement in a round narrow space, as shown in Fig. 1. It was made of a pump, a filter, a needle valve, a flow meter, a confined array jet test section, a plate heat exchanger, an industrial chiller, a liquid reservoir, some valves and stainless steel tubes, as well as some test apparatus, such as thermocouples, pressure transducers and pressure difference transducers. The coolant of the present study was deionized water.

The experimental setup began running after filling the working fluid. The deionized water was pressurized by pump, and passed through the filter, stop valve and flow meter. In the test section, the deionized water was jetted to a round heated surface in a narrow space, and absorbed the heat, which was transferred to the water of the industrial thriller through the plate heat exchanger. Then the deionized water flowed through the liquid reservoir and the pump to complete a cycle. The needle valve installed on bypass cycle was used to adjust the volumetric flow rate of the test section. The function of the liquid reservoir was to keep the entrance pressure of the pump positive and to control the back pressure of the cycle. The inlet deionized water temperature of the test section could be regulated by changing the water temperature of the industrial chiller.

A turbine flow meter ranging from 0 to 10 L/min was amounted to measure the volumetric flow rate of the test section, and the precision was ±1%. The pressure transducer with the measurement range of 0~2MPa and the accuracy of 1‰ was used to test the system back pressure. Two pressure difference transducers with the measurement range of 0~1MPa and 0~60kPa and the accuracy of 1‰ were amounted on both sides of the pump and the test section, respectively. Three armored Pt100 resistances with an accuracy of 0.1% were installed before and after the test section and the outlet of the plate heat exchanger to measure the fluid temperatures. Five T-type thermocouples with 0.2% accuracy were arranged at the test section to measure the wall temperatures. All of the test signals were acquired by the data acquisition and collected by the computer for data reduction.

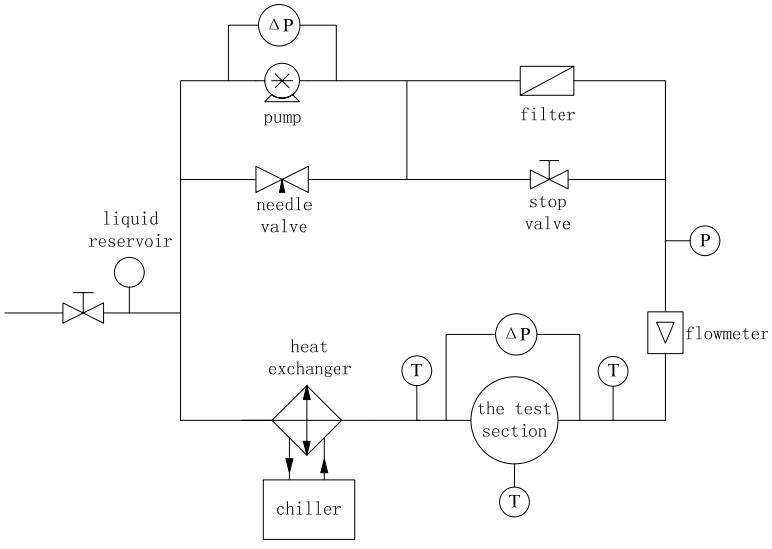


Fig. 1 Schematic of experimental setup

The test section was shown in Fig. 2. It consisted of outer part, middle part, copper block, heating rods, insulated material and supporting setting. The outer part, middle part, and copper block were welded one another to stand over 1.0 MPa pressure. The middle part was designed to separate the inlets and outlets of the test section. The insulated material was to prevent the heat leak to the environment.

Twenty-one heating rods were inserted into copper block to make the heat source, as shown in Fig. 3 (a). Each heating rod could generate 250W heat quantity at an alternating voltage of 240V. The thick line in Fig. 2 represented the round heated surface with an area of 7.1 cm^2. The temperature measurement holes around the heated surface 2mm was used to install the thermocouples.

150 orifices with diameter of 0.5mm in three rows were designed on the bottom of the middle part in Fig. 3 (b), which was exactly corresponded to the heated surface. The orifices in three rows were arranged in staggered pattern. The ratio of jet-to-target spacing to orifice diameter H/d was equal to 4, and the ratio of jet-to-jet spacing to orifice diameter S/d was 4.5. Besides, four outlets around the top of the middle part were designed to make the deionized water flow out uniformly.

Four inlets of the test section were designed around the top of the outer part in Fig. 3 (c) to make sure that the deionized water could enter the test section uniformly so as to achieve uniform water jet at the orifices.

The circular tubes in Fig. 4 were connected with the inlets of the test section in Fig. 2. Its function was to distribute the water into four brunches to enter the test section averagely. Also the same form of the tubes was connected with the outlets to flow out of the test section averagely.

During the experiments, the water entered the test section averagely from four inlets, then flowing along the arrows in the narrow space between the middle part and the outer part in Fig. 2. The water was shot upon the round heated surface of the heat source at the orifices, then flowing in the narrow space between the middle part and the heat source. The water finally flowed out of the test section from the four outlets of the middle part averagely.

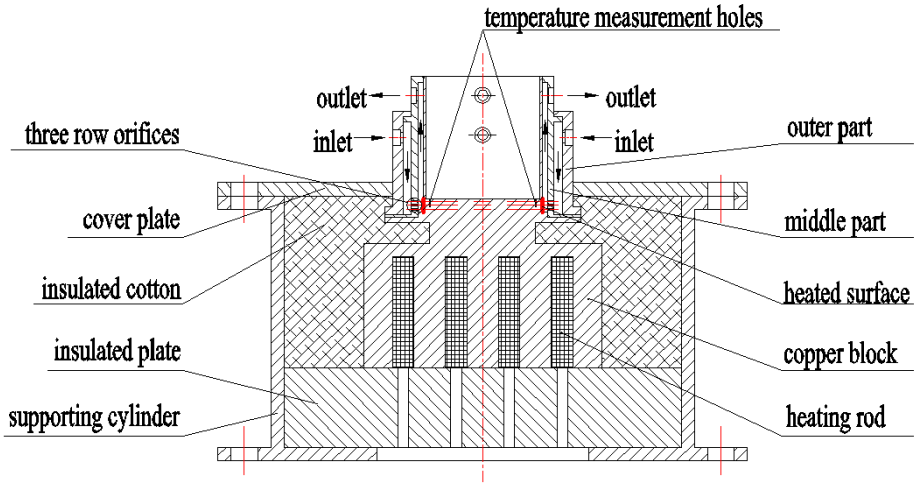


Fig. 2 The test section

(a) heat source (b) middle part (c) outer part

Fig. 3 The components of the test section

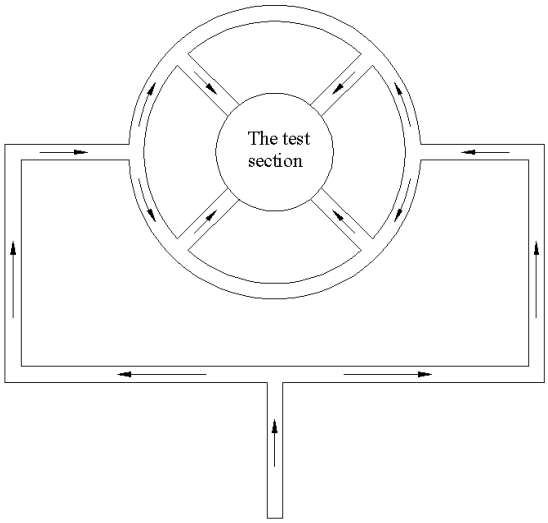


Fig. 4 The circular tubes

3. Data reduction

The average heat transfer coefficient was defined as:

$$h = \frac{Q}{A\left(t_f - t_w\right)} = \frac{q}{t_f - t_w} \tag{1}$$

in which Q was heat quantity, q was heat flux, A was area of heated surface, t_f was the average fluid temperature and t_w was the average wall temperature.

The heat quantity Q was calculated as:

$$Q = \left(Q_h + Q_f\right)/2 \tag{2}$$

where Q_h was the sum of input power of each heating rod, and Q_f was heat quantity of water temperature rise, which could be calculated by

$$Q_f = c_p m\left(t_{f,out} - t_{f,in}\right) \tag{3}$$

in which c_p was specific heat, m was mass flow rate, $t_{f,in}$ and $t_{f,out}$ were inlet water temperature and outlet water temperature of the test section, respectively.

The average fluid temperature was calculated by

$$t_f = \left(t_{f,in} + t_{f,out}\right)/2 \tag{4}$$

The average wall temperature t_w was calculated as:

$$t_w = t_m - \frac{qr_w \ln(r_w / r_m)}{2\lambda} \tag{5}$$

in which t_m was the arithmetic mean value of measured wall temperatures, r_w and r_m were the radii of the heated surface and temperature measured points, respectively, λ was the thermal conductivity of the copper block.

Heat balance was compared with the preliminary experimental data between the electrical power and water temperature rise in Fig. 5. The error was within ±10%, which indicated that heat insulation of the test section was conducted well. The maximal uncertainty of the heat transfer coefficient in the present study was 18.8%.

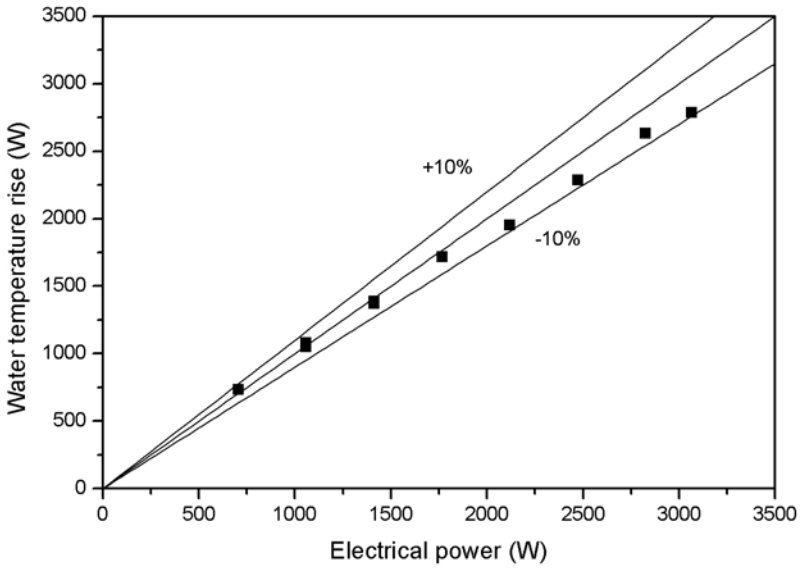


Fig. 5 Heat balance between electrical power and water temperature

4. Results and discussion

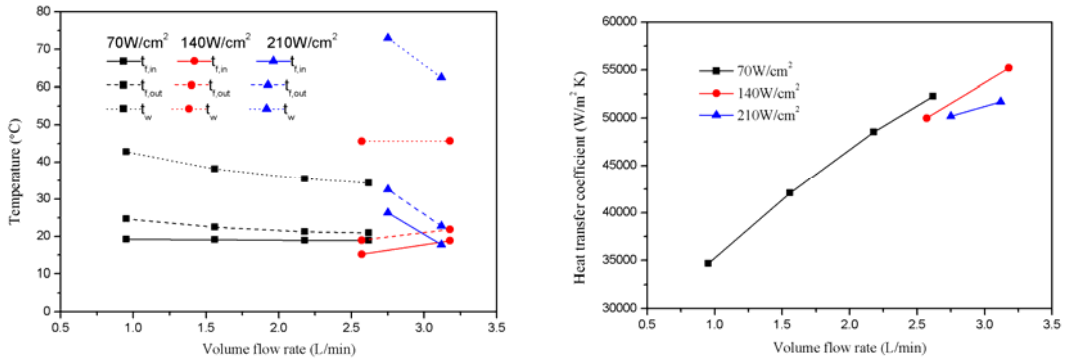
4.1 Uniform water jet test

Prior to the heat transfer experiments, it was essential to test the uniformity of the round water jet to obtain the uniform heat reject around the round heated surface. The test was performed by using the middle part, the outer part and the circular tubes. As shown in Fig. 6, the water flowed into the test section from two entries. Then it was averagely distributed to four branches corresponding to four inlets of the test section. It could be seen that the shooting water from each orifice concentrated in the center to form an ascending water column, which proved that the water at each orifice was well distributed in the test section, and the round heated surface was uniformly cooled.

Fig. 6 Uniform water jet test

4.2 Heat transfer characteristics

Figure 7(a) and 7(b) presented the temperature versus volume flow rate and heat transfer coefficient versus volume flow rate at 70, 140 and 210 W/cm^2, respectively. The system pressure was 6.0 bar or so. As could be seen that, the wall temperature decreased while the heat transfer coefficient increased with the increase of volume flow rate for given heat flux. This was because with the volume flow rate increased, the jet velocity increased and the jet impingement became stronger, thus the boundary layer became thinner, and the heat transfer enhanced. Besides, the wall temperature was relative with the inlet water temperature. The wall temperature was almost the same at 140 W/cm^2, the reason was that the inlet water temperature was 15.2 ℃ at 2.57 L/min, and 18.8℃ at 3.18 L/min.

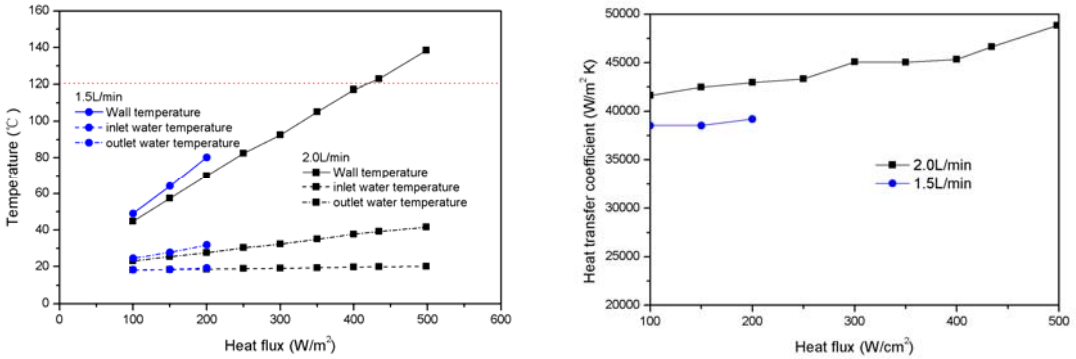


(a) temperature (b) heat transfer coefficient

Fig. 7 Heat transfer at different heat fluxes

Figure 8(a) and 8(b) presented the temperature versus heat flux and heat transfer coefficient versus heat flux at 1.5 L/min and 2.0 L/min, respectively. The system pressure was 2.0 bar or so. It could be seen that the wall temperature and outlet water temperature increased with the increase of heat flux for given volume flow rate. The dotted line in Fig. 8(a) represented the saturated temperature (120.4℃) at 2.0 bar. It was obvious that the conditions below the dotted line were single-phase heat transfer. The boiling occurred on the heated wall under the conditions that volume flow rate was at 2 L/min, heat flux at 434 and 498.5 W/cm^2. The outlet water temperature at 498.5 W/cm^2 was 41.5℃, indicating that the boiling water near the wall occupied a few and was rapidly cooled by the shooting water nearby. The single-phase heat transfer coefficient trended to increase

with heat flux increased in Fig. 8(b). The authors thought the phenomena might be caused by the experimental error. This could be explained that the single-phase jet impingement was related to volume flow rate, thermo-physical properties and structural parameters. That was to say heat transfer coefficient was affected by Re, Pr and structural parameters, such as d, N, H/d and S/d, but regardless of heat flux. In the experiments, the heat reject of jet array impingement reached as high as 498.5W/cm^2 and the pressure drop was 0.4 bar or so under the conditions that the mass flow rate was 2.0L/min and the system pressure was 2.0 bar. At this condition, the jet was at the stage of onset of nucleate boiling (ONB), far away from CHF, which indicated that the system still had sufficient heat transfer margin.

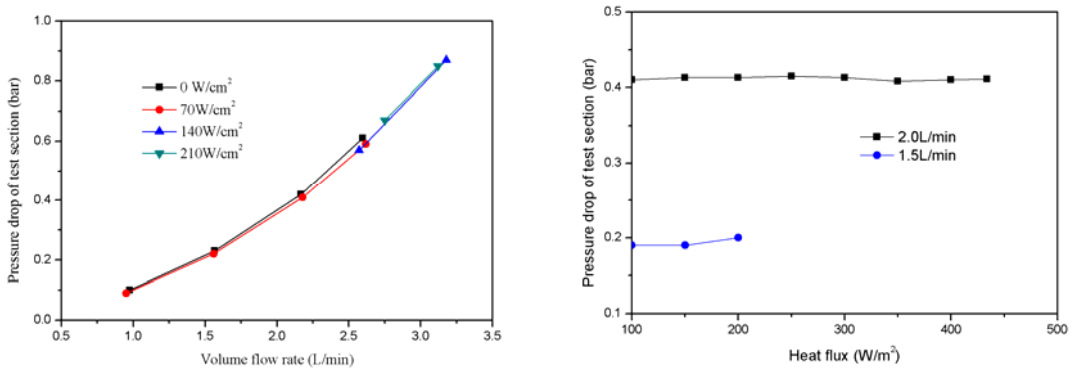


(a) temperature (b) heat transfer coefficient

Fig. 8 Heat transfer at different volume flow rates

4.3 pressure drop

Pressure drop of the test section was presented in Fig. 9(a) and 9(b) at different heat fluxes and volume flow rates. It could be seen that pressure drop increased with the volume flow rate increased, but was unaffected by the heat flux. This was consistent with the results of Rau and Garimella [7]. It could be explained that the increasing heat flux did not lead to the variations of the thermo-physical properties, and pressure drop in single phase was related to the flow pattern or Re.



(a) Effect of heat flux (b) Effect of volume flow rate

Fig. 9 Pressure drop of the test section

5. Conclusions

A closed loop experimental setup was established for studying the single-phase heat transfer characteristics and pressure drop of confined jet array on a round heated surface in narrow space by using deionized water in the present paper. The conclusions achieved were:

1) Uniform water jet test proved that the deionized water at each orifice was well distributed in the test section, thus making the round heated surface uniformly cooled.

2) The wall temperature decreased and the heat transfer coefficient increased with the increase of volume flow rate, and the pressure drop increased with the increase of mass flow rate, but was unaffected by the heat flux.

3) The heat reject of jet array impingement reached as high as 498.5W/cm^2 and the pressure drop was 0.4 bar or so under the conditions that the mass flow rate was 2.0L/min and the system pressure was 2.0 bar. The condition was far away from CHF, indicating that the system still had sufficient heat transfer margin.

References

[1] Garimella, S.V., 2000, Heat Transfer and Flow Fields in Confined Jet Impingement, Annual Review of Heat Transfer, 11, 413-494.

[2] Chang, C., Kojasoy, G., Landis, F., Downing, S., 1995, Confined single- and multiple-jet impingement heat transfer-Ⅰ. Turbulent submerged liquid jets, International Journal of Heat and Mass Transfer, 38, 833-842.

[3] Sung, M., Mudawar, I., 2006, Experimental and numerical investigation of single-phase heat transfer using a hybrid jet-impingement/micro-channel cooling scheme, International Journal of Heat and Mass Transfer, 49, 682-694.

[4] Meyer, M., Mudawar, I., Boyack, C., Hale, C., 2006, Single-phase and two-phase cooling with an array of rectangular jets, International Journal of Heat and Mass Transfer, 49, 17-19.

[5] Hong, F., Zhang, C., He, W., Cheng, P., Chen, G., 2013, The local and average heat transfer characteristics of confined jet array impingement boiling of aqueous ethylene glycol solutions, Proceedings of the ASME 2013 4th International Conference on Micro/Nanoscale Heat and Mass Transfer, Hong Kong, China, December, 11-14, 1-8.

[6] Hong, F., Zhang, C., He, W., Cheng, P., Chen, G., 2014, Confined jet array impingement boiling of subcooled aqueous ethylene glycol solution, International Communications in Heat and Mass Transfer, 56, 165-173.

[7] Rau, M., Garimella, S., 2013, Local two-phase heat transfer from arrays of confined and submerged impinging jets, International Journal of Heat and Mass Transfer, 67, 487-498.

ELECTRIC PROPULSION APPLICATION IN CHINESE TELECOMMUNICATION SATELLITES

Zhou Zhicheng, Wang Min，Zhong Xiaoqing

Institute of Telecommunication Satellite, China Academy of Space Technology，Beijing 100094，China

Abstract: Ion Electric propulsion system (IPS) is the most attractive way to improve the communication satellites capability. To meet Chinese new telecommunication satellites performance requirement, China Academy of Space Technology (CAST) had successfully developed its own IPS and completed system level integration test. Test was divided into six important parts, using the hardware and software of flight configuration. IPS capability was demonstrated and the high risk was mitigated. This paper describes the IPS architecture of Chinese first GEO application, and focuses on the IPS integration test results.

1 Introduction

April 24th 1970, DFH-1 (Dongfanghong-1), the first satellite of China was launched successfully into orbit, which represented the start of Chinese space exploration[1]. After forty years' continuous improvement and enduring development, DFH satellite series have played an important role in the international market of telecommunication satellite. Several satellites launched for many foreign customs including Venezuela, Nigeria, Pakistan and Bolivia[2], which were all manufactured based on DFH satellite platform, and are all operating smoothly in orbit.

To satisfy the emerging business demand for increasing carrying capacity of more payloads on a satellite, China Academy of Space Technology (CAST, the manufacturer of DFH satellite), started the application of electric propulsion technology and independently developed the electric propulsion system for the new DFH satellite platform[3~5].

This article will introduce the design procedures of electric propulsion system on DFH satellite, and illustrate the progress of test in detail.

2 Electric propulsion system for communication satellite

2.1 Overview

Nowadays, the reliability of propulsion system directly determines the performance of a Telecommunication satellites, since it usually operated in geosynchronous orbit (≥36000 km) and should have over 15 years of lifetime. So more advanced propulsion system is needed to accomplish such missions.

The propulsion system had played a key role in the evolution of the DFH satellite platform. In early days, solid propellant propulsion systems were used to perform orbit maneuver for DFH satellite, and monopropellant chemical propulsion systems were used to accomplish the attitude control and orbit maintenance missions. DFH-3 satellite used an integrated bipropellant chemical propulsion system to perform the orbit maneuver and attitude control in the orbit. DFH-4 satellite inherited the configuration of propulsion system from DFH-3 but improved the specific impulse and life of the thrusters.

In recent years, space communication application is calling for new satellite platform which has higher caring capacity and lower cost. This trend firstly resulted from the great commercial demand of growing communication business such as high capacity communication, broadband multimedia access and multimedia broadcast. Secondly, the continuing falls in unit price of transponders pressurize performance upgrade of communication satellite to maintain its commercial profit.

The specific impulse of electric propulsion is 5 to 10 times higher than traditional propulsion technology. It's such a huge advantage that can result in at least three aspects of benefit of longer service time, more payload and lighter launch mass. This advantage seems to be customized for communication satellite development and many manufacturers endeavor to make progress in this field[6~10].

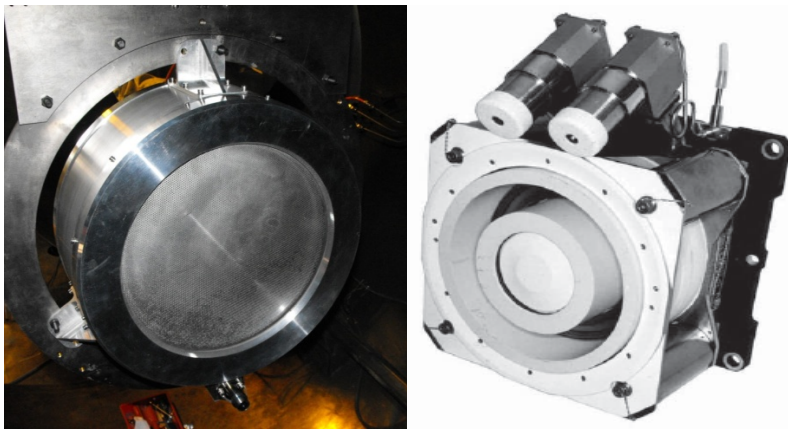


Figure 44 Picture of Ion(left) and Hall(right) thruster

2.2 The electric propulsion system on DFH satellite series

CAST has started the research work in electric propulsion system which is focusing on the realization of an ion thruster since the second half of 20th century, and the ion thruster was launched successfully in space in 2012[5]. With the technology accumulation, CAST revised its application plan of electric propulsion system, which evolved from north-south station keeping application to all-electric propulsion system[3~5].

The whole system contains four ion thrusters that installed on the north and south side of a satellite. It also comprises two xenon tanks, one Pressure Regulating Module (PRM) and four flow controlling module (FCM) which controls the micro flow very precisely. In addition, four power processing unit are employed in the system to convert the power supplied from the bus into multi-channel power with different voltage or current output. There is also a thrust pointing assembly mechanism (TPAM) on both side of the satellite that can regulate the vector of the thruster according to the control command.

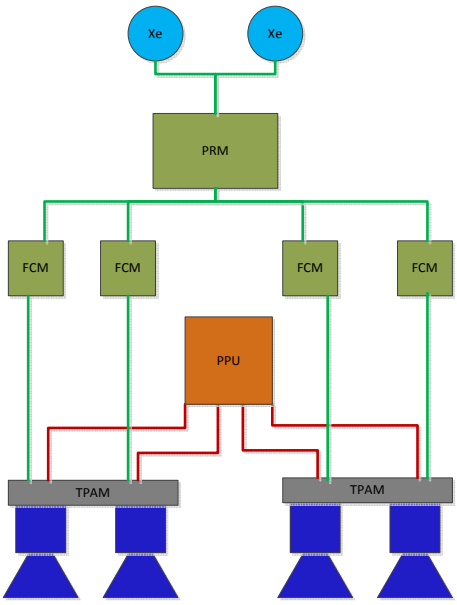


Figure 45 EPS diagram

2.3 The systematic validation test program for electric propulsion

CAST worked out a systematic validation test program in order to verify the performance of the electric propulsion system and its compatibility with the satellite platform. The test program consisted of such six following procedures:

(1) Firing test of the electric propulsion system

The firing test validated power and signal interfaces between every two units, including ion thruster, PPU and FCM. Moreover, and this test conducted performance characterization of specific impulse and thrust.

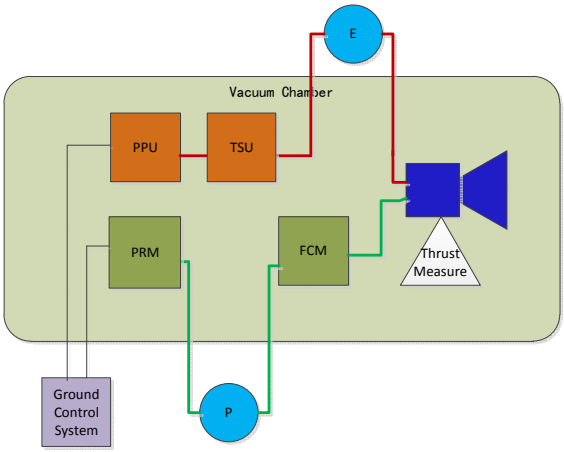


Figure 46 Block diagram of the firing test

(2) Life test of the thruster

Life test validated whether the life time of the ion thruster can satisfy the requirement of the satellite through continuous firing and start-shutdown operation of the thruster. The test will complete until the ion thruster fail to work.

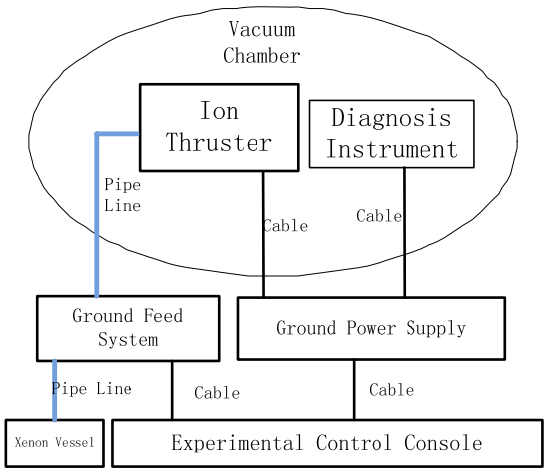


Figure 47 Block diagram of the life test

(3) Xenon filling test

filling test was conducted in accordance with the same process and regulation in the launch site to meet the high purity requirement of the xenon propellant.

(4) Plume test

Plume test diagnosed many parameters in relative to the thruster plume and contamination and deposition around the thruster to evaluate the acceptance of several effects caused by ion thruster including force, thermal and contamination.

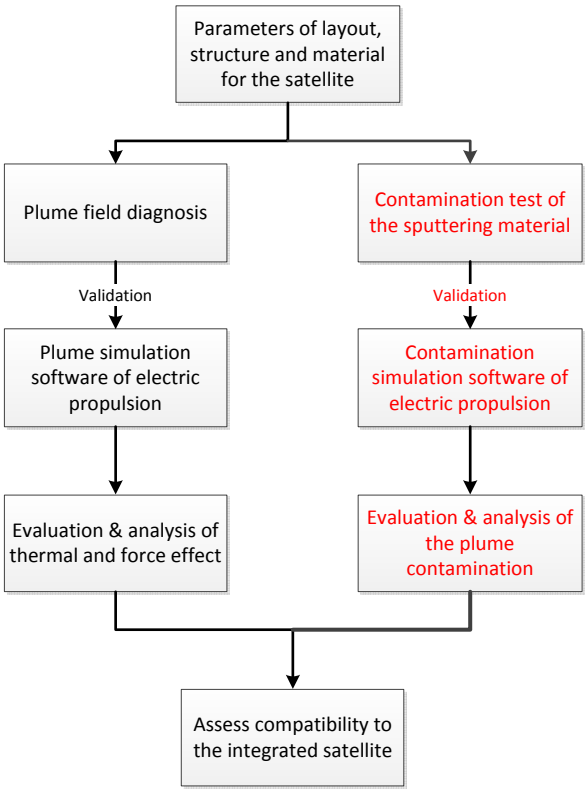


Figure 48 the diagram of the plume test plan

(5) EMC test

Electromagnetic interference and the variation of amplitude and phase of the electromagnetic wave after it was transmitted through the thruster plume to make sure the plume wouldn't disturb the communication of the satellite.

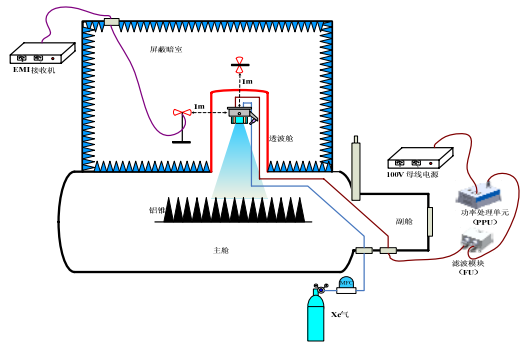


Figure 49 Block Diagram of the EMC test

(6) Firing test of the whole satellite

After the satellite was completed assembled, the electric propulsion system would be fired to validate its performance character and compatibility to other units on the satellite.

3 Analysis of the test results

3.1 Firing test of the electric propulsion system

Test goal. This test validated the interface connectivity between each two units, verifying the thrust, specific impulse and efficiency of the thruster according to the configuration with the full integration of all the hardware and software of the system. Finally, the automatic execution of the control commands and failure process function are verified in the test, too.

Test approach. All the units including ion thruster, PPU and xenon feed unit were located in a vacuum chamber. An electric parameter test instrument was connected between PPU and the ion thruster. A micro-thrust measurement system was employed. There are also two mass flow controllers integrated in the pipeline that connected the PRM and FCM to directly measure the flow rate, which would provide an important parameter for the calculation of the specific impulse.

The test results of thrust, specific impulse and power were compared with the data that collected from the former flying thruster in the orbit, all the parameter was coincident with the flying data[5, 11].

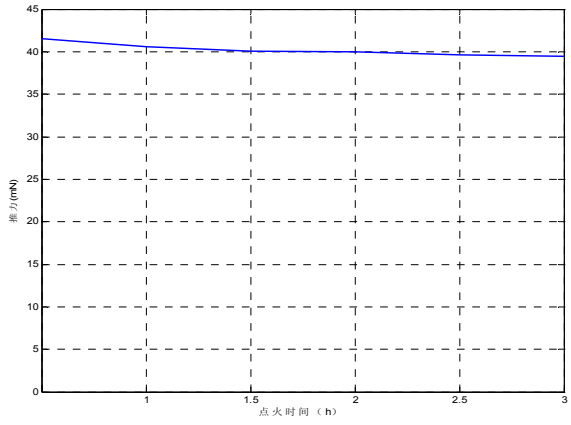


Figure 50 Curve of the thrust in the firing test

3.2 Life test of the ion thruster

Test goal. The Life time of the ion thruster is in order to verify the reliability of the thruster in the whole service life of the satellite. The test has been conducted in 2 different operation model-continuous firing and start-shutdown operation mode until the ion thruster fail to work.

Test approach. To simulate the orbital environment as real as possible, the ion thruster is fitted in a big vacuum chamber, and ground test facilities are employed to supply the power and xenon for the thruster.

Currently, the thruster has been operated continuously over 8500 hours. The performance characterization of the thruster during the test demonstrates its compliance with all beginning of life (BOL) specifications and the thruster doesn't reach its life margin. Some real-time tests are also conducted to gather the electric and flow related parameter, and an optical measurement device is employed to determine the sputtering rate of the ion optics whose result is shown in Fig. 5.

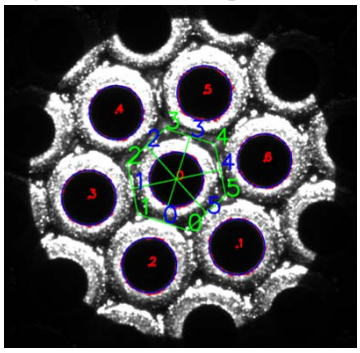


Figure 51 the sputtering photo after 6500 hours life test

The life test is still caring on to validate the life margin of the thruster.

3.3 Xenon refilling test

Test Goal. This test was conducted in the same way as the process and regulation in the launch site to verify whether the refilling approach can meet the high purity propellant specification.

Test approach. A new refilling test device was developed to fill the xenon as the same way that the satellite employs. The purity of the xenon was measured online and verified again after the filling procedure.

The test results demonstrate that the purity of the xenon which illustrated by the concentration of oxygen and water is compliance with the requirement of the thruster.

3.4 Plume test

Test goal. Plume test diagnosed many parameters in relative to the thruster plume and evaluate the acceptance of contamination and deposition caused by the ion thruster around the satellite.

Test approach. The ground test was conducted in combination with the numerical method to evaluate the plume effect in a vacuum circumstance. Different kinds of plasma diagnosis instrument were employed in the test to contour the distribution of the plume particles; in addition, thermal effect to the satellite caused by the ion thruster was also simulated through numerical model. To evaluate the influence of the plume to the star trigger and sun panels, a certain contamination test was organized with a solid model of the satellite and the test result was compared to the simulated result to improve the numerical model.

3.5 EMC test

Test goal. Electromagnetic Emission and the variation of amplitude and phase of the electromagnetic wave after it was transmitted through the thruster plume to make sure the plume wouldn't disturb the communication of the satellite.

Test approach. The test program, by the thought which was 'extrapolate the product test results to the design of simulator', was design according to the EMC simulator of the ion thruster. The normal operation condition of the ion thruster was contained into the regular procedure of the satellite EMC test. The ion thruster was located in a vacuum chamber that had affiliated

wave-transparent chamber. The electromagnetic emission character was measured while the thruster was firing. Then the EMC test data was analyzed to perfect the design of the EMC simulator of the ion thruster which would be used in the ultimate EMC test of the satellite.

The test results demonstrated that the ion thruster can be compliance with the satellite in the orbit.

3.6 Firing test of the whole satellite

Test goal. After the satellite was completed assembled, the electric propulsion system would be operated through the complete process in the orbit to validate its performance character and compatibility to other units on the satellite.

Test approach. All the other units of qualification model or electrical model which have the same function as the flying model were integrated in the test. The related sub-system such as integrated electronics, power supply, control and chemical propulsion would be configured as needed.

The test result demonstrated that the electric propulsion system have good compatibility with other sub-system on the satellite.

4 Conclusion

Electric propulsion is a technology aimed at achieving thruster with high specific impulse, which results in a reduction in the amount of propellant required for the communication satellite. Reduced propellant mass can significantly decrease the launch mass of the satellite, leading to lower costs from use of smaller launch vehicles to deliver a desired mass. Through the systematic and satellite on-board test of the electric propulsion system, CAST has validated the TRL of the electric propulsion technology and reduce its risk of engineering application.

References

[1] Huang Zhicheng, China Celebrates 45th Anniversary of Successful Launch of DFH-1, Space International, 2015-4

[2] www. space.skyrocket.de

[3] Zhou zhicheng, Wang Min, Li Feng. An Overview of Electric Propulsion Activities in Chinese Telecommunication Satellite, Space International, 2013-6

[4] Zhong Xiaoqing. Electric Propulsion DFH-3B Communication Satellite Platform, Aerospace China, 14(4), 2013

[5] Zhang Weiweng, Zhang Tianping. Technological Development and Application of Electrically Powered Spacecraft Propulsion, Space International, 2015-3

[6] John R. Beattie, XIPS KeIPS Satellites on Track, The Industrial Physicist, vol 4 issue 2, American Institute of Physics, June 1998.

[7] Ronald L. Corey, David J. Pidgeon. Electric Propulsion at Space System/Loral. IEPC-2009-270.

[8] J. Gonzalez del Amo, G. Saccoccia. Electric Propulsion Activities at ESA. IEPC-2009-237.

[9] Semenkin A.V. Overview of Electric Propulsion Activity in Russia. IEPC-2007-275.

[10] Steven A. Feuerborn1 and David A. Neary. Finding a way: Boeing's "All Electric Propulsion Satellite". AIAA 2013-4126

[11] Sun Mingming, Zhang Tianping, Chen Juanjuan. Thermal analysis of LIPS-200 ion thruster, HIGH PoWER LASER AND PARTICLE BEAMS, 26(8), 2014

Research in the Conductor Brazing Technology for the Hall Thruster in High-Temperature Working Environment

Yang Jian[1], Chen Bo[2]

(1. Beijing Institute of Control Engineering, Beijing 100190, China;
2. Beijing Institute of Aeronautical Materials, Beijing 100095, China)

Abstract: A Hall thruster is composed of an accelerator, a cathode and a bracket. The accelerator and the cathode should be electrically connected to the conductors reliably and stably since they work at a high temperature during the thruster operation. Because of the space limitation for jointing, the local heating induction brazing technique is used to connect conductors to the accelerator and the cathode.

The article describes the content and technical data of the local heating induction brazing technology for the hall thruster with details. Through the structure design and lots of tests, the stand holder for brazing was manufactured. The geometric structure for the conduction part and the brazing data was verified. The final test showed that the acceptance rate for brazing increased to 85%. This technological research solves the key problem during the manufacture of the Hall thruster and could be widely used in similar conductor connection cases in high-temperature working environment.

Keywords: Hall thruster; Electrically connected; Induction brazing; High-temperature; Technological research

1 Introduction

The Hall thruster is a world-prevailing electric thruster for its outstanding performance, simple structure, and high reliability. It is widely used for spacecrafts in America, Europe and Russia. NSSK for synchronous orbit communication satellites and main thruster for deep space exploration are successful applications. Being the most promising thruster for future spacecrafts, the Hall thruster has a large thrust power rate and a high thrust density. This design helps to increase the effective loads of a spacecraft, decrease its launching cost, and prolong its lifetime. It is an effective means to improve commercial satellites efficiency and enhance competitive power. Thus, the Hall thruster, a more promising thruster for NSSK and orbit lifting, has a huge potential market.

A Hall thruster is composed of two parts: an accelerator and a cathode, and a bracket as the carrier, which also provides circuit and air interfaces to connect the thruster to satellites.

The excitation conductor of the accelerator is coiled on the magnetizer and then connected to a common outer lead, to ensure that the electromagnetic coil produces a magnetic field once the power is on. The electromagnetic conductor withstanding up to 500℃ high temperature has been adopted for the excitation system since its working temperature is around 400℃. The cathode is used to emit electrons and provide startup temperature for electron emission. Since its working temperature is around 500℃, the cathode is also connected to a common outer lead. Both parts work properly to guarantee the normal operation of the Hall thruster.(1)

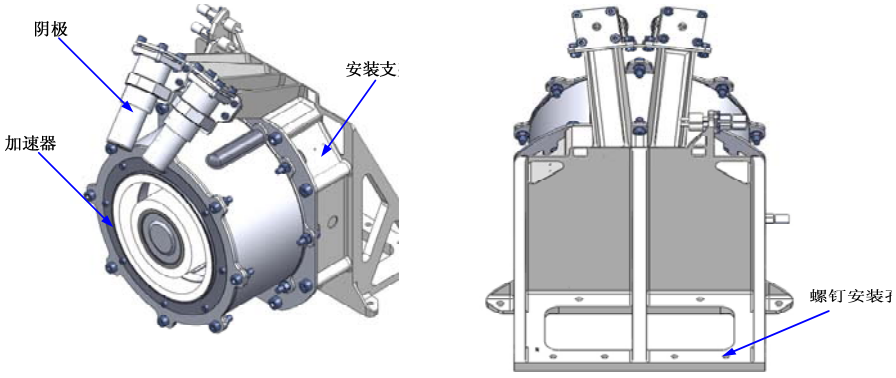


Figure 1 Structure of A Hall Thruster

2 Subject Selection

2.1 Background

The following figures are respectively the Accelerator Electric Connection and the Cathode Electric Connection. Figure 2 shows that on the accelerator, one single-head conductor from the additional coil has to be connected to one that from the inner coil, and that each single-head conductor from another additional coil and an inner coil are respectively conducted to two equidirectional multiple-head outer leads. The cathode base is to be connected to the lead. As the figure shows that two conductors, three conductors and seven conductors are respectively connected to two-head, three-head and seven-head filling holes. The whole wiring secures the power supply to the Hall thruster for normal operation.

There are normally two options for engineering conductor connection, namely crimping and brazing. A single wire is used as the electromagnetic conductor on the accelerator while stranded wires are used for the connection of the accelerator and the outer lead of the cathode. According to related standards on wire crimping, the standard Q/W732A-2004, Crimping Process Specification of Crimped-mode Electric Connectors Terminals for Spacecrafts and the standard Q/W731-97, Crimping Process Specification of O-type Terminals for Spacecrafts, made by China Academy of Space Technology (also called CAST), crimping conductors should be silver-gilt stranded copper conductors and the single conductor should not be crimped in principle. Given the fact that we are adopting brazing for the conduction of the cathode base and the outer lead, brazing becomes a feasible technique for the Hall thruster connection.

The accelerator is connected to the outer lead to provide a magnetic field once the power is on. The cathode base is connected to the outer lead to provide the startup temperature for electron emission. Either connection has direct influence on the Hall thruster's operation. It is crucial to the successful operation of the thruster. For this reason, we propose to carry out research in the conductor brazing technology for the Hall thruster in high-temperature working environment.

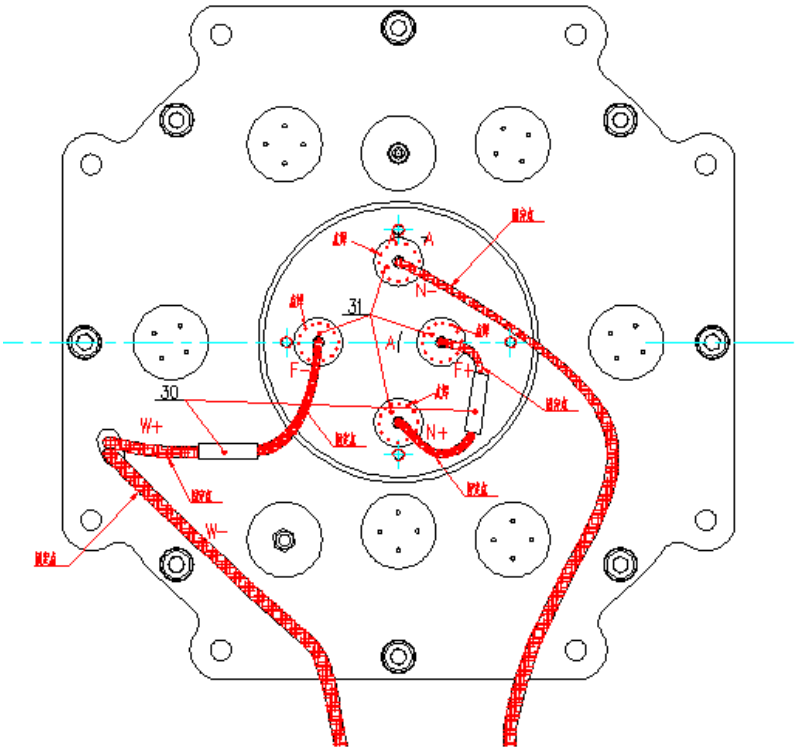


Figure 2 Accelerator Electric Connection

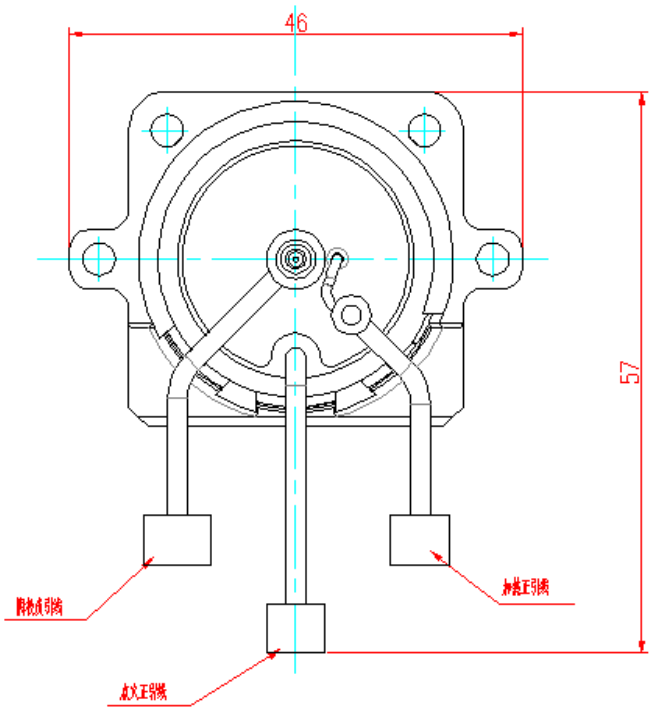


Figure 3 Cathode Electric Conduction

2.2 Crucial Problems of the Conductor Brazing Technology

Given that the working temperature of the accelerator and the cathode of the Hall thruster is around 400 to 500℃, the conductors used should be high temperature resistant and highly reliable. It should have small resistivity change after brazing connection. Thus it ruled out soft soldering since it only stand the temperature up to 183℃. Reading through various national and international references didn't help much in this respect. Therefore, we summed up the following crucial problems in the brazing connection of the accelerator and cathode conductors:

(1) The design of the conductor cooling unit and the selection of the location for the filler metal installation have great influence on the protection of the surface of the outer leads due to high working temperature.

(2) The design of the brazing interfaces on the accelerator and the cathode of the Hall thruster and the optimization of the brazing data are key to guarantee the brazing quality.

(3) If the brazing technology is adopted for the connection, the filler metal should be able to withstand a high temperature over 500℃, and the installation of the filler metal for brazing should be carefully designed.

2.3 Subject Confirmed

Based on the foresaid, we decided to do the Research in Conductor Brazing Technology for the Hall Thruster in high-temperature working environment.

3 Raised Research Proposals and the Selected

3.1 Research Proposal

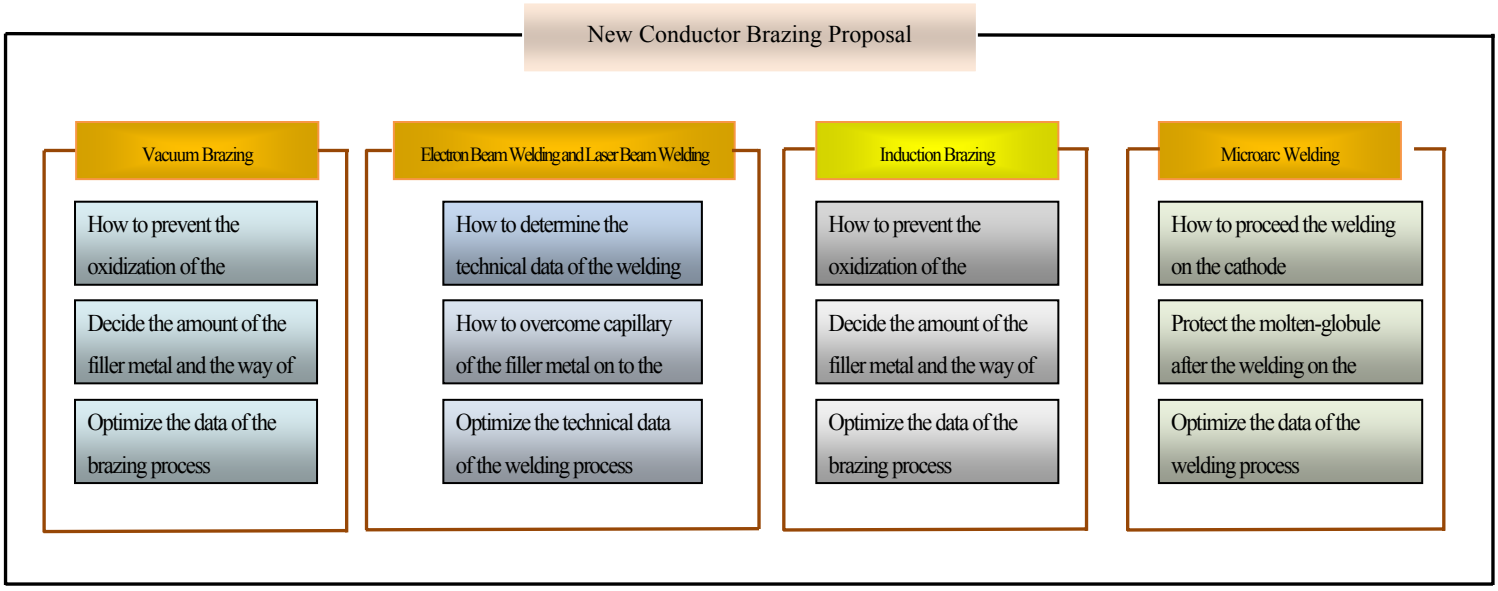


Figure 4 Technical Proposals for New Conductor Brazing Process[2]

Table 1 Features of Various Brazing Proposals[3]

Type	Technical Specification
Vacuum Brazing	It is a process of heating the workpieces and the filler metal in vacuum . In vacuum , the evaporation of the metal, especially the metallic oxides can effectively damage the surface oxidation film so that fluxless brazing becomes possible.
Induction Brazing	It is a brazing process where the power frequency current, the high-frequency current and the medium-frequency current act as the thermal source. The workpieces when put between the induction coils which would turn into a magnetic field with power on will be heated by the resistance heat produced by the induction current in the alternating magnetic field. And this local heating can be accurately controlled.
Electron Beam Welding and Laser Beam Welding	Cast the laser beam on a very small area through optical system for a very short period so as to form a high-density thermal source on the welding parts. Thus, the workpieces shall melt and form rigid welding points and seams. For electron beam welding, the cathode of an electron gun when directly or indirectly heated will emit electrons. Hit the workpieces with electron beams. When the huge kinetic energy turns into thermal energy, heating the welding workpieces till it melt into a pool and thus the workpieces shall be welded.
Microarc welding	It was a welding process where the workpieces and the filler metal (if any) are melt by the microarc thermal produced between the tungsten electron and the workpieces under the protection of inert gas(mainly argon). This welding presents nice welding seams.

3.2 Selection of the Subject Research Proposal

Workpieces were put through the above four processes respectively and the best proposal shall be selected with the following two standards:

The conductor with smooth and no-oxidation surface after the treatment indicates the better proposal it underwent.

Judge by the reliability of the jointing parts. An electric drawbench is used to test the tensile strength of the conductor and the one with bigger tensile strength indicates a better proposal it underwent.

Table 2 below shows the after-treatment pictures under each proposals.

Table 2 Result of the Four Proposals

Type	Result
Vacuum Brazing	

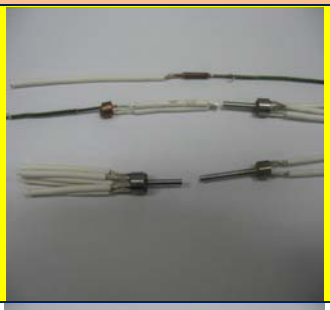
Type	Result
Induction Brazing	
Electron Beam or Laser Beam Welding	
Micro-arc Welding	

Table 3 Review of the Four Results

Proposal	Surface	Tensile strength	Yes or No
Vacuum Brazing	The surface of the conductor outer surface turned black and was oxidized after the trial brazing.	150N	No
Induction Brazing	The outer surface of the conductor was smooth and not oxidized after the trial brazing.	220N	Yes
Electron beam or laser beam welding	The outer surface of the conductor was smooth and not oxidized after the trial brazing.	Capillary action happened on the conductor during trial welding resulting fracture at the welding point.	No
Micro-arc Welding	The outer surface of the conductor was smooth and not oxidized after the trial brazing. But the cathode base cannot be connected to the conductor in this way.	178N	No

3.3 Determine the Best Solution

The above tests and analyses show that the best solution would be the local heating induction brazing.

4 Presetting of the Subject Goal

Table 4 below shows the quality requirements within our institute and the acceptance rate under this solution should reach over 85%.

Table 4 Technical Data Required

Item	Data
Brazing Rate	≥75%
Max. Linear Measurement per Radial hole	≤0.7mm
Max. Linear Measurement per axial hole	≤1.5mm

5 Optimization and Selection of the Subject Technical Data

When we selected the best solution, we analyzed three factors that affect the quality of the induction brazing in details shown as follows, namely power, air condition and filling type of the filler metals.

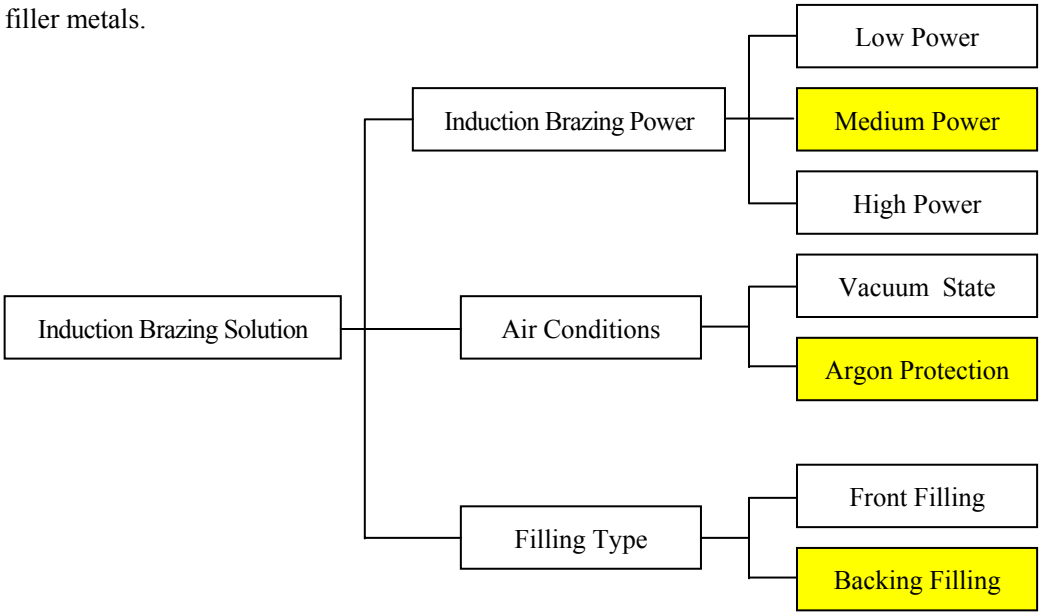


Table 5 Factors Affecting the Quality of Induction Brazing[2]

5.1 Selection of the Induction Brazing Power

Features of induction brazing under three powers shown in Table 5

Table 5 Features of Induction Brazing under Three Powers

Induction Brazing Power	Features
Low Power (0-4KW)	Slow heating, stable flow of the filler metal, no overheating
Medium Power (4-8KW)	Faster heating, improved efficiency, protecting the conductor outer surface
High Power (8-20KW)	Fastest heating, best efficiency and short time under high temperature

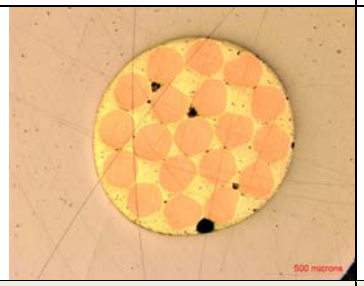
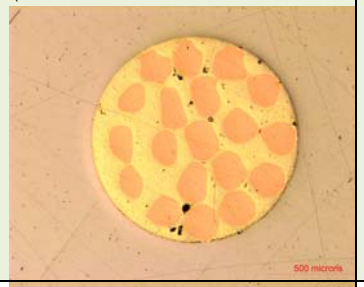
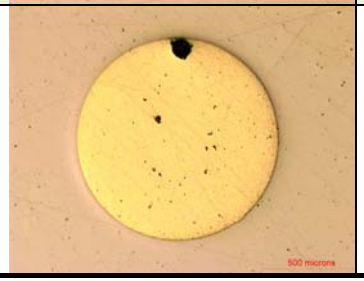
Factors that used for determining the proper power, cord quality, filler metal conditions, conductor quality, operation etc. are listed in Table 6 below with the assessment standards.

Table 6 Assessment Standards for Factors Affecting Induction Brazing

Score	Cord Quality	Filler Metal Conditions	Conductor Quality	Operation
5	Clear cord number and boundaries	Homogeneous filler metal, no hole	Conductor outer surface not oxidized	Very easy
4	Clear cord number and boundaries, very few holes on the boundaries	Homogeneous filler metal, a few holes, the brazing rate and the size of holes up to the required standards	Conductor outer surface slightly oxidized	Much easy
3	Clear cord number and boundaries, a few holes on the boundaries	Homogeneous filler metal, some holes, the brazing rate and the size of holes up to the required standards	Conductor outer surface slightly oxidized	Easy
2	Clear cord number and boundaries, many holes on the boundaries	Homogeneous filler metal, many holes, the brazing rate and the size of holes not up to the required standards	Conductor outer surface partially oxidized	Difficult
1	Conductor cords were melt with the filler metal so that the cord number could no longer be told clearly.	Homogeneous filler metal, the brazing area and the size of holes not up to the required standards	Conductor outer surface severely oxidized	Very difficult

Based on the above standards, we conducted three experiments and the results are listed in table 7. Finally the medium power was adopted with the total time of induction brazing is 50s.

Table 7 Selection of Power for Induction Brazing

Power	Result Show	Cord Quality	Filler metal	Conductor Quality	Operation	Score	Yes or No
Low Power (0-4KW)		3 (Clear cord number, 4 holes on adjacent boundaries)	3 (Homogeneous filler metal, six holes, the brazing rate and the size of holes up to required standards)	4 (Conductor outer surface slightly oxidized)	4 (Relatively easy)	14	No
Medium Power (4-8KW)	\ 	4 (Clear cord number and boundaries, very small holes on adjacent boundaries)	4 (Homogeneous filler metal, tiny holes, the brazing rate and the size of holes up to required standards)	5 (Conductor outer surface no oxidized)	4 (Relatively easy)	17	Yes
High Power (8-20KW)		1 (The conductor cords melt with the filler metal, the cord number could not be told.)	2 (Homogeneous filler metal, conductor cords melt with the filler metal during heating, the brazing rate and the size of holes not up to required standards)	2 (Conductor outer surface partially oxidized)	4 (Relatively easy)	9	No

5.2 Selection of Working Air and Filling Type

The filler metal is added horizontally during the brazing of a single conductor of the accelerator as shown in figure 6. The process data of induction brazing of stranded conductors shall also apply in this application.

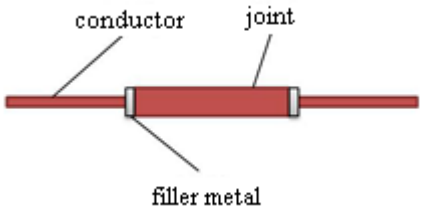
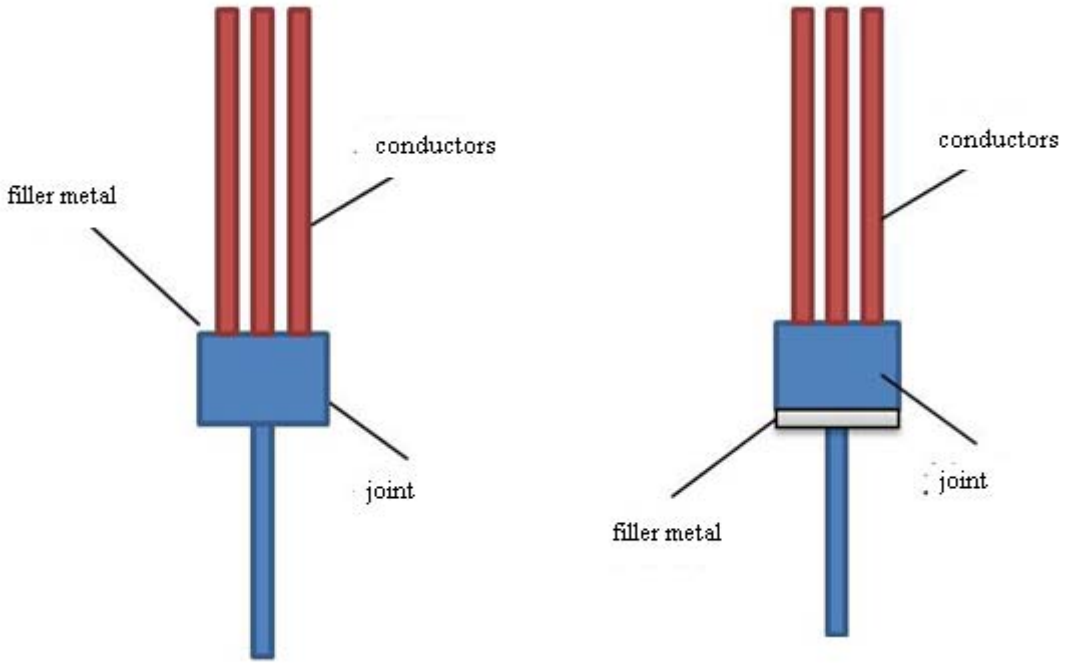


Figure 6 Induction Brazing of A Single Conductor of the Accelerator

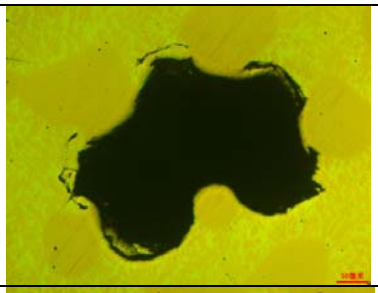
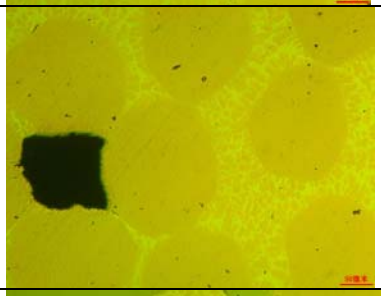
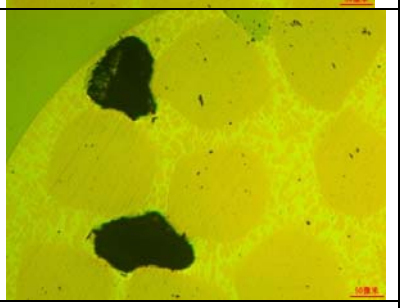
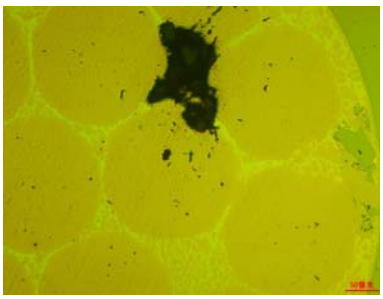
Now we focus on research in the brazing of stranded conductors to the alloy cathode base made of Mo-Cu-Ni hermetic whereas the working air and the filling type of the filler metal are two important factors affecting the brazing. There are two conditions for each factor, working air (0-Vacuum , 1-argon) and filling type (0-front filling, 1-back filling). The following figure shows that the two filling types and tests with both factors under two conditions were designed and the test schedule is as follows.



Front Filling Back Filling

Figure 7 Filler Metal Filling Scheme

Table 8 Experiment Schedule

Item	Working Air	Filling Type	Results
1	0	0	
2	0	1	
3	1	0	
4	1	1	

Assessment standards were made based on factors affecting assessment include joint-strength, filler metal filling, conductor quality, productivity, economic efficiency etc. Shown in Table 9.

Table 9 Assessment Standards on Factors Affecting Assessment on Working Air and Filling Type

Score	Joint Strength	Filler Metal Filling	Conductor Quality	Economic Efficiency	Productivity
5	P≥250N	Homogeneous filler metal, no holes	No oxidation on conductor outer surface	Small consumption of workpieces during brazing	Very high
4	200N≤P≤250N	Homogeneous filler metal, few holes, the brazing rate and the size of holes up to required standards	Extremely slight oxidation on conductor outer surface	Relatively small consumption of workpieces during brazing	Medium high
3	150N≤P≤200N	Homogeneous filler metal, a few holes, the brazing rate and the size of holes up to required standards	Slight oxidation on conductor outer surface	Normal consumption of workpieces during brazing	High
2	100N≤P≤150N	Homogeneous filler metal, many holes, the brazing rate and the size of holes not up to required standards	Partial oxidation on conductor outer surface	Medium high consumption of workpieces during brazing	Normal
1	P≤100N	Homogeneous filler metal, lots of holes, the brazing rate and the size of holes not up to required standards	Severe oxidation on conductor outer surface	High consumption of workpieces during brazing	Not high

Tests were done based on the above standards and the results were listed in Table 10 showing that the No. 4 test had the highest score. It means that the induction brazing receives best effect when working under argon environment and the filler metal filled from the back .

Table 10 Selection of Working Air and Filling Type

Item	Joint Strength	Filler Metal Filling	Conductor Quality	Economic Efficiency	Productivity	Score	Yes or No
1	3 The tensile strength is 187N.	2 Homogenous filler metal, one big hole, the brazing rate and the size of holes not up to preset standards	4 Extremely tiny oxidation on conductor outer surface	2 Relatively high consumption of workpieces during brazing	2 Normal	13	No
2	4 The tensile strength is 221N.	4 Homogenous filler metal, one hole, the brazing rate and the size of holes up to preset standards	5 No oxidation on conductor outer surface	2 Relatively high consumption of workpieces during brazing	2 Normal	17	No
3	4 The tensile strength is 202N.	3 Homogenous filler metal, two holes, the brazing rate and the size of holes up to preset standards	4 Extremely tiny oxidation on conductor outer surface	4 Relatively low consumption of workpieces during brazing	4 Relatively high	19	No
4	4 The tensile strength is 236N.	4 Homogenous filler metal, one hole, the brazing rate and the size of holes up to preset standards	5 No oxidation on conductor outer surface	4 Relatively low consumption of workpieces during brazing	4 Relatively high	21	Yes

5.3 Selection of Best Data for Induction Brazing

Based on the above test results and analyses, medium power was adopted for the local heating induction brazing under argon protection with the filler metal filled from the back.

5.3.1 Selection of Best Data under Medium Power Operation

5.3.1.1 Selection of Best Power Data for Stranded Conductors Brazing

Test factors: power regulation period, power duration (power regulation period, 50s)

Value set at: 10 15 20 25

The test results for stranded conductors cases were listed in table 11. It reviews that the workpiece in No. 3 test has the min. internal defects and has the best quality. This is to say, during the four tests on stranded conductors, when the power regulation period is 20s and the max. power duration is 30s, we got the best joint quality through induction brazing.

5.3.1.2 Selection of Best Power Data for Single Conductor Brazing

Test factors: power regulation period, power duration (power regulation period, 50s)

Value set at: 10 15 20 25

The test results for single conductor cases were listed in table 12. It reviews that the workpiece in No. 6 test has the min. internal defects and has the best quality. This is to say, during the tests on single conductor brazing, when the power regulation period is 15s and the max. power duration is 35s, we got the best joint quality through induction brazing.

5.3.2 Result Analysis

The power data for single-conductor brazing and stranded-conductor brazing are different. The reason is that the joint diameter of stranded conductors is bigger than that of the single conductor, thus it takes longer time to spread the heat through induction heating and to slowly regulate the power is good for stable temperature rise which will allow the filler metal to be fully melt and fill the whole filling clearance with less defects. The final induction brazing data is selected as follows: the induction power for stranded conductors is 7kW with the power regulation period of 20s and the max. power duration of 30s for back filling case and the induction power for single conductor is 5kW with the power regulation period of 15s and the max. power duration of 35s for side filling case.

Table 11 Test Results on Stranded Conductors

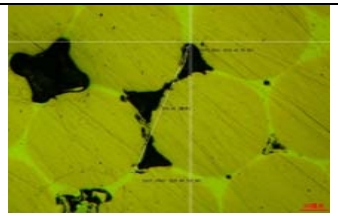
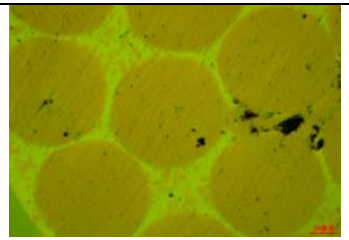
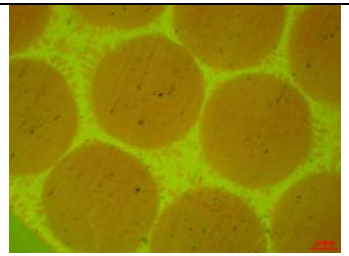
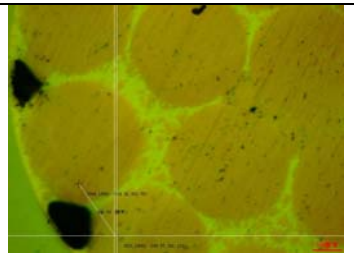
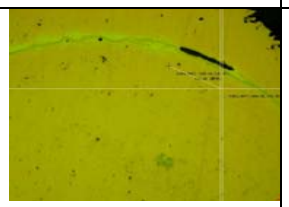
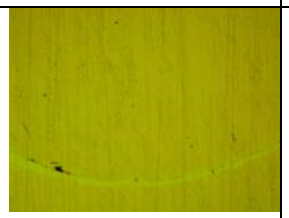
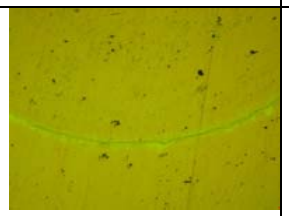
Item	**Conductor Type**	**Power Regulating Period**	**Power Duration**	**Test Result**	**Max. linear dimension of each hole on radial plane≤0.7mm**	**Max. linear dimension of each hole on axial plane≤1.5mm**
1	Stranded conductors, including two-head, three-head and seven-head conductors shall be connected to the Mo-Cu-Ni hermetic alloy joint through induction brazing.	10s	40s		0.6	1.3
2		15s	35s		0.3	0.4
3		20s	30s		0.1	0.2
4		25s	25s		0.5	0.87

Table 12 Test Results on Single Conductor

Item	**Conductor Type**	**Power Regulating Period**	**Power Duration**	**Test Result**	**Max. linear dimension of each hole on radial plane≤0.7mm**	**Max. linear dimension of each hole on axial plane≤1.5mm**
5	Single conductor brazing	10s	40s		0.6	1.2
6		15s	35s		0.15	0.6
7		20s	30s		0.24	0.9
8		25s	25s		0.31	0.75

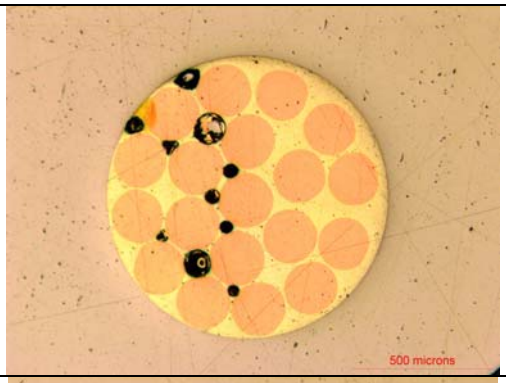
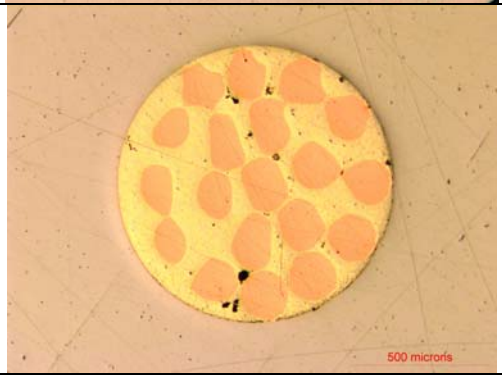
5.3.2 Argon Pressure Determination

Figure 8 The Induction Brazing Furnace and Coils

The above drawing shows the induction brazing furnace and the induction coils. If the argon pressure is bigger than one air pressure, we shall not be able to use the glove to reach the workpieces for operation. Therefore, we only compare the brazing qualities under the argon pressure equal to and smaller than one air pressure.

We tried brazing on workpieces respectively under the argon presser lower than one air pressure and equal to one air pressure and then took metallographic cutting test on the brazing workpieces to inspect its brazing quality. Details are given in the following table.

Table 13 Brazing Quality Result List of Trial Workpieces

Argon Pressure	Test Result	Analysis
Argon pressure < one air pressure		The metallographic picture shows that there were many big holes.
Argon pressure = one air pressure		The metallographic picture shows that there were only a few small holes

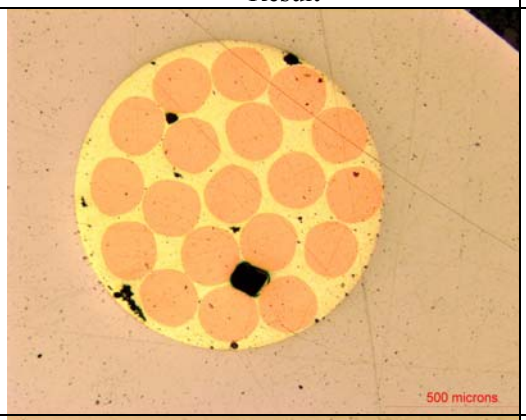
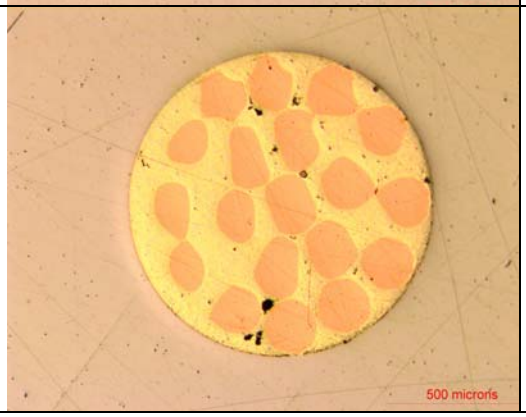
The above table shows that when the argon pressure is lower than one air pressure, there would be many big holes after brazing, which could not meet the quality requirements . But when the argon pressure equals to one air pressure, there would be only a few small holes, which actually would meet the quality requirements.

Based on the test results, we decided to adopt one air pressure for the argon pressure. In this way, the argon pressure inside the furnace would equal to the air pressure outside of the furnace, reaching a balance to prevent air exchange and ensure the purity of argon for better quality of the brazing process. If the argon pressure is lower than one air pressure, there would be a difference between the argon pressure inside the furnace and the air pressure outside the furnace and thus the air would come into the operation room and lower the argon purity. When air exists during brazing, the conductor outer surface would be easily oxidized during brazing operation. Besides, the flow of the filler metal in this situation would not be so good as it was under argon protection of higher purity. In this case, it is not good for the brazing quality.

5.3.3 Determination on Adopting Backing Filling of the Filler Metal

We conducted tests on the filler metal pressing against the back of cathode base and not against the back of the cathode base and then put them through metallographic cutting tests to see the brazing results.

Table 14 Brazing Quality List on Back Filling of the Filler Metal

Item	Back Filling Type	Result	Brazing rate
1	The filler metal did not press against the base brazing seam.		72%
2	The filler metal pressed against the base brazing seam.		92%

The test results show that during back filling brazing, the filler metal pressing against the cathode base brazing seam is good for the internal air inside the seam to getting out. Under the capillary action, the filler metal could fill in the whole seam thus making the brazing rate up to the required standards. Otherwise, the filler metal would be of poor flowability leading to more holes. In that case, the brazing rate will fail to meet the required standards.

6 Subject Goal Check

Tests on ten more products has been conducted based on the optimized brazing process data and the test results are listed in table 15 below.

Table 15 Quality Inspection List on Trial Workpieces

No.	Inspection Item			Result
	Brazing rate≥75%	Max. linear dimension for each hole on radial plane≤0.7mm	Max. linear dimension for each hole on axial plane≤1.5mm	
1	98%	0.2	1.2	Qualified
2	87%	0.6	1	Qualified
3	83%	0.1	1.3	Qualified
4	78%	0.3	0.9	Qualified
5	77%	0.5	0.8	Qualified
6	84%	0.4	0.6	Qualified
7	81%	0.7	1.4	Qualified
8	96%	0.6	1.5	Qualified
9	91%	0.4	1.1	Qualified
10	89%	0.2	1.6	Not qualified
Acceptance rate		90%		

Table 15 shows that the 10 products tested according to the optimized brazing data, nine are acceptable. The internal quality and the brazing rate meet the subject presetting requirements. Thus the acceptance rate reached 90% while the preset acceptance rate is 85%. Therefore, the subject goal is realized.

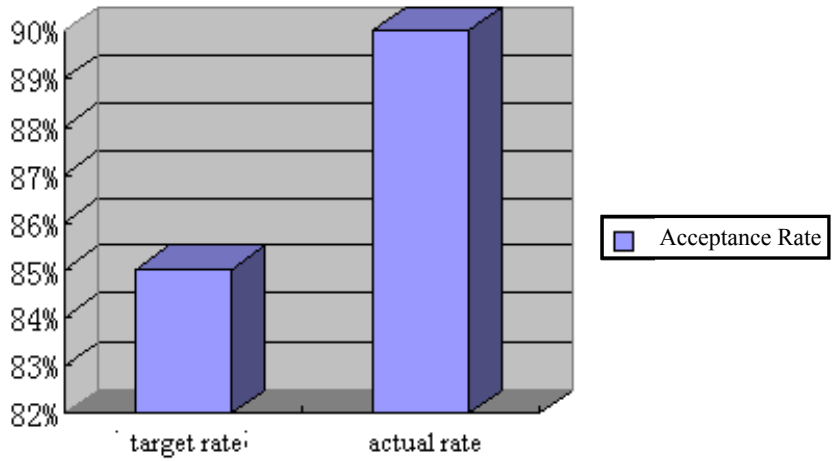


Chart 9 Bar Chart on Subject Target/Actual Acceptance Rate

7 Conclusion

The induction brazing technology under this subject research solves the problem of the conductor connection to the Hall thruster in high-temperature working environment. It is the key technology for electric Hall thrusters and plays a crucial role in the application of Hall thrusters on satellites. Besides, this technology can also be used in similar cases of conductor connection in high-temperature working environment and has a broad application potential in future engineering field.

References

[1] Fundamentals of Space Electric Propulsion edited by Yu Daren and others and published by Harbin Engineering University Press in 2014:157-176.

[2] General Knowledge on Quality Management edited by China Association for Quality and published by China Metrology Publishing House in 2011:157-176.

[3] Technology and Application of Special Welding edited by Li Yajiang etc. and published by Chemical Industry Press in 2008:157-176.

Development of Non-Explosively Actuators

Liu Shiyi, Chen Jin

(Beijing Institute of Space Mechanics & Electricity)
(Mailbox 5142,239 branch, Youyi Road, Haidian District, Beijing, 100094, China)

Abstract: Non-explosively actuators (NEAs) on spacecraft are developed associated with pyrotechnic devices. Because of the shortcoming of functional shock, safety, and overall system cost issues, pyrotechnic devices were limited their continued use on space missions. With the emergence and application of new technology, America and Europe launched the promotion of NEAs alternative to pyrotechnics. Recent years, NEAs have been required further development in the demand of detection online and low impact mechanisms. The developed and potentially related technologies were introduced in this paper, which contributes to the design of mechanism in spacecraft.
Keywords: Non-explosively Actuator Development Application

1 Introduction

Over the past 50 years, pyrotechnic powered devices provided the greatest power-to-weight ratio for all space mechanisms. However, because of the shortcoming of functional shock, safety, and overall system cost issues, pyrotechnic devices were limited their continued use on space missions. With the emergence and application of new technology, America and Europe launched the promotion of non-explosively actuators (NEAs) alternative to pyrotechnics. Upon the request from NASA, Langley Research Center conducted the surveys to evaluate the mature of alternative technology, and listed a number of potential devices [1]. ESA summarized the key technologies and compared the level of shock with pyrotechnics [2]. Recent years, NEAs have been required further development in the demand of detection online and low impact mechanisms, such as solar array deployment, antenna release, and load hold-down and release, etc. The developed and potentially related technologies were introduced in this paper, which contribute to the design of mechanism in spacecraft.

2 Torque motor

Hi-Shear Technology Corporation has designed an electrically powered separation devices using mature release technology [3]. Figure 1 shows the appearance and configuration of the device. The configuration of pyrotechnic separation nut was inherited. Three threaded segments are held radially together by a locking ring. In the locked position, the cam drive ring is forcing the rollers through the slots in the outer helical ramp into the pockets in the inner helical ramp. When the cam ring is allowed to rotate by the torque motor, the radial force on the rollers rotates the cam ring, releasing the rollers from the pockets in the inner ramp. When the bolt load release, the radial force on the segments is removed and the wave spring lifts the locking ring off of the segments. The segments are driven outward, releasing the bolt.

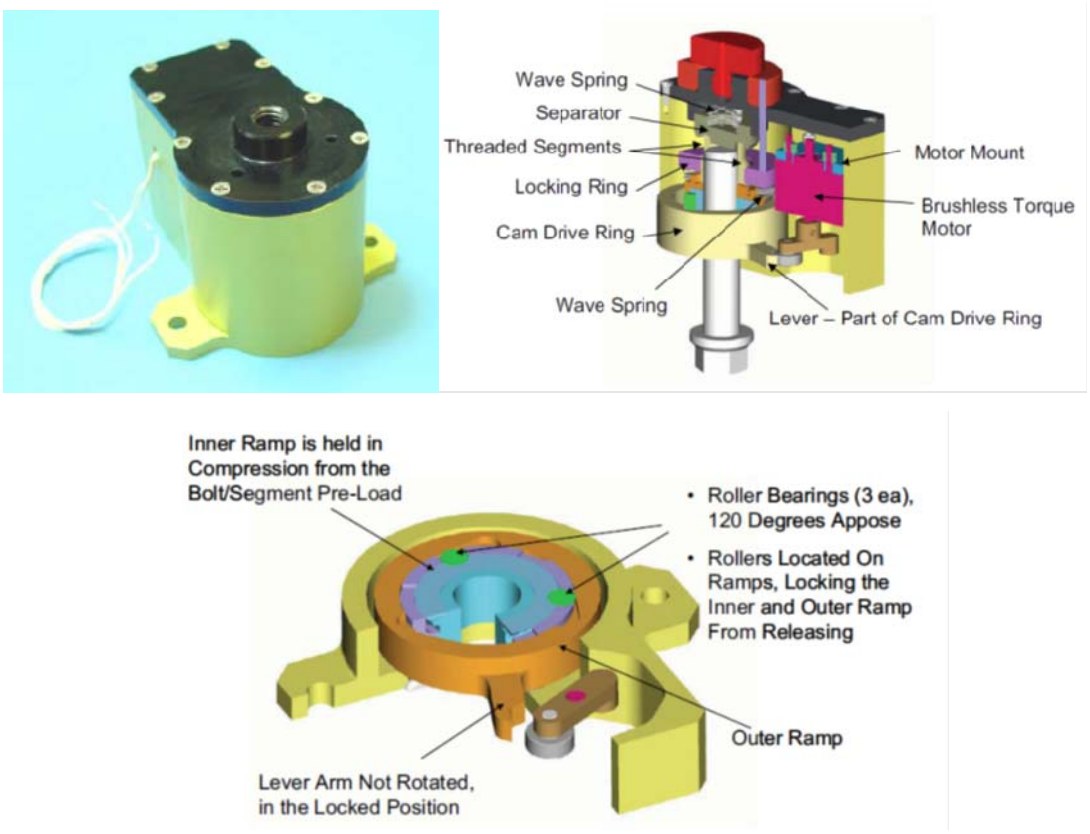
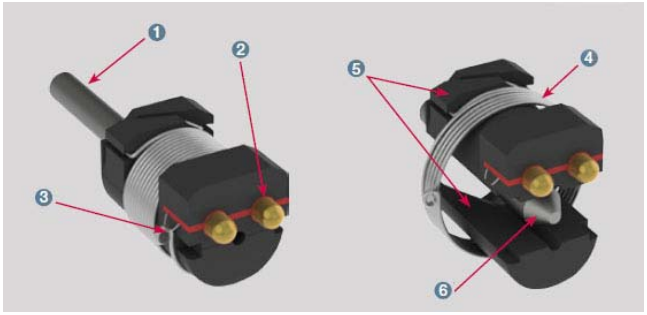


Figure 1. Cross view of separation devices

The devices have advantages over pyro-powered ones, some of which are resettable, no out-gassing and ultra-low shock, etc. However, the driving of torque motor also has some disadvantages. The complex structure of motor reduces the reliability of products, which restricts its further application.

3 Split spool

Split spool technology is the heart of various Cooper Interconnect's release mechanisms and electrical switches. The electrical characteristics of split spool are chosen to mimic a pyrotechnic initiator [4]. Thereby the power requirement is compatible with pyrotechnic actuator circuitry. Figure 2 depicts an armed split-spool assembly, which is armed by applying compression spring-force to the plunger. At actuation, electrical current causes the link wire to open and release the tensile wire wrap. With radial expansion, the tensile-wire wrap releases tension on the split-spool bobbin, and separation of the two halves is facilitated by the forward movement of the spring-loaded plunger.



1. Plunger
2. Electrical Contacts
3. Linkwire
4. Restraining Wire
5. Spool Halves

Figure 2. Configuration of Split spool

Split spools can be replaced initiators in all the pyrotechnic devices. Figure 3 and Figure 4 shows the separation nut and pin-puller, these devices have proven having flown on a variety of spaceflight missions. They are highly reliable, provide fast actuation, possess high energy output for limited power input, and are reusable following refurbishment. The primary disadvantage is mechanical shock. The releases stored mechanical energy in the spring loaded portion of the mechanism to effect primary device functioning [1].

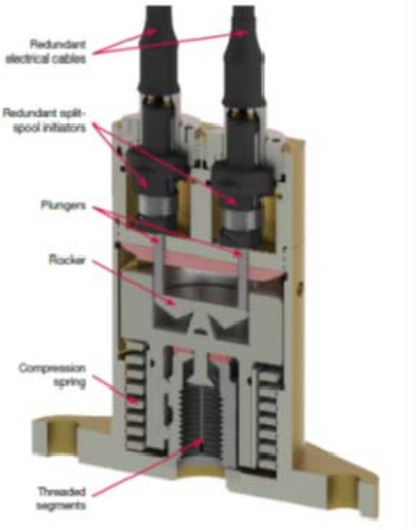


Figure 3. Separation Nut by using Split-spool

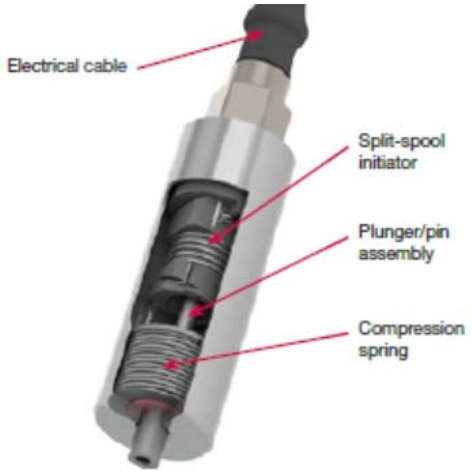


Figure 4. Pin-puller by using Split-spool

4 Shaped Memory Alloy

Shaped memory alloys refer to a group of materials which have the ability to return to a predetermined shape when heated. This phenomenon can be harnessed to provide unique and powerful actuators. The shape memory effect is caused by a temperature dependent crystal structure. When an SMA is below its phase transformation temperature, it possesses a low yield strength crystallography. While in this state, the material can be deformed into other shapes with relatively little force. The new shape is retained provided the material is kept below its transformation temperature. When heated above this temperature, the material reverts to its parent structure causing it to return to its original shape [5].

4.1 SMA direct actuator

TiNi Aerospace Corporation Provides a kind of SMA direct actuators named frangibolt, which is a generic assembly that may be readily adapted to a broad array of applications [5]. See Figure 5, the bolt includes a slit on threaded end to allow for removal and replacement of broken segment after actuation. After frangibolt actuator was heated for about few seconds of power, it increased in length. During the process which caused the bolt to break, Integrated cup-cone feature resists lateral loads. Panel Separation may be adjusted to suit each individual application by changing the cone height, whereas this does not affect the basic joint design.

The user must avoid bolt bending loads. Another concern was maintaining heater contact with the SMA slug during heating as the slug diameter shrinks. Special attention must be paid to heater design to insure sufficient contact.

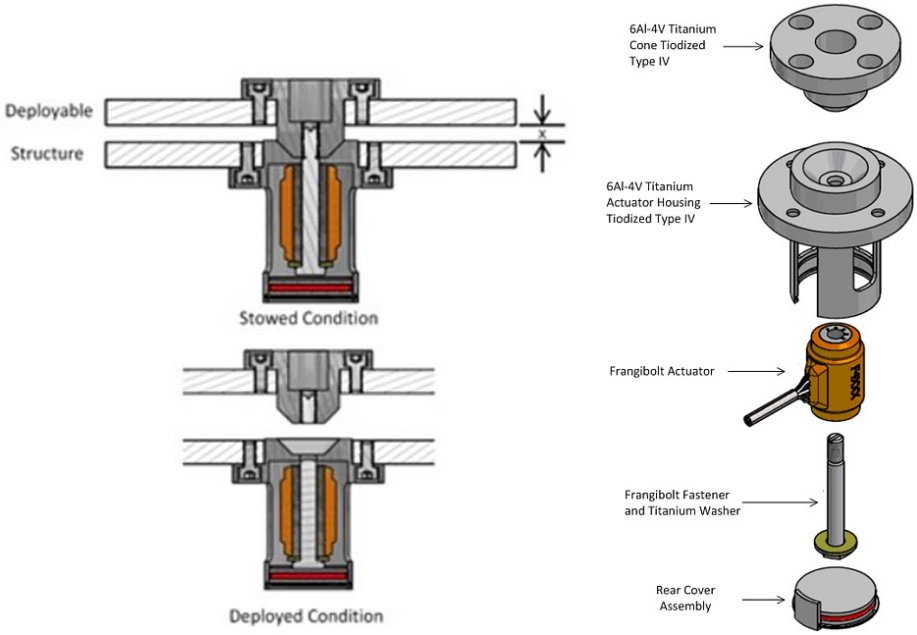


Figure 5. Frangibolt assembly

4.2 SMA trigger actuator

TiNi Aerospace Corporation and Space Nevada Corporation all own their SMA trigger actuators. The devices utilize a roller-bearing type geometry to achieve force reduction and to maximize mechanical efficiency [5][6][7]. When the tensile preload is applied to the bolt, the threads react against the nut segments in the release device. The contact angle of the threads creates both radial and axial loads against the segments. The axial loads are reacted into the base plate of the mechanism and into the spacecraft structure. In the latched position the redial loads are reacted into the rollers and into the notched outer race. In this position, all loads are carried through the structure and do not act through the SMA wire. Rotation of the race allows the rollers to drop into the notches. The radial load path is no longer intact and the nut segments are free to move outward. Figure 6 shows the basic configuration of the segments, the rollers, and the outer race.

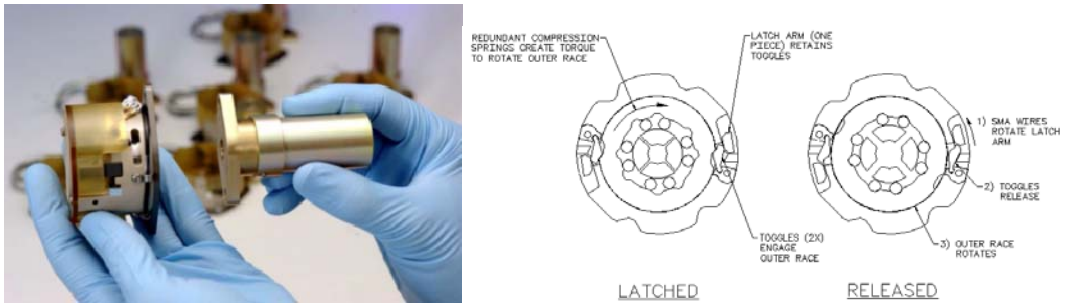


Figure 6. Latch geometry and operation of SMA trigger Separation Nut

ARQUIMEA company has been designed a new Resettable Hold Down and Release Actuator based on SMA technology [8]. This design meets a high preload requirement. REACT is electrically connected from outside the structure to SMA actuator using proper supply cables, crimps, and insulation pieces. Electrical crimps provide the connection among the power supply cables and SMA actuators in a compact and robust way. The SMA wires are fixed by commercial crimps, one to a screw in the SMA Actuator Column and the other to a screw in the crown, as shown in Figure 7. One independent pair of power supply cables is used for each of the two redundant SMA actuators and another pair is used for the release switch connection. The switch presents an infinite resistance when the REACT is in its initial position and a zero resistance when the preload is released at the final position. Finally, all the cables are extracted from the REACT device using proper slots in the bottom of the housing. This jacket is also used to keep the wire isolated in case of SMA break.

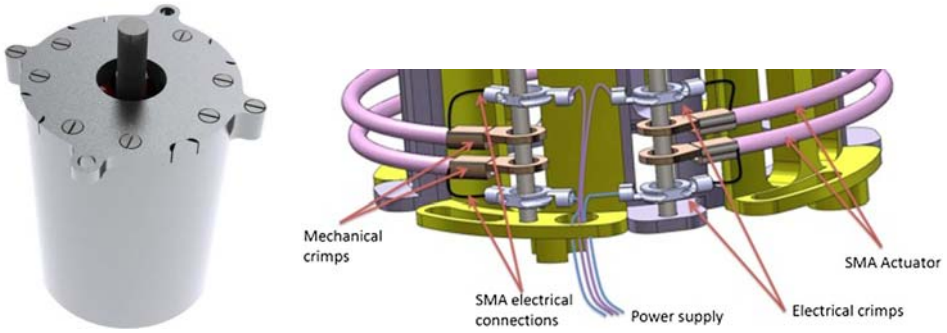


Figure 7. SMA arrangement in the REACT mechanism

Korea Advanced Institute of Science and Technology has developed a newly designed SMA trigger actuator [9]. The trigger operation for the bolt release is performed by the contraction of the SMA wire. This activation is fully redundant, as it is provided with two SMA wires electrically connected to different circuits. Once the SMA wires contract, the trigger block moves in the upward direction, from the position shown in Figure 8, after which the balls move radially in a grooved slot in the trigger block due to the annular surface formed on the inside of the locking sleeve loaded by the compression spring. The locking sleeve and rollers then move upwardly so that the segmented nuts are separated into three pieces and the bolt is released. Figure 8(b) shows the released position of the device. The proposed non-explosive actuator was shown to be fully reusable after all tests. It provides a very easy resetting mechanism without a change of any of its components.

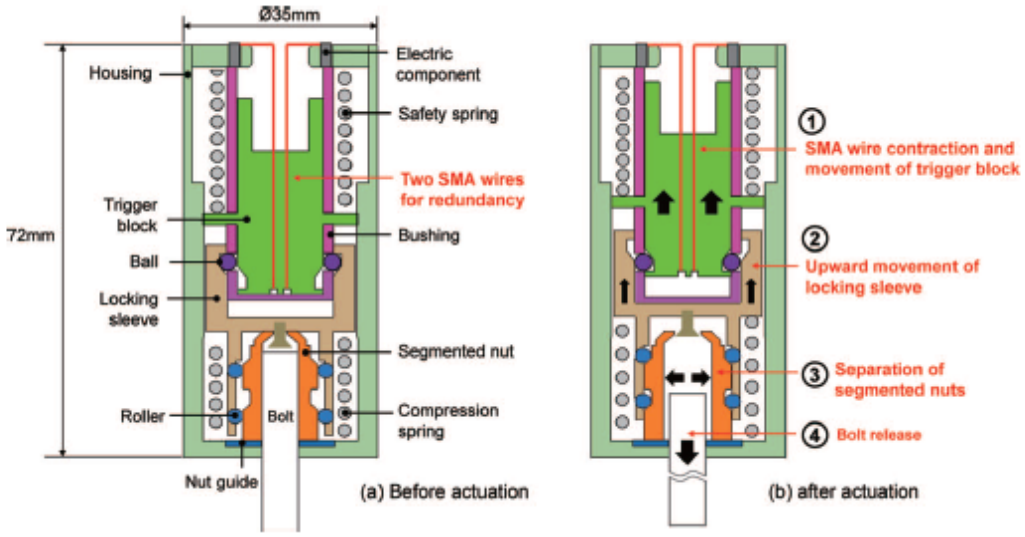


Figure 8. Configuration of release device

5 Paraffin actuator

Sierra Nevada Corporation has developed the Paraffin actuator [10]. The principle of paraffin linear actuator is based on the thermal expansion of a paraffin with a solid to liquid phase change at a specific temperature, which producing a smooth displacement of a piston. See Figure 9, as the actuator is heated to the actuation temperature, the paraffin melts, expands and creates hydrostatic pressure, which is transformed into a gentle, high-force shaft extension, causing the actuator rod to extend from the actuator. As the paraffin cools and solidifies, it contracts, allowing the rod to be pushed back into the actuator.

Paraffin actuator is used in applications where weight savings are of great importance. With small dimensions and weight in comparison to the large available applied load, able to produce also a long stroke. The actuators are also designed for extreme cleanliness applications and zero-outgassing conditions. Therefore, paraffin actuator technology could be adapted to the specific requirements of the aerospace mechanisms, where light weight, simplicity and reliability are prime concerns.

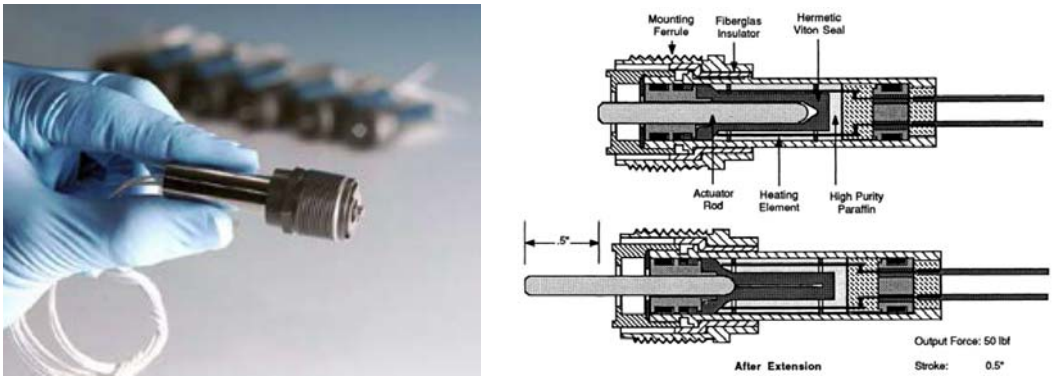


Figure 9. Configuration of Paraffin actuator

6 Thermal knife

The standard Dutch Space solar array release system employs the principle of cutting an aramid hold-down cable by an electrically heated Thermal Knife [11][12]. This principle has significant advantages over pyrotechnic devices. It is insensitive to electro-magnetic disturbances, emits a very low release shocks and the Thermal Knife, as actuator, can be used a number of times. Also the electrical circuit can be tested safely at system level. The flights proved thermal knife hold-down

and release mechanism is a simple, effective device based on thermal degradation of Kevlar/Aramid cable. It is extensively used by the Europeans and can release all deployable spacecraft appendages.

The Thermal Knife is a heater element that is mounted on a piston in an aluminum cylinder. The piston is actuated by a spring. The resistance element is a Silicon Nitride substrate with a Platinum resistance pattern on both sides, glass coated for protection. When the knife is electrically powered via the 2 leads the substrate edge runs hot and melts the branches of the coiled, one after each other. A dedicated retraction wire keeps the heater plate of the Thermal Knife separated from the Reel. The wire, which is the same Dyneema cord as used for Reel production, melts during the first seconds of powering. Contamination of the heater blade or damage of the resistance pattern of the redundant knife is there with avoided. This retraction wire improves the testability; the non powered Thermal Knife is thus kept away from the melting zone of the primary Thermal Knife.

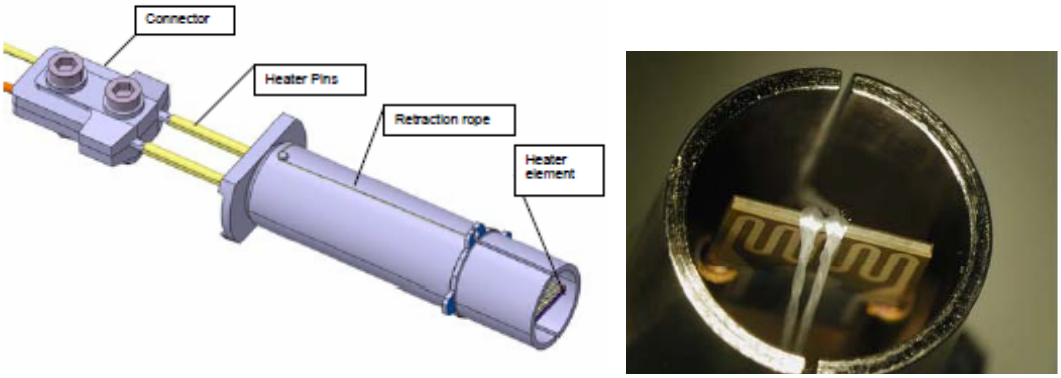


Figure 10. Impression of Thermal Knife

7 Magnetic clamp

RUAG corporation developed a No Shock Hold Down and Release Mechanism based on a magnetic clamp under an ESA contract suitable for space applications [13]. The device holds the appendage via permanent magnetic force, and releases via a counter-oriented electromagnetic force pulse. It is based on a magnetic clamp which is driven by a permanent magnet and therefore has no power consumption during hold down. The release is performed by an electric current pulse through one of the two implemented redundant coils. The permanent magnet is located in one part and the coils in the other part as can be seen from Figure 11. According to the need for electric connections the part with the coils will be the fixed part. As the other part with the permanent magnet needs no power supply, this will therefore be the moveable part.

The HDRM developed is dedicated to small deployable appendages and payloads of future spacecraft. An engineering model has been manufactured and tested.

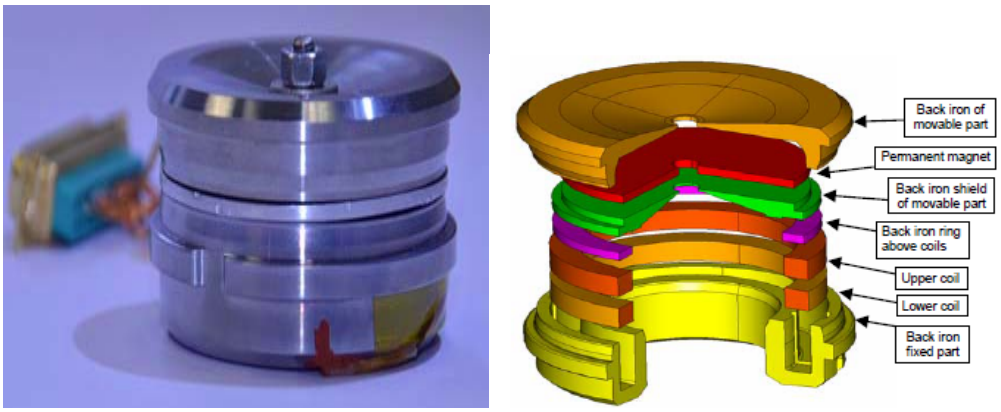


Figure 11. Parts of Magnetic clamp HDRM

8 Conclusion

NEAs have been used on spacecraft to accomplish the functions of traditional pyrotechnic devices. NEA mechanisms are typically either reusable or refurbishable to reduce overall system costs, especially in today's environment of smaller spacecraft and the need of low-shock. However, not all technologies alleviate functional shock, and alternate technologies require further development.

Split spool, paraffin actuator and thermal knife can provide gentle, shock-free release, and are considered mature, flight proven technologies. SMA devices possess significant potential for reducing functional shock, and continue development in progress.

Conference

[1] M.H. Lucy，R.C. Hardy＆E.H. Kist，et．Report on Alternative Devices to Pyrotechnics on Spacecraft[R]. 10th Annual AIAA/USU Conference on Small Satellites，Logan，17-19, Sept. 1996

[2] http://www.esa.int/Our_Activities/Space_Engineering_Technology/Mechanisms/HoldDown_and_Separation_Systems

[3] Barney Little, Electrically Powered Separation Nuts. the 39th Aerospace Mechanisms Symposium, NASA Marshall Space Flight Center, May 7-9, 2008

[4] http://www.cooperindustries.com/non_explosive_actuator.html

[5] http://www.tiniaerospace.com/

[6] http://www.neaelectronics.com/products/

[7] Shawn H. Smith, David Dowen, Eugene Fossness, Andrew Peffer, Development of Shape Memory Alloy (SMA) Actuated Mechanisms for Spacecraft Release Applications, 13th AIAA/USU Conference on Small Satellites

[8] Nestor Nava, Marcelo Collado, and Ramiro Cabas，REACT: Resettable Hold Down and Release Actuator for Space Applications，Journal of Materials Engineering and Performance, 23(7) , July 2014

[9] Young Ik Yoo, Ju Won Jeong, et. Development of a non-explosive release actuator using shape memory alloy wire, REVIEW OF SCIENTIFIC INSTRUMENTS, 015005 (2013)

[10] Scott Tibbitts, HIGH OUTPUT PARAFFIN ACTUATORS:UTILIZATION IN AEROSPACE MECHANISMS, NASA. Langley Research Center, The 22nd Aerospace Mechanisms Symposium; p 13-28, 1988

[11] Alphonso C. Stewart and Jason H. Hair, Intricacies of Using Kevlar Cord and Thermal Knives in a Deployable Release System: Issues and Solutions, Proceedings of the 36th Aerospace Mechanisms Symposium, Glenn Research Center, May 15-17, 2002

[12] Theo Konink, Gerard Kester, MULTIPURPOSE HOLDDOWN AND RELEASE MECHANISM(MHRM), 13th European Space Mechanisms and Tribology Symposium, Vienna, Austria, Sept. 2009

[13] Christian Laa, Theodor Nitschko, et. UNLIMITED RESETTABLE NO-SHOCK HOLD-DOWN AND RELEASE MECHANISM, 14th European Space Mechanisms & Tribology Symposium, Constance, Germany, Sept. 2011

Integrating Design of Lightweight Flat Cable

WANG Zhibin, CHEN Mengjiong, HUANG sanbo, MA Jusha, CAO Jiaye,

LEI Gang，LU Jianfeng

Shanghai Institute of Space Power Sources, NO.2965 Dong-chuan Road Shanghai, 200245, China.

Email：wzb.hit@163.com

Abstract：The availability of low-cost, lightweight cable is an important component in reducing the cost of solar array, especially used in satellites. In addition, future high-power spacecraft will require lightweight photovoltaic (PV) arrays with reduced stowage volume, thus, to develop lightweight energy transimission cable is high on the agenda. This paper presents a flat conductor cable which has great advantages in terms of lightweight designing.

Keywords: flat cable; lightweight; power transmission

1 Introduction

Energy is one of the basic elements to support the world and human life. With the progress of society and exploding population, energy demand is rising fast globally. Due to its sufficient cleanliness, absolute security, PV has been widely used all over the world. Photovoltaic solar array power system is an important part of the energy systems, while, the main problem encountered in solar array power transmission technology is the higher weight. Conventional round wire cable could not meet the demand for higher power transimission capacity. Thus, to develop lower weight energy supply system is high on the agenda which brooks no delay.

Flat conductor cable (FCC) has higher packaging density and power mass density[1,2], which consists of flat parallel conductors laminated between thin flexible plastic insulating films. Its thickness is less than 0.2mm (which fell by up to 80percent comparing with the round cable). Because of this fact, in most occasions, FCC is the only adopted cable for a particular project, such as flexible solar array.

2 Material selection

The commonly materials used in solar array are gold, silver, copper, aluminum, etc.[3], the main performance of these materials are showed in table 1. Aluminum has a lower density, with a big advantage in terms of lightweight. But aluminum has a lower tensile strength. Copper has higher tensile strength and lower electrical resistivity, which is a good conductor for energy tranmission. While, the density of the copper is more than thress times of the aluminum, there is no advantage in terms of lightweight desiging.

Table1. Physical and mechanical properties of several conductor materials

Property	Material			
	gold	silver	aluminum	copper
Specific weight, g/cm^3	19.3	10.5	2.7	8.3
Tensile strength, MPa	150	170	60	230
resistivity,$10^{-6}\Omega$.cm	2.1	1.66	2.7	1.72

Considering the carrier, mechanical properties and specific weight, a new electrical cable meterial was selected which names as copper-aluminum wire[4]. This wire has a weight density of

3.63g/cm^3, its specific weight reduced by 40% than copper wire, its resistivity was 2.4×10^{-6}Ω.cm, 11% lower than pure aluminum wire.

3 Designing analysis of FCC

3.1 Thickness of FCC

Thin solar cells, folm solar cell(CIGS, amorphous silicon, gallium arsenide film) thickness is only several tens of microns, thick triple -juntion GaAs solar cell having a thickness of 150μm~300μm. For this reason, thin thickness of FCC is designed. The thickness of FCC is less than 0.15mm, which is depicted in Fig.2.

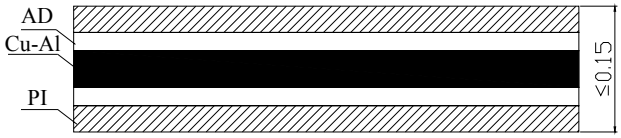


Fig.1. Thickness of the FCC

3.2 carrier performance of FCC

The material of the FCC conductor is copper-aluminum, its current-carrying capacity can be calculated as follows:

$$I = 1.02\frac{A^{0.5}p^{0.39}\theta^{0.61}}{[(1+\alpha\theta)\rho]^{0.5}} \quad (-1)$$

where:

A: cross-sectional area, mm^2;

p: conductor cross section perimeter, mm;

Θ: temperature difference between conductor and the environment, ℃;

α: temperature coefficient of resistivity, per ℃;

ρ: resistivity of the conductor.

Under the same operating conditions, FCC can, due to its larger surgace area to volume ratio and thinner insulation, carry much more current than round cable[5].

3.3 Insulation performance of FCC

The main factors affectiong the insulation performance of the flat cable has: an insulating gap, the surface condition, medium insulation performance, and with the distribution of the electric field between the contact menber concerned, as shown in fig.2. According to the actual situation of the product of insulation resistance simplified calculation model of the FCC is depicted in fig.3.

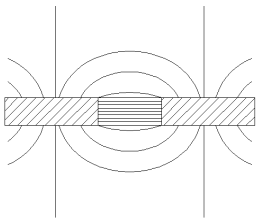


Fig.2. electric field distribution of the FCC

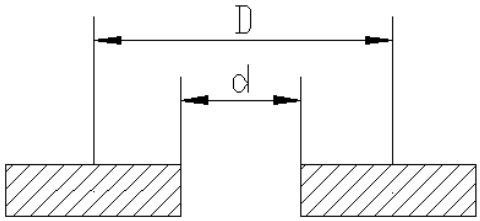


Fig.3. Calculation model of the insulation performance

d is the shorted distance between the insulating contacts, D is the line center pitch, the insulation resistance R can calculated by the formula(2) .

$$R=\rho \cdot L/S. \quad (-2)$$

where:

ρ: resistivity of the wire;

L: length of the parrallel carrier;

S: cross-sectional area of the insulation gap;

Polyimide(PI) has high volume resistivity, so, FCC has a good insulaton performance.

4 Test of FCC

In order to verify the performance of the FCC, A flat conductor cable was designed, manufactured and tested, as shown in fig.4, fig.5, and table2. which was featured with good carrier, insulation and mechanical bending performance.

Fig.4. Photograph of the FCC

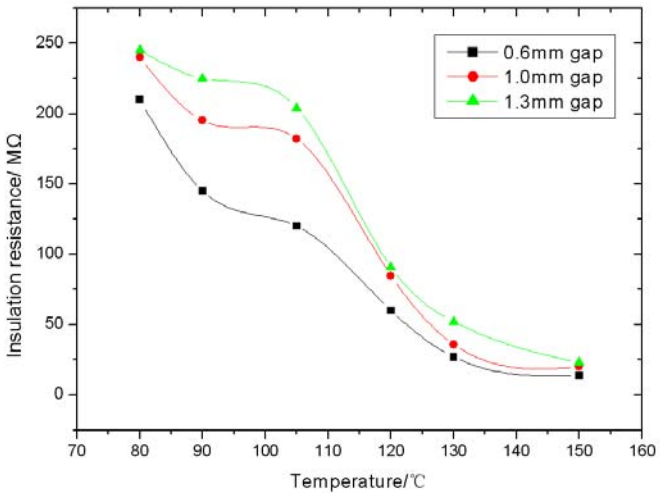


Fig.5. Insulation resistance test of the FCC

Insulation testing and data analysis proved that FCC could optimize large solar array power transmission with lighter weight.

Table2. Comparation between FCC and round cable

Property	FCC	RWC-22
Current-carrying capacity	4.5A	4.5A
Equal weight	1.5g/m	4.17g/m
thickness（diameter）	≤0.2mm	1.09mm
Dielectric strength	3300V	2500V
Temperature range	-100℃～100℃	-55℃～150℃

As can be seen from table 2, the FCC has higher power transmission capacity, its specific weight is only 1.5g/m, weight reduced more than 200% to traditional RWC round cable. Therfore, FCC has great advantages in terms of lightweight designing.

5 Conclusion

Copper-aulninum conductor was choiced for FCC, which has good current carrier capacity, also, FCC has good insulation resistance, which can be used for high-voltage solar array power transimission. This new FCC, which reduces weight more than 200% and with a promising application prospect, will be used in a wide range of fields, particularly in aerospace industry.

References:

[1]. Mc Donnell Douglas, Handbook for flat conductor cable system design and manufacturing, 1968, 2-7.
[2]. J.P. Morris, Airocraft flat conductor cable power feeders, Boeing, IPC meeting, 1969.
[3]. Wil.Helm Angele, Flat conductor cable technology, 1969, 4-5.

Structural Analysis of the Interconnector used in Flexible Solar Array

WANG Zhibin, CHEN Mengjiong, HUANG sanbo, MA Jusha, CAO Jiaye, CHI Weiying, LEI Gang

Shanghai Institute of Space Power Sources, NO.2965 Dong-chuan Road Shanghai,200245, China.
Email：wzb.hit@163.com

Abstract：With the deepening of China's manned space missions, the project of founding China's independent research and construct space station has been gradually put on the agenda. Through stress simulation analyses and fatigue experiment, a brief analysis of the interconnector used in flexible solar array was made. These works will make positive contribution to the design of flexible solar array in the future.

Keywords: interconnector; flexible solar array; stress analysis

1. Introduction

According to China's manned space agenda, China's independent-research space station will be setup about 2020. This new station is different from the traditional satellites, which is larger, more powerful, and has a longer life[1]. In order to supply enough power, a large area of flexible solar array is needed. The high effective solar cells connected by interconnectors are bonding to the flexible solar panel, which will endure lots times of temperature cycling, low energy protons and high energy electrons .etc. For these environments, the designing of solar circuits must be compatible with them, especially for the interconnectors, which must be compatible with the flexible solar panel and endure temperature cycling up to 50000 cycles, over the -100℃to+100℃range, or even more[2-3]. Thus, the importance of solar cell interconnector design has become more emphasized.

2. Requirements

The interconnector structure used in flexible solar array must be compatble with the following requriments on the solar array.

1)Temperature cycling: up to 50000 cycling, over the -100℃ to +100℃ rang;

2)Solar array substrate: flexible solar panel;

3)Launch environment: must survive the launch vibration without failure.

In addition, the interconnector design must lend itself to economic solar array fabrication, could not be too heavy, and may be restricted to nonmagnetic materials or low atomic number materials.

3. Designing analysis of the interconnector

These new requirements are followed with a lot of technology difficulties, such as fatigue designing of the interconnector. Therefore, people need to analysis the structure of interconnectors. When solar cells are bonded to a substrate, the gap between cells will change its dimensions when the assembly temperature is varied[4]. The variation in gap dimension is associated with the

temperature and the substrate material. Fig.1 illustrates this condition, showing the change in gap dimensions between solar cells when mounted on flexible solar panel, epoxy-fiberglass and kapton. Thus, when an assembly of cells is temperature cycled, an interconnector attached to both cells across the gap will alternately be stretched and compressed. The cyclic stretching and compression of an interconnector will eventually lead to its failure because of metal fatigue.

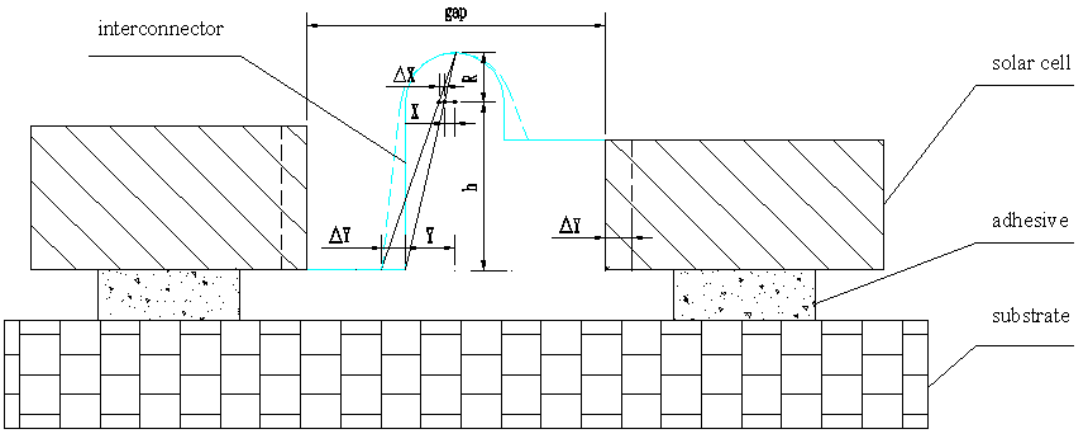


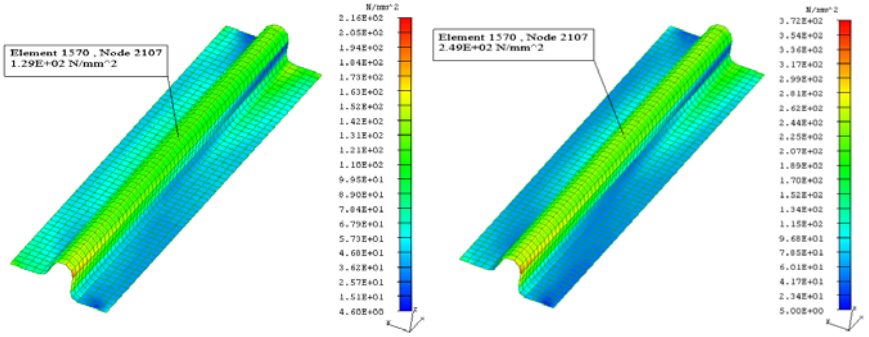
Fig.1. Change in gap between solar cells when the temperature is varied

In order to evaluate the structure of the interconnectors, stress analyses were carried out for structural analysis of the flexible solar array module. Analytical parameters were determined based on the obtained material parameter and the environment temperature. Physical and mechanical properties of materials uesd in flexible solar array are shown in table1.

The calculated stress strength of the interconnector is 216MPa and 372MPa, which is depicted in Fig.2.

Table1. Physical and mechanical properties of materials uesd in flexible solar array

Property	Material			
	Interconnector	Solar cell	Substrate	Adhesive
Specific weight, kg/cm3	0.008	0.0053	0.0025	0.0011
Tensile strength, MPa	475	283	-	2
Coefficient of thermal expansion,cm/cm/℃	5.2	6.1	57	201
Modulus of elasticity, GPa	324	105	3.2	2.1



a）high temperature b）low temperature

Fig.2. Stress strength analysis of the interconnector

Stress analysis explains that the max stress is lower than the material strength, which is suitable for the flexible solar panel.

4. Fatigue experiment of the interconnector

In order to verify the conclusion of the stress analysis, a fatige test was set up, as shown in Fig.3.

Fig.3.Test configuration for measuring fatigue strength

After 5000 times fatigue test, no crack happened, which indicated that the interconnector has a good adaptive for amplitude test.

5. Conclusion

Base on the simulation analysis and experiment research in interconnector structure, the structural of interconnector is available. These works will make positive contribution to the design of flexible solar array in the future.

References:

[1] W.LUFT,Solar Cell Interconnector Design, IEEE, 1971, 781-791.

[2] Hans.S.Rauschenbach, Solar array design handbook, the principles and technology of photovoltaic energy conversion, Van Nostrand Reinhold Company, 1980.

Study on Design Method for Distribution Systems Based on Maintainability

Du Zhanchao; Chen Dan

Institute of Manned Space System Engineering; China Academy of Space Technology; Beijing 100094 China

Abstract: Distribution system is the transmission network of electrical power on spacecraft, it has high failure severity, on-orbit maintainability of the distribution system can improve the reliability and lifetime of the spacecraft significantly. The necessity and feasibility of on-orbit maintainability of the spacecraft distribution system were discussed in the paper. A typical spacecraft distribution system was taken as an example, without changing the power system and other external input constrains, to achieve on-orbit maintenance, the topology of the distribution system was optimized, operating modes and maintenance items of the improved topology were proposed in detail. The topology and design method in the paper can be used for reference in the design of the on-orbit maintainability for spacecraft distribution system.

Key words: Spacecraft; Distribution System; On-orbit Maintainability; Design

I. Introduction

For a long time, the development and operation of spacecraft have the character of high investment and high risk, but generally the spacecraft on orbit faults are just problems with individual components, if the whole star scrapped, the loss is too big. Space powers such as The United States and Russia have started the research of on orbit maintenance since 1960s, and have accumulated a wealthy experience on spacecraft on orbit maintenance from the ISS maintenance. In China, during twice Shenzhou spacecraft with Tiangong-1 manned rendezvous and docking missions, the astronauts successfully implemented equipment on orbit maintenance. With the requirements of high reliability and long life of the spacecraft, whether the spacecraft's various components can be on orbit maintenance is increasingly concerned.

As an important part of the spacecraft, distribution system provides energy transmission and control, it determines whether the spacecraft can perform flight missions normally, the length of its life also directly determines the life of the spacecraft, during the 527 spacecraft on orbit faults announced in recent years, almost 14% have relations with the power supply and distribution system, so it belongs to the high risk area. According to the spacecraft life analysis and FMEA (failure mode and effect analysis), spacecraft power distribution system fault is severe and disastrous grade, so if the fault severity can be reduced through on orbit maintenance, the spacecraft’s reliability will increase, even prolonging the spacecraft’s life.

From the perspective of on orbit maintenance, A typical fundamental unit of spacecraft distribution system was taken as an example in this paper, the topology of the distribution system was optimized, and the key points of the design were given, nowadays the on orbit maintenance is generally completed by the astronauts, but how the astronauts get to the spacecraft requiring maintenance is not within the scope of this research.

II. The analysis of spacecraft distribution system maintainability

The power distribution system of the spacecraft generally receives electricity power from the power supply system, and then transmits the power to each electrical equipment, including critical loads, so a reasonable set of distributor and distribution channel is very important.

Spacecraft distribution system can be classified into three types: centralized, decentralized and hierarchical distribution system, centralized distribution is defined by the spacecraft systems providing only the overall level distributor, responsible for the entire electrical equipment on the spacecraft; decentralized distribution is defined by subsystems conduct their own power control; and hierarchical distribution is defined by the spacecraft systems providing the master distributor, while subsystems conduct their own power control.

A typical fundamental unit of spacecraft distribution system was taken as an example in this paper, in Figure 1, the distributor receives electricity power from the power supply system, and in the hierarchical distribution system is also likely to be the lower grade distributor receiving the power provided by the upper grade distributor. And then transmits to the load through the bus bar expansion. Sometimes the power transformers like DC/DC will be set in the distributor according to the voltage. Any spacecraft power distribution system can be decomposed to several such fundamental units.

During on-orbit flight, when the distribution system fails or needs regular maintenance, for the fundamental unit in figure 1, there is only one distributor, and the GNC(Navigation Guidance and Control)system can't be power off, so the maintenance can't be implemented; but from the perspective of the entire distribution system, if there are more than one distributor , and distribution redundancy is also taken into account, then the fault can be isolated by switching the distribution channel, thereby performing on orbit maintenance.

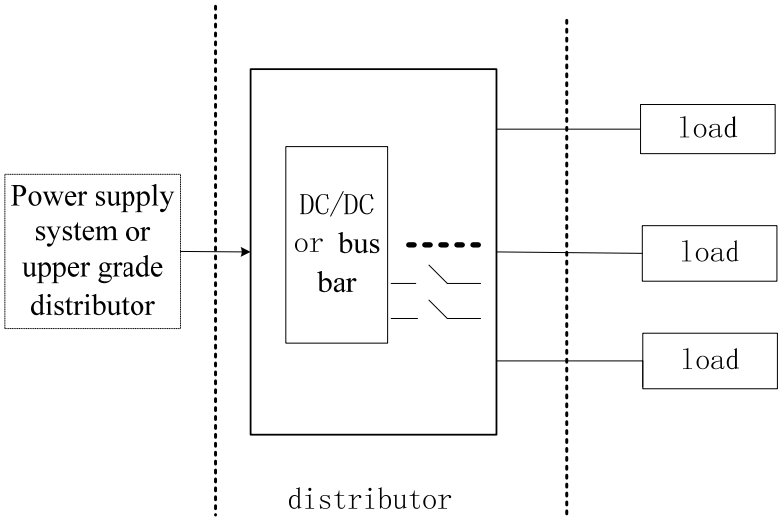


Fig 1 A typical fundamental unit of spacecraft distribution system

III. Design method for distribution systems based on maintainability

Fault diagnosis and localization

To achieve on orbit maintenance of spacecraft power distribution system, first of all the fault should be able to locate in one distributor, therefore the power distribution system firstly requires fault diagnosis capability.

Fault diagnosis is to determine whether the system fails, to locate the fault location, and to recognize the size and occurrence time of the fault. Currently spacecraft on orbit monitoring is still artificial. As computers become increasingly intelligent, fault diagnosis gradually developed as part of the autonomous management of the spacecraft, the diagnosis process will eventually be fully completed by the computer. But whether artificial or with computers, input information for fault diagnosis are still critical spacecraft telemetry information and expected failure modes, therefore achieving power distribution system fault diagnosis function, the key is to ensure the testability of the critical information and the integrity of the fault mode.

Spacecraft power distribution system consists of the distributor (including power converters, power switches, etc.) and cable. Distributor generally set up sensors to monitor voltage and current, switching state power channel is also monitored by telemetry parameters, these parameters can be hard-wired telemetry, or be sent through the bus transmission, the working status of the distribution system will be monitored by interpreting these parameters, some advanced distributors have the state flag in its internal procedures, which can quickly determine whether they work properly.

Distributors' common failure modes include the following categories:

Table 1. Typical failure mode in distribution systems

NO	Failue mode	Commonly used detection method	Distributors' failure	Failure affecting	Possibility of maintenance
1	Voltage abnormal	Voltage sensor	Sensor failure, power conversion module failure	Load device input voltage is too high, it does not work or even be damaged Load device input voltage is too low, does not work	possible
2	Current abnormal	Current sensor	Sensor failure, power conversion module failure	Current is too large, load device does not work and even be damaged Current is too small, load device does not work	possible
3	Switch abnormal	Switch status telemetry circuit	Telemetry circuit fault, the power switch malfunction	The power switch cannot be turned on, the load device cannot work, so a function cannot be implemented The power switch cannot be turned on, the load device will working continually until damaged	Possible, but if the input and output terminal of the switch cannot be disconnected ,the maintenance cannot be implemented
4	Communication abnormal	software settings status flag set by software, or judgment by all the telemetry parameters	Communication breakdown	working status of distribution system can not be detected, distribution controlling function fails	Possible, but it needs to be judged by the relevant telemetry parameters, there is a certain risk, if the switch cannot be turned on and off, the maintenance cannot be implemented

When the above failures are detected, and the anomalies still can't be ruled out by using established plans even disconnect the load equipment, then the distributor can be determined abnormal and should be temporarily isolated. Now the backup distributor is needed, and on orbit maintenance is on the schedule. From the above table, among the various failure modes of the distribution system, during any failure modes associated with the power channel switch, the possibility of on orbit maintenance needs to be analyzed in the specific circumstances, but the final criteria is whether the fault power distributor can be power off.

The isolation of fault device

When the fault device is posited, it should be isolated before maintenance to prevent the spread of the fault.

Distribution system is the power network of the spacecraft, the distributor is always in long-term working mode, power can not be easily isolated, and even during the process of on orbit maintenance, fundamental power and information devices still need to work.

So the basic power supply must be guaranteed during the process of on orbit maintenance, besides that, from a security point of view, it can only provide the basic functions required for spacecraft power maintenance, to avoid big power load carried out which may be dangerous. This requires the identification of spacecraft minimum operating mode.

Spacecraft minimum operating mode generally includes the following systems:

- GNC: to ensure spacecraft attitude stabilization on orbit;
- Thermal control system: to ensure thermal environment of the equipment;
- Energy systems; to ensure the power supply of the minimum load;
- Information Systems: to ensure the telemetry data upload and down.

In the above four kinds of sub systems, three kinds receive electrical power from the distribution system except the energy system, and before isolating the fault distributor, should ensure that there is no such key device in the loads of this distributor, or the power of such key device has been switched to other distributors. Therefore, the critical device should be power distributed redundantly, namely power the primary and backup device by two distributors, and power the critical load devices by switch. So that when carried out the distributor on orbit maintenance, the critical device under the fault distributor can be switched to the normal one, then after disconnecting the original channel, the isolation of the fault distributor is achieved.

To sum up, to achieve on orbit maintenance of the fundamental unit of spacecraft distribution system, a distributor responsible for powering the critical device in the minimum operating mode must be set. In order to reduce the impact on the entire spacecraft weight and power consumption, the distributor can be individually designed given to the minimum operating mode power and normally as a cold backup, If there are more than one distributor in the spacecraft, distributing power redundantly should also be considered.

Maintenance Safety Design

Similar to maintaining the ground power distribution system, the maintenance of the spacecraft distribution system also has certain risk to astronauts, therefore safety design is a must.

Spacecraft distribution system is always located in the subtle position with complex surrounding environment, like many equipment and cables around. therefore the device with electricity, especially high-voltage equipment and cables should be clearly marked to warn astronaut be careful.

Hot plugging cable is the key problem when implementing the maintenance of the distribution system, for the power supply cable, hot plugging is forbidden, such security risks can be avoided by setting switch in the distributor's input terminals and output terminals, see Figure 2. Theoretically, when the two switches are cut off, no matter whether the connector is power off or not, there is no current loop, and input terminals and output terminals are set by the forefront and the most end, there will be no damage to the internal parts of the distributor.

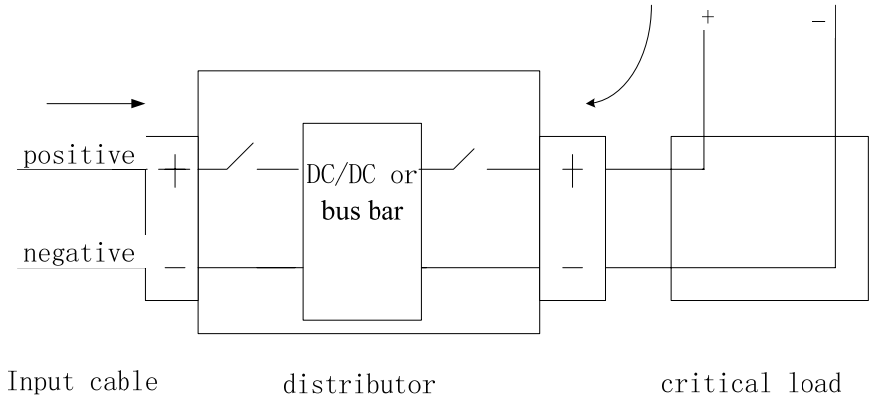


Fig 2 the switch set in the input and output terminals of the distributor

The reason of setting output switches is to prevent the load equipment, especially the critical load equipment, still live work when implementing the maintenance. If the faulty distributor's connectors are still electrified, it may cause further damage to the distributor, even cause ignition and other issues, such risks can be avoided by setting a switch while cutting off the current path.

If possible, there should be enough output switches in upstream device, while the primary and backup power supply are designed isolated in downstream critical loads, to ensure the operation of the cable is fully powered down.

After the cable is powered off, the cable will be suspended before maintenance is completed, antistatic protective cover should be put on to avoid the extra material getting into the exposed electrical connector. Besides that, to prevent the impact, all other operations is forbidden during on-orbit maintenance of the distribution system.

distribution system topology after optimized

To summarize the above, several aspects need to be improved to make the fundamental distribution system possible to be on orbit maintained.

1) After FMEA, reasonably set enough telemetry parameters, instructions, etc, to achieve the fault diagnosis and localization of distribution system;

2) through the optimization of the distribution channel configuration, especially the equipment working in spacecraft minimum operating mode must be redundant power supplied, so that the faulty distributor can be isolated;

3) setting switches and other control circuit in the downstream and upstream device of the distribution channel to ensure the astronaut's safety.

the first and the third article can be achieved by optimizing the design of the distributor, while the second article can be achieved by optimizing the topology of the distribution system, the improved distribution system topology is shown in Figure 3.

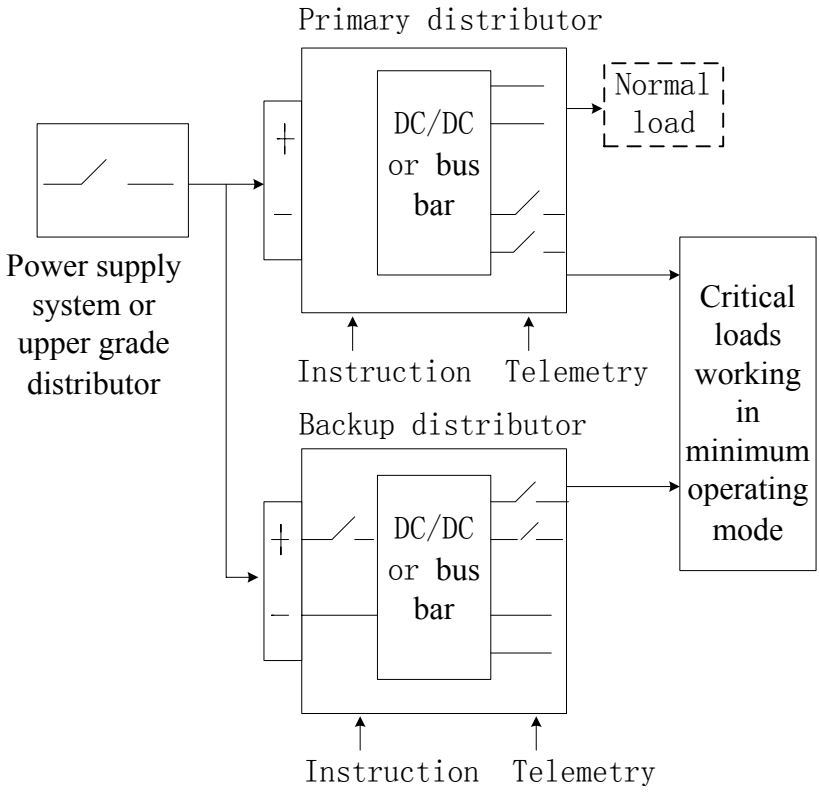


Fig 3 The optimized topology of the fundamental distribution unit

In the improved distribution system topology, one backup distributor is increased to receive the power from power systems or upstream distributor, the designing state and operating mode of this distributor are analyzed as follows:

- If the backup distributor only supply power to critical equipment working in spacecraft minimum operating mode, it can be designed according to the spacecraft minimum power consumption, and work in cold backup mode, the internal output switches are all off, and start working only when the main distribution is faulty and need to carry out maintenance and replacement work;
- If the backup distributor not only supply power to critical equipment working in spacecraft minimum operating mode,, but also other load, from the universal point of view, the backup distributor can be designed the same to the main distributor with the normal mode is hot backup, the internal input switches are on, the output switches are on or off according to need.

As can be seen from the optimized topology map, if one distributor is faulty and need on orbit maintenance, firstly in turn switch the critical load power supply channel, disconnect the power supply and other operations on the downstream device with electrical isolation, then the rest maintenance procedures are unanimous with the general equipment maintenance program. It can be said that the optimized distribution system base unit has been maintainable after designing on two aspects of equipment and system. For complex distribution system, which can be divided into several basic unit distribution system, design each basic unit with the above method, the backup distributors for different units can be combined selectively, after comprehensive consideration and

optimization of many aspects such as weight, power, redundancy, so as to achieve the entire distribution system maintenance of optimal designs.

IV. Conclusion

A typical spacecraft distribution system was taken as an example, without changing the power system and other external input constrains, to achieve on-orbit maintenance, the topology of the distribution system was optimized, operating modes and maintenance items of the improved topology were proposed in detail. The necessity and feasibility of on-orbit maintainability of the spacecraft distribution system were discussed in the paper. The topology and design method in the paper can be used for reference in the design of the on-orbit maintainability for spacecraft distribution system.

Reference

[1]Tan Chunlin，Hu Taibin，Wang Dapeng，Analysis on foreign spacecraft In orbit failures[J].Spacecraft Engineering, 2011, 20（4）:130-136

[2] Li Zhe, Zhu Yali, Design and verification of on-orbit maintenance of pipe equipment in manned spacecraft pressure module [J] Spacecraft Environment Engineering, 2014 (6): 321-323

[3] Wang Shiqing, Wang Jing, spacecraft in orbit maintenality[J] quality and reliability, 2012 (4): 4-8

[4] Wei Pengwei, Research of Maintainability Technology in Manned Spacecraft[J] Spacecraft Engineering 2008,17 (6): 68-72

[5] Wang Dapeng，Tan Chunlin，Zhang Bainan, On-orbit maintainability system design for manned spacecraft[J].Chinese Space Science and Technology，2010，30(5):16-21

[6] Nan Hongtao，Liu Hongtai，Wang Lintao, Safety design of high voltage power supply for manned spacecraft[J].Spacecraft Environment Engineering，2013，30(5):16-21

[7] Zhang Wei, Xia Qiaoli, study on design and verification method of space station maintainability[J] manned space flight, 2014,20 (2): 135-137

[8] Zhu Ye, Cheng Kai, satellite autonomous health management techniques outlined technology wind 201205

A Study on the Thermal Performance of the Motorized Thermal Shade System

Dong La, Han Chen, Zaiteng Zhai, JiangFeng Lu

Laboratory of Satellite Mechanical and Thermal Integrative Technology, Shanghai Institute of Satellite Engineering, Shanghai 200240, China

Abstract: The motorized thermal shade system, which is advantageous in providing variable heat radiation capability, being capable of high turn down ratio and operating in various modes, has been recognized as a promising technology for the spacecraft on the inclined orbit, geosynchronous orbit, deep space, etc. In this paper, the thermal performance of the motorized thermal shade system is studied. The motorized thermal shade system is formed by integrating the side plate of the spacecraft and a thermal shade together. The side plate of the spacecraft is coated with white paint, and multi layer insulation is adopted as the thermal shade. The radiating surface of the spacecraft can be adjusted by deploying the shade to cover the radiator and stowing to expose the radiator. The effects of the open of the thermal shade, the orbital heat flux and the temperature of the spacecraft on the thermal control capability of the motorized thermal shade system have been analyzed. Besides, investigations on the adaptability of the motorized thermal shade system have been performed. The objective of this paper is to obtain useful information for practical application of the motorized thermal shade system. It is found that the open of the shade impacts the effective emissivity of the motorized thermal shade system greatly, thereby achieving a great effect on the temperature of the spacecraft. Moreover, the motorized thermal shade system is found to be applicable to a wide range of conditions, such as the sun side, the dark side, and that on the sun side and the dark side alternately (i.e. inclined orbit spacecraft).

Keywords: Motorized thermal shade system, Thermal performance, Thermal control, Spacecraft

1. Introduction

Radiator is one of the key thermal control measures for spacecraft. Conventionally, internal component heat is dissipated through radiator passively. Yet, the passive radiator is limited by adaptability and power consumption, and couldn't satisfy the thermal control requirements of inclined orbit spacecraft, deep space probe, geosynchronous orbit spacecraft, etc. How to control the radiator actively, thereby meeting the demands of heat dissipation and preservation for various thermal environments, has become an urgent problem [1-5].

The motorized thermal shade system, which is advantageous in providing variable heat radiation capability, being capable of high turn down ratio and operating in various modes, has been recognized as a promising technology for the spacecraft thermal control [5]. In this paper, the thermal performance of the motorized thermal shade system is studied. The effects of the open of the thermal shade, the orbital heat flux and the temperature of the spacecraft on the thermal control capability of the motorized thermal shade system have been analyzed. Besides, investigations on the adaptability of the motorized thermal shade system have been performed. The objective of this paper is to obtain useful information for practical application of the motorized thermal shade system.

2. System Description

The motorized thermal shade system is formed by integrating the side plate of the spacecraft and a thermal shade together. The side plate of the spacecraft is coated with white paint, and multi layer insulation is adopted as the thermal shade. The radiating surface of the spacecraft can be adjusted by deploying the shade to cover the radiator and stowing to expose the radiator.

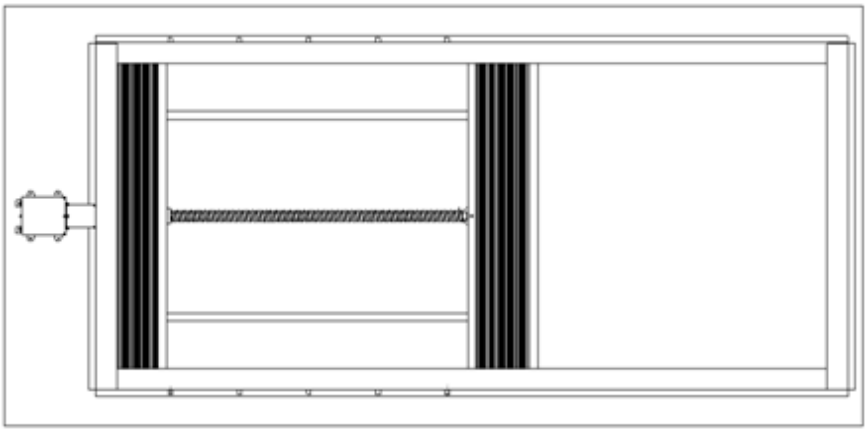


Figure 1 Motorized thermal shade (100% open)

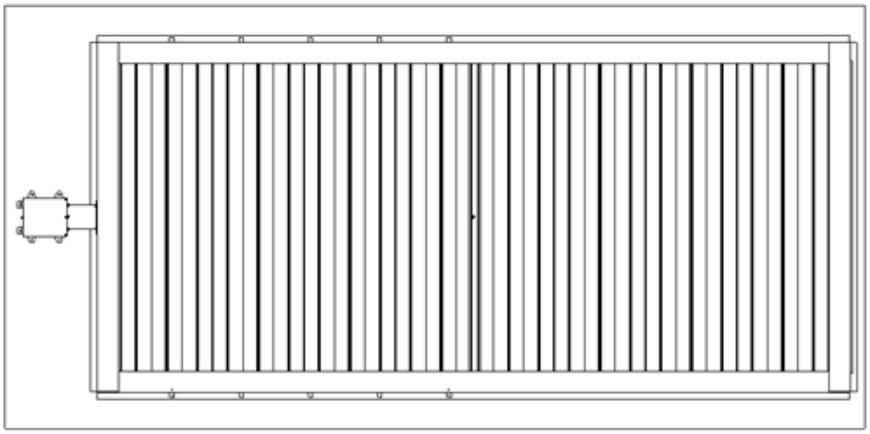


Figure 2 Motorized thermal shade (close)

3. Results and Discussion

In this section, the effects of the open of the motorized thermal shade and internal heat dissipation are analyzed based on test results. The effective emissivity of the motorized thermal shade has been tested at various radiator temperatures. When being closed, the effective emissivity should be as low as possible, which means less heat loss. When 100% open, the effective emissivity should better approximate the emissivity of the radiator, which means minimal blockage and high radiation capacity. Furthermore, an implementation has been carried out. The performance of the motorized thermal shade and the effect of external heat flux are discussed.

3.1 Effect of thermal shade open

Figure 3 and Figure 4 shows the effect of thermal shade open on effective emissivity and heat dissipating capacity. It can be seen that, the open of the thermal shade impacts the effective emissivity of the motorized thermal shade system significantly. As the open of the thermal shade increases, the effective emissivity increases greatly from 0.08 to 0.57. Besides, the heat dissipation capacity increases from 20.2 W to 386.2 W. It is indicated that the temperature of the spacecraft can

be controlled effectively by the motorized thermal shade system.

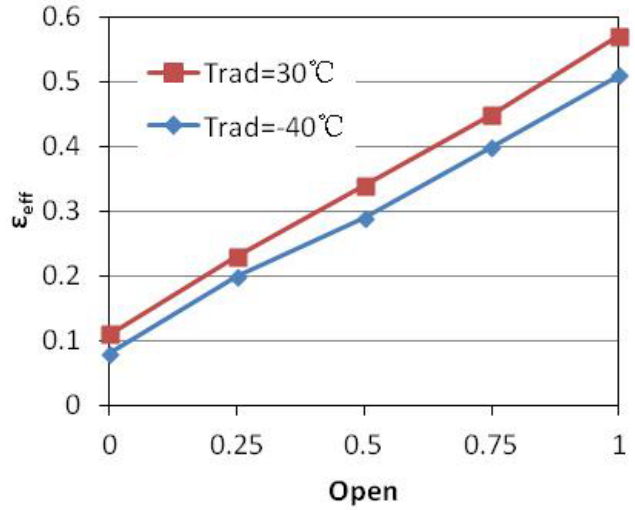


Figure 3 Effect of thermal shade open on effective emissivity

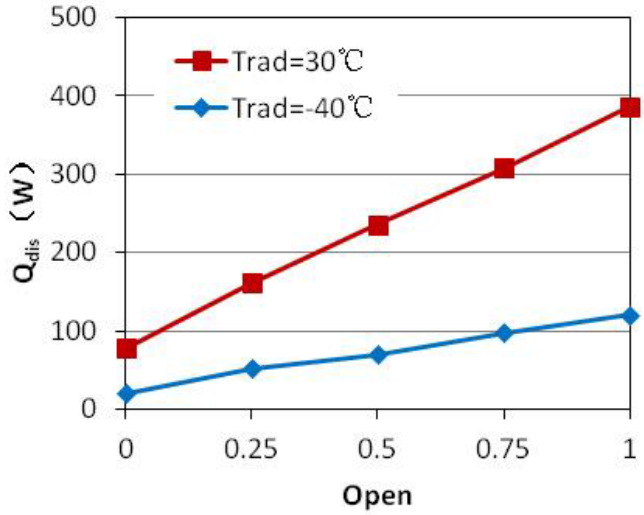


Figure 4 Effect of thermal shade open on heat dissipating capacity

3.2 Effect of internal heat dissipation

The effect of the internal heat dissipation on the effective emissivity is depicted in Figure 5. As can be seen, the temperature of the radiator impacts the effective emissivity slightly. When the thermal shade is closed, the effective emissivity is 0.08~0.12. When being 100% open, the effective emissivity increases from 0.51 to 0.57 as the radiator temperature increases from -40 ^{o}C to 30 ^{o}C. However, the heat dissipating capacity is greatly affected by the radiator temperature, as illustrated in Figure 6. Especially when the thermal shade is 100% open, the heat dissipating capacity increases from 120.4 W to 386.2 W. This indicates that, the motorized thermal shade is of great benefit for the spacecraft with high heat load.

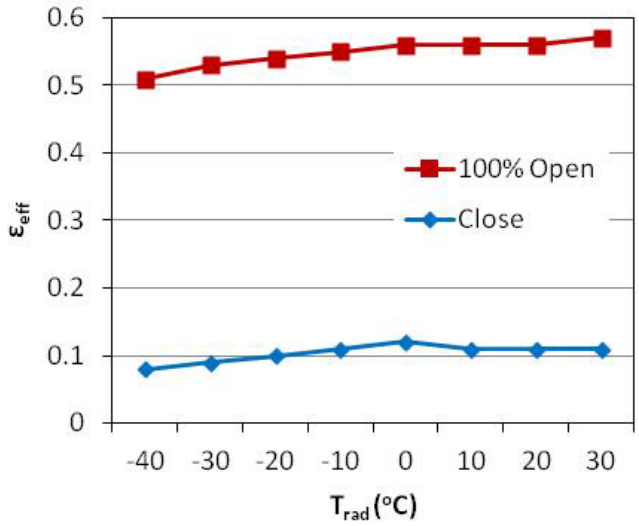


Figure 5 Effect of internal heat dissipation on effective emissivity

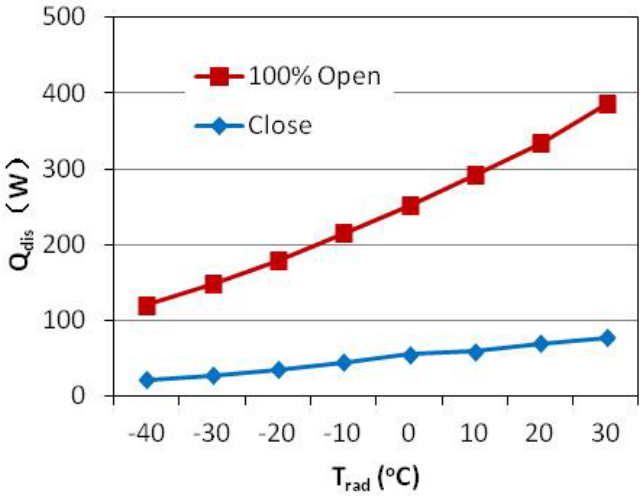


Figure 6 Effect of internal heat dissipation on heat dissipating capacity

3.3 Implementation

The following discussion is based on an application of the motorized thermal shade to a typical spacecraft. The test has been performed under vacuum thermal environment. For the worst cold case, when the thermal shade is closed, the radiator area is reduced. As a result, the average temperature of the equipment in the bay section increases 4.0 ^{o}C, with a highest temperature increment of 7.5 ^{o}C. Simultaneously, the power consumption for heating the equipment is reduced by 15.5 W. For the worst hot case, when 100% open, the average temperature of the equipment in the bay section increases 1.7 ^{o}C, with a highest temperature increment of 8.1 ^{o}C. Furthermore, the power consumption for heating the equipment is reduced by 15.7 W. For the case that the temperature in the bay section is maintained the same, the heat consumption for component heating can be reduced by about 130 W. It is suggested that, the thermal shade has better performance when the external heat flux is relatively low, and when the external heat flux is high, the external flux can be used for heating. Hence, properly setting the working mode of the motorized thermal shade can contribute greatly to the thermal control of the spacecraft.

4. Adaptability Analysis

According to the configuration of the spacecraft and the external heat flux, the location of the motorized thermal shade can be classified into three types, namely, the sun side, the dark side, and that on the sun side and the dark side alternately (i.e. inclined orbit spacecraft). Based on this, the adaptability of the motorized thermal shade is analyzed.

When being located on the dark side, the thermal shade can be deployed with a big open when the temperature of the spacecraft is high, so that the temperature of the spacecraft and the equipments can be decreased. On the contrary, when the temperature is low, the thermal shade can be stowed to reduce the radiator area. As a result, the temperature of the spacecraft and the equipments can be maintained at normal level. Compared with conventional fixed area radiator, the main feature of the motorized thermal shade located on the dark side is providing sufficient radiator area with relatively low power consumption by controlling the radiator area actively.

For the case on the sun side, when the temperature of the spacecraft is high, the thermal shade can be stowed to reduce the external heat flux, thereby decreasing the temperature of the spacecraft and the equipments. Conversely, when the temperature is low, the thermal shade can be deployed to increase the area under sunshine, and then the temperature would be increased. The main feature of the motorized thermal shade located on the sun side is fully utilizing external resource to heat the spacecraft, and thus the power of the spacecraft can be saved.

Table 1 Adaptability analysis of the motorized thermal shade system

Location	Spacecraft temperature	The open of thermal shade	Function	Main feature
Dark side	High	Increase	Heat dissipation	Providing sufficient radiator area with low power consumption by controlling the radiator area actively
	Low	Decrease	Power consumption saving	
Sun side	High	Decrease	Heat insulation	Fully utilizing external resource to heat the spacecraft
	Low	Increase	Power consumption saving	
Location	Side in sunshine	Working mode	Radiator	Main feature
Sun side and dark side alternately	+Y side	+Y close -Y open	-Y side	Solving the problem of no fixed radiator for the inclined orbit spacecraft
	-Y side	-Y close +Y open	+Y side	

For the inclined orbit spacecraft, due to the perturbation of the orbital plane is in contrast with the revolution direction of the earth, all the sides of the spacecraft would be exposed in the sunshine. As a result, the external heat flux changes greatly and no ideal radiator can be achieved. For this case, two motorized thermal shades could be considered, which can be installed on two opposite sides simultaneously. In other words, the motorized thermal shade system is installed on two radiators that expose in sunshine in turn. For example, the motorized thermal shade system is located on +Y and –Y sides of the spacecraft. When the +Y side is in sunshine, the –Y side is used

for radiating, with the thermal shade on the +Y side close and that on the –Y side open. Otherwise, when the –Y side is in sunshine, the +Y side is adopted as radiator, with the thermal shade on the –Y side open and that on the +Y side close. The power consumption can also be reduced by adjusting the open of the thermal shade. Moreover, the motorized thermal shade located on the sun side and the dark side alternately has solved the problem of no fixed radiator for the inclined orbit spacecraft and provides technical support for the development of modern spacecraft.

5. Conclusions

In this paper, the thermal performance of the motorized thermal shade system is studied. It is found that the open of the shade impacts the effective emissivity of the motorized thermal shade system greatly, thereby achieving a great effect on the temperature of the spacecraft. Moreover, the motorized thermal shade system is found to be applicable to a wide range of conditions, such as the sun side, the dark side, and that on the sun side and the dark side alternately.

References

[1] Hou ZQ, Hu JG. Spacecraft thermal control. China Science Press, 2007.

[2] Shen C. Adaptability development of the thermal control system of satellite common platform. 5th Space Thermal Physics Conference Proceedings, 2001; 1-5.

[3] Zhou SH. Conceptual design summary of a motorized louver. 5th Space Thermal Physics Conference Proceedings, 2001; 477-483.

[4] Shen S, Zhang WS. Research and development of a motorized thermal shade. 6th Space Thermal Physics Conference Proceedings, 2003; 579-582.

[5] Baumann JE. Performance Testing of the Motorized Thermal Shade System. AIAA-87-1569.

Surface modification of RDX using conducting polymer: Preparation, properties and application

Zhiquan Shi*, Yali Lu，Hongyong Huang，Rui Li

Shanghai Aerospace Chemical Engineering Institute; Huzhou 313000; China

* Corresponding author; e-mail:shizhi163@163.com

Abstract: To decrease the sensitivity of RDX and improve the tensile strength of composite solid propellant containing RDX, conducting polymer was introduced to modify RDX. The microstructure and morphology of coated RDX were characterized by FTIR, XRD, TGA and SEM measurements. The impact sensitivity, friction sensitivity were also investigated. By modification, the impact and friction sensitivity of RDX decreased obviously; the impact sensitivity H_{50} was changed from 39.0cm to 61.2cm, and the friction probability was reduced from 90% to 30%. The surface conductivity of RDX had greatly increased from insulating to 5×10^{-4}S/cm, which may benefit to remove static electricity buildup and improve the safety of composite solid propellant formulation which contains RDX particles. In addition, the tensile strength of PET (poly epoxy–tetrahydrofura) composite solid propellant containing RDX could be improved obviously by surface modification of RDX.

Keywords: RDX, Surface modification, Conducting polymer，Safety Properties, Composite solid propellant

1 Introduction

Explosive compounds with Reduced Sensitivity for Insensitive Munitions is desired for reasons of safe load, assemble, and pack operations and to meet the improved performance requirements of the energetic. RDX (1, 3, 5-Trinitro-1, 3, 5-triazacyclohexane) has been widely used in gun powders and composite solid propellants due to its high energy performance, good stability and low cost. The most widely used energetic nitramines RDX is very sensitive to impact and friction stimuli. For improving energy performance, near to 60% RDX were added to many composite propellants such as composite modified double base (CMDB) propellants, however, the propellant with a high proportion of RDX could increase mechanical sensitivity. To lower the sensitivity toward mechanical stimuli, many materials have been used to coat nitramines explosive, such as wax [1], graphite [2], stearic acid [3, 4], polymers [5, 6] and insensitive explosives [7-9]. On the other hand, RDX is the dielectric material, it may accumulate electric charges due to the rubbing effect around from the material contact, leading to explosions, fires or personal injury shock. Therefore, using a simply strategy to realize RDX surface conductive remains a scientific and technical challenge. In the present work, we report a simple strategy to coat RDX with CP to realize surface conductivity and decrease mechanical insensitivity. To the best of our knowledge, the researches on RDX coated with conducting polymer are not available in earlier publications.

2 Experimental

2.1 Materials

RDX was provided by Jiangsu Hongguang Chemistry Factory (China.). PET, A3 (mixture of BDNPF and BDNPA) and were supplied by Liming Research institute of Chemical Industry (China). Bonding Agent（BA-1）was fabricated by ourselves. Conducting Polymer (CP) monomer (Shanghai chemical Co.) was distilled under reduced pressure. Ammonium peroxydisulfate (APS, $(NH_4)_2S_2O_8$) was of analytical grade and used as received without further purification.

2.2 Equipment and Characterization

Scanning electron microscopy (SEM) images were obtained on a HITACHI S-4800. Elements analyses were carried on the Elementar Vario EI Cube instrument. X-ray diffraction (XRD) patterns were measured on a Brucker D8 Advance instrument using Cu Kα radiation. Thermo gravimetric analysis (TGA) was carried out on s Shimadzu TGA-50 instrument at a heating rate of 10℃min^{-1} in air. Electrical conductivity measurements were carried out by using four-point conductivity method at room temperature. Fourier transform infrared (FTIR) spectra were recorded on a Bruker IFS 66/s spectrometer.

Impact sensitivity and friction sensitivity of samples were measured and evaluated with National Army Standard of China. Impact sensitivity was surveyed by the drop hammer apparatus made by De Kong Corporation. The experimental conditions were as follows: drop hammer weight, 2kg; sample mass, 30±1mg. Friction sensitivity was tested by MGY-1 pendulum friction apparatus made by 213 Research Institute of China. The experimental conditions were as follows: pendulum weight,1.5kg; swaying angle, 90°, sample mass, 20±1mg.

Propellant mixtures were prepared in vertical mixing system and mechanical tensile testing was conducted with an 810MTS (Material Testing System) according to the procedure of STANAG 4506.

2.3 Preparation of Samples

2.3.1 Preparation of BA-1/RDX Sample

After the Bonding Agent BA-1 dissolved in DMF solvent, RDX powder was added into the above solution with vigorous stirring for one hour at room temperature. Then, the solution was evaporation in vacuum at 50℃ for about 48 hours, the BA-1/RDX Samples were fabricated.

2.3.2 Preparation of CP/RDX Sample

RDX powder was dispersed in de-ionized water with vigorous stirring. Then, APS were put into the solution under continuous stirring. The polymerization was allowed to proceed for 8h at room temperature. After that the deposited precipitate was separated and washed by de-ionized water for three times. Finally, the precipitate was dry at 50℃ for 48h and the CP/RDX Samples was fabricated.

2.3.3 Preparation of CP/BA-1/RDX Sample

In a typical procedure, BA-1/RDX was dispersed in the de-ionized water. Then, APS were put into the solution under continuous stirring. The polymerization was allowed to proceed for 8h at room temperature. After that the deposited precipitate was separated and washed by de-ionized water for three times. Finally, the precipitate was dry at 50℃ for 48h and the CP/BA-1/RDX Samples were fabricated.

3 Results and Discussion

3.1 SEM Characterization

RDX, BA-1/RDX and CP/BA-1/RDX samples were characterized by SEM and the results were shown in Fig1, Fig2 and Fig3. As can be seen from Fig1, RDX particles were dispersed evenly and

their surface was quite smooth and clean. It can be seen that from the optical photographs in the insets of Fig1b, RDX is white powder. When BA-1 was used to modify RDX, RDX surface was coated with a film and attached many small particles (Fig2). Fig.3 shows the SEM image of CP/BA-1/RDX that the surface of RDX becomes coarse. A lot of particles deposited on the surface of RDX. The composite particles are shown a dark green color (see the inset optical photographs of Fig3b), indicating that the aniline monomers have successfully polymerized on the surface of BA-1/RDX.

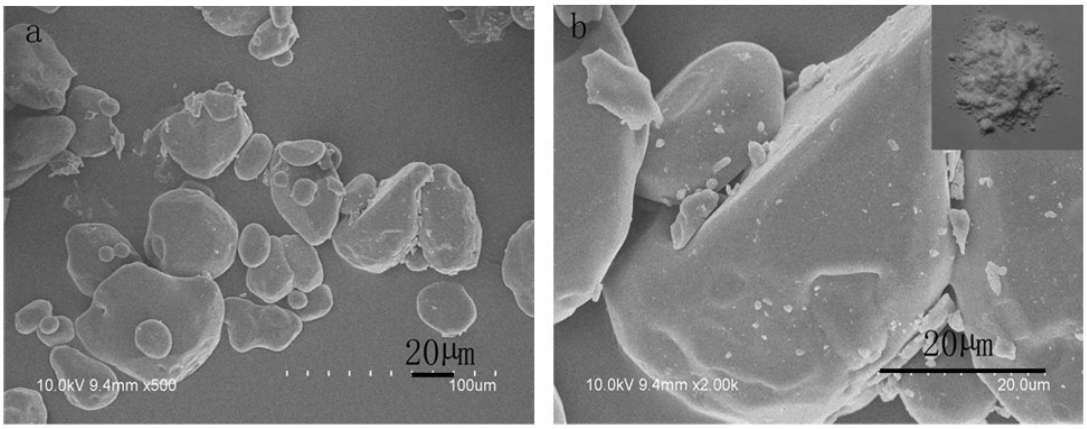


Fig.1 SEM images of unprocessed RDX（b shows the higher magnification SEM of a，The insets showed the optical photographs of RDX）

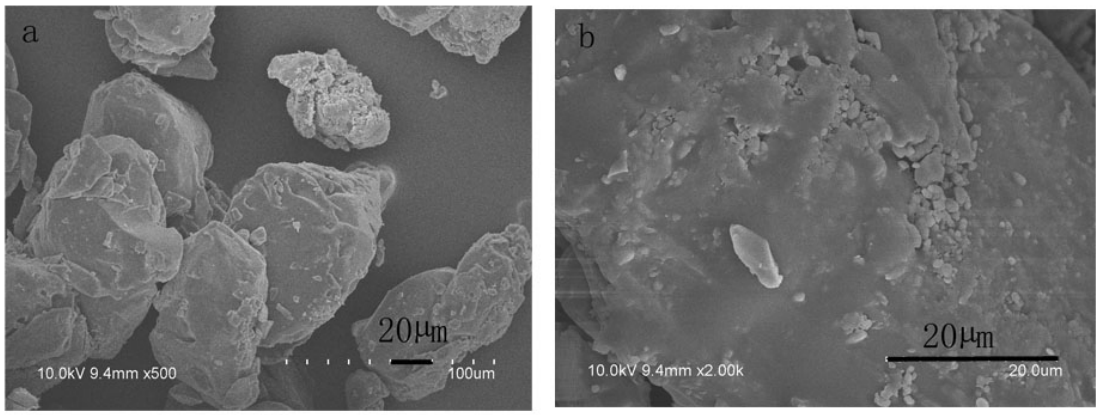


Fig.2 SEM images of BA-1/RDX（b shows the higher magnification SEM of a）

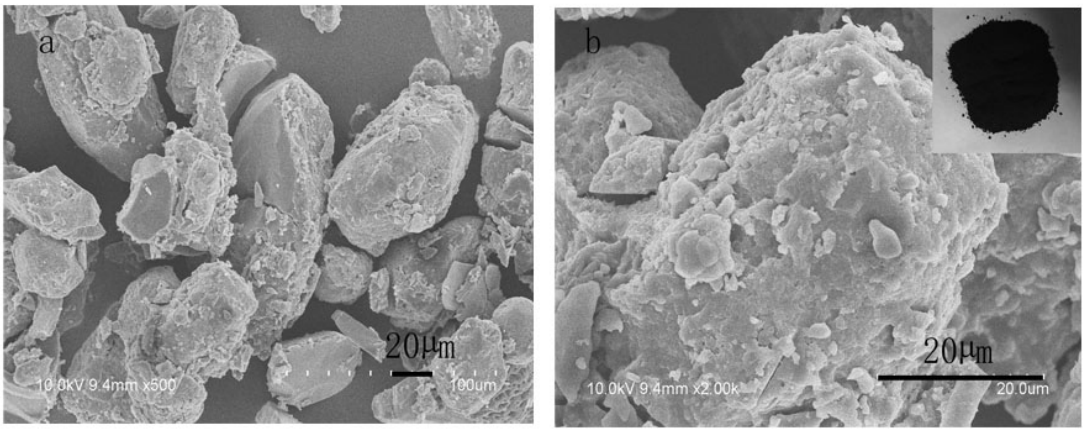


Fig.3 SEM images of CP/BA-1/RDX（b shows the higher magnification SEM of a，The insets showed the optical photographs of CP/BA-1/RDX）

3.2 Elements Analysis

The surface elementary content of BA-1/RDX and CP/BA-1/RDX samples was characterized by elements analysis. The results of elements analysis are shown in table 1.The surface elementary mass fraction of coating samples was also obtained by elements analysis. The surface elementary mass fraction w, a parameter indicating the coating effect, is estimated according to the mass fraction of C. C_{mix} and C_{RDX} are the mass fraction of C of the coated sample and raw samples, respectively. C_m is the mass fraction of C of the coating material.

$$w = \frac{C_{mix} - C_{RDX}}{C_m - C_{RDX}} \tag{1}$$

The elements analysis showed that the mass fraction of C increases from 16.25 to 18.56% after surface coating, which could be due to the higher C content in BA-1 and CP than that in RDX. Moreover, the surface elementary mass fraction (w %) of BA-1/RDX and CP/BA-1/RDX are calculated as 2.7% and 1.84%, which indicated that BA-1 and CP were successfully coated on the surface of RDX. The results were almost consistent to cast material mass fraction of BA-1 (3%) and CP (2%).

Table 1 Test results of elements analysis

Sample	C wt %	w （%）
RDX	16.25	-
BA-1	62.47	-
CP	51.62	-
BA-1/RDX	17.94	2.7
CP/BA-1/RDX	18.56	1.84

3.3 Particle size distribution

In Fig.4, particle size distributions of RDX and CP/BA-1/RDX were shown in upside figure and downside figure respectively. The d10, d50 and d90 of RDX were 8.2μm, 40.9μm and 163.9μm. The d10, d50 and d90 of CP/BA-1/RDX were 9.13μm, 50.7μm and 109.3μm. Figure 4 showed that the particle size of CP/BA-1/RDX was bigger than RDX but the particle size distribution was narrower than RDX. The results were also indicating that the aniline monomers and BA-1 had been successfully coated on the surface of RDX.

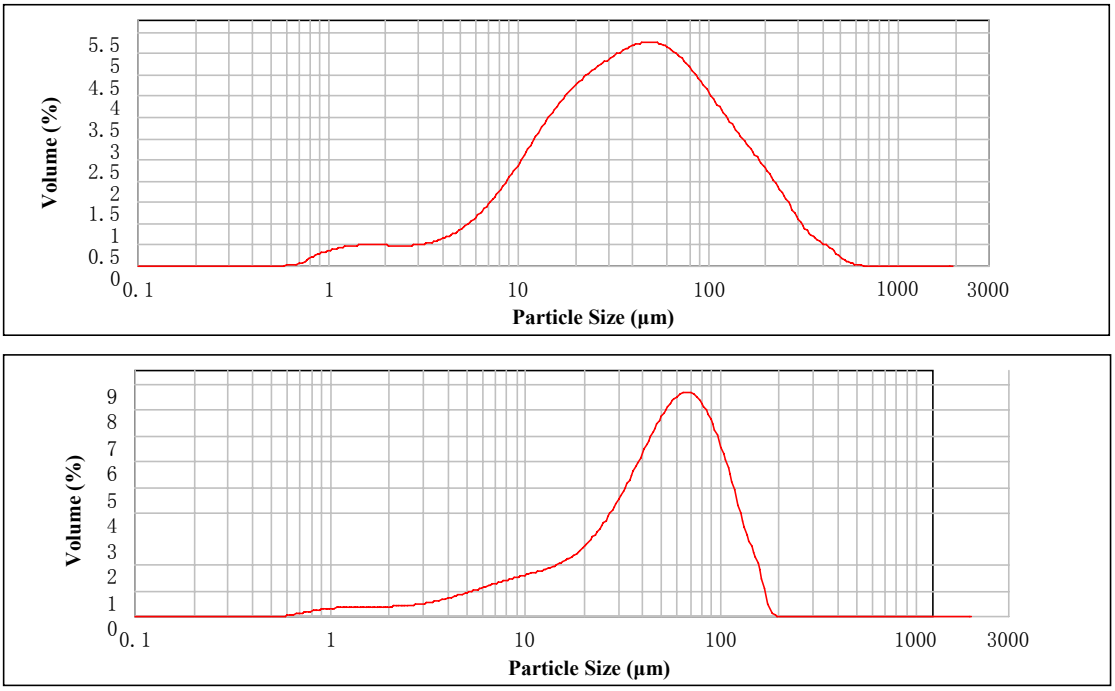


Fig.4 Particle size distribution of RDX (upside figure) and CP/BA-1/RDX (downside figure)

3.4 XRD Character

The crystalline structures of RDX, BA-1/RDX and CP/BA-1/RDX were examined by XRD. In Fig.5, the characteristic narrow peaks were observed at the same diffraction angles for all the samples. All the diffraction peaks were indexed to RDX, indicating that the crystalline structure of RDX was retained with coated BA-1 and CP. The CP and BA-1 coated on RDX were in amorphous nature. The reductions in the peak intensities of the BA-1/RDX and CP/BA-1/RDX powder may be attributed to successful surface modification of RDX.

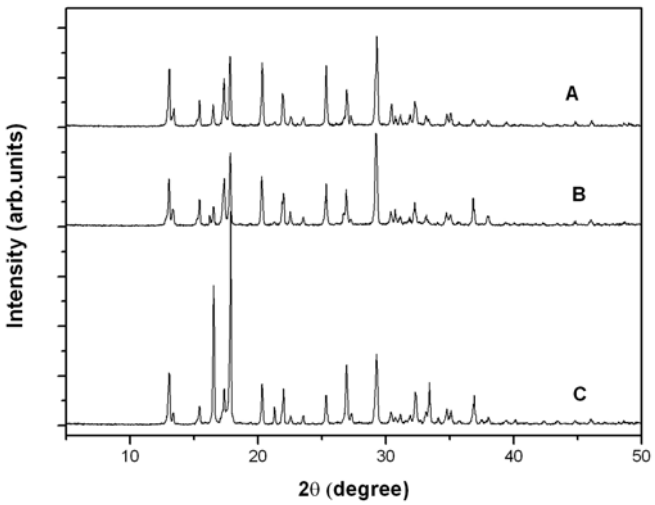


Fig.5 XRD patterns of （A）CP/BA-1/RDX, (B) BA-1/RDX and (C) RDX

3.5 Mechanical Sensitivity

Impact and friction sensitivity were measured and the results were listed in Table 2. The results of friction sensitivity were expressed by the explosion probability (P) and its confidence interval (P_1, Pu). The greater P is, the higher sensitivity is. The impact sensitivity of RDX was character by the

drop height (H_{50} and standard deviation (S)). The impact sensitivity (H_{50}) and friction sensitivity (P) of RDX were 39cm and 90%, respectively. When coated by CP and BA-1, the impact sensitivity (H_{50}) and friction sensitivity (P) of RDX were reached to 61.2cm and 30%, respectively. The BA-1 plays a buffer and lubrication role when external forces were acted on the coated sample, therefore leading to a reduction in the probability of formation of hot spots [10, 11]. These sensitivity results were slight higher than the report about RDX coated with HP-1/TNT [12]. As can be seen form table 2, the surface conductivity had great increase from insulating to 5×10^{-4}S/cm. The conductive property of RDX may benefit to remove static electricity buildup and improve the safety of propellants formulation which contains RDX particles.

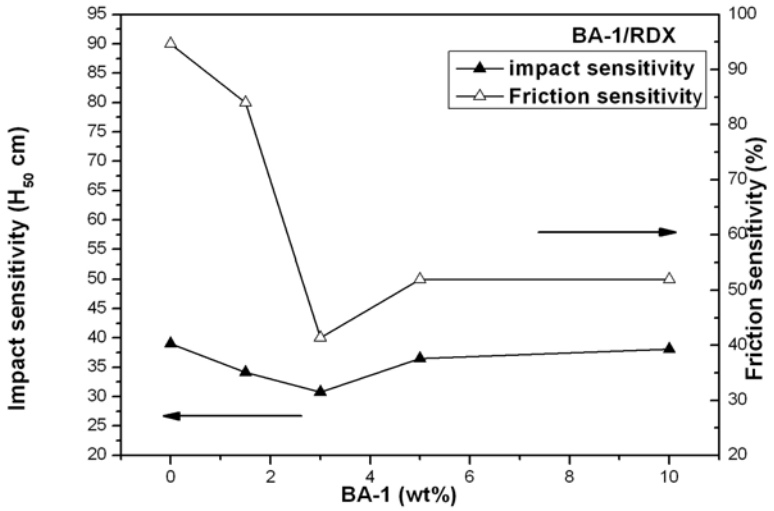


Fig.6 Impact sensitivity and friction sensitivity of BA-1/RDX

To evaluate the effect of BA-1 on the sensitivity of RDX, different content of BA-1 were used to coat on the surface of RDX (Fig.6). The BA-1 did not appear to be as effective for both impact and friction sensitivity. It could be seen that with the increase content of BA-1, the friction sensitivity of BA-1/RDX had the tendency to decline. When the content of BA-1 arrived at 3wt%, the friction sensitivity was decreased at the lowest point. However, the BA-1 made RDX more sensitive to impact. When RDX coated by BA-1, the impact sensitivity of BA-1/RDX were obviously increased. Thus, it could be concluded that the BA-1 on the surface of RDX affected the friction sensitivity only.

Table 2

Sample	Impact Sensitivity		Friction Sensitivity		Electrical Conductivity
	H_{50} (cm)	S	P%	[P_1, Pu]	(S/cm)
RDX	39.0	0.70	90%	[70%,98%]	-
CP/BA-1/RDX	61.2	5.37	30%	[13%,52%]	5×10^{-4}

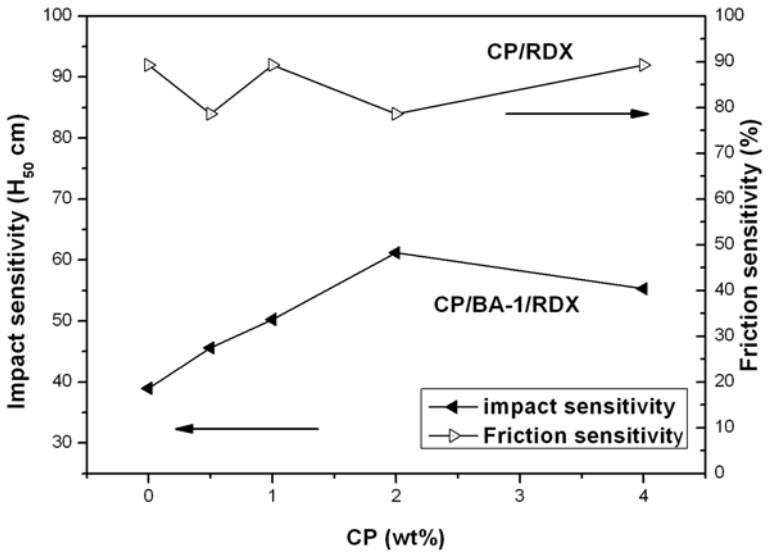


Fig.7 Impact sensitivity and friction sensitivity of CP/BA-1/RDX

As shown in Fig.7, with increasing of CP content, the friction sensitivity of CP/RDX was almost unchanged, which indicated that the CP did not affect friction sensitivity remarkably. However, the impact properties were improved significantly when CP was coated on the surface of BA-1/RDX, the impact sensitivity H_{50} reached the highest 61.2cm with 2% CP. In addition, as shown in Table 2, the friction sensitivity was reduced to the lowest value of 30%. Therefore, both impact sensitivity and frication sensitivity of RDX were decreased after surface modification with CP/BA-1, moreover, only the BA-1 and CP used together to modify RDX, the impact and friction propertied could be improved simultaneously.

3.6 DTA and TGA Analysis

The thermal characteristics of BA-1/RDX and CP/BA-1/RDX in comparison to RDX had been studied by DTA and TGA in air atmosphere. The exothermic peak at approximately 232.27℃ observed in raw RDX (shown in Fig.8(a)), which was lower than that of BA-1/RDX and CP/BA-1/RDX. The results indicated that the thermal stability of RDX was enhanced by surface coating with BA-1 and CP. The decomposition temperature of RDX had been increased about 7℃. According to hot spots mechanism [10, 11], the process decomposition of CP/BA-1/RDX need absorb more energy to improve the temperature to arrived at the hot spots temperature, which may be the reason of lower sensitivity of CP/BA-1/RDX than raw RDX.

It is possible to observe in Fig.8 (b) that the TGA shows, RDX showed one step weight loss, which was attributed to RDX decomposition. By comparison, two steps of weight loss for BA-1/RDX start, respectively at 172℃ and 245℃, corresponded to removal of the decomposition of RDX and degradation of BA-1. The CP/BA-1/RDX also showed two steps of weight loss and the sec-step lost for CP/BA-1/RDX was more than that for BA-1/RDX. The 5% weight loss in sec-steps could be referenced to the content of BA-1 and CP, which indicated that CP was successfully coated on the surface of BA-1/RDX.

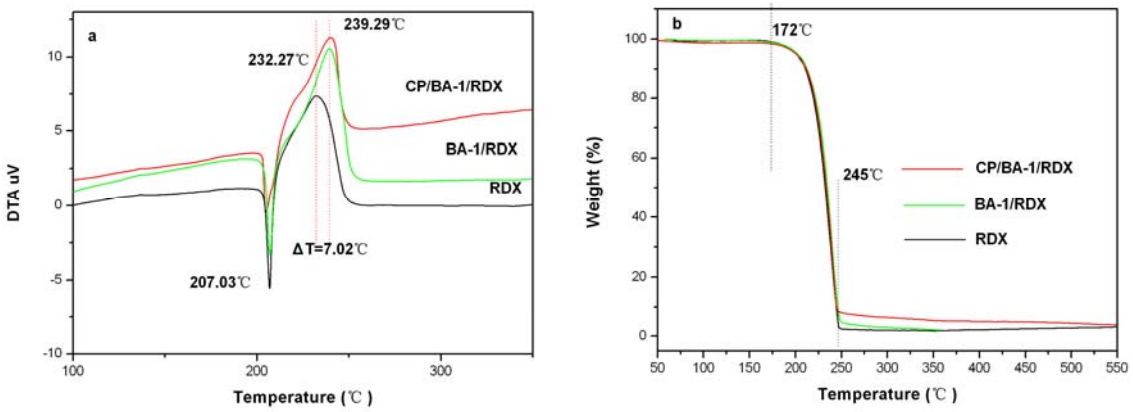


Fig.8 (a) DTA measurement, (b) TGA measurement

3.7 Application

The coated and raw RDX were used to prepare composite solid propellant. The formulations of composite solid propellant were listed in table 3. The propellant based on RDX (15%), AP filler (47%), PET binder, Al fuel and A3 plasticizer were mixed on a 1300g scale. The total solid loading was 80wt%. Mechanical properties of propellant were presented in table 4. The composite solid propellant mixed with raw RDX filler gave a tensile strength of 294KPa and an elongation at break of 105.4% at 70℃. However, after replacement of the raw RDX with coated RDX, the tensile strength at 70℃ was increased to 345KPa. In addition, under the temperature of 20℃ and -45℃, the tensile strength were increased from 612KPa, 2081KPa to 950KPa,3279KPa, respectively.

Table3 Coated RDX application in composite solid propellant

PET	AP	RDX	A3	Al
8%	47%	15%	12%	18%

Table 4 the mechanical properties of composite solid propellant

Number	70℃			20℃			-45℃		
	σ_m KPa	ε_b %	ε_m %	σ_m KPa	ε_b %	ε_m %	σ_m KPa	ε_b %	ε_m %
1	294	105.4	101.2	612	121.7	85.5	2081	90.3	18.7
2	345	102.9	98	950	123.6	99.2	3279	53.5	40.4

1 RDX, 2 CP/BA-1/RDX, σ_m (Max tensile strength), ε_b (elongation at break), ε_m (elongation at σ_m)

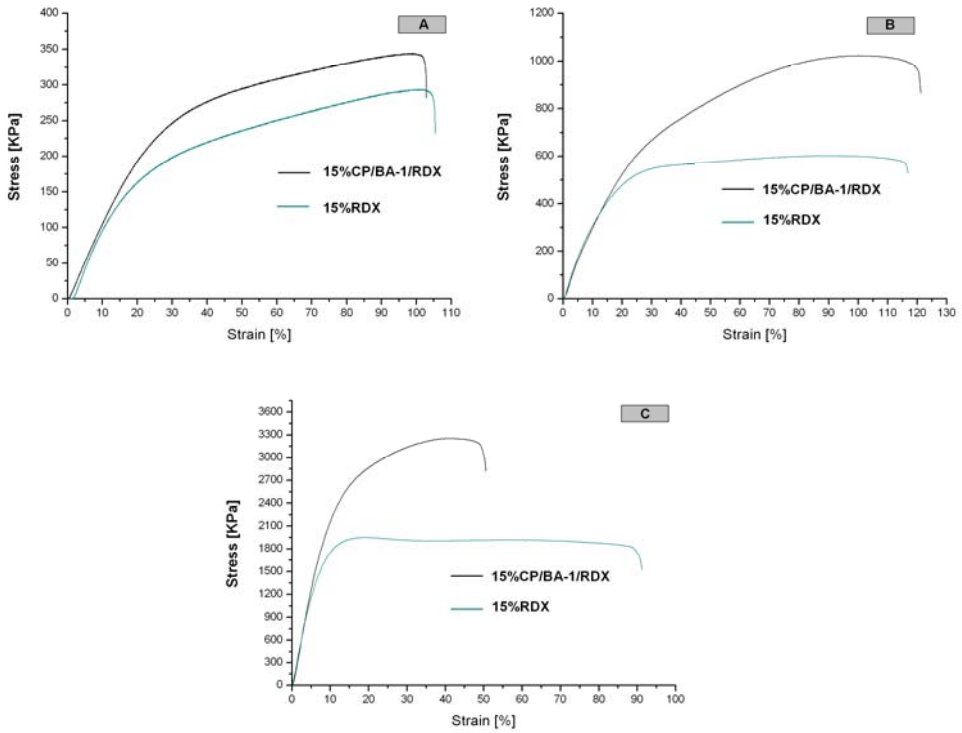


Fig.9 Tensile testing of PET composite propellants, A, 70℃, B, 20℃, C, -45℃

From Fig.9, the tensile strength of propellant mixed with CP/BA-1/RDX was higher than the propellant mixed with raw RDX. Furthermore, the propellant mixed with raw RDX has the "dewet" phenomena at lower temperature (Fig.9-c), but the phenomenon disappeared after replacement of the raw RDX with CP/BA-1/RDX. Therefore, the mechanical properties of the composite solid propellant containing RDX could be improved obviously by surface modification of RDX. The BA-1 is a macromolecular bonding agent .It is possible to synthesize a polar bonding agent with a favorable adsorption onto the RDX surface. Reactive pendant hydroxyl functionalities are incorporated through copolymerization, and the assumption is that these react with isocyanate curing agent. Therefore, coating RDX with BA-1 and CP could improve the tensile strength of composite solid propellant.

4 Conclusions

In summary, the CP and BA-1 had been successfully coated on the surface of RDX. The impact and friction sensitivity of RDX could decrease obviously when its surface coated with 2% CP and 3% BA-1. The impact sensitivity H_{50} was changed from 39.0cm to 61.2cm, which thanks to the effect of CP. The friction probability was reduced from 90% to 30%, which mainly due to the coated BA-1. The surface conductivity of RDX had greatly increased from insulating to 5×10^{-4}S/cm.The modified RDX could also improve the tensile strength and safety of PET composite solid propellant containing RDX.

References

[1] R. C. L. Dutra, B. G. Soares, E. A. Campos, J. D. G. Melo, J. L. G. Silva, Composite Materials Constituted by a Modified Polypropylene Fiber and Epoxy Resin, J. Appl. Polym. Sci. 1999, 73, 69.

[2] C. Hohmann, B. Tipton, Jr., Viton's Impact on NASA Standard Initiator Propellant Properties, Tech. Report 2000, NASA/TP 210187, Houston, Texas.
[3] P. Wanninger, R. Wild, E. Kleinschmidt, H. Späth, Pressable Explosive Granular Product and Pressad Explosive Charge, U. S. Patant 5547526, 1996, Daimler-Benz Aerospace AG, Germany.
[4] T. Urbanski, Chemistry and Tecnolgy of Explosive, Pergamon Press, Great Britain 1967, Vol.3, P. 391.
[5] G. Silva, E.C. Mattos, N. M. Nakamura, K. Iha, Aplicacão da Calorimetria Exploratória Diferencial no Estudo da Cinética de Transicão alfa-delta HMX, Quím. Nova 2004, 27, 889.
[6] U. Teipel, Energetic Materials –Processing and Characterization, Wiley- VCH Verlag, Weinheim, Germany 2006.
[7] G. Silva, E.C. Mattos, N. M. Nakamura, K. Iha, Estudo Cinético da Decomposicão Térmica do alfa HMX por Calorimetria Exploratória Diferencial, An, Assoc, Bras, Quím, 2003,52, 81.
[8] A. L. Smith, Applied Infrared Spectroscopy, John Wiley & Sons, New York 1979, p. 286.
[9] M. S. Campbell, D. Garcia, D. Idar, Effects of Temperature and Pressure on the Glass Transitons of Plastic Bonded Explosive, Thermochim, Acta 2000, 357-358, 89.
[10] M. S. Kirshenbaum, Reactivity of Explosive-Binder/Sediment Mixtures, Technical Report ARLCD-TR-82007, US Army ARDCOM, 1982, Dover, NJ.
[11] J. E. Field, Hot Spot Ignition Mechanisms for Explosives, Acc, Chem, Res. 1992,25,489.
[12] C. W. An, F. S. Li, X. X. Song, Y. Wang, X. D. Guo, Surface coated of RDX with a Composite of TNT and an Energeitic-Polymer and its Saftey Investigation, Propellants Explos. Pyrotech. 2009, 34, 400-405

Acknowledgements

This work was sponsored by Natural Science Foundation of Shanghai (No.14ZR1420500)

A TOOL FOR SIMULATION AND ANALYSIS OF SPACE PROPULSION SYSTEMES TRANSIENT PERFORMANCE

Chen Hongyu[1] , Liu Hongjun[2] , Hong Liu[3]

1. National Key Laboratory of Science and Technology on Liquid Rocket Engines, China, chy08211@163.com
2. National Key Laboratory of Science and Technology on Liquid Rocket Engines, China
3. National Key Laboratory of Science and Technology on Liquid Rocket Engines, China

ABSTRACT: Modelling and simulation of liquid propellant space propulsion system is one of the essentials to develop advanced propulsion system with lower cost, less risk and higher efficiency. But, constructing flexible simulation tool capable of evaluation of integrated propulsion system architectures is hard work. This paper presents a liquid propellant propulsion system model library in Modelica language, containing component models such as pipes, valves, tanks, turbo-pumps, combustion chambers, nozzles, injectors, gas generators, etc. These component models can be applied to establish a variety of space propulsion systems with the capability of predicting propulsion system transient performance during start-up and shutdown process. The liquid propellant propulsion system model library has been applied to the Chinese space bus. By comparing its outputs with combustion test data, the validity of the library has been checked and improved. All the modelling and simulations are implemented in MWorks, which is a modelling and simulation platform full supporting Modelica.

NOMENCLATURE

A	areas	C_d	flow rate factor
$\dot{m}$	Mass flow rate	f	Darcy–Weisbach friction factor
K_m	propellants mixture ratio	t	Time
RT	product of gas constant and temperature	n	rotational speed
V	volume	q	volume flow rate
v	velocity	η_p	pump efficiency
P	power consumption	ω	angular velocity
g	acceleration of gravity	ρ	density of fluid
p	pressure	M	Mass matrix
D	pipe inner diameter	γ	specific heat ratio
L	pipe length	ε	nozzle expansion ratio
F	thrust force	τ	time

Subscripts

c	chamber	in	input
e	exit	out	output
f	fuel	ox	oxidant
g	gas	t	throat

I. INTRODUCTION

Modeling and simulation of liquid propulsion system is one of the essentials to develop advanced space propulsion system with lower cost, less risk and higher efficiency. Typically, the start-up and shutdown process of a liquid propellant propulsion system involves non-linear interactions between multiple components with phenomena such as flow resistance, heat transfer, phase change, combustion, and turbo-pump operation. Furthermore, the physical properties of liquid propellants and combustion gases in such systems vary widely. Therefore, developing tools for predicting the dynamical behavior of space propulsion systems with such characteristics is a both challenging and important task for engineers and researchers [1].

Worldwide, various tools have been developed to simulate the transient behavior of propulsion systems. Many countries and areas including America, Europe, Russia, Japan and China, which keep ahead in aerospace field, have respective owned software. Moreover, any software has achieved commercial application. The Rocket Engine Transient Simulator (ROCETS) [2] was designed and developed during the 90s by Pratt & Whitney for NASA-MSFC, allows for cost-effective computer predictions of liquid rocket engine transient performance. The most popular application of ROCETS is the RL-10A-3-3A rocket engine[3,4]. Another powerful tool created in the United States is the Generalized Fluid System Simulation Program (GFSSP) for modelling cryogenic fluids in a complex flow circuit [5].

In Europe, CNES tested a dedicated library for the modelling and simulation of propulsion system dynamics developed in the AMESim platform [5]. The platform carries multiple sub-systems of a propulsion system, such as tanks, pneumatic lines, turbo-pumps, regenerative circuit, combustion chamber, and starters. Recently, a complete set of models able to simulate liquid propulsion system components called European Space Propulsion System Simulation (ESPSS) has been developed by a joint European team in the frame of a GSTP Program for the European Space Agency [6].

In Japan, the Visual Integrated Simulator for Rocket Engine Cycle (VISREC)[7] was developed by Mitsubishi Heavy Industries. VISREC is a one dimensional flow and heat analysis program using the lumped parameter approach. Together with the LE-7A rocket engine, the Rocket Engine Dynamic Simulator (REDS) [1] in 2004 was developed and applied to start-up and shut-down transient analyses.

This paper presents a generic modular-modelled library for the modelling and analysis of the hydro-dynamical and thermo-dynamical events inside a liquid propellant space propulsion system, called LiquidRocketEngine. It can analyse start-up and shut-down transient behaviour of many types of engine cycles such as pressure fed cycle, staged combustion cycle, and gas generator cycle.

II. IMPLEMENTATION OF COMPONENT MODELS

We try to utilize capabilities of Modelica[8] to the fullest when developing the liquid propellant rocket engine library. Firstly, we determine the objective of the library. The system model is expected to predict the flow rate and pressure of components in the liquid propellant space propulsion system during the whole running time. The dynamic phenomenon of propulsion system mainly includes fluid dynamics, heat transfer, thermal dynamics and combustion, which should be taken into consideration.

According nature border in liquid propellant space propulsion system and object orientation, we divide system into interacting components. Object orientation is viewed as a structuring tool to

handle topological structure description of propulsion system. In order to decrease complexity, the assumption is made that decomposed components are independent physical functional objects, which are pipes, valves, tanks, turbo-pumps, combustion chambers, nozzles, injectors, gas generators and bottles. While subject decomposition is also performed to obtain base models, which collect common properties of a class of models and are physical phenomenon units. Main base models in LPRE library are control volume, flow model, ideal gas property, heat convection, heat conduction and combustion model. These base models can't be simulated directly, and are inherited and aggregated by more than one component models for reusing.

The components models exchange information through connectors. The connectors ensure that components are independent of each other and work under a set of boundary conditions provided by connectors. Connectors of component models should present properties of interaction between these components in the real physical propulsion system. Thus it would be easy and nature to connect components. LPRE library contains four connectors for fluid flow, heat flow, 2D rotation and 2D translation respectively. In order to make liquid propellant space propulsion system library consistent with Modelica Standard Library (MSL) and increase versatility, four kinds of the connectors in propulsion library are the same with that in MSL. Using base models, connectors and mathematical model of physical components, detailed component models are developed and implemented.

III. DESCRIPTION OF MATHEMATICAL MODELS

In this chapter, mathematical models of some most interesting components are presented, including combustion chamber, nozzle, pipe, valve, pump and turbine.

III. I COMBUSTION CHAMBER

The thrust chamber is the key subsystem of a liquid propellant space propulsion system. The combustion chamber is the part of a thrust chamber where the chemical reaction of the propellant takes place to generate hot gas products. It's assumed that liquid propellants react and change to hot gas after a constant delay time and gas flow in combustion chamber is adiabatic. The control equations of a bipropellant combustion chamber are described as following.

$$\frac{dm_{ox}}{dt} = \dot{m}_{oxin} - \frac{m_{ox}}{\tau_{ox}} - \dot{m}_{oxout} \tag{1}$$

$$\frac{dm_f}{dt} = \dot{m}_{fin} - \frac{m_f}{\tau_f} - \dot{m}_{fout} \tag{2}$$

$$\frac{dm_g}{dt} = \frac{m_f}{\tau_f} + \frac{m_{ox}}{\tau_{ox}} - \dot{m}_{gout} \tag{3}$$

$$\frac{dK_m}{dt} = \frac{1+K_m}{m_g}\left(\frac{m_{ox}}{\tau_{ox}} - \frac{m_f}{\tau_f}\right) \tag{4}$$

$$\frac{dp}{dt} = \frac{RT}{V}\frac{dm_g}{dt} + \frac{p}{RT}\frac{d(RT)}{dt} + \frac{p}{V}\frac{dV}{dt} \tag{5}$$

$$\tau_g \frac{dRT}{dt} = RT_i(p, K_m) - RT \tag{6}$$

Where, m_{ox} is oxidizer mass, m_f is fuel mass, m_g is product gas mass; $\dot{m}_{oxin}$ and $\dot{m}_{oxout}$ are inlet and outlet mass flow rate of oxidizer, $\dot{m}_{fin}$ and $\dot{m}_{fout}$ are inlet and outlet mass flow rate of fuel; K_m is propellants mixture ratio, p is pressure inside chamber, RT is the product of gas constant and temperature of gas staying in combustion chamber, $RT_i(p, K_m)$ is the product of gas constant and temperature of gas products defined as the function of pressure and mixture ratio, V is the volume of chamber; τ_g is the stay time of gas in chamber, τ_{ox} and τ_f are transformation time of oxidizer and fuel, given by Equation 7 (and is a function of the chamber pressure to the power -m).

$$\tau = p^{-m} e^{\frac{E_{prop}}{RT}} \tag{7}$$

The control equations of gas generator resemble that of combustion chamber.

III. II NOZZLE

Nozzle is also the part of a thrust chamber, whose function is to accelerate gases and create high exhaust velocity. It is assumed that the gas flow through the nozzle is an isentropic expansion. Nozzle expansion ratio is defined as follow:

$$\varepsilon = \frac{A_e}{A_t} = \frac{p_c}{p_e}\left(\frac{2}{\gamma+1}\right)^{\frac{1}{\gamma-1}} \left\{\frac{\gamma+1}{\gamma-1}\left[1-\left(\frac{p_e}{p_c}\right)^{\frac{\gamma}{\gamma-1}}\right]\right\}^{-1/2} \tag{8}$$

Here, A_e and A_t are flow areas at nozzle exit and throat, p_c and p_e are pressure at chamber and nozzle exit, γ is specific heat ratio.

Velocity at nozzle exit:

$$v_e = \sqrt{\frac{2g\gamma}{\gamma-1} RT_c \left[1-\left(\frac{p_e}{p_c}\right)^{\frac{\gamma-1}{\gamma}}\right]} \tag{9}$$

Here, R is gas constant, T_c is gas temperature in combustion chamber.

The mass flow through a nozzle is obtained:

$$\dot{m} = \frac{A_t p_c \gamma}{\sqrt{\gamma RT_c}} \sqrt{\left(\frac{2}{\gamma+1}\right)^{\frac{\gamma+1}{\gamma-1}}} \tag{10}$$

The thrust force is defined as follow:

$$F = \dot{m} v_e + A_e(p_e - p_a) \tag{11}$$

III. III PIPES

Pipes are interconnect components that carry fluid to the intended components.For one-dimensional transient flow, based on the Newton's second law and the mass conservation principle, the basic control equations of transient flow which includes the momentum conservation

equation and continuity equation can be finally deduced and expressed as two partial differential equations [9].

$$\frac{\partial p}{\partial t}+v\frac{\partial p}{\partial x}+\rho a^2\frac{\partial v}{\partial x}=0,\ x\in[0,L] \tag{12}$$

$$\frac{\partial v}{\partial t}+v\frac{\partial v}{\partial x}+\frac{1}{\rho}\frac{\partial p}{\partial x}+\frac{f}{2D}v|v|+g\cos\theta=0,\ x\in[0,L] \tag{13}$$

Where, x is the space coordinate, t the time, v the velocity in pipe, p the pressure, f the Darcy–Weisbach friction factor, D the inner diameter of pipe, ρ the density of fluid, g the acceleration of gravity, L the pipeline length, and a the wave speeds.

Transients flow in pipelines is formulated by a set of nonlinear, hyperbolic partial differential equations (PDE). The equations can seldom be solved analytically. In this paper, we use the Chebyshev super spectral viscosity (SSV) method [10] to solve the equations.

III. IV VALVES

Valves control fluid flows and any one LPRE will use some of them. Liquid valves are governed by the following familiar equation, where flow rate is the function of pressure drop and flow area:

$$\dot{m}=C_d A\sqrt{2\rho \mathrm{V} p} \tag{14}$$

Here C_d is the flow rate factor.

For a gas valve, when $\frac{p_o}{p_i}\geq\left(\frac{2}{\gamma+1}\right)^{\frac{\gamma}{\gamma-1}}$:

$$\dot{m}=C_d A\sqrt{\frac{2\gamma p_i\rho_i}{\gamma-1}\left[\left(\frac{p_o}{p_i}\right)^{\frac{2}{\gamma}}-\left(\frac{p_o}{p_i}\right)^{\frac{\gamma+1}{\gamma}}\right]} \tag{15}$$

When $\frac{p_o}{p_i}\leq\left(\frac{2}{\gamma+1}\right)^{\frac{\gamma}{\gamma-1}}$, then:

$$\dot{m}=C_d A\sqrt{\frac{2\gamma p_i\rho_i}{\gamma+1}\left(\frac{2}{\gamma+1}\right)^{\frac{2}{\gamma-1}}} \tag{16}$$

III. V PUMP

The pump pressurize propellants and deliver them to extended components in a turbo-pump propellant feed system. Centrifugal pump is the most widely used pump type. Performance maps for head and power are used in the pump model. The head h is evaluated by following equation:

$$h=\left(\frac{n}{n_{ref}}\right)^2 h_q\left(q\frac{n_{ref}}{n}\right) \tag{17}$$

$$P=\left(\frac{n}{n_{ref}}\right)^3\frac{\rho}{\rho_{ref}}P_q\left(q\frac{n_{ref}}{n}\right) \tag{18}$$

$$P=\omega\tau \tag{19}$$

$$\eta_p = \frac{\mathrm{V}pq}{P} \tag{20}$$

Here, n is rotational speed, q is volume flow rate, P is power consumption, ω is angular velocity, τ is torque, η_p is pump efficiency, $h_q\left(q\frac{n_{ref}}{n}\right)$ and $P_q\left(q\frac{n_{ref}}{n}\right)$ are functions obtained from performance map, Subscript ref is the reference value.

III. VI TURBINE

The turbine gets energy from the expansion of high temperature and high pressure gas, and provides power to the pump. The expansion is assumed to be isentropic. The control equations of turbine is presented as follows:

$$W = \frac{\gamma}{\gamma - 1} RT_i \left[1 - \left(\frac{p_o}{p_i}\right)^{\frac{\gamma-1}{\gamma}}\right] \tag{21}$$

$$\eta_t = \eta_t\left(\frac{u}{C}\right) = \frac{\omega\tau}{W\dot{m}} \tag{22}$$

Here, W is the power of gas expansion, $\eta_t\left(\frac{u}{C}\right)$ is the turbine efficiency defined as a function of velocity ratio obtained from performance map. The mass flow rate in a turbine is evaluated by the familiar equation in gas valve model.

IV. SIMULATION AND ANALYSIS OF SYSTEM

A propulsion system is modelled as a system of pipes with various hydraulic elements such as pumps, turbines, valves, orifices, tanks. Flows of fuel, oxidizer and combustion gas are simulated by solving conservation equations taking into account the materials' phase changes and physical properties. The graphic image of analysis model of the LiquidRocketEngine library is shown in Fig. 1. It is implemented in MWorks[11].

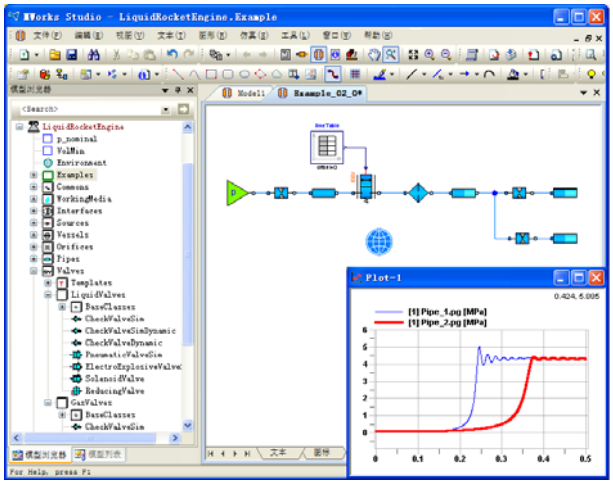


Fig. 1 Graphic modelling environment

The LiquidRocketEngine library has been developed to simulate unsteady behaviour such as the start-up process of an upper stage engine, and has been applied to the Chinese space bus. The analysis model is shown in Fig. 2.

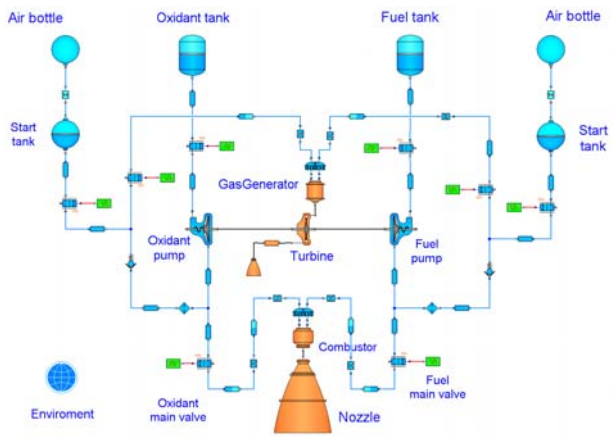
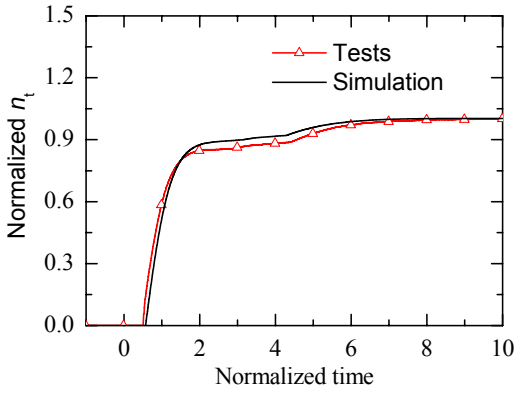
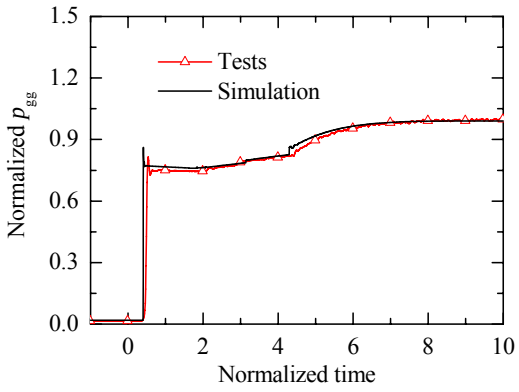


Fig. 2 Example of an upper stage engine analysis model

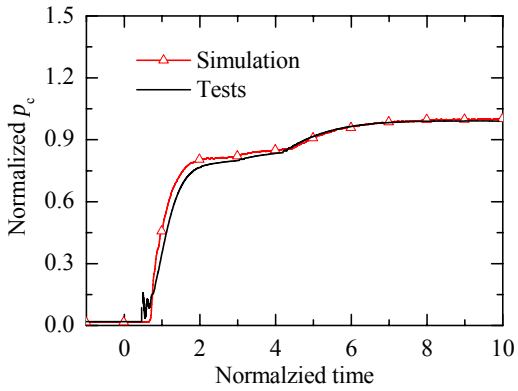
By comparing its outputs with combustion test data, the validity of LiquidRocketEngine library has been checked and improved. The sample analysis results with test data are given in Fig.3. This shows the LiquidRocketEngine library does well in reproducing the unsteady responses occurring both transient and steady-state working condition.



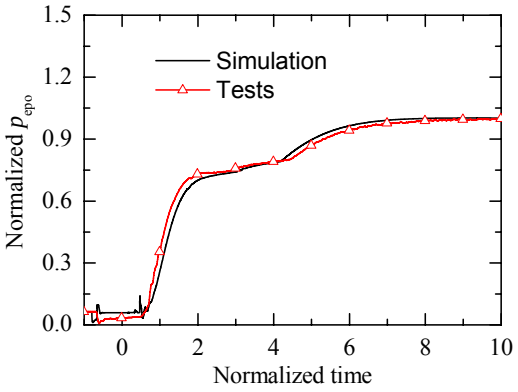
(a) turbine rotation number



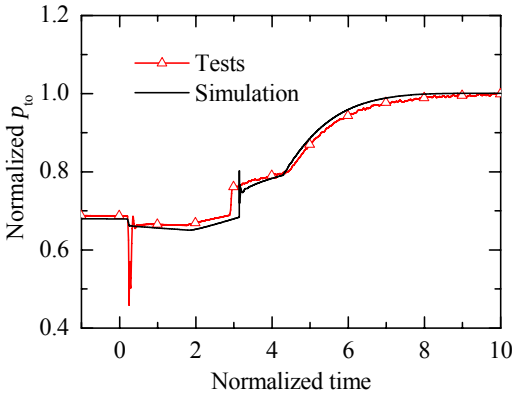
(b) gas generator pressure



(c) thrust chamber pressure



(d) outlet pressure of oxidant pump



(e) oxidant start tank pressure

Fig. 3 Simulation results compare with ground test results of an upper stage engine start-up process

VI. CONCLUSIONS

We have established a component model library for liquid propellant space propulsion system, which is able to build space propulsion system model efficiently and simulate engine transient performance. In this paper, control equations of some most interesting components are given. The general method for applying the characteristics of Modelica, especially object-orientation and connection mechanism, to modeling procedure of space propulsion system is presented. An upper stage engine model is build using component model from established LiquidRocketEngine library. The transients during engine start-up are simulated and analysed. Due to extreme working conditions and uncertainty, the transient process are very complex. Our LiquidRocketEngine library

provides an efficient tool to study transient properties. In the future, we will validate the system model with existing experimental results to improve accuracy.

REFERENCES

[1] N. Yamanishi, T. Kimura, M. Takahashi, K. Okita, and H. Negishi, “Transient Analysis of the LE-7A Rocket Engine Using the Rocket Engine Dynamic Simulator (REDS).” AIAA Paper 2004-3850, July 2004.

[2] J. R. Mason, and R. D. Southwick, “Large Liquid Rocket Engine Transient Performance Simulation System.” NASA-CR-184099, prepared for NASA-Marshall Space Flight Center.

[3] M. P. Binder, “An RL10A-3-3A Rocket Engine Model Using the Rocket Engine Transient Simulator (ROCETS) Software.” AIAA Paper 93-2357, June 1993.

[4] M. P. Binder, “A Transient Model of the RL10A-3-3A Rocket Engine.” AIAA Paper 95-2968, July 1995.

[5] A. LeClair and A. K. Majumdar. “Computational Model of the Chilldown and Propellant Loading of the Space Shuttle External Tank.” AIAA Paper 2010-6561, July 2010.

[6] M. De Rosa, J. Steelant, and J. Moral. “ESPSS: European Space Propulsion System Simulation.” In Space Propulsion Conference, May 2008.

[7] M. Atsumi, A. Ogawara, K. Akazawa, and K. Takeishi, “Development of Visual Integrated Simulator for Rocket Engine Cycle.” In 4th International Conference on Launcher Technology, Space Launcher Liquid Propulsion, Belgium, 2002. CNES.

[8] Fritzson, Peter. Principles of object-oriented modeling and simulation with Modelica 2.1. John Wiley & Sons, 2010.

[9]Wylie, E.B., and Streeter, V.L., Fluid Transient in Systems, Engle-wood Cliffs, USA: Prentice Hall, 1993.

[10] H.Y. Chen, H.J. Liu, J.H.Chen, L.J. Wu. “Chebyshev super spectral viscosity method for water hammer analysis.” Propulsion and Power Research, 2013, 2(3):201–207.

[11] F. L. Zhou, et al. "MWorks: a modern IDE for modeling and simulation of multi-domain physical systems based on Modelica." Proceedings of the 5th International Modelica Conference. Vol. 2. 2006.

Design and Fabrication of High Transparent and Conductive Optical Thin Films with Atomic Oxygen Resistant

FAN Xiang, CHI Weiying, LEI Gang, YANG Hongdong, SU Baofa

(Department of Physical Power-sources, Shanghai Institute of Space Power-sources, Shanghai, 200245)

All solar arrays have biased surfaces which can be exposed to the space environment. It has been observed that when the array bias is less than a few hundred volts negative, then the exposed conductive surfaces may undergo arcing in the space plasma, might lead to permanent short-circuit in the array circuit and the arc current keeps flowing until thermal breakdown of insulative layer occurs[1]. Now, There are several solutions have been presented, including increased the gap between solar array series-parallel connection[2], designed double mirror mapping compensation current mesh loops[3], and coated the surface of Glass with Indium Tin Oxide (ITO) films[2] that may prevent damage from ESDs and sustained arcs. In these ways, ITO-coated Glass was effective method which mitigation radically ESD with respect to high bus voltage or plasma. But ITO-coated Glass has been applied and reported which low transmittance in visible light range, poor Atomic Oxygen (AO) resistant capability and difficult metal interconnectors encapsulation, etc [4].

To improve the electrostatic resistant and AO resistant capability and optical property, SiO_2 single layer film was designed on the Glass with ITO films. The left area of Fig. 1 was shown the structure of Glass. The thickness and optical properties of Glass/ITO/SiO_2 multi-layer films were simulated by software of TFcale. The transmittance of Glass with optimized multi-layer films was greater than 92.5% from 400 nm to 1700 nm. The designed thickness of ITO and SiO_2 layer is about 12nm, 88nm, respectively. Figure 2 is the simulated transmitted spectrum of Glass with optimized multi-layer films.

Based on the designed multi-layer films, ITO and SiO_2 films were prepared by electron beam evaporation on the ordinary glass slide, with high purity (99.99%) of the ITO (wt.90% In_2O_3+ wt.10% SnO_2) and SiO_2 ceramic grains, while coating thickness and deposition rate were monitored and demonstrated by quartz crystal oscillation. Different processing conditions on the optical and electrical properties are discussed in detail. Properties of thin films were characterized by UV-visible spectrophotometer and four-point probe, respectively. Optimal conditions of ITO thin film obtained by experiment is the substrate temperature 300~400℃, the oxygen pressure 2.0×10^{-2}Pa~2.5×10^{-2}Pa, Under the optimal conditions the result is the transmittance 89.8% from 400 nm to 1200 nm, surface resistance is about $10^5\Omega/\square$. Optimal conditions of SiO_2 thin film obtained by experiment is the substrate temperature 100~150℃, working pressure $\leq3.0\times10^{-3}$Pa, average transmittance ≥90.5% from 400nm to 1800nm. Ground on above optimal conditions. A new Glass (Glass/ITO/SiO_2, the structure is shown in Fig. 1) has been fabricated by EB. The tested results reveal that the new Glass has the high transmittance (≥90.0%) from 400 nm to 1700 nm and more better Atomic Oxygen(AO) resistant capability than others Glass. Fig. 3 is the testing transmitted spectrum of Glass with multi-layer films.

Key words: High voltage solar array; ITO/SiO_2 films; ESD; electron beam evaporation;

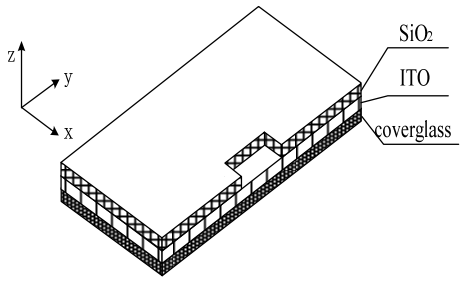


Figure 1 the structure of cover-glass with multi-layer films

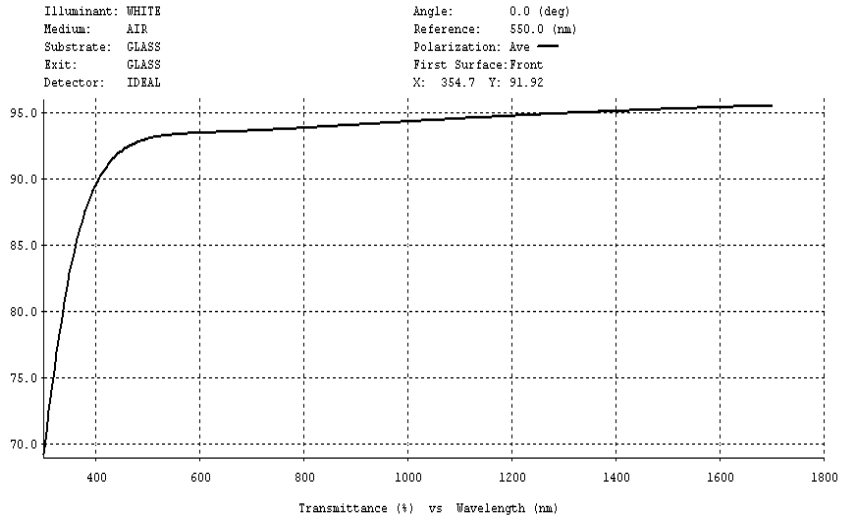


Figure 2 Transmission of Glass/ITO/SiO_2 were simulated by software of TFcale

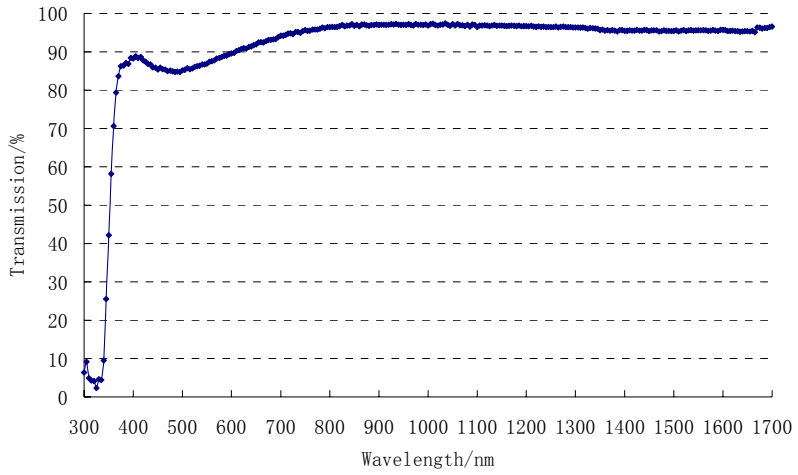


Figure 3 The transmission of Glass with multi-layer films was measured by UV-IR spectroscope measurements

Acknowledgements

This work was partially supported by the Natural Science Foundation of Shanghai (No.13ZR1457800).

References

[1] Cho, M., Kitamura, T., Ose, T., Masui, H., and Toyoda, K.: Statistical Number of Primary Discharges Required for Solar Array Secondary Arc Tests [J] Spacecraft and Rockets, 2009(46): 438-448.

[2] Teppei, O., Jeongho, K., Mengu, C., and Kazuhiro, T.: Development Of High Voltage Solar Array For Large Space Platforms[C]// 5th International Astronautical Congress, 2004: 612~613.

[3] Li, G. X, The technology introduction of spacecraft electrical source system [M] China Aerospace Press, Beijing, 2008: 269~412.

[4] Granqvist, C.G., Hultaker, A.: Transparent and conducting ITO films: new developments and applications [J] Thin Solid Films, 2002(411): 1~5.

Thermal Analysis and Thermal Design Research of Space Power Control Unit Based on The Finite Element Method

Jiangyue，Tubo

Abstract: In this paper, the basic concepts of the thermal design and thermal analysis methods were introduced, and the impact of space environment on high-power electronic product was analyzed. The paper introduced the major methods of the thermal design of a certain type space power control unit, including the structural design of the product and cooling measures of electronic components. At the end of this paper, the results of thermal design were verified by using NASTRAN, a finite element method analysis software, to determine whether the thermal design of the power control unit would meet the high-reliability, long-life design requirements.
Keywords: Finite Element Method, Power Control Unit, Thermal Design

Introduction

The power subsystem plays an important role in the system of a spacecraft. Other than the previous lower-power short-term manned spacecraft and the small and medium power space laboratory, this certain type of spacecraft is provided with the new technical requirements such as high-power, long-life, maintainable and high-reliability. Since the space environment is to the disadvantage of heat dissipation and the space capacity is limited, the heat emission condition is harsh for the high-power electronic equipment. Moreover the electronic components and PCBs are working in a constant state for years. Thus the reasonable thermal design is an important way to ensure that all the components' junction temperature or case temperature can meet the requirements and to enhance the reliability of product.

The calculation of traditional thermal analysis methods is complicated, time-consuming and needs to go through repeated test trials. However, rational use of thermal analysis software can shorten the development cycle and can accurately reflect the actual temperature of components and each points within the product. The paper mainly introduced the thermal analysis and thermal design of the space high-power power control unit based on the finite element analysis software-MSC NASTRAN.

1 The methods of thermal design and thermal analysis

1.1 Thermal Design

Thermal design is a kind of reliability design which takes the temperature influence on electronic products into consideration, to ensure that products work properly in accordance with the requirements under specified environmental conditions. With obtaining and taking full advantage of various thermal parameters, the aim of thermal design is to obtain high reliability electronic products with less cooling cost.

1.1.1 the aim of thermal design

With the development of electronic technology and the design ability of aerospace electronic devices, the volume of aerospace electronic equipments and components is increasing miniaturization, which makes volumetric power density of electronic devices has greatly increased.

In particular, in the space environment, the reduction of air convection limits heat transfer paths. Hence, the reasonable thermal design is necessary to the electronic devices in space. By configuring the appropriate cooling system which can provide a low thermal resistance path between the heat source and the heat sink, the heat produced by the electronic devices can be transferred out smoothly and the internal components will not exceed the maximum allowable temperature.

1.1.2 basic requirements of thermal design

The basic requirements for the thermal design of electronic devices, including:

a. Thermal design should meet the reliability requirements of electronic equipments. High temperature will have a serious impact on most electronic components, which will lead to the failure of electronic components, or even the failure of electronic devices;

b. Thermal design should meet the requirements of the expected thermal environment of the devices, which includes the hot ambient temperature and pressure limits, the rate of change in ambient temperature and pressure, radiant heat load around the sun or other objects and the valuable heat sink.

1.2 methods of thermal analysis

Thermal analysis methods include analytical method, experimentation and simulation.

(a) The analytical method is based on the mathematical to solve the problem and draw the solution expressed by function. This function shows a continuous temperature distribution within a certain location. This method can solve the problem accurately and reliable, however, the range of problem that analytical method can solve is limited, for it can not to carry out complex thermal design such as the power control unit in this text.

(b) experimental analysis method obtains the temperature of the electronic devices by direct measurement with the thermal infrared imager. It can be carried out by thermal test of the product. The shortcoming of experimental analysis method is that the obtained temperatures are commonly the surface temperature of the product and few components, cause that it is difficult to measure the internal junction temperature of the components. Experimental analysis method is usually carried out after completion of the whole electronic device design, resulting a long cycle overall thermal design process and an iterative design process.

(c) numerical simulation is a numerical method for solving the heat conduction equations based on discrete mathematics. The heat conduction equations describe the temperature inside and outside heat field of electronic devices, heat flow paths and thermal resistance demand and a series of partial differential equations. By using three-dimensional numerical thermal analysis method, complex models can be solved and the internal temperature of any point can be obtained. This method has advantages such as a short analytical cycle, a visualized and accurate simulation results. With the continuous development of computer technology, it is wiser to choose a finite element software to simulate and analyze the heat field of a electronic product.

2 thermal design of power control unit

2.1 analysis of heat transfer path

There are three basic methods of heat transfer, namely heat conduction, convection and radiation. Due to the special environment of micro-gravity, the heat transfer path is limited to conduction and radiation. Conductive heat transfer of cooling electronic devices is commonly used in space, which means the surface between electronic devices and thermal control system must be

high-pressure and smooth, so that the temperature rise of each junction surface is the smallest and the heat produced by electronic devices will be transferred easily.

2.2 method of thermal design

Two methods of thermal design are commonly used when designing electronic devices, one is to select electronic components and materials which have high heat resistance. Since this method can greatly increase the design costs of electronic devices and makes designing electronic devices limited by choosing components, thermal design method described in this text is the heat controlling method. By using rational structural design, optimization of electronic components placement and the use of proper installation techniques, the temperature of electronic components can meet the high reliability design specifications.

2.2.1 structural design of power control unit

During the installation process of the certain type of space power control unit, it can be seen that the heat power control unit produced mainly be transferred to temperature regulator of thermal control system by way of conduction, as shown in Figure 1. To maximize the heat conduction, the flatness of power control unit's base plate must strictly meet the requirements, which is not less than0.1mm/100mm×100mm. In addition, to reduce the thermal contact resistance between the product and the temperature regulator, thermally conductive silicone must be coated between the contact surface.

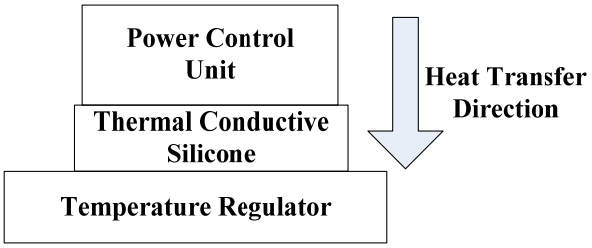


Figure 1 Installation Method

The main compositions of power control unit are PCBs, electronic components installed in case, cables, chassis and other parts of the product. The structure of the product uses a modular design, which means greater-heat-generating modules can be placed interval according to the product actual operating features, so that the interior heat radiation and conduction can be minimized. Connections between each module uses the frame projections and recesses reinforcement bite stitching design is conductive to heat transfer within the modules.

2.2.2 PCB thermal design

The PCB thermal design method of power control unit includes the following three aspects:

(a) The PCB layout design should be uniformly distributed to ensure a balanced distribution of heat dissipation, thus avoiding local overheating situation.

(b) For the components mounted on PCBs, the main heat transfer path is through components' leads and the contact between components and printed circuit board. By shortening the length of components' pins, the thermal resistance can be reduced and the heat conductive effect can be enhanced.

(c) In PCB deisgns, PCB copper area can be expanded to increase the heat radiation area of the printed circuit board in order to speed up PCB cooling rate. Figure 2 shows a schematic diagram of the internal printed circuit board of the power supply control unit , by increasing the copper area on the PCB to accelerate heat transfer of PCB.

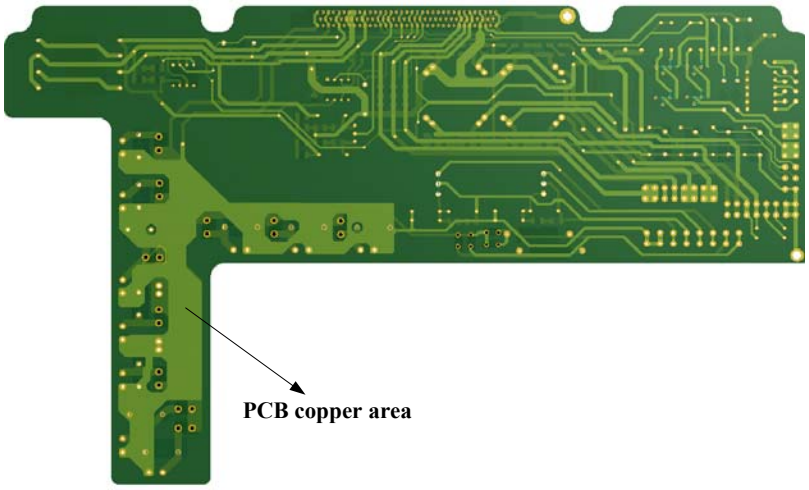


Figure 2 Schematic diagram of PCB copper area

2.2.3 electronic components thermal design

Thermal design of electronic components mainly through the layout and installation of components. At the beginning of thermal design of components, we should deal with the components' heat consumption statistics in order to master the device's heat source distribution. After that, we can decide the installation of various components according to their heat consumption. It shows several kinds of installation of components which have lager heat consumption.

(a) inductive components

the installation of inductive components as shown in Figure 3, the heat conductive components produced can be passed out through the bottom plate of the device, thereby transferred to the temperature controlling unit of the thermal control system. Meanwhile, the pins of inductive components are connected with other electronic components in the device by wires in which way the heat can be transferred to other components.

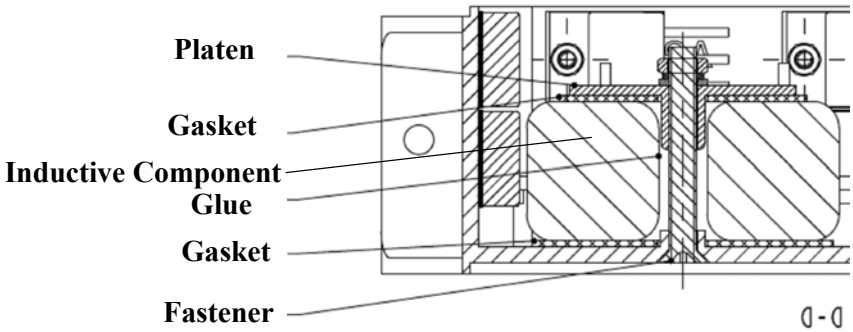


Figure 3 Schematic diagram of mounting method for Inductive Component

(b) power components

due to the large output power of the power control unit, the power consumption generated by the internal power components cannot be ignored. Power components such as MOSFETs, diodes and so on belong to the class of temperature-sensitive electronic components, therefore thermal failure is the main failure mode of power components. The installation of this kind components uses

screws and nuts to fixed to the device's bottom plate, as shown in Figure 4. The contact surface between the power components and cooling pad structure has thermal insulation pads, furthermore the thermal insulation pads between the bottom plate and power components are coated with thermal grease to conduct heat.

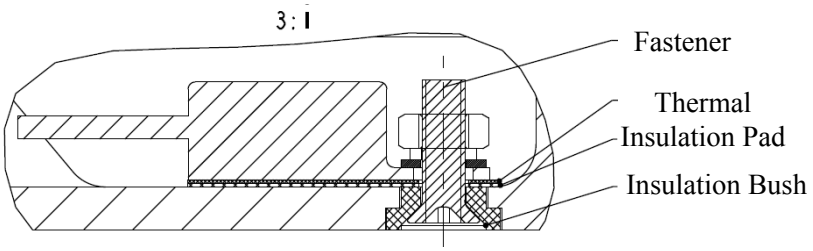


Figure 4 Schematic diagram of mounting method for Power Component

2.2.4 thermal design of structure

Thermal design of the power control unit's structure includes the following methods:

(a) We should try to increase the area of the power control unit's bottom plate without increasing the weight and influence on the mechanical properties of the device in order to increase the cooling area between the device and the cooling plate.

(b) The contact surface between the device and the temperature control unit of thermal control system is coated with thermal grease in order to reduce the thermal resistance and increase the cooling capacity.

(c) The outer surface of the device is black anodized to increase the surface emissivity to not less than 0.85, aiming at increasing the ability of the device to radiate the heat to outer space.

(d) verification of thermal design

Section 1.2 shows that using finite element analysis method is the optimal method at present for electronic devices' thermal analysis and thermal design verification. In this passage, we use MSC NASTRAN finite element software for power control unit's thermal analysis, calculation and verification, to judge that the internal components' junction temperature or shell temperature can meet the requirements.

3.1 simplification of thermal simulation model

The finite element analysis model is indicated by discrete grids or nodes, which are connected by various types of structural elements together. When calculating and analyzing complex model, the composition grid is higher and the computation time to obtain the results is longer. In order to facilitate analysis and calculation, taking the choice of heat transfer into account, appropriate simplification is used in the establishment of thermal simulation model of the power control unit, as follows:

(a) Deleting or simplifying the unnecessary structural components, such as screws, connectors and gaskets, etc.

(b) In the establishment of thermal simulation model, the heat produced by wires, free convection and forced convection are not taken into account.

(c) In the first place, the heat radiation is not considered in the simulation and calculation. If the maximum gap between the simulation temperature and the ambient temperature is 20℃, then consider radiation area based on the device's actual work environment to recalculate.

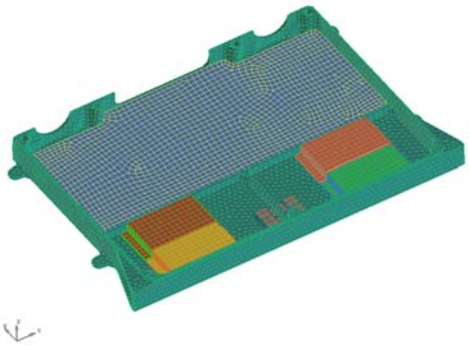
(d) The PCBs are modeling by uniform distribution of the heat source.

3.2 description of thermal simulation model

In the MSC NASTRAN , the finite element model of power control unit is shown in Figure 5. The device's shell is modeled by low-order tetrahedral elements(CTETRA4), and the regular components such as PCBs and electronic components are modeled by low-order hexahedral elements(CHEXA8) or pentahedron units(CPENTA). Overall, the number of nodes is 413869 and the number of units is 503042.

Since the heat consumption of electronic components are the main heat source of the power control unit, the electronic components' finite element model use volumetric power model to simulate and calculate.

(a) Finite element model of power control unit



(b) Finite element model of internal module

Figure5 Finite element models for thermal simulation

3.3 simulation parameter value

(a)boundary temperature settings: the setting of finite element mathematics environment condition is the temperature which is the device's actual working temperature and the maximum temperature of environmental tests, namely 70℃. Ambient temperature entered directly in the solution setup. As the boundary conditions of the cooling plate , 70℃ is the junction temperature constraint and applied directly to the node of the device in contact with the cooling plate, shown in Figure 6.

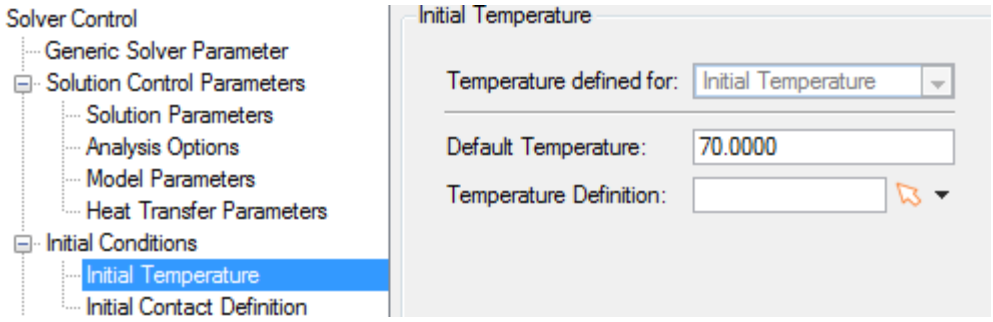


Figure6 settings of junction temperature constraint

(b)contact thermal resistance settings

in NASTRAN, we use the contact function to implement thermal contact. The heat conduction between different components is carried out through the actual geometry contact region. For each pair of contact simulation, it is divided into two steps: (a) need to specify which parts may occurs contact; (b) because NASTRAN detect heat in accordance with the spatial position, it is necessary to specify the appropriate distance tolerance to make the heat contact region is the actual geometry contact region. Contact state(that is, whether the contact occurs and where the contact occurs) can be seen by using cloud picture to verify the results.

The heat transfer capability in thermal contact area of different components is defined by the heat transfer coefficient(i.e. equivalent to thermal resistance). Contact transfer coefficient between PCBs and structure is defined 1000W/m^2 ℃ with reference to the parameters in “Spacecraft Electronic Device Reliability Design”. Furthermore, thermal resistance of power electronic components and mounting surface is selected 0.3℃/W according to the measurement by craft technology.

（c）material parameters settings: PCB’s heat production is modeled based on different copper layer area. The relevant parameters of all the PCBs in the power control unit are shown in Table 1, wherein the copper layer thickness of each PCB is 70μm and the thickness is 2mm.

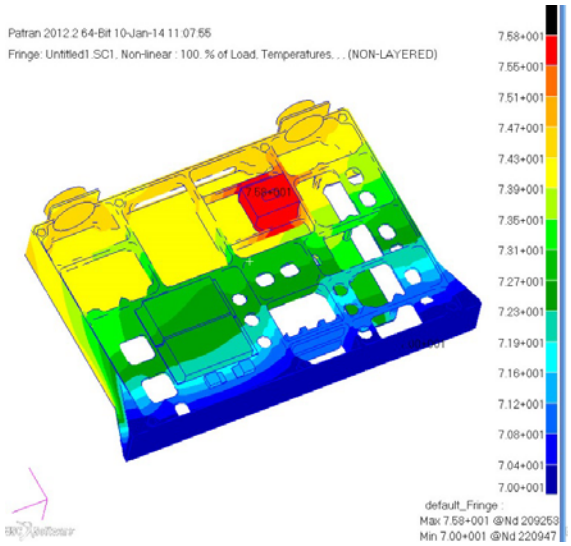
Table1 PCB parameters of power control unit

PCB Name	Copper Layer Thickness	Copper Layer Area
PCB1	70μm	L1 12%,L2 15%
PCB2	70μm	L1 20%,L2 27%
PCB3	70μm	L1 8%,L2 6%,L3 4%,L4 9%
PCB4	70μm	L1 17%,L2 16%
PCB5	70μm	L1 9%,L2 88%,L3 7%,L4 10%
PCB6	70μm	L1 14%,L2 10%
PCB7	70μm	L1 18%,L2 6%,L3 94%,L4 15%
PCB8	70μm	L1 11%,L2 13%
PCB9	70μm	L1 10%,L2 8%
PCB10	70μm	L1 50%,L2 50%,L310%,L4 50%,L5 50%,L6 40%
PCB11	70μm	L1 4%,L2 7%,L3 7%,L4 9%,L5 6%,L6 13%
PCB12	70μm	L1 4%,L2 7%,L3 7%,L4 9%,L5 6%,L6 13%

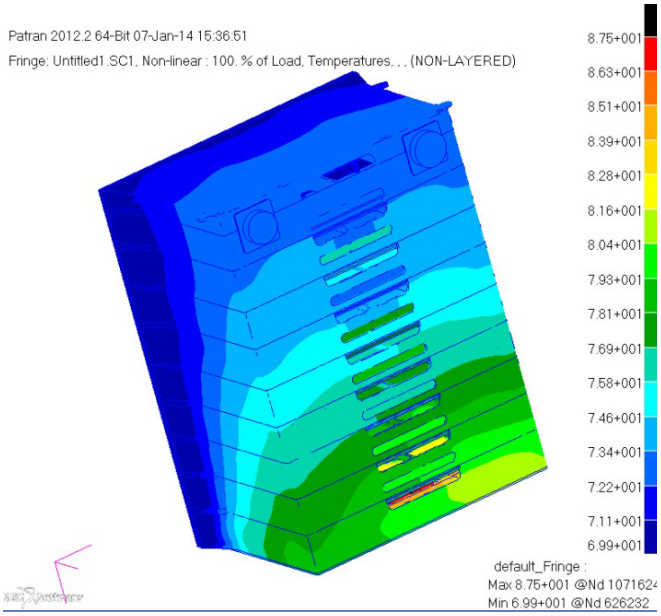
3.4 analysis of simulation results

According to boundary conditions set by NASTRAN, the finite element thermal simulation results are shown in Figure 7 and the temperatures of each PCB and main electronic components are

shown in Table 3. The temperature rises of main power electronic components and PCBs are at around 10℃ in accordance with thermal simulation results, which can meet the space application requirements.



（a）Temperature field cloud picture of power control unit



（b）Temperature field cloud picture of internal module

Figure 7 Temperature field cloud pictures

4 conclusion

In this paper, thermal analysis and thermal simulation based on the finite element software NASTRAN were carried out for a certain type of space-based power control unit. We can draw the conclusion that, by reasonable and effective thermal design, including structure design, various installation methods and different kinds of cooling measures, the internal electronic components' case temperatures or junction temperatures can meet the space application requirements.

The actual temperature distribution can be obtained in the follow-up thermal tests, which can verify the results of thermal design and increase the reliability of the power control unit.

A Bandwidth Efficient Phase Demodulator for On-Board Satellite Receivers

Bilal Hassan, Humayun Zahid, Zia ul Haq, Atif Noushad

Pakistan Space and Upper Atmosphere Research Commission, SUPRACO, Lahore, Pakistan

Abstract: The angle demodulators employed in satellite receivers for recovering the sub-carrier signal calls for optimizing two contradictory performance parameters i.e. bandwidth and SNR. In satellite communications, bandwidth is a premium resource and is used astutely; at the same time, it is imperative to achieve high SNR of the sub-carrier to ensure enhanced sensitivity of the receiver. Since smaller bandwidth adversely affects the output SNR of the demodulated sub-carrier, the design of the demodulator to achieve small bandwidth and high SNR is a challenging task. This paper presents the design of a Phase Locked Loop based Phase Demodulator which efficiently incorporates and optimizes the above mentioned critical design parameters. A phase modulated signal with a very small phase deviation of 0.5~0.6 radians (corresponding to a very small occupied bandwidth) has been successfully recovered with a very high output SNR. Moreover, a wide Lock range of phase demodulator has also been achieved using this design technique. A phase demodulator having a wide lock range is an essential requirement for counteracting Doppler's shift in satellite receivers. The designed demodulator is capable of recovering the sub-carrier even at very low power levels of the input phase modulated signal around -25 ~ -30 dBm. The paper commences with a comparison of different design strategies to solve the stated problems, and then moves on to present a customized PLL design technique to solve these. After establishing the importance of a customized PLL design technique, different optimization strategies are presented with supporting mathematical calculations. The measured results of a phase demodulator designed by incorporating the optimizations mentioned above are discussed.

Keywords: Signal-to-Noise Ratio, Bandwidth, Phase Deviation, PLL

I. Introduction

Phase Demodulators are used for data recovery in Tele-Command and Data receivers of various contemporary communication satellites. Bandwidth and SNR are the most important performance parameters of any satellite receiver since, if chosen wrongly, they can degrade the performance of a receiver unit which in turns affects the performance of the entire satellite. From the basics of communications theory, it is known that bandwidth and SNR are in opposition to each other. Bandwidth is usually kept at a minimum to improve the output SNR at the IF stage of the satellite receivers. Due to this tradeoff, the angle demodulators used in satellite receiver use a small phase deviation to keep bandwidth requirement at a minimum thus keeping noise floor at an acceptable level throughout the receiver. Standard practice involves the use of phase deviation between 0.8 ~ 1.2 radians. This phase deviation, being quite small, presents its own problems during the sub-carrier demodulation process. The SNR of the demodulated signal is directly proportional to the extent of the modulation i.e. the phase deviation. A catastrophic situation arises due to the very poor SNR and the receiver performance can be totally compromised.

To above discussion can be summarized in the form of the following statement:

“The angle demodulator should be capable of recovering the sub-carrier signal which has been modulated on a carrier frequency (Intermediate Frequency i.e. IF in case of satellite receivers) with very low phase deviation (corresponding to a small bandwidth) with an appreciable output SNR.”

The IF Carrier frequencies generally used in communication receivers are 10.7 MHz, 12.8 MHz and 21.4 MHz; but there is no restriction on the choice of IF frequencies. This paper presents the design of a demodulator at an IF frequency of 26 MHz and a sub-carrier frequency of 8 KHz. The design can be easily configured to the any new IF frequency. The discussion below compares different design methodologies available to solve the stated design problem.

II. Comparison of Design Methodologies

There are many existing solutions available in the form of commercial-of-the-shelf products. These COTS products (mainly IC based solutions) can be used to obtain satisfactory performance if certain specific conditions are met. The IC based solutions presented below can be used at the above mentioned IF frequencies:

1. NE/SE564 from Philips Semiconductors
2. LM3189 from Texas Instruments

NE/SE564 (Philips Semiconductors)

The first option, NE/SE564 is a PLL based solution designed for working upto 50 MHz of input intermediate frequency. The downside of the IC is its requirement of very large phase deviation. **NE564 is only capable of supporting minimum carrier deviation of about 1% [1].** For 26 MHz, 260 KHz frequency deviation is required otherwise the demodulated output will be very noisy and will be difficult to recover even with the help of a good low pass filter. The frequency deviation of 260 KHz corresponds to a phase deviation of around 33 radians at the mentioned IF and sub-carrier frequency. NE564 will output the same signal for the both the Frequency and Phase modulated signal keeping the frequency deviation to 260 KHz and phase deviation to 33 radians. The measured results presented in Figures 1 and 2 below using this approach illustrate this situation graphically:

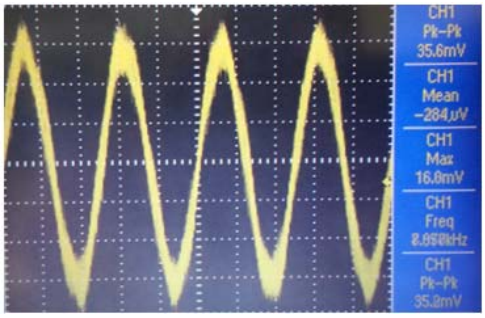


Figure 52: Frequency Demodulated Output at 260 KHz frequency deviation

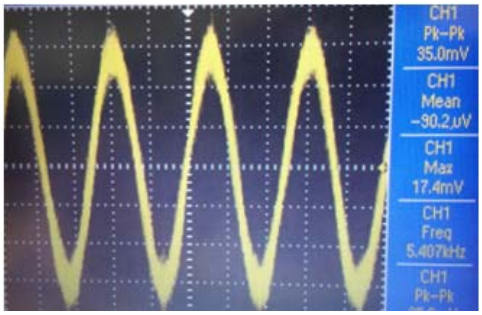


Figure 53: Phase Demodulated Output at 33 radians phase deviation

In above figures, sub-carrier of 8 KHz has been recovered from an IF of 26 MHz. As it is evident in the figures, the output SNR is very low despite large frequency/phase deviation. The Table 7 below summarizes the measured results:

Table 7: Measured Results of NE/SE564

Parameter	**Frequency Demodulation**	**Phase Demodulation**
Output Voltage	35 mV_{pp}	30 mV_{pp}
Frequency Deviation	260 KHz	33 radians
Sub-Carrier Frequency	8 KHz	8 KHz

LM3189(Texas Instruments)

The second option, LM3189 is a FM/IF system designed for a carrier frequency of 10.7 MHz. This system is rigid in terms of its frequency i.e. it only demodulates the sub-carrier riding on 10.7 MHz. To use this IC as a phase demodulator, an integrator can be used at the sub-carrier FM output of the IC. The datasheet of the IC reveals that it also requires a very large phase deviation. The output voltage of about 500 mV will be obtained when a 400 Hz tone is modulated on 10.7 MHz IF with a frequency deviation of ±75 KHz [2]. Translating this frequency deviation to phase deviation keeping the frequencies same represents a very large phase deviation.

There are other solutions available in the form of IC's such as CD4046 and LM565 but these are not designed for handling large IF frequencies mentioned in Section I. CD4046 can work up to 2.7 MHz [3] while LM565 has an upper frequency limit of 500 KHz [4].

In the light of above discussion, this paper presents a customized design of highly sensitive and bandwidth efficient (requiring low phase deviation) PLL based Phase Demodulator design which is capable of working under a very small phase deviation while at the same time giving a very good SNR performance at the sub-carrier level.

III. Design Of A Customized Pll For Phase Demodulation

The problems encountered in the previous COTS based solutions are related to the sensitivity of the Voltage Controlled Oscillator (VCO) – an integral part of the PLL. Thus selecting an appropriate VCO (as discussed in section V) plays an important role in mitigating the above mentioned problems. Thus, customization of the PLL becomes the need of the hour for making demodulators operating at the IF frequencies of contemporary satellite receivers.

Frequency synthesizer or Local Oscillators can be used to design a customized phase locked loop. This customized phase locked loop can be used as a Phase/Frequency demodulator, a local oscillator of any desired frequency or a carrier tracker.

The basic building blocks used in the design of a PLL using frequency synthesizers (shown in Figure 54 below) are:

Analog Section:

- Frequency Synthesizer Integrated Circuit
- Voltage Controlled Oscillator
- Modulated Signal
- Loop Filter

Digital Section:

- Microcontroller (for programming the registers of frequency synthesizer)

The details of the individual blocks of the analog and digital section are presented below.

The block diagram of a complete customized PLL loop is given below in Figure 54:

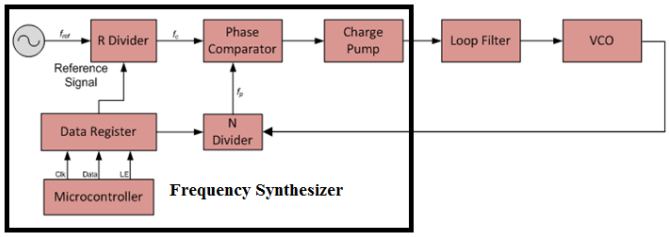


Figure 54: Basic Block Diagram of PLL

Analog Section

Frequency Synthesizer IC:

A monolithic Frequency Synthesizer has been used in the design of the demodulator. The device has been selected after critically scrutinizing different synthesizer chips available in the market. The selection criteria of the synthesizer IC will be discussed in section IV. More information can be obtained from [5].

Frequency Synthesizer consists of R Divider (reference signal divider), N divider (VCO feedback signal divider), Phase Frequency Detector (PFD) and a Charge Pump.

The value of R Divider has been set to 1 and that of N divider has been set to

= 220

For the purpose of demodulation, the frequency range and the charge pump gain of the phase frequency detector of the IC are important in optimizing the design. This aspect will be discussed later in the paper under design optimization.

Voltage Controlled Oscillator:

VCO has an important role in locking the incoming signal by generating the same frequency as the incoming signal's frequency. C band VCO has been used in the present design.

The choice of VCO is also very critical for ensuring high sensitivity. For the purpose of demodulation, frequency of operation (not to be confused with frequency range) and tuning sensitivity of the VCO are important in optimizing the design. This aspect will be discussed later in the paper under design optimization.

The selection criteria of the VCO will be discussed in section IV.

Reference Signal / Modulated Input:

Phase locked loop uses the reference signal (being very stable)

for locking onto the incoming VCO signal. The choice of this reference depends upon the application for which synthesizer is being used. For making Phase Demodulators, modulating signal is given to this reference.

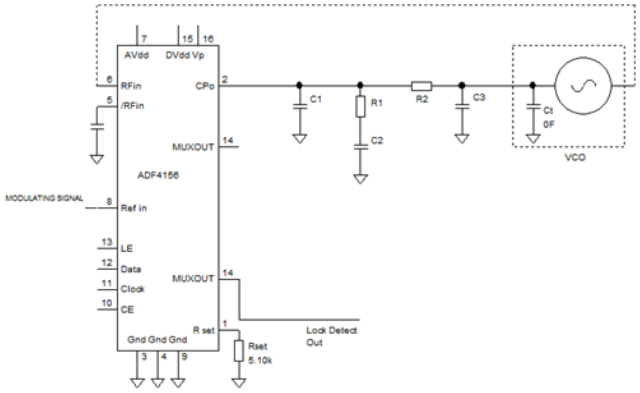


Figure 5: Complete Design Schematic

Loop Filter:

A Loop Filter follows the Phase Comparator, which is of passive type RC Low Pass 3rd order filter. The Loop Filter tracks the VCO, converts the charge pump output into a DC controlled voltage to be given to VCO, which in turn produces an appropriate microwave signal at its output. In case of FM demodulator, the output is taken from the output of the loop filter whereas the input to the loop filter is the phase demodulated output.

Loop filter design is very critical in PLL applications. It determines the overall dynamic of the PLL feedback loop. Apart from providing a ripple free dc voltage at the input of the loop filter, it also determines the lock range, settling time, phase margin etc. To satisfy all these requirements, a 3rd order loop filter with a band width of 10 KHz was used. The schematic of the loop filter is shown in Figure 55 below [6]:

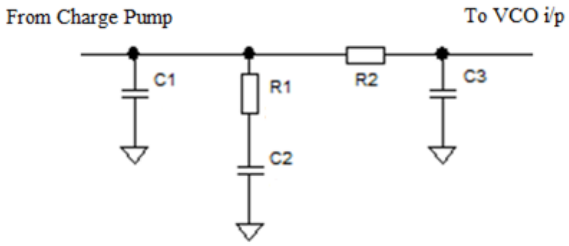


Figure 55: 3rd Order Loop Filter

Digital Section

Microcontroller:

Microcontroller is used with Frequency Synthesizers to program its internal registers. There are five registers common to Frequency Synthesizers namely:

- CLK DIV Register
- Function Register
- MOD/R Register
- Phase Register
- FRAC/N Register

These registers need to be programmed in the same order as given above.

The design schematic [7][8] consisting of a frequency synthesizer IC, loop filter, VCO and Microcontroller is shown in Figure 5 below:

The basic design of a phase demodulator has been presented in section III. It is important to note that there are many key parameters in the whole design process which dictates the demodulated sub-carrier output. In other words, there are some optimization steps which when applied to the basic design, optimize the requirement of phase deviation and output sub-carrier SNR. The next section of the paper will discuss in detail these optimization steps.

IV. Optimization Strategies

As Phase Demodulators are responsible for recovery of the baseband data, therefore, they are at the very heart of any satellite receiver (i.e. both data and tele-command receiver). Sensitivity of phase demodulator is very important. Higher the sensitivity, higher will be the output voltage for a small change in the input frequency i.e. even low phase deviations will suffice. Phase demodulators having poor sensitivity requires very larger occupied bandwidth or a very large phase deviation which in turn degrades the overall SNR of the receiver. The sensitivity can be improved by properly selecting frequency synthesizer IC, VCO, phase detector/reference frequency and loop filter bandwidth. Thus selection criteria of different Phase Demodulator components for achieving optimized performance will be presented in the following text.

Choice Of Frequency Synthesizer:

The only important thing worth seeing in the datasheet of a frequency synthesizer from sensitivity point of view is the gain of charge pump. Frequency synthesizers have a built-in PFD (phase frequency detector). The output of PFD is fed to the charge pump which boosts the signal from PFD. So, a higher charge pump gain is very important for making highly sensitive demodulators. The gain of the charge pump is determined by its current sourcing/sinking capability. Consider two frequency synthesizers A and B. Suppose the charge pump of A has a current sourcing/sinking capability of 10 mA and B has the same capability of 1 mA. If these two devices are used to make (say) Phase Demodulator, the one which employs A will be inherently more sensitive i.e. have higher output signal on the charge pump output (the sub-carrier in fact). So, the first step is to choose a synthesizer which has a powerful charge pump.

Choice Of Voltage Conterolled Oscillator:

It may seem surprising that the choice of VCO can affect the sensitivity of phase demodulator, but, it is in fact the case here. The following two parameters of the VCO play a decisive role in determining the sensitivity of phase demodulator:

- Tuning Sensitivity
- Frequency of Operation

The following three equations will be used in the numerical examples given below:

To find the value of N Divider:

$$\frac{\text{VCO Output frequency}}{\text{Reference Frequency}} \quad \ldots (1)$$

To find the frequency of the VCO:

$$\text{IF Frequency} \times \text{N Divider} \quad \ldots (2)$$

Change in the tune voltage of the VCO (filtered Charge Pump Voltage i.e. Phase Demodulated Output):

$$\frac{\text{Change In Output Frequency of VCO}}{\text{Tuning Sensitivity of VCO}} \quad \ldots \quad (3)$$

The mathematical formulation below shows the quantitative effect of the above parameters on the sensitivity of Phase Demodulator.

Numerical Example:

For the purpose of analysis, a tone signal has been considered at the reference input of the synthesizer. The frequency of this tone will be changed through a small frequency offset (much like the case in which modulating signal is present) and the effect on the charge pump output (in fact the phase demodulated output) will be compared.

Case 1: VCO's having different frequencies of operation and same tuning sensitivities

Considering two VCO's A11 and A12 with the following characteristic parameters:

The Table 8 below represents VCO's A11 and A12 with the characteristic parameters:

Table 8: Choice of VCO's with same Tuning Sensitivity

Parameter	VCO A11	VCO A12
Frequency Range	1000 ~ 1200 MHz	5400 ~ 5600 MHz
Tuning Sensitivity	60 MHz/V	60 MHz/V

The sensitivity of both the phase demodulators will be different since the characteristics of the two VCO's are different.

Let's assume that the PLL has locked onto the incoming 26 MHz reference signal. The charge pump produces a constant DC output voltage at its output say V_o. Further assume that at V_o, both the VCO's output their center frequencies i.e. 1100 MHz and 5500 MHz. To determine the sensitivity, the following question needs to be answered:

By how much the voltage V_o will change if a small offset of 10 KHz is introduced in the reference frequency?

Consider the VCO A11 first, and suppose that the value of R divider is 1 and that on N divider is:

From (1),

The reference frequency after 10 KHz offset now becomes 26.01 MHz. It is clear that the frequency of the feedback signal should be the same as reference to achieve the locked state again. From (2), the VCO output the frequency will be equal to:

$$26.01 \times 42.30 = 1100.423 \ MHz$$

in order to achieve the lock again. In this case, we have a frequency change of 423 KHz. From (3), we can find the net voltage change as:

$$\frac{0.423}{60} = 7 \ mV$$

Consider the VCO A12, and suppose that the value of R divider is 1 and that on N divider is:

$$\frac{5500}{26} = 211.54$$

Following the similar steps as before, the new VCO frequency will be:

$$26.01 \times 211.54 = 5502.11 \ MHz$$

in order to achieve the lock again. In this case, we have a frequency change of 2.11 MHz. From the tuning sensitivity, we can find the net voltage change as:

$$\frac{2.11}{60} = 35\ mV$$

If we have two VCO's having almost the same tuning sensitivities, the one with higher frequency of operation should be used to increase the sensitivity of the Phase Demodulator.

Case 2: VCO's having same frequencies of operation and different tuning sensitivities

The Table 9 below represents VCO's A21 and A22 with the characteristic parameters:

Table 9: Choice of VCO's with same frequency ranges

Parameter	VCO A21	VCO A22
Frequency Range	1000 ~ 1200 MHz	1000 ~ 1200 MHz
Tuning Sensitivity	30 MHz/V	60 MHz/V

Following the calculations presented above, we will calculate the net voltage change for a small frequency offset of 10 KHz in the IF Frequency for VCO's A21 and A22.

The net voltage change for VCO A21 will be:

$$\frac{0.423}{60} = 7\ mV$$

The net voltage change for VCO A22 will be:

$$\frac{0.423}{30} = 14\ mV$$

If we have two VCO's having almost the same frequency ranges, the one with higher tuning sensitivity should be used to increase the sensitivity of the Phase Demodulator.

Choice Of Reference/If Frequency:

From (3) it can be seen that higher frequency change results in larger output voltage change. Higher frequency change results from larger N divider value as evident from (2). Larger N divider value means lower IF or Reference Frequency (all this has been viewed in the context of same VCO). Let's find the output voltage for the Reference Frequencies of 10 MHz, 20 MHz and 30 MHz keeping the VCO A11 and 10 KHz offset.

For 10 MHz,

$$N\ Divider = \frac{1100}{10} = 110$$

The new VCO output frequency will be:

$$10.01 \times 110 = 1101.1\ MHz$$

Change in output voltage will be:

$$\frac{1.1}{60} - 18.33\ mV$$

For 20 MHz,

$$N\ Divider = \frac{1100}{20} = 55$$

The new VCO output frequency will be:

$$20.01 \times 55 = 1100.55\ MHz$$

Change in output voltage will be:

$$\frac{0.55}{60} = 9.16\ mV$$

For 30 MHz,

$$N\ Divider = \frac{1100}{30} = 36.67$$

The new VCO output frequency will be:

$$30.01\ x\ 36.67 = 1100.367\ MHz$$

Change in output voltage will be:

$$\frac{0.367}{60} = 6.1\ mV$$

Choice Of Loop Filter Bandwidth:

Sub-carrier frequency is one of the most important factors determining the loop filter bandwidth. The bandwidth should be greater than the sub-carrier frequency. The voltage present at the input of VCO i.e. V_{tune} is given by [6]:

$$V_{tune}(t) = V_{charge\text{-}pump}(t) * h_{loop\text{-}filter}(t)$$

Or in Laplace domain,

$$V(s) = I_{charge\text{-}pump}(s) \times Z(s)$$

Where Z(s) is the impedance of Loop Filter and I(s) is the current output of the charge pump

If the loop bandwidth is smaller than the sub-carrier frequency, the voltage at the loop filter output (i.e. VCO input) will be different from loop filter input (charge pump output). The PLL loop will not be able to track the ac signal variations correctly. The frequency range in which the VCO now operates will become smaller since the ac variations on the V_{tune} signal decreases. It can be said that the effective phase deviation decreases in this case and the output phase demodulated voltage decreases and so is the sensitivity.

V. Measured Results

A phase demodulator has been designed using the basic design presented in section III. The basic design was then optimized under the methodology presented in section V. The input parameters of the phase demodulator are given in Table 10 below:

Table 10: Input Parameters of Phase Demodulator

Parameters	Value
Carrier Frequency:	26 MHz
Sub-Carrier Frequency:	8 KHz
Phase deviation:	0.8 radians

The figure 7 below shows the designed and developed module of Phase Demodulator:

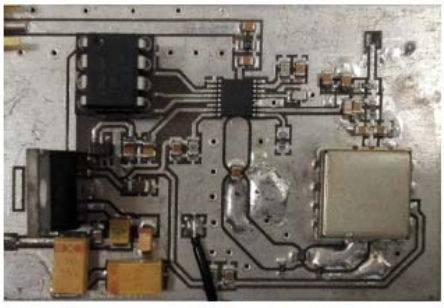


Figure 7: Designed and Developed Phase Demodulator

The measured results in the Figure 6 below shows the 8 KHz sub-carrier recovery with a high SNR.

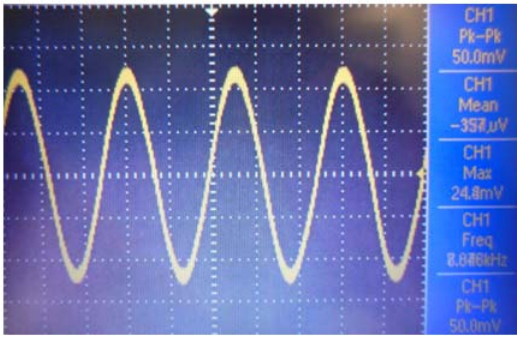


Figure 6: Phase Demodulated Output at 0.8 radians phase deviation

The table 5 below shows the measured results of the designed Phase Demodulator in tabular form:

Table 11: Measured Results of a customized PLL based Phase Demodulator

Parameter	Value
Output Voltage	50 mV_{pp}
Lock Range	1 MHz
Minimum Detectable IF Input	-25 dBm
Sub-Carrier Frequency	8 KHz

The output SNR as already mentioned is crucial to the satisfactory performance of the receiver. A very low SNR at the output of the demodulator, in general, can severely affect the proper functioning of the receiver. The designed and developed Phase Demodulator shows remarkable output SNR at both the:

- very low modulation indices
- very low input modulating signal level

The table 6 below shows the measured values of output SNR at various modulation indices keeping the modulating input level at -25 dBm.

Table 6: Output SNR measurement of Designed and Developed Phase Demodulator

Modulation Index (radians)	Output SNR (decibels)
0.1	34
0.2	39
0.3	41
0.4	44
0.5	48
0.6	51
0.7	53
0.8	55
0.9	56
1.0	57

The figure 8 below shows the spectrum of the demodulated 8 KHz signal (modulation index of 0.5 radians and modulating input level of -25 dBm). The SNR can be easily read by the two markers M1 and D2 showing the SNR of around 48 dB's.

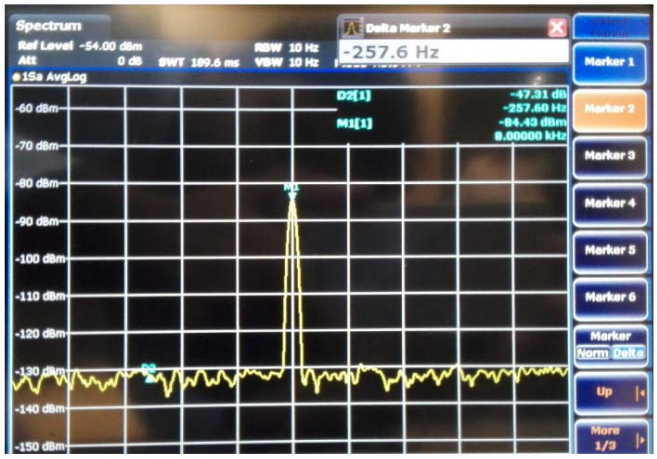


Figure 8: Spectrum of Demodulated 8 KHz Signal

The measured dynamic range of the designed phase demodulator is approximately 35 dB with the lower limit of input power at -25 dBm and the upper limit at +10 dBm (upper limit is dictated by the maximum input power of the PLL Chip).

The ability of the designed Phase Demodulator to work at very low input power levels is a highly desirable feature in the case of satellite receivers. Satellite receivers usually employ AGC (automatic Gain Control) loop to keep the input power level of the IF frequency, which is fed to the demodulators, within a certain limit (say 0 dBm with a tolerance of ±1 dB). In case the AGC loops fails and the input power level of the demodulator starts falling, the high dynamic range of the demodulator allows the receiver to compensate for a drop of approximately 35 dB of input power in the IF chain.

The wide lock range of the PLL (dictated mainly by the VCO frequency range) is also an essential parameter for satellite receivers to compensate the Doppler Frequency shift. The presented design has a measured lock range of about 1 MHz (25.5 ~ 26.5 MHz).

VI. Conclusion

The paper presents the design of a PLL based Phase Demodulator for contemporary satellite receivers. The capability of the Phase Demodulator to work at very low phase deviation is quite remarkable. The optimization steps can be applied to the basic design to increase the sensitivity of the Phase Demodulator. The numerical example strongly justifies the optimization techniques mentioned in the paper. The customization approach described in the paper can be tailored according to the evolving requirements of the modern satellite receivers.

References

[1] Philips Semiconductors, “Phase-Locked Loop”, NE/SE564 datasheet, Aug 31, 1994.

[2] National Semiconductor, “LM3189 FM IF System”, LM3189 Datasheet, April 1997.

[3] NXP Semiconductors, “HEF4046B Phase-Locked Loop”, HEF4046 Datasheet, Rev. 5-18 November 2011.

[4] Texas Instruments, “LM565/565C Phase Locked Loop”, LM565/565C Datasheet, May 1999, Revised April 2013.

[5] Analog Devices, “6.2 GHz Fractional N-Synthesizer”, ADF4156 Datasheet, Rev.E.

[6] Dean Banerjee, PLL Performance, Simulation and Design, 4th Edition, 2006.

[7] G. Kolumban, “Phase-Locked Loops,” article in the Encyclopedia of RF and Microwave Engineering, K. Chang, (Ed), vol. 4, pp. 3735-3767, Wiley, New York, 2005.

[8] G. Kolumban, “Phase-Locked Loops,” article in the Encyclopedia of RF and Microwave Engineering, J. G. Webster, (Ed), vol. 16, pp. 158-188, Wiley, New York, 1999.

Design of Annular Ring Patch Antenna for Circular Polarization Operation Using Dual Dielectric Substrate Layers

Zeenat Rajar[1] Saima Siddiqui[2]

GSC&O-K, SCF-K, SUPARCO, Pakistan

1 rajar_zeenat@yahoo.com ; 2 saimasiddiqui333@gmail.com

Abstract: With the revolution in space industry much attention is given to the methods to reduce mission cost to make access to the space affordable by introducing small satellites which offer distinct features such as light weight and low cost. In this paper, light weight annular ring patch antenna with circular polarization has been designed. Many methods are being used to achieve circular polarization such as truncated corners scheme for square patch, indentations for circular patch or by using dual feed mechanisms etc. The former two methods suffer from very low Axial Ratio (AR) bandwidth for thin substrates, whereas using dual feed mechanism for thin substrates can enhance the AR bandwidth at the expense of complex feed circuitry. In this paper, circular polarization has been achieved by attaching an L shaped strip with the inner circle of single feed annular ring patch antenna which offers good AR bandwidth for thin substrates. To improve performance characteristics further as well as to keep the weight light, two thin layers of dielectric substrates have been used. The proposed antenna operates at a resonant frequency of 2.45GHz and having a peak gain of 8.43 dBi, impedance bandwidth of 190 MHz and a good Axial Ratio (AR) bandwidth for AR ≤ 3dB has been achieved. The simulation results show good impedance matching over the desired frequency and a 3 dB beam width equal to 60^o. The proposed antenna is simple, light weight and low cost and can be used for S-band LEO satellite and small satellites applications.
Keywords: Annular ring patch, S-band, circular polarization

I. Introduction

The advancement in technology has brought compactness in every electronic component. This has led to the development of miniature products in variety of applications including satellite communications, WLAN and broadcasting networks.

The antenna used in these applications must be small, efficient and being capable of bearing up the electrical and thermal environment. For such purposes microstrip patch antennas are used as these antennas have advantages of less weight and low cost [1], but having disadvantages of low gain and less bandwidth [2]. Bandwidth and gain can be improved by using comparatively large patch antenna as there is a tradeoff between the size of patch antenna, bandwidth and gain.

Several methods have been proposed to improve gain and bandwidth of patch antennas with small size and light weight including use of impedance matching methods and increasing substrate thickness [3], stacked patches [1], using multiple dielectric substrate layers [4], L probe feeding, using substrate of low dielectric constant etc.

For space applications, antenna should be designed while considering following characteristics: Firstly, the antenna must be very reliable, being capable of supporting mechanical variations and be able to survive in harsh radiation environment [5]. Therefore, the materials for manufacturing of such antenna must be carefully chosen so as to keep antenna light weight while maintaining good conductivity and thermal control properties [6]. Secondly, as circularly polarized wave abstain less distortion in ionosphere than linearly polarized wave [7] therefore; circular polarization is preferred. Different methods have been proposed to achieve circular polarization (CP) for microstrip patch antenna that includes truncated corners for square patch, indentations for circular patch, use of two or more feeds, padding strips and use of paired slits at various angles with probe or line feeding mechanisms [2][8]. The former two methods suffer from very low Axial Ratio (AR) bandwidth for thin substrates, whereas using dual feed mechanism can enhance the AR bandwidth at the expense of complex feed circuitry.

In this paper, design of right hand circularly polarized annular ring patch antenna using L shaped strip with inner circle of annular ring with single feed has been proposed. All simulations have been carried out using High Frequency Structure Simulator (HFSS) version 13.The rest of this paper is organized in following three sections: Section II represents proposed antenna configuration. Simulation results of antenna configuration are discussed in section III. Future extensions and conclusions are presented in Section IV.

II. Proposed Antenna Configuration

The proposed antenna structure and configuration is shown in Fig. 1(a) and Fig. 1(b). Antenna consists of annular ring patch with L shaped strips, excited by probe feeding mechanism. It is designed using commonly known performance enhancement method i.e. use of multiple dielectric substrate layers. Two dielectric substrate layers with different thickness have been used. The two layers are separated by an air gap of 5mm.Rogers RT/duroid 5880 is chosen as substrate for both layers having relative permittivity of 2.2.

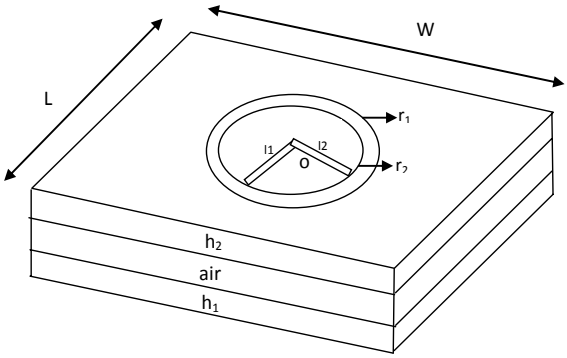


Fig. 1(a) Antenna Structure

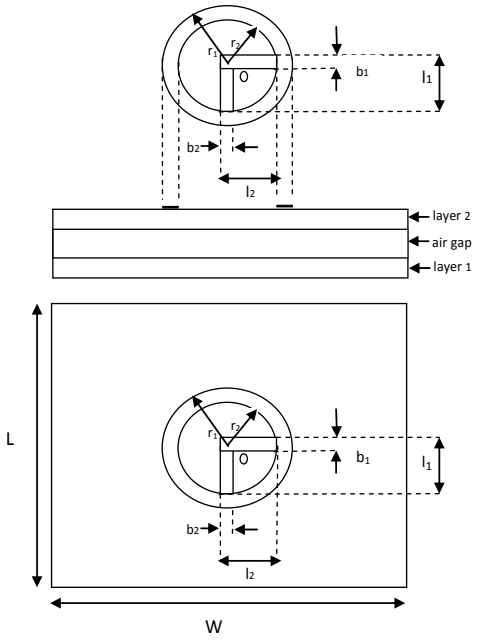


Fig. 1(b) Antenna Configuration

Fig.56. Proposed Antenna Configuration

To achieve circular polarization, L shaped strips are attached to inner circle of annular ring. The L shaped strips are formed by combining two strips placed at 90 degrees. This configuration provides good AR bandwidth for thin substrates. The dimensions of substrate and patch are listed in Table I

Table1 Antenna Dimensions

Antenna Parameters	Dimensions (mm)
Substrate(Rogers RT/duroid 5880): Layer 1 Layer 2	 $h_1 = 1.575$ $h_2 = 1$
Annular Ring: Inner Circle Outer Circle	 $r_1 = 16$ $r_2 = 22$
L Slit: Slit 1 Slit 2	 $l_1 = 16$, $w_1 = 1.3$ $l_2 = 16$, $w_2 = 1.3$
Feed Position	(x,y) = (-0.5, 0)

III. Simulation Results

All simulations have been performed in High Frequency Structure Simulator (HFSS) version 13. The proposed antenna has been optimized at 2.45 GHz having return loss of -36dB and peak gain of 8.43dBi. The simulated return loss and radiation pattern are depicted in Fig. 2 and Fig. 3 respectively. Impedance bandwidth of 190 MHz is achieved as shown in Fig. 4; the impedance has reached to 50 Ohms with almost zero reactance.

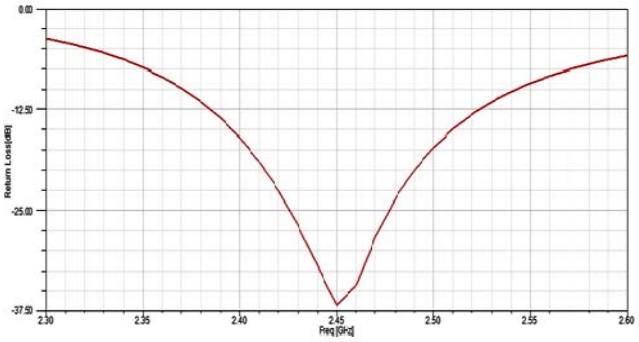


Fig. 2 Return Loss

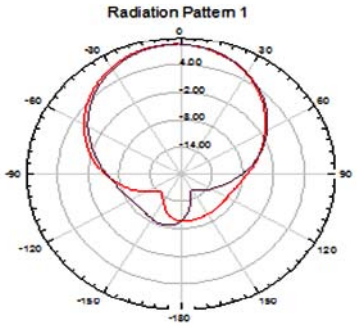


Fig.3 Radiation Pattern

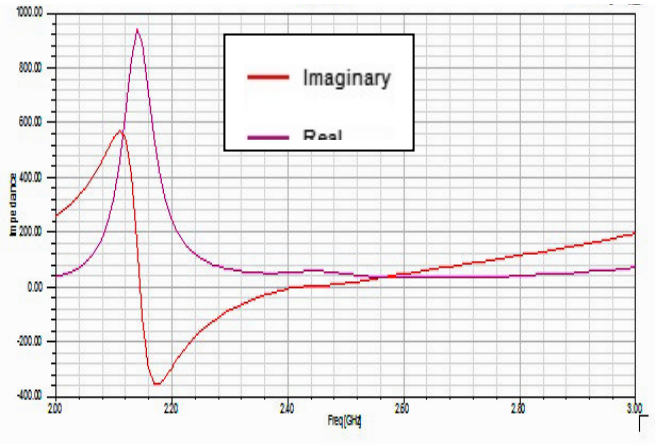


Fig. 4 Characteristic Impedance

Peak gain Vs frequency results are given in Fig. 5.The figure shows a gain of more than 6dBi in the frequency range of 2GHz to 3GHz. Right hand circular polarization has been achieved with AR bandwidth 50 MHz for AR $\leq$ 3dB as shown in Fig. 6. Fig. 7 shows 3dB beam width equal to 60^{o}.

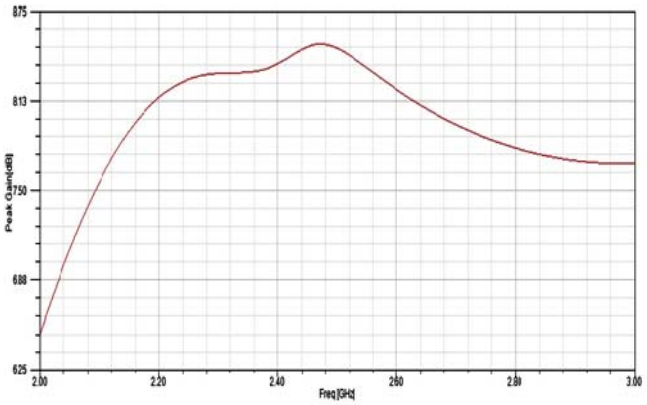


Fig. 5 Peak Gain

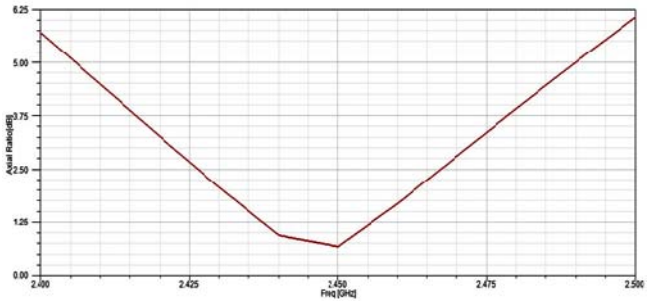


Fig. 6 Axial Ratio

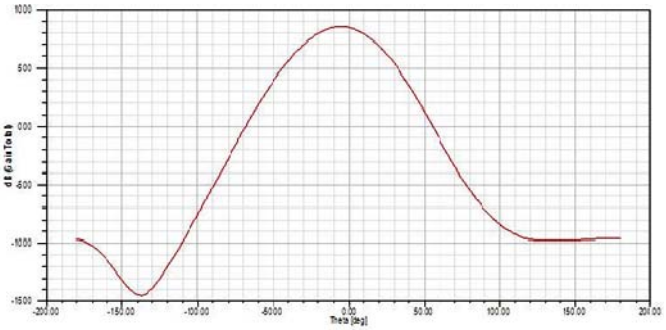


Fig. 7 3dB beam width

IV. Conclusion

Right hand circularly polarized annular ring patch antenna for S band has been designed and simulated. The proposed antenna operates at a center frequency of 2.45 GHz and has a gain of 8.43 dBi. The antenna can be used for small satellite applications or LEO satellite applications such as TTC downlink purpose. Future design work may include forming array of this antenna for wideband applications.

References

[1] S. Yang , Q. Y. Feng , and B. Dong, “Broadband stacked annular ring patch antenna for CP operation,” College of Information Science and Technology, Southwest Jiao Tong University Chengdu, China, 2009.

[2] Constantine A.Balanis, “Antenna theory: Analysis and design book”.

[3] R. S. Kushwaha, D.K. Srivatsava, J.P. Saini. “A design of H-shape slot loaded wide band microstrip patch antenna,” International Journal of Electronicsand Computer science Engineering, ISSN-2277-1956.

[4] Ji-Hyuk Kim, Hyeon Cheol Kimand K. Chun, “Performance enhancements of a microstrip antenna with multiple layer substrates," International Symposium on Signals, Systems and Electronics 2007 (ISSSE '07).

[5] L. H. Abderrahmane, and A. Brahimi, “A new design of dual band fractal antenna for LEO applications,”The Eighth International Conference on Systems and Networks Communications, ICSNC 2013.

[6] S. Gao, K. Clarke, M. Unwin, J. Zackrisson, W. Shiroma, J. Akagi, K. Maynard, P. Garner, L. Boccia, G. Amendola and M. Sweeting, “Antennas for modern small satellites,” IEEE Antennas and PropagationMagazine. Vol. 51, No. 4, Nov. 2009, pp. 40–56.

[7] J. Li, Q. Chen, Y. Zhang, C. Wu, “Design of TTC antenna for mini-satellite application,” School of Electronic and Information Northwestern Polytechnical University Xi’an, P. R. China, IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), Sep. 2011.

[8] J. S. Row,”Design of aperture-coupled annular-ring microstrip antennas for circular polarization,” IEEE Transactions on Antennas and Propagation, vol. 53, NO. 5, May 2005

Design of Multi-Bus Communication Framework Using Multi-level 1553B Bus and RS422 Multi-Bus for Digital Systems in Satellites

Yixin Geng, Xiaogang Hu, Ye Rui

Shanghai Institute of Aerospace Electronic Technology
1777 Zhongchun Rd., Minghang, Shanghai, China
gyx0576@hotmail.com

ABSTRACT: Based on observations of the growing complexity of satellites, a multi-bus framework of the bus control network was designed to meet the requirements of rapid communication among different kinds of digital systems in a satellite. The framework was architected as a hub of different levels of bus, with realization of a connection between a multi-level 1553B bus and RS422 multi-bus. Such a bus control network under the management of a Bus Control Component in a subsystem satisfies the requirement of different-speed data transfer: real-time and high-quality data transfer with 1553B bus, while point to point communication using a RS422 bus network at low costs. With implementation of a static priority-based task distribution, payload limit has been extended, a real-time response to a task guaranteed, as well as the robustness of the data transfer enhanced.

1 INTRODUCTION

Satellites, both their standalones and subsystems, have grown more complex and applications of satellites are more diversified. The communication protocols have evolved from single-bus control, not necessarily applicable, to multi-level bus control. Thus, there is an urge need for a bus control component that satisfies standalones and subsystems with different data transfer requirements.

At current, 1553B bus and RS422 bus are in majority of bus deployment in satellites, due to their stability of data transfer. MIL_STD_1553B, a digital time-division command / response multiplexing data bus[1], is widely applied to astronautic digital systems because of advantages like high-stability, qualified data transfer, good extendibility, easy maintenance and so forth. RS422 interface, that can be adopted at ease with a low cost and thus is the most widely used for serial communication in embedded systems, provides data transfer at a good maximal distance and speed, which satisfy almost all devices designed for satellites.

This research aims to design a framework of a multi-level bus communication and control architecture to guarantee long-range data transfers between satellite standalones at different rates, and to be continued with assess the applications of this framework.

2 METHODS

2.1 System Design

The framework was designed as the bus controller in a satellite, a core control unit that controls every bus at different levels of all connected subsystems. The bus controller takes charge of the systematic data transfer and processing between different layers, coordinates and manages the unit data transfer in standalones, and conducts the point-to-point communication tasks.

Based on analysis of systematic communication in satellites, data transfers can be roughly divided to three levels of speed, as well as subsystems. A block diagram that describes the system constitution is shown as Figure 1. The bus controller is consisted of three types of bus network at three levels of data transfer rates.

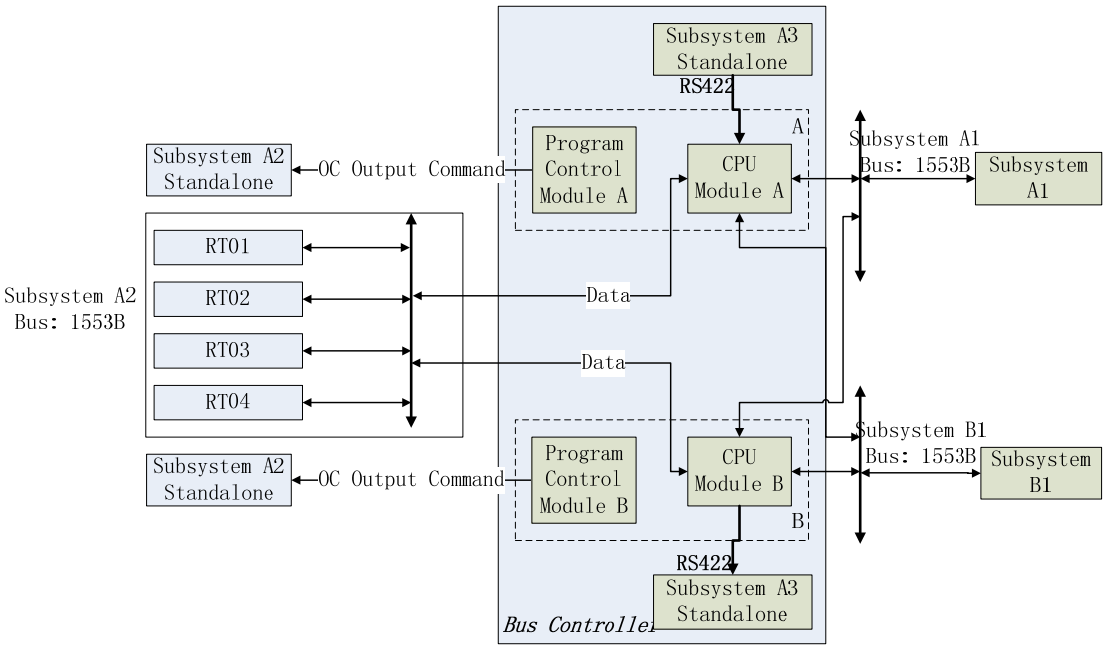


Figure 1: System Chart of a Multi-Bus Controller

Subsystem A1 and B1 are connected to the first layer of network, while subsystem A2 is to the second, referred to Figure 1. Devices in both networks require a capacity for large data flows with high stability, thus are constructed and connected under 1553B protocols. Standalones in subsystem A3, joining in at the third layer, satisfy with a relatively low-speed and small-size data transfer, thus are deployed using RS422 serial bus.

The bus controller runs at duplicated cold-redundancy. Its manager, embedded into the CPU modules, conducts all the functionality related to central control. Duplicated 1553B bus interfaces connected, as shown in Figure 1, subsystem A1, B1, and A2 at duplicated cold-redundancy together.

The system is divided to three layers: two layers of 1553B bus networks on top of a RS422 multi-bus network, in the design of the multi-bus communication framework. On the first layer, subsystem A1 and B1 works as BCs of a 1553B network, while the Bus Controller is treated as RTs; when it comes to the second layer, the Bus Controller is the BC, and the standalones in subsystem A2 are the RTs; the Bus Controller, or to be specific, its RS422 interface is connected with the Test Equipments or subsystem A3 by RS422 multi-bus on the third layer (see Figure 2).

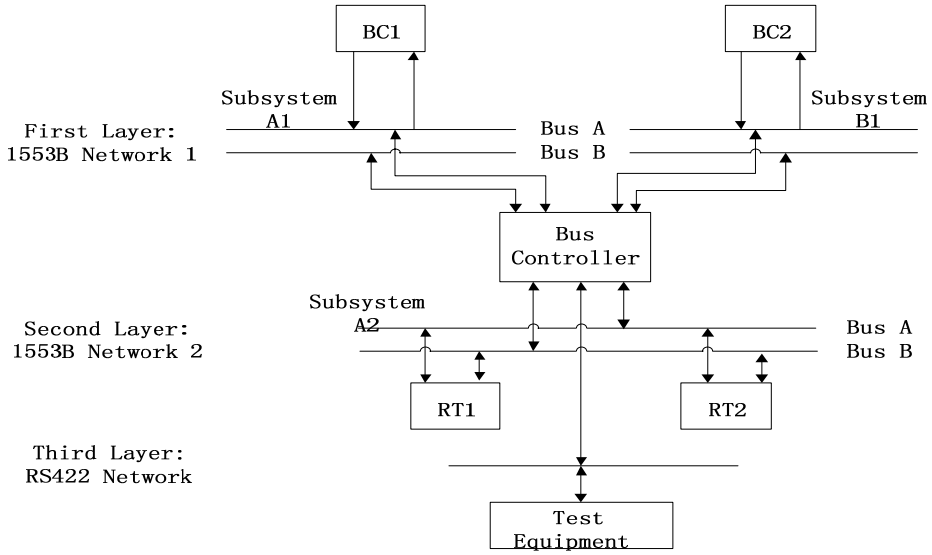


Figure 2: Topology of the Multi-Bus Control System

Therefore, the Bus Controller needs four master control chips: two are picked as RTs to correspondingly "speak" with BCs in subsystem A1 and B1; another is configured to a BC that manages and "commands" to RTs dispatched in subsystem A2; the rest one works as a RS422 interface and provides a commute to data coming and going at a relatively low speed.

Please refer to Figure 2, the priority of the bus network drops from top to bottom. The first 1553B network is allocated to data that is most urgent, dealing with it at its immediate arrival, thus plan subsystem A1 and B1 with time-sensitive devices. By parity of reasoning, RS422 network can tolerate slow responses, thus arrange time-insensitive standalones or test equipments as subsystem A3.

2.2 Software Design

To realize rational hierarchical data processing, which means to respond according to relative importance, the task dispatching strategy[3] of software was designed on the basis of graded static priority. Based on the observation of 8051 micro-control and embedded system[2], taking advantage of software and hardware interrupts, the interrupt source can be extended.

The system uses five interrupts, apparently, including the one from the internal timer and each from the four subsystems connected, whose resource scheduling are shown in Table 1.

Table 1: Resource Scheduling for Interrupts as how the CPU reads

Interrupt Source	Input Address	Circumstance
External_Intrrupt_0	0003H	Subsystem A1 Interrupts
Timer/Counter_T0	000BH	Internal Timer Interrupts
External_Intrrupt_1	0013H	Subsystem B1 Interrupts
Timer/Counter_T1	001BH	Subsystem A2 Interrupts
Serial_COM_Interrupt	0023H	RS422 (Subsystem A3 or Test Equipments) Interrupts

Subsystem A1 and B1 interrupts rank the highest thus allocate to the first layer in the bus network hardware interrupts as external events to the embedded software; in particular, connect the mono-chip's – the CPU's – external interrupt pins respectively to signal pins of chips in A1 and B1.

The framework has three levels of priority; however, the embedded software, designed for the CPU, deals with interrupts of four classes. Interrupts of subsystem A1 and B1 have the highest priority; followed those of A2; RS422 interrupts go next, those from internal timers will be responded at last. Thus, to improve the real-time instantaneity of data transfer, the software realizes a priority stack with five grades. A task scheduling algorithm is implemented to ensure functional modules to execute in a well-organized manner abiding by the priority ranking. To be specific, the software always directly performs most urgent tasks through mono-chip hardware interrupts, while leaving other tasks to the task dispatching module based on the mono-chip's internal timer.

2.3 Software Modules

The task dispatching module is an interrupt service routine with a timer (not three timers to timing the three sorts of less urgent tasks), a counter (not four counters for the tasks from four subsystems), and a FIFO queue. When an interrupt with a high priority arrives before a lower one being processed, its represented task enters the queue first; the task dispatching module, taking in account of predefined ranking, periodically queries every task – its current state – being executed and to be executed, therefore determines whether to begin execution of the head task in the queue. As shown in Figure 3, the timer decides when to restart a task query period by throws in an internal timer interrupt, which is used as a time slice for the task dispatching module.

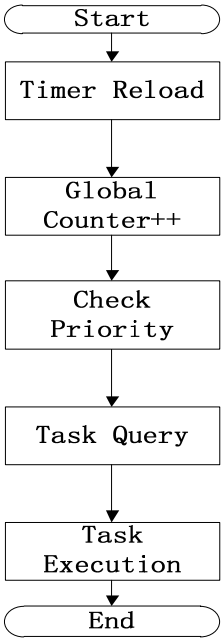


Figure 3: Process of a Task Dispatch

A flow chart of how the task dispatching module distributes resources to tasks is described as Figure 3. The module does the task execution state query one task after another from the highest priority to the lowest. The higher the priority, the shorter the time interval between two queries. In a word, prior tasks are queried at a higher frequency, which is predefined as a configuration, static, to the software and loaded at the initialization, thus their responding time is made sure. For instance, tasks from subsystem A1 is responded directly; as to tasks from subsystem B1, the module queries

every timer interrupt comes, while tasks on the second and third layers are queries every n time slices; the second layer is where n is around 5; thinking of tasks from subsystem A3 or Test Equipments, n should be greater than 10. If several tasks are ready to be executed, they are scheduled execution successively from the top priority to the bottom.

Hardware interrupts are naturally higher-ranked to software interrupts, which are shared by subsystem A2, A3, and internal timers, thus are divided into three classes. Figure 4 demonstrates how the interrupt service routine works, realizing interrupt nesting and responding to advanced software interrupts. Tasks with a low ranking join in the queue at the end, while those with a higher ranking always jump into the queue at the right position. A sorting algorithm is adopted to find the exact position.

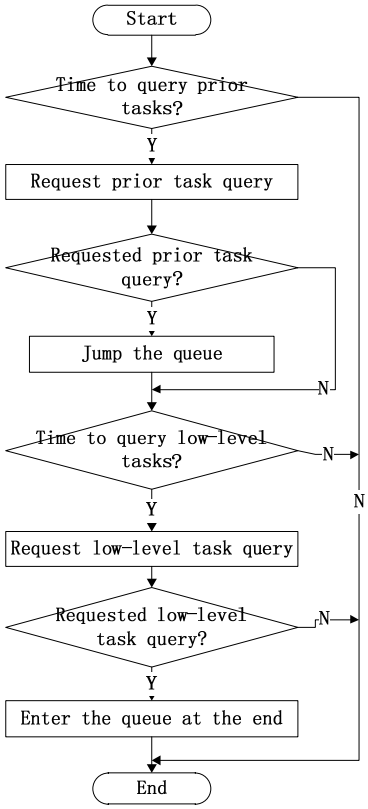


Figure 4: Process of a Bus Data Dispatch

Despite of the task dispatching module, the software have five functional modules, marked from T1 to T5, individually represented by interrupts from five sources, as described in Table2.

Table 2 Task Priority

No.	Task	Priority	Interrupt	Description
1	T1	I	Subsystem A1 Interrupts	Subsystem A1 Data Transfer
2	T2	I	Subsystem B1 Interrupts	Subsystem B1 Data Transfer
3	T3	II	Subsystem A2 Interrupts	Subsystem A2 Data Transfer
4	T4	III	RS422 Interrupts	RS422 Data Transfer
5	T5	IV	Internal Timer Interrupts	1553B Bus Periodic Check

To improve the CPU usage rate, and reduce the idle time, as long as the queue of to-be-executed tasks is empty, the task dispatching module allocates the CPU to brief and periodic tasks, for example, bus error detection at regular time. Due to that external tasks always have a higher priority to external regular tasks, resource competition is effectively lessened.

2.4 Software Protocols

2.4.1 Conflict Avoidance

To ensure that tasks from different layers do not conflict, a task is designed and programmed to a set of atom operations, thus a lower-ranking task can always be safely paused to give in the CPU and resource control to a prior task.

Utilizing that different tasks spawn at different bus communication period, the embedded software can sort irrelative tasks.

2.4.2 Bus Polling

A BC does bus polling, by broadcasting a self-check message, at a regular time to acquire its RTs' work states. The pattern mark of the self-check message is independent from the systematic data transfer, and only helps with the hardware "health" and work information management.

The bus polling starts every one second, checks either Bus A or Bus B, and in turns, see Figure 5. The system chooses Bus A to communicate at default, but switches to Bus B only when Bus B succeeded the bus polling while Bus A failed.

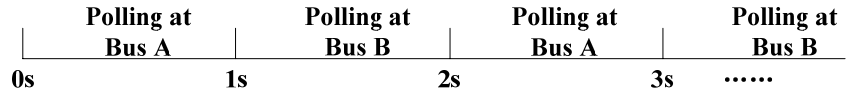


Figure 5: Bus Polling Timing Sequence

2.4.3 Handshake Protocol

To synchronize BCs and RTs, a handshake protocol is designed: at the beginning of a cycle, a RT requests a data transfer; the BC replies with a message inserted after all the data collection messages, and sends them out to RTs altogether; the next data transfer cycle does not start until the BC confirms. On the other hand, before a BC starts a data transfer task, it checks whether a RT has reset its "Data Transfer Request" bit; if reset, start the data transfer; if not, drop the task and wait a cycle.

2.5 Assessment of the System

The system was taken preliminary assessment by applying to a practical power management system. The protocols at physical and link layers of the two 1553B bus networks abide by GJB289A-97[1], while application layer is left to software in support of 1553B interface chips like Bu-61580. RS422 interface is designed as universal asynchronous serial data transfer module in accordance with 80C32, the best standard selected for CPU.

When it comes to the application, we recommend SC80C32E-30SV by ATMEL used in the assessment as the CPU. Such 8-bit processor[2] with four 8-bit I/O ports, three 16-bit registers as timers or counters, a 256-byte RAM, six channels of 2-class interrupt, and two pieces of 64KB memory space: one for programs and the other for data, is perfect suit for this framework.

As to that carries out the communication under 1553B protocol, Bu-65170 and Bu-61580 by DDC are good choices. Bu-65170 offers functionality as both RT and BM, but not of BC, thus chips of this model are planted in the Bus Controller as RTs to "talk" to subsystem A1 and B1. Bu-61580 embodied BC, RT, BM all three 1553B control modules, thus this model is equipped as BC to communicate with subsystem A2. Bu-61580 contains a micro-processor and interfaces supporting

the complete 1553B protocols[1], and can be switched among BC, RT and BM work-modes at ease. Software-logically, this model integrated dual-transceiver (receive-transmit component), CODEC (coder and decoder), protocol logic, memory manager, and interrupt controller. Besides micro-processor, Bu-61580 provides 4K-word (16-bit) internal shared static RAM and buffers between every two buses. Its software interface consists of 17 internal operated registers, 8 test registers, and 64K-word shared memory address space.

The system was assessed in loop task scheduling strategy and dispatcher task scheduling strategy. CPU usage rate, delays and other critical parameters are collected in the tests.

3 RESULTS

With applying dispatcher task scheduling strategy, this multi-bus control system effectively improves CPU usage and reduced the wait time, compared to serial task scheduling. Moreover, when a task with higher priority arrived, a complete task, can be dived into several blocks and is able o be interrupted by the control system, in case a prior task arrived. That is how the real-time performance is guaranteed.

3.1 Sample

Based on sequence scheduling strategy, T1~T4 (as described in Table 2) were executed after arrival successively, and T5 was randomly spawned. The software worked circularly. The timing sequence is shown in Figure 6.

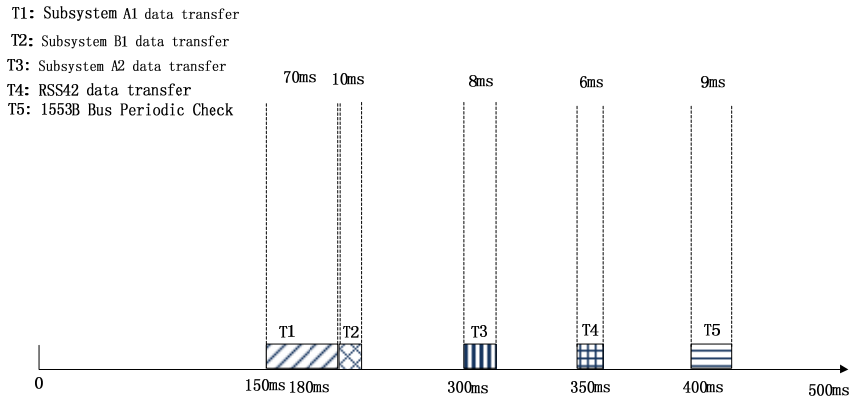


Figure 6: Timing Sequence based on Sequential Strategy

Based on static priority based dispatcher scheduling strategy, T1 was executed at immediate arrival, while T2~T4 were dispatched by the task dispatching module, and T5 was regulated at 500ms±1ms, and carried out only when CPU starts to idle. The software worked circularly. The timing sequence is shown in Figure 7.

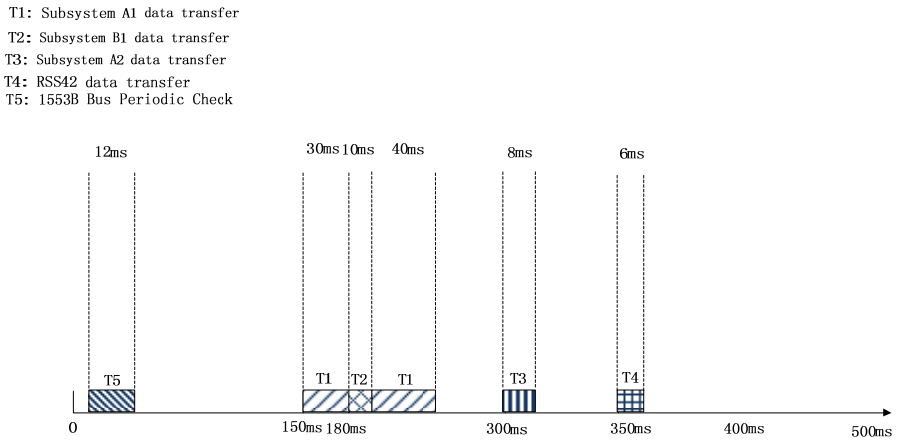


Figure 7: Timing Sequence based on Static Priority

The two strategies are measured by CPU usage and execution delay, and are compared in Table 3.

Table 3: Performance Comparison

No.	Strategy	CPU Usage (%)	Delay (ms)
1	Sequential task scheduling	25.8	70
2	Static priority based dispatcher task scheduling	30.3	10

4 CONCLUSION

The multi-bus communication framework fuses two layers of 1553B bus and a layer of RS422 multi-bus. An application of the framework, a bus controller for the Power Manager of a satellite, using 8051 micro-controller as the primary control unit, perfectly conducted data transfers encrypted with different protocols through different layers of bus networks, under the control of the software that adopted the task-scheduling algorithm based on static priority stack. The strategy, to assign a priority in account of the importance and sensitiveness to responding time, to balance the priorities between the three layers of bus networks guaranteed instant responses to real-time tasks and the robustness of systematic communication.

The multi-bus communication framework, designed for the connection and communication of subsystems in a satellite, realizes data transfer channels at different speeds in a complex system, by emerging multiple bus networks. With the application, a bus controller, the original limit on the number of online devices and the limit on data payload are gone. Besides above, this research also offers a reference solution to the multi-level network connection in satellites and other similar astronautic products.

REFERENCES

[1] GJB 289A-97. Digital Time Division Command / Response Multiplex Data Bus. 1997

[2]Muha Ali Mazidi, Janice Gillispie, Mazidi Rolin, 8051 Micro-Control and Embedded System. Mechanical Industry Press, 2007

[3] Min Fang, Yining Quan, Changshan Wang, Computer Operating System. University of Xi'an Electronic Science and Technology Press, 2004

Fuel Cell Power System with High Reliability for Space Applications

Xiangan You and Chengan Wan

China Academy of Space Technology Beijing Spacecrafts, Beijing, 100190.

Abstract: Proton Exchange Membrane Fuel Cells (PEMFC), which is the most competitive power source used in the space missions, have the advantages of high specific energy, high specific power, and environment friendly. The output voltage of a single cell is about 0.4V at full load and a large number of cells have to be connected in series to satisfy the power requirement of a spacecraft. However, bad or mal-performing cells will lead to breakdown of the whole stack, which decreases the reliability of fuel cell power system. In this paper, a modular fuel cell architecture is proposed which can effectively solve the failure issues caused by bad cells, and avoid problems caused by the multi-stack parallel connection, such as large volume, overweight, and high cost. The framework, control scheme and failure isolation method of the proposed architecture are analyzed on the basis of a three-module, 3kW prototype. Besides, simulation results obtained by Simulink software demonstrate that, by removing only a few cells instead of the whole stack, fuel cell system with this architecture, which can be taken as a reference design for fuel cell systems used in space missions, is able to work properly under failure modes and consumes less cells than systems with multi-stack architecture.

Index Terms: Fuel cell, high reliability, modular architecture, failure isolation.

I. INTRODUCTION

With the technical improvements of proton exchange membranes as well as high pressure gas cylinders, efforts has been made to explore new proton exchange membrane fuel cell(PEMFC) system used in space missions [1]. PEMFC, which is a kind of environmental friendly electrochemical equipment, is able to convert chemical energy into electrical energy [2]. In a PEMFC, H_2 and O_2 are taken as reactants and only electrical power, heat, and water will be provided [3]. In space missions launched to date, the most significant energy storage devices are primary and rechargeable batteries [4]. However, devices with better performance, such as higher specific energy, higher specific power, and longer mission life, are badly in need due to the rapidly increasing space activities. Compared to primary and secondary batteries, PEMFC technology has a potential specific energy of 500Wh/kg, a potential specific power of 500W/kg, and a more safety and longer mission life [5][6], which make it more competitive using as either a main power supply or an auxiliary one in solar systems [7].

Several approaches to reliability improvement must be taken in order to satisfy the requirements of a space mission [8]. In traditional designs, a fuel cell system consists of the fuel cell stack, gas supply unit, power regulation unit, thermal control unit, water management unit, and etc. [3]. Actually, a main factor of low reliability is traditional fuel cell stack architecture. With a voltage of 1.0V at no load and 0.4V at full load [9], one fuel cell would not meet the requirements of spacecrafts. Several fuel cells need to be connected in series to form a stack with sufficiently high

voltage. Only one stack and one dc-dc converter are employed in traditional designs. Consequently, a bad or mal-performing cell will breakdown the whole stack and finally influence fuel cell system. To solve this problem, multiple fuel cell stack should be connected in parallel, so that remaining stacks can keep working when a certain stack has been broken. Despite of reliability enhancement, the amount of cells increase rapidly in pace with the number of stacks, which implies larger volume, heavier weight, and sometimes higher cost.

In [10], a concept of modular fuel cell, modular dc-dc converter has been proposed, and a 12V/150W prototype with a three-section fuel cell stack and a modular dc-dc converter has been designed. A conventional fuel cell stack has been divided into three individual sections by two center-tapped electrodes. Each section is powered by a dc-dc converter. It has been proved that the architecture proposed by [10] is capable of achieving failure isolation by shutting down the ill section instead of the whole stack, which leads to less cells requirement than multi-stack systems. Moreover, all stack sections utilize a single set of auxiliary equipment such as gas supply unit. All these factors contribute to a more compact volume and a lighter weight, which implies a more economical system. Nevertheless, either the voltage or the power level of the prototype described in [10] is too low to satisfy the demands of a spacecraft. Additionally, new control scheme is needed to eliminate the fluctuation caused by the traditional power allocation method.

In this paper, a modular fuel cell system with high reliability is designed on the basis of [10]. To demonstrate the method, a 3kW prototype has been designed and the simulation results have been provided. The remainder of this paper is organized as follows. The detail of the proposed architecture, including the system connection and dc-dc converter topology, is described in Section II. In Section III, a decoupling control scheme is proposed for the purpose of power allocation and fluctuation elimination. The failure isolation method suitable for the proposed system comes in Section IV. Simulation results are presented in Section V for demonstration and Section VI comes to the conclusion.

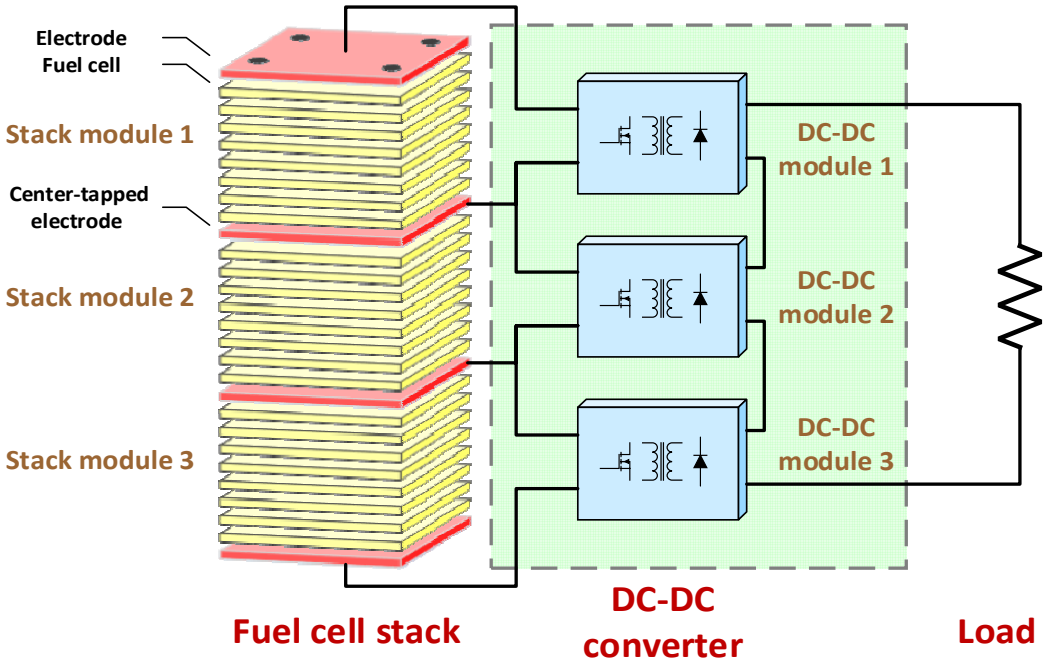


Fig. 1. A modular fuel cell architecture proposed in this paper.

II. Modular Architecture

Under the condition of safety and high reliability, fuel cell stack should have a high output voltage to meet power requirements of a spacecraft. In terms of a single stack, low stack voltage implies high output current, which is prone to a shorter operating life and low reliability. Meanwhile, excess current stress will make converter design a heavy burden and leads to a low converting efficiency. In conventional designs, a fuel cell system is always containing one stack so that the failure from any single cell would break the whole system down. In practice, an N+1 redundancy can be used for reliability enhancement, by which the impact of bad cells can be reduced. For example, assume that three fuel cell stacks had been connected in parallel, which realize a 2+1 redundancy, and a reasonable power redundancy had been taken into account in each stack, then rated power would be guaranteed even if a failure stack occurred. However, multi-stack architecture bring a lot of trouble to system volume and weight. For a similar parallel voltage, all stacks should contain the same number of fuel cells. Consequently, the amount of cells increase rapidly, along with which is a large system volume and heavy weight.

To avoid the stack shutdown induced by a certain bad cell, a modular architecture used for space applications, shown in Fig.1, is proposed by this paper. Center-tapped electrodes are inserted, by which traditional fuel cell stack is divide into several sections. In order to generate the expected bus voltage, each module is connected to a dc-dc module which makes the corresponding stack section an individual module. With all modules sharing the same set of auxiliary equipment such as gas supply unit and water management unit, the stack volume and weight depend only on the power requirement. The proposed architecture has several advantages over the multi-stack one. Firstly, with the independence of each module, the performance of a certain module can be controlled by adjusting corresponding dc-dc converter. In reliability, bypass method can be taken when a certain module, either stack or dc-dc module, is broken. Moreover, the proposed structure is able to achieve the same voltage level by employing less cells than multi-stack system, which leads to a higher power density. For example, under the circumstance with 2+1 redundancy, the proposed architecture consumes 110 fuel cells to realize a 44V full-load stack voltage, which is only 1/3 of multi-stack demand.

Due to the series connection between stack modules, non-isolated topologies cannot be selected as the following dc-dc modules. The equivalent circuit of the proposed architecture with boost converters is shown in Fig.2. It can be seen that dc-dc modules No.2 and No.3 have been bypassed through the common ground line of each boost converter, so that only the No.1 module can work properly. This problem can be avoided only by using isolated topologies.

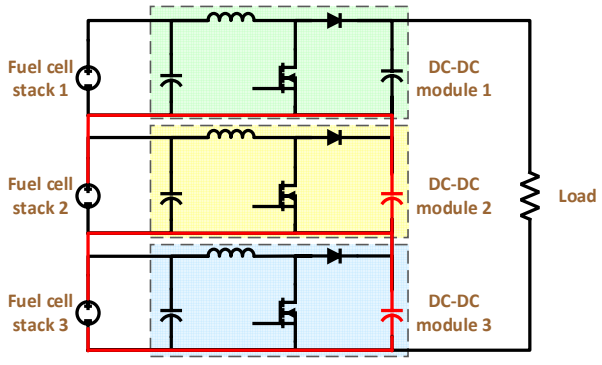


Fig. 2. Equivalent circuit of modular architecture with non-isolated dc-dc converters.

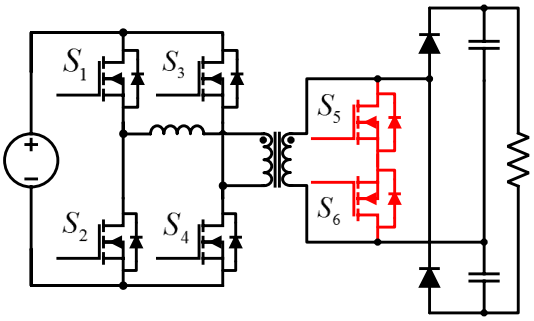
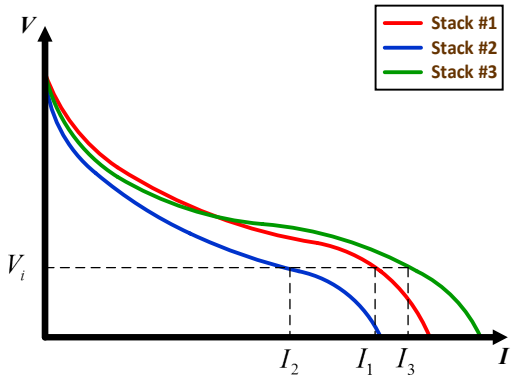
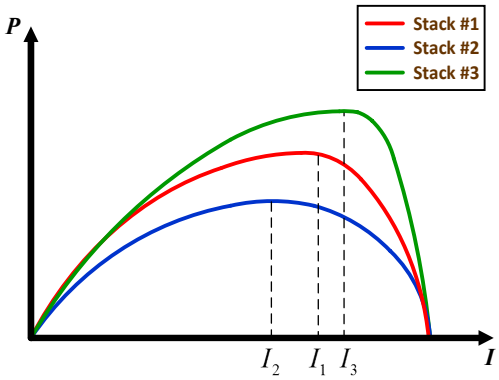


Fig. 3. Equivalent circuit of modular architecture with non-isolated dc-dc converters.



(a)



(b)

Fig. 4. Examples of individual PEMFC stack output characteristics: (a) Stack voltage vs. current, (b) stack power vs. current.

TABLE I Parameters of prototype

Components	Parameters
Stack voltage(U_{stack})	33~66V
Output voltage(U_{bus})	102V
Number of stacks(N)	3
Rated power	3kW
Maximum power	6kW
Minimum stack voltage(U_{th})	9V

This paper selects a soft-switching full-bridge converter based on a phase-shift-controlled

active boost rectifier, as shown in Fig.3, as dc-dc converter modules [11]. The proposed topology has the ability of achieving a high step-up ratio with a relatively low transformer ratio. Phase shift control has been employed and the soft-switching characteristic can be kept in a wide operation range. Fault influence from stack to load can be eliminated by isolation and vice versa. In addition, the probability of short circuit caused by single event latchup effects can be reduced by the full-bridge option and hence improves the system reliability.

III. Power Control Scheme

There exist two purposes that should be achieved by system power control scheme: appropriate power allocation between stack modules and output voltage stabilization. In ground applications, the total power can be allocated based on the proportion of each stack voltage, i.e., stack with higher voltage bears more output power. It seems that output power of each stack module has been balanced by above allocation strategy. In fact, the output of stack module has been coupled with each other, which can easily lead to output fluctuate of dc-dc modules. As a result, the traditional power allocation method is not appropriate to be used in space applications.

The single cell voltage can be affected by fuel pressure, membrane humidity, and state of the catalyst [10]. In practice, the performance of stack modules will not be identical and will fluctuate over time. For example, the characteristics of a set of stack modules has been drawn in Fig.4. From Fig.4 (b), the maximum power is restricted by the limited current of the stack. Combined with Fig.4 (a), it can be inferred that, in high power area, stack with a higher limited current outputs more power when all stacks output a same voltage. Although this phenomenon is only hold in high power area, the imbalance power allocation in low power area has little influence to system safety and reliability, so an appropriate power allocation can be realized through controlling all stack modules output the same voltage.

The proposed power control scheme has been shown in Fig.5. There are three control loops: stack power sharing loop, output voltage control loop, and current loop. Suppose that there exist N stack modules, each with voltage of U_{stack1} and current of I_{stack1}. The voltage of series fuel cell modules is U_{stack}. The output voltage is U_{bus} and the reference voltage is U_{ref}. All stack voltages are forced to be equal through stack power sharing loop, by which balance power distribution will be realized. The purpose of output voltage control loop is stabilizing the bus voltage. Current loop can avoid starvation effect by limiting the variation of stack current. Stack voltages have been decoupled by the proposed control scheme and the fluctuation caused by conventional method has been eliminated. With such an uncomplicated structure, the proposed power control scheme can be realized either by digital controller or analog circuitry.

IV. Failure Isolation Method

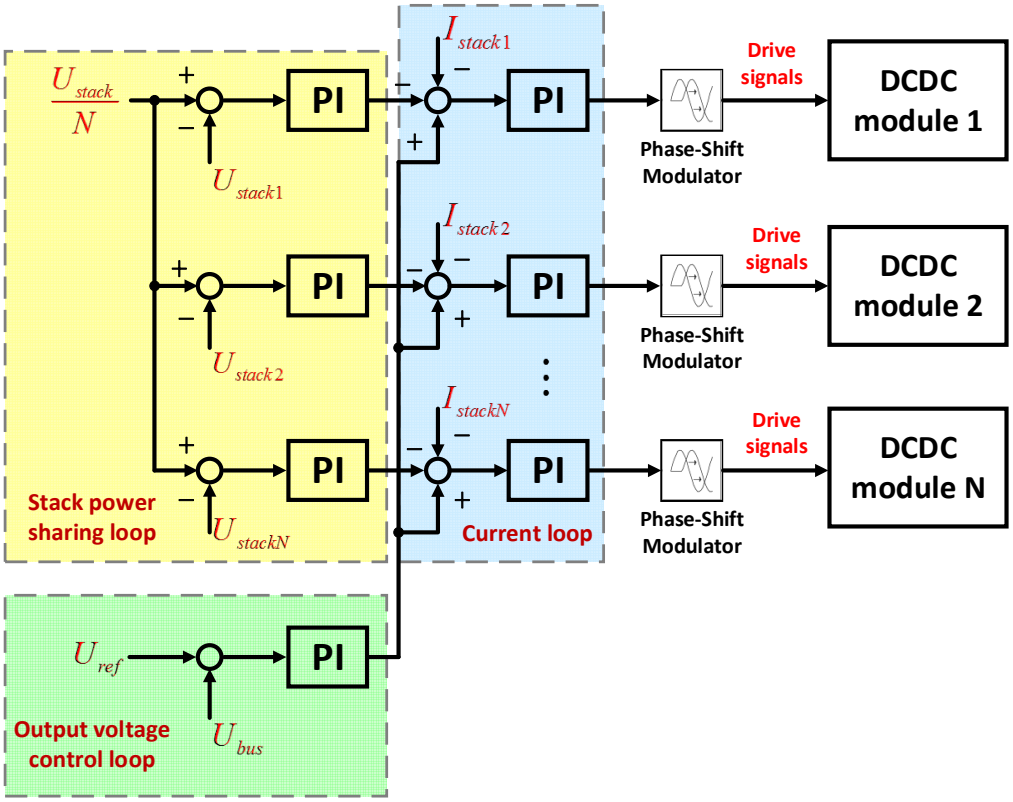
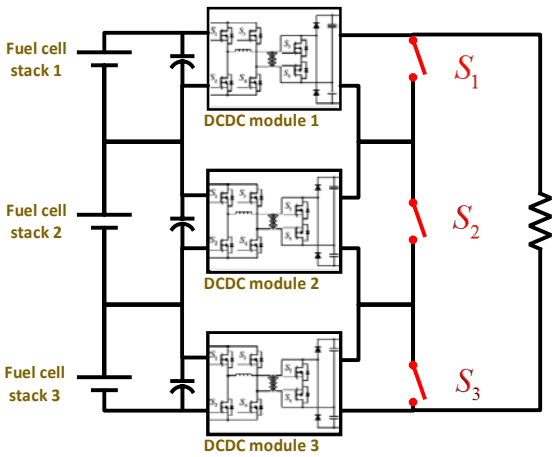


Fig. 5. Proposed system power control scheme.

In multi-stack fuel cell systems, failure stacks can be isolated by removing driving signals of corresponding dc-dc converters. However, only few bad cells exist in a failure stack and a large amount of cells will be waste if the whole stack is removed. One of advantages of the proposed modular architecture is achieving failure isolation by removing a few cells, not the entire stack, so that the efficiency of system will be improved.

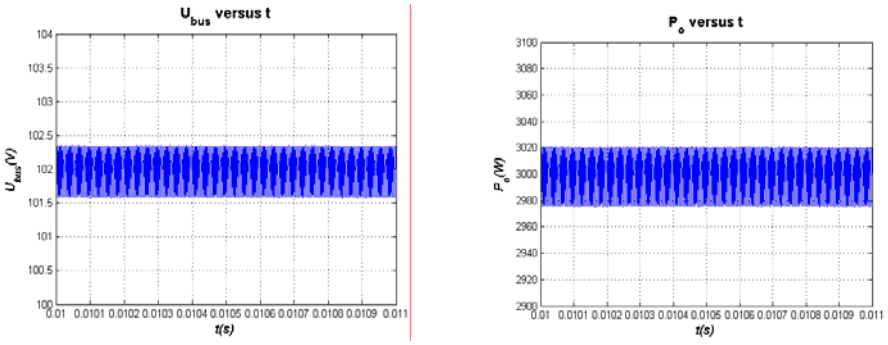


The proposed modular fuel cell architecture uses short-circuit switches to deal with bad cells, as shown in Fig.6. There are four kinds of failure mode: stack module output short circuit, connection of stack module and corresponding dc-dc converter open circuit, dc-dc module output short circuit, and dc-dc module output terminal open circuit. Short-circuit switches can be closed to bypass corresponding modules while the system has fallen into any failure mode. Meanwhile, driving signals of the corresponding converter module should be removed and gas supply of the corresponding stack should be cut out as well. Consequently, other modules in stack can operate normally and would not be affected by a certain bad cell.

When short circuit of either stack modules or dc-dc modules has happened, several cells in

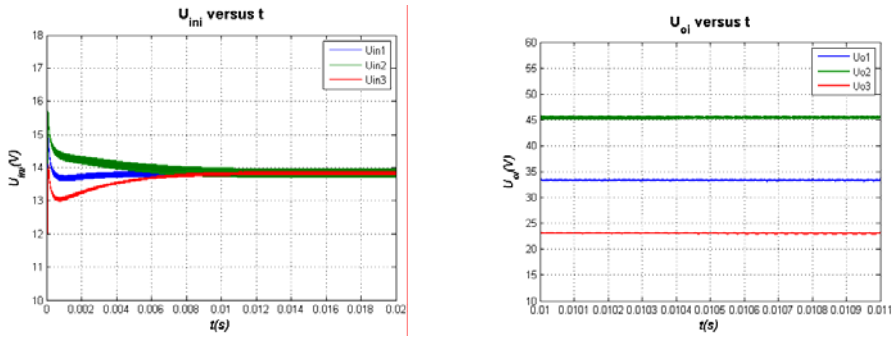
stack will be tend to be breakdown which will give rise to a large internal stack resistor and lead to a rapid decrease of stack voltage. Therefore, stack voltage can be treated as a sign of short circuit failure modes and measures could be done while stack voltage is lower than the preset threshold . Failures caused by open circuit can be diagnosed off line and be handled by telecontrol approach.

V. Simulation Results



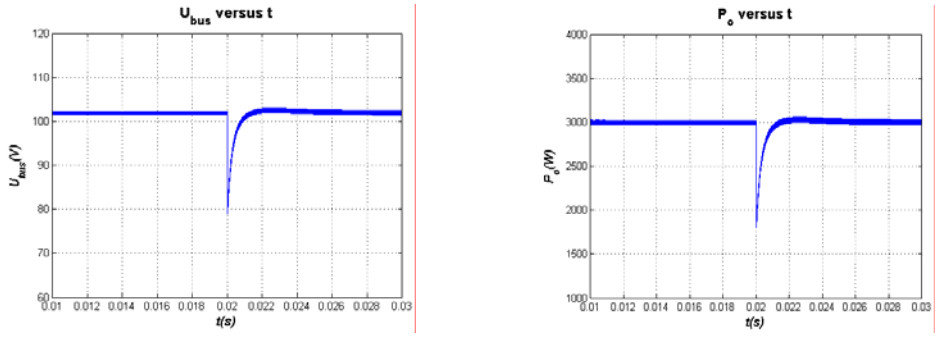
(a) (b)

Fig. 7. Bus voltage and power in steady state: (a) Voltage, (b) power.



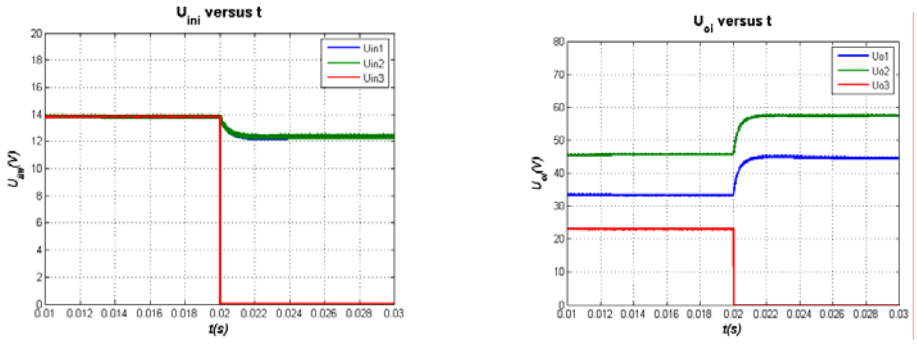
(a) (b)

Fig.8. Input and output voltages of dc-dc modules: (a) Input voltages, (b) output voltages.



(a) (b)

Fig. 9. Bus voltage and power in failure mode: (a) Voltage, (b) power.



(a) (b)

Fig. 10. Input and output voltages of dc-dc modules in failure mode: (a) Input voltages, (b) output voltages.

To demonstrate the effectiveness in reliability enhancement, a prototype with 102V/3kW has been designed and simulation experiments have been carried out in Simlink® software. Characteristics of the prototype are shown in TABEL I. The fuel cell model in SPS toolbox has been employed to form proposed stack.

Output voltage of the proposed system in steady state is shown in Fig.7. It can be seen that the proposed fuel cell system architecture is able to output preset voltage and power stably. In Fig.8, the purpose of that stack with higher limited current output more power has been realized by forcing the stack voltages to be equal, through which the effectiveness of power control method discussed above has been proved.

To force the system falling into a failure mode, we can make the fuel cell module 3 short circuit at a certain time. The output voltage and power of fuel cell system in transient are shown in Fig.9 while stack voltages and dc-dc module outputs are in Fig.10. It can be inferred from Fig.10 that the proposed system is able to achieve fault tolerance by removing only a few cells instead of the whole stack, and power balance can be maintained even in failure mode.

VI. Conclusion

A modular fuel cell architecture, which has the advantages of lighter weight and smaller volume than traditional multi-stack structures, has been proposed in this paper. This architecture has improved power density and reliability of fuel cell systems and is suitable to be used in space applications. A 3kW prototype has been designed in this paper and the power capacity, control scheme, and failure isolation method have been proved through simulation. The results of simulation have shown that bad or malperforming cells can be bypassed without isolating the whole stack.

REFERENCES

[1] Kenneth A. Burke, “Fuel Cells For Space Science Applications,” 1st International Energy Conversion Engineering Conference, 2003, pp. 17–20.

[2] Lu-Ying Chiu, Bill Diong, and Randall S. Gemmen, “An Improved Small-Signal Model of the Dynamic Behavior of PEM Fuel Cells,” IEEE Transactions on Industry Applications, vol. 40, no. 4, pp. 970–977, July/August 2004.

[3] Fuel Cell Handbook, 7th ed, EGG Services Parsons Inc., U.S. Department of Energy, Nov. 2004.

[4] Surampudi S., “Overview of the space power conversion and Energy storage technologies,” http://www.lpi.usra.edu/sbag/meetings/jan2011/ presentations/day1/d1_1200_Surampudi.pdf.

[5] Perez - Davis ME, Loyselle PL, Hoberecht MA, Manzo MA, Kohout LL, Burke KA, et al., “Energy storage for aerospace applications,” 36th intersociety Energy conversion engineering conference, 2001.

[6] Orazio Barbera, Filippo Mailland, Scott Hovland, and Giosue Giacoppo, “Energy and provision management study: A research activity on fuel cell design and breadboarding for lunar surface applications supported by European Space Agency,” International Journal of Hydrogen Energy, vol. 39, pp. 14079–14096, 2014.

[7] Yoshitsugu Sone, “A 100-W class regenerative fuel cell system for lunar and planetary missions,” Journal of Power Sources, vol. 196, pp. 9076–9080, 2011.

[8] Haijiang Wang, Hui Li, and Xiao-Zi Yuan, PEM Fuel Cell Failure Mode Analysis, CRC Press, pp. 1–2, 2012.

[9] Colleen Spiegel, PEM Fuel Cell Modeling and Simulation Using MATLAB, Elsevier, pp. 44–46, 2012.

[10] Leonardo Palma, and Prasad N. Enjeti, “A Modular Fuel Cell, Modular DC–DC Converter Concept for High Performance and Enhanced Reliability,” IEEE Transactions on Power Electronics, vol. 24, no. 6, pp. 1437–1443, June 2009.

[11] Hongfei Wu, Yangjun Lu, Tiantian Mu, and Yan Xing, “A Family of Soft-Switching DC–DC Converters Based on a Phase-Shift-Controlled Active Boost Rectifier,” IEEE Transactions on Power Electronics, vol. 30, no. 2, pp. 657–667, February 2015.

Research on Discharge Plasma Characteristics of An Ablative Pulsed Plasma Thruster via Emission Spectrometry

Lei Yang[1,2,3], Yuping Huang[1], Xiangyang Liu[2], Haibin Tang[3], Chunming Wang[1], Zaiping Zheng[1]

1 Beijing Research Institute of Precise Mechatronic Controls, China;
2 Beijing Institute of Technology, China; 3 Beihang University, China;

Abstract: Pulsed Plasma Thruster (PPT) is a well-developed thruster used on micro-satellites by virtue of its low average electric power, high specific impulse, and simple structure. However, PPT has extremely low efficiency, which is related to the ablation and ionization of its solid polymer propellant. To better understand its working mechanism, this paper carried out the emission spectroscopy diagnostics of a 20J PPT prototype discharge chamber, and revealed the components and characteristics of plasma generated by PTFE ablation process under the excitation of microsecond pulse currents. In the assumption of local thermodynamic equilibrium (LTE), the paper further used Boltzmann distribution relationships and Stark broadening method to obtain characteristic parameters of discharge plasma, such as plasma electron temperature and electron number density. The results showed that discharge plasma mostly detected around ablated surface, which confirmed the observations of higher emission intensity. The plasma temperature was found to be about 2.2eV, whereas number density showed maximum values of more than $10^{17}cm^{-3}$. Finally, plasma streaming velocity in the discharge chamber was also investigated.

Keywords: electric propulsion; pulsed plasma thruster; plasma characteristics; emission spectroscopy

Nomenclature

ε	=	light intensity
g	=	degeneracy
A	=	transition probability
λ	=	wavelength
k	=	Boltzmann constant
T	=	temperature
N	=	number density
E	=	energy
Z	=	partition function
w	=	half the width of the electronic collision
d	=	spectral lines shifts
α	=	ion broadening parameters
S	=	slit area
$\Delta\lambda_{Ins}$	=	instrument pulse width
k_{Ins}	=	proportionality coefficient
V	=	plasma streaming velocity

Subscript

m = upper state
n = lower state
mn = a transition from upper state to lower state
s = Stark broadening

I. Introduction

Today's trend towards small spacecraft, which is not only mass limited but also power limited, has produced a strong interest in development of micropropulsion devices. The significance in reducing launch masses has attracted growing interests in regard to a decrease in mission cost and an increase in launch rate. Although, in the past, many very small spacecraft lacked propulsion systems, future microspacecraft will require significant propulsion capability to provide a high degree of maneuverability and capability. The benefit of using electric propulsion for the reduction of spacecraft mass will likely be even more significant for mass limited microspacecraft missions[1,2]. Among electric propulsion systems, the pulsed plasma thruster (PPT) can offer a wide range of performance depending on design and energy input, and therefore, might be suited for various applications, especially with regard to small satellites down to cubesats[3]. Several PPT were applied in the past to serve, e.g., as attitude control system or for North-South station keeping on a fair number of satellites[4,5].

A schematic of the most common ablative PPT is illustrated in Fig.1[6]. The basic mechanism of the system is that the energy stored in a capacitance is discharged across the surface of the solid propellant (most often polytetrafluoroethylene, PTFE) in order to disaggregate it into a plasma and accelerate the charged particles by the self induced magnetic field. The only moving part is a constant force spring, which assures that an edge of the propellant is held against the fuel retaining shoulder. For initiation of the discharge, an igniter plug is used to shortcut the open oscillation circuit[7].

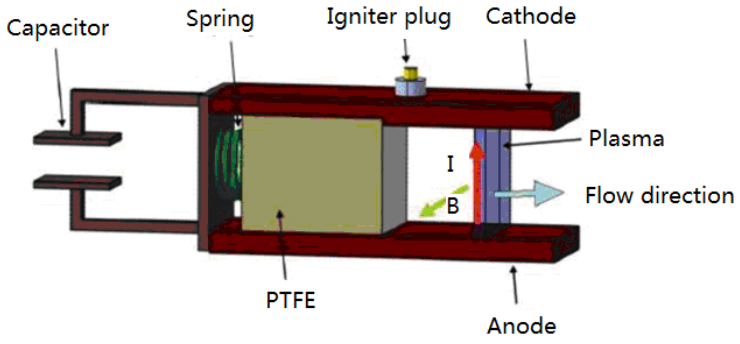


Figure 1. Schematic representation of the ablative PPT.

Previous research showed that the ablated and ionized particles (i.e., plasma) have significant effect on the discharge behavior, the plasma acceleration, and the efficiency of the thruster alike[8,9]. In order to clarify the discharge mechanism and improve the efficiency, it is necessary to study plasma parameters in the discharge chamber. Electrostatic probes were the first diagnostics developed to take measurements inside the plasma. Langmuir Probe theory was first introduced and implemented by Chen and Sekiguch, which was a method to derive the plasma temperature and density within the plasma over time at the specific locations[10]. Endeavors to apply Langmuir probes

in the PPT plume test have been underway for several years. Double Langmuir probe was used by W. J. Guman and M. Begun to study the plasma temperature and density variations of the exhaust of a PPT at different locations[11]. Triple Langmuir probe was used by Gatsonis et al. to derive the average maximum plasma temperature and density at plume [12]. However, the theory of Langmuir probe is valid assuming the theory of the current mode. Moreover, the probe tips would disturb plasma to some extent, particularly, in the discharge chamber, which result in uncertainty of electron density measurement.

In order to evaluate how the plasma properties, i.e., plasma temperature and electron density, are distributed in the discharge channel, optical method was applied in this study. The most common optical method is the emission spectroscopy, which has been applied in various plasma sources because of its simple equipment and nonintrusive characteristic. Measurement of plasma temperature and density was performed by spectral profile method in a 20J PPT prototype discharge chamber. And the velocity of the plasma at some specific locations was also evaluated. These results can help to understand better the nature of the plasma during the discharge of the PPT.

II. Experimental Apparatus

A. Vacuum system

All the experiments were taken in a vacuum system, shown in Fig. 2, which including the vacuum chamber, pumps, control system and other assistant systems. The vacuum chamber is 1.8 m in diameter and 3.2m in length. The base pressure in the vacuum chamber can maintain at 1×10^{-3} Pa during the experiment

Figure 2. Vacuum system.

B. Emission spectroscopy and thruster

In our experiment, the spectral lines were obtained by a monochromator (Sofn Ins-7ISW301) with 10 μm~3mm entrance slit width and with 14 μm entrance slit height. The monochromator has a 70G1200-500 optical grating (wavelength range of 330-828nm) and grating size is 55nm×55nm. A collimation system with two lenses collects the light signal of end-on view which is perpendicular to the fiber.

A stepper motor driven translation system was used to move the fiber axially from the ablated surface to the exit of the PPT discharge chamber, allowing a range of measurements at a set of positions. The translation system includes a moving stand and a control stand, the moving stand was set inside the vacuum chamber and the control stand outside the vacuum chamber allowing for reposition during test.

The thruster used in this experiment has a parallel electrode configuration which uses PTFE as

propellant. A 10±5% μ F capacitor capable of storing approximately 20J is used before discharge. A solid bar of PTFE propellant is spring fed to the electrodes. The power conditioner for the laboratory model PPT mainly includes the adjustable main discharge circuit, the 500 V trigger circuit and the trigger control circuit. The power conditioner accepts a commonly used AC 220V 50 Hz voltage and converts it to a 0-2 kV adjustable DC for main discharge and 500 V DC needed by the igniter plug.

The thruster was mounted on the moving stand horizontally inside the vacuum chamber and the fiber was mounted axially with the center line of the thruster. The setup of the test Facilities such as monochromator, fiber, PPT, moving stand and the vacuum chamber were shown in Fig. 3. And the corresponding PPT discharge is shown Fig. 4.

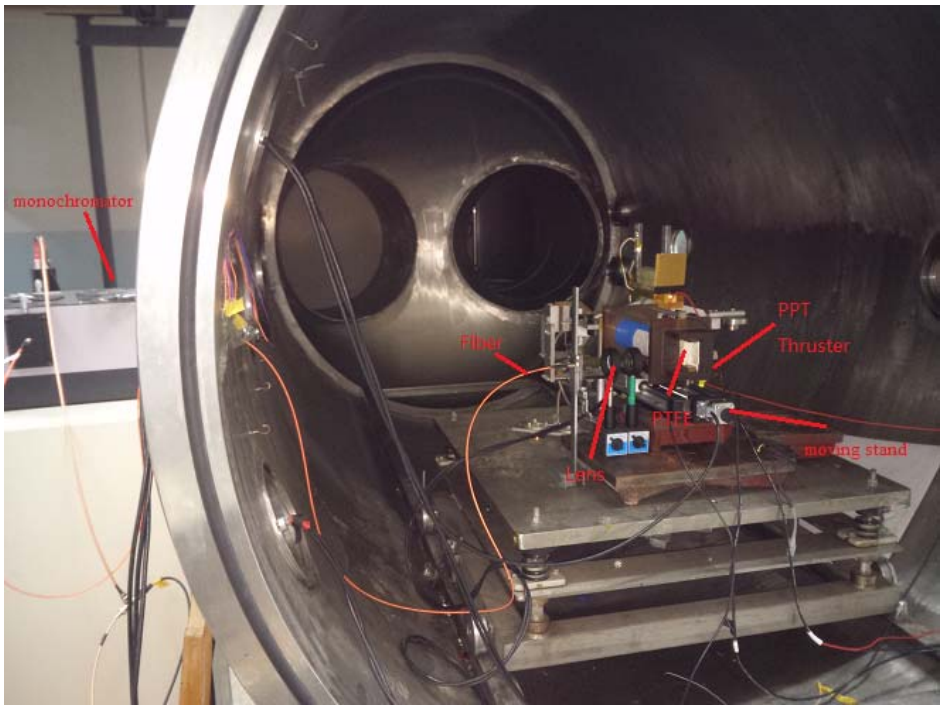


Figure 3. Schematic of apparatus for PPT and emission spectroscopy system.

Figure 4. PPT discharge photograph.

III. Experimental Results

A. Identification of the emissions from the plasma

Emission spectrograms of the PPT discharges were mainly obtained at PTFE surfac. The prominent lines are labeled to identify species, by reference to the National Institute of Standards and Technology Atomic Spectra Database[13], which shown in Fig.5. The most lines originate in singly ionized carbon. There are a few weak lines originating in doubly ionized carbons or fluorine. Neutral carbon atoms were not observed in the present work because the strongest lines of neutral carbon are outside of the measurement range. Based on these spectrograms, emissions around CII wavelengths (i.e., 391.898nm, 588.977nm or 658.288nm) were selected as representatives of PPT plasma.

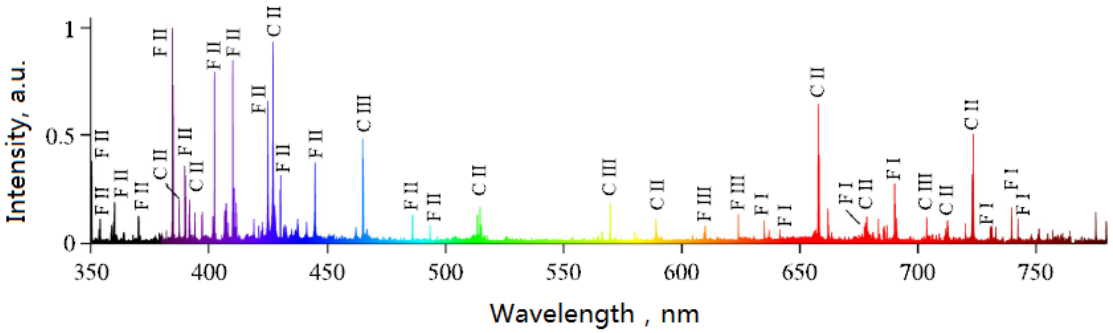


Figure 5. Emission spectrogram emitted from the discharge plasma in the PPT.

B. Plasma temperature and density

When the conditions of a plasma are such that local thermodynamic equilibrium (LTE) applies, the populations of the bound states follow a Boltzmann distribution. It is uncertain whether the prevailing conditions in a transient discharge such as the PPT reach a state of thermal equilibrium, or if measured excitation equilibrium necessarily implies complete local thermodynamic equilibrium (CLTE). In our studies, the assumption of LTE is made a priori, and explored by referring to the measured data.

Relative emissivity of singly ionized carbon will be used to calculate plasma temperature. The lines must be well resolved for accurate emissivity (ε_{mn}), and the transition probabilities (A_{mn}) must be known. Since the populations of the excited states are given by the Boltzmann distribution, Eq.1 describes their relative emissivity:

$$\ln\left(\frac{\varepsilon_{mn}\lambda_{mn}}{g_m A_{mn}}\right) = \ln\frac{N}{Z} - \frac{E_m}{kT} \tag{1}$$

A plot of the quantity on the left-hand side of Eq. (1) vs. E_m has a slope of $-1/T$. Therefore, plasma temperature T can be obtained without knowing the values of N and Z. Information on CII spectral lines is listed in table 1. Unit conversion is used for calculation (e. g. cm^{-1} to J).

Table 1. Parameters for plasma temperature determination of the different wavelength

λ/nm	A_m/s^{-1}	E_m/ cm^{-1}	g_m
391.898	6.36E+07	157234.07	2
392.069	1.27E+08	157234.07	2
588.977	3.15E+07	162524.57	4
658.288	3.66E+07	131724.37	2

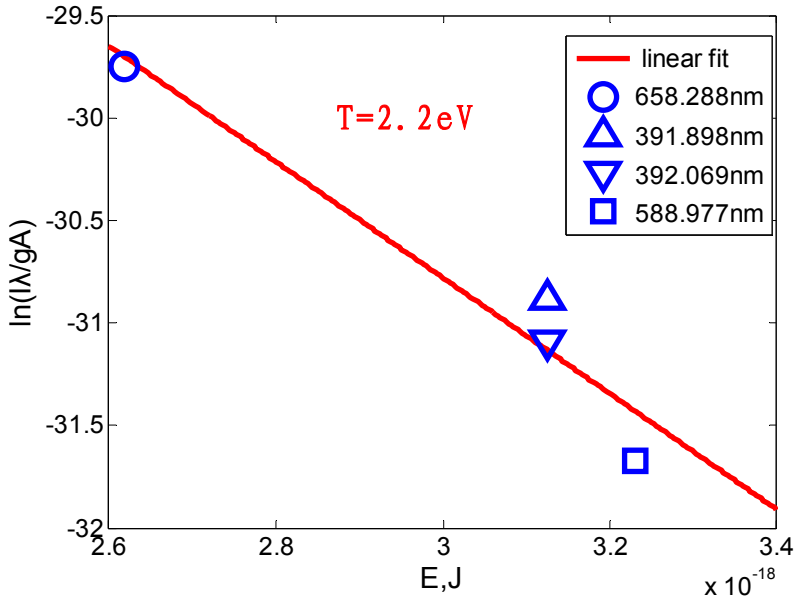


Figure 6. Relative emissivities corrected for degeneracy and radiative rate vs. energy of upper state.

A plot of the left-hand-side of Eq. 1 using the experimentally acquired lines given in Table 1 is shown in Figure 6, where each of the points represents data averaged over five shots. A least squares linear fit is also illustrated. The linear fit indicates a plasma temperature of 2.2 eV (~ 25000 K). The experimental data does not fall perfectly on the straight line, nor does it deviate in any ordered manner; therefore, no strong argument for or against LTE can be made. However, since the deviation from linearity is not large, we can safely assume that the calculated temperature is representative of the mean thermal energy of the electrons.

In our studies, the plasma density was determined by Stark FWHM of C II 658nm spectral line. Profiles of the spectral lines emitted from plasma are subject to various broadening and shifting mechanisms such as natural broadening, resonance, van der Waals, and Stark broadening. The most important broadening effect for our experimental conditions is Stark broadening. The plasma density is derived by using the FWHM of the Stark broadening effect of the spectral emission lines:

$$\Delta\lambda_s = \left[(d/w) + 2.0\times10^{-4} N^{1/4} \alpha \left(1 - 0.068 N^{1/6} T^{-1/2}\right)\right] 10^{-16} wN \tag{2}$$

where plasma density N has to satisfy the following conditions:

$$10^{-4} \times \alpha N^{1/4} < 0.5 \tag{3}$$

$$8.0\times10^{-2} \times \omega\lambda^{-2} \left(T/\mu\right)^{-1/2} N^{2/3} > 1 \tag{4}$$

$$9.0\times10^{-2} \times N^{1/6} T^{-1/2} < 0.8 \tag{5}$$

The parameters W, d and α are taken from literature and interpolated for the given temperature. The results are plotted in Fig. 7. Plasma densities are in the order of 10^{17} cm^3, which is high compared to other electric propulsion systems, e.g., steady-state magnetoplasmadynamic (MPD) thruster (10^{13} cm^3), Hall thruster (10^{12} cm^3), or ion thruster($5\times10^{11}$$cm^3$). Nevertheless, the values agree well with other researchers as well as with the previously shown findings although being slightly higher in average as shown in Table 2.

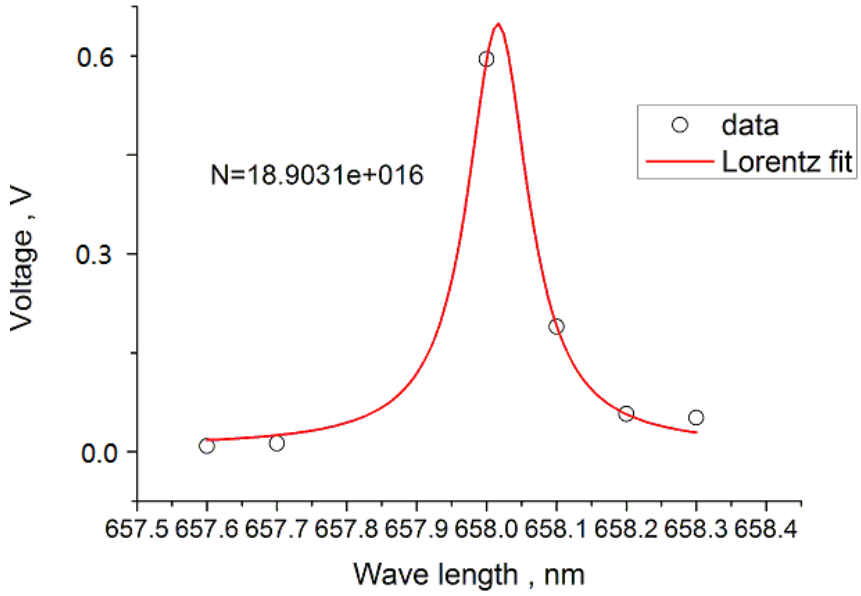


Figure 7. Results for 658nm spectral line and Lorentzian profile.

Table 2. Measured plasma density is compared with other experimental results

Research	E_0/J	N range/10^{16} cm^3
This work	20	18.9
Komurasaki et al. [14]	68	2.7-37.3
Liu et al. [15]	3.15E+07	2.534
Kazeev et al. [16]	3.66E+07	0.4-1.3

C. Plasma streaming velocity

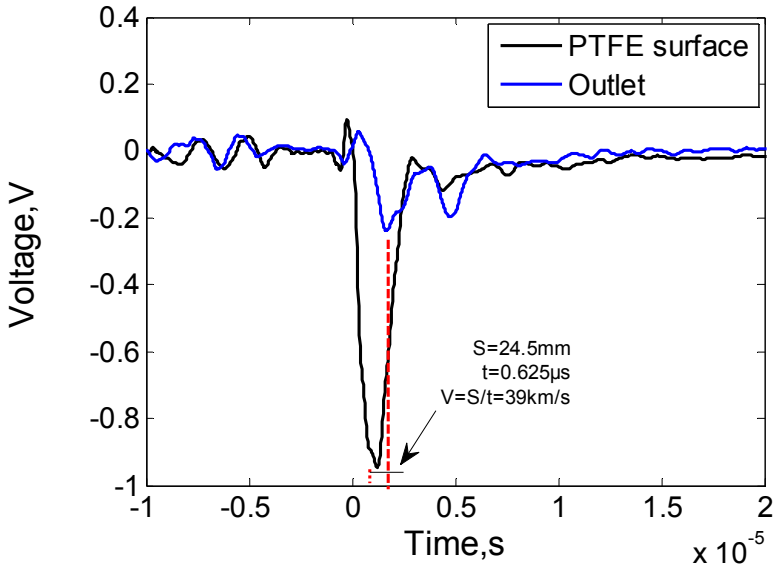


Figure 8. The voltage curves of C II 426nm spectral line at different locations.

The velocity of the ejected plasma varies with position and time in the PPT. In this section we use fiber to determinethe speed of the propellant immediately after it is liberated from the solid PTFE surface during discharge. The fiber was placed respectively at the PTFE surface and the outlet area of discharge chamber. The voltage signals of the C II 426nm spectral line at different locations

were collected to identify the time delay t during the distance S. The velocity of the plasma V will be derived from Eq.6.

$$V = \frac{S}{t} \tag{6}$$

The voltage curves of the C II 426nm spectral line at different locations are s hown in Fig. 8. The voltage of the spectral line at different positions indicates the characters of the charged particles in the plasma; they all go sharply down to -1~-0.2V at the beginning of the curves, then goes upwards mildly to zero and oscillate several times. The first negative peak value decrease with further distance because the density and velocity of electrons are decreased. From the Fig. 8, the average value of the plasma streaming velocity in discharge chamber, i.e., $V = 39\text{km/s}$.

IV. Conclusion

By means of emission spectroscopy, the plasma composition, density and temperature in the discharge plasma of a 20J pulsed plasma thruster were determined during the discharge. Values for the plasma density lie in the order of 10^{17} cm^3 determined by Stark broadening from the spectral emission lines and plasma temperature is about 2.2eV, which is in excellent agreement with other experimental and theoretical data. In addition, the velocity of the plasma has been measured depending on the voltage curves of the C II 426nm spectral line at the locations of PTFE surface and the outlet area of discharge chamber. The average velocity with the arriving time of the spectral line signals is about 39 km/s.

Acknowledgments

This work was supported by China Postdoctoral Science Foundation (No. 2015M570918).

References

[1] Popov, G.A. and Antropov, N.N., “Ablative PPT. New quality, new perspectives”, Acta Astronautica, Vol. 59 No.1-5, 2006, pp. 175-180.

[2] Ozaki, J., Ikeda, T., Fujiwara, T., Nishizawa, M., Araki, S., Tahara, H., and Watanabe, Y.,“ Development of Osaka Institute of Technology Nano-Satellite “PROITERES” with Electrothermal Pulsed Plasma Thrusters,” 32nd International Electric Propulsion Conference, Paper No. IEPC-2011-035, Wiesbaden, Germany, 2011.

[3] Naka, M., Hosotani, R., Tahara, H., and Watanabe, Y., “Development of Electrothermal Pulsed Plasma Thruster System Flight-Model for the PROITERES Satellite,” 32nd International Electric Propulsion Conference, Paper No. IEPC-2011-034, Wiesbaden, Germany, 2011.

[4] Zhang, Z., and Tang, H. B., “Investigations on the Flowing Pulsed Plasma of a 20J Pulsed Plasma Thruster,” 33rd International Electric Propulsion Conference, Paper No. IEPC-2013-147, Washington, D.C., USA, 2013.

[5] Burton, R. L., and Turchi, P. J., “Pulsed Plasma Thruster,” Journal of Propulsion and Power, Vol.14, No.5, 1998, pp.716-735.

[6] Hiroyuki K., Ryosuke N., Kimiya K., and Yoshihiro A., “Plasma acceleration processes in an ablative pulsed plasma thruster”, 43rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, Paper No.AIAA-2007-5226, Cincinnati, OH, USA, 2007.

[7] Mikellides P. G., “Theoretical Modeling and Optimization of Ablation-fed Pulsed Plasma Thruster,” Ph.D. Dissertation, Ohio State Univ., USA, 1999.

[8] Keidar M., Boyd I. D., “Ionization and Ablation Phenomena in an Ablative Plasma Accelerator,” Journal of Applied Physics, Vol.96, No.10, 2004, pp.5420-5428.

[9] Schönherr, T., Investigation of Performance and Plasma Dynamics of the Pulsed Plasma Thruster ADD SIMP-LEX, Ph.d. thesis, The University of Tokyo, Kashiwa, Japan, September 2011.

[10] Sin-Li Chen, and T. Sekiguchi, “Instantaneous Direct-Display System of Plasma Parameters by Means of Triple Probe,” Journal of Applied Physics, Vol. 26, No. 8, 1965, pp. 2363–2375.

[11] Guman W. J., and Begun M., “Exhaust plume studies of a pulsed plasma thruster,” AIAA/DGLR 13th International Electric Propulsion Conference, USA, 1978.

[12] Gatsonis N. A., “Current-mode triple and quadruple Langmuir probe methods with applications to flowing pulsed plasmas,” IEEE Trans. Plasma Sci. ,32, 2118, 2004.

[13] http://physics.nist.gov/cgi-bin/ASD/lines1.pl

[14] Schönherr, T., Komurasaki K., “ Mass and Plasma Characteristics in the Current Sheet of a Pulsed Plasma Thruster,” 32nd International Electric Propulsion Conference, Paper No. IEPC-2011-301, Wiesbaden, Germany, 2011.

[15] Liu F., Nie Z., Xu X., Zhou Q., Li L., and Liang R., “Measurement of electron density by Stark broadening in an ablative pulsed plasma thruster,” Appl. Phys. Lett. 93,111502, 2008.

[16] Kazeev M. N., Popov G. A., Antropov N. N., Diakonov G. A., Orlov M. M., Posokhin V. S., Tyutin V. K., and Yakovlev V. N., “Dynamics and distribution of electron density in the channel of pulsed plasma thruster,” 38th AIAA/ASME/SAE/ASEE Joint Propulsion Conference, AIAA Paper2002-4119, 2002.

Analysis on the Out-of-Plane Mechanical Properties of Honeycomb Sandwich Structures and Experimental Validation

Xiaoyu Liu, Shuang Liu, Chaoqi Xu, Lei Wang

The 41st Institute of the Sixth Academy of CASIC, Hohhot 010010, xiaoyu_run@163.com

Abstract: The thin-wall metal honeycomb sandwich structures(HSS) are widely developed and used in aerospace and aeronautic devices, such as, critical parts of spacecraft cabin, aircraft wings, and thermal-protection system because of significant advantages compared with traditional materials. In this study, the elastic mechanical parameters and the equivalent strength of HSS under out-of-plane compression were investigated. According to the test results and the failure mechanism, the conjecture that the elastic stage of unequal cell-wall thickness HSS should include the pre-elastic stage and the post-elastic stage shown in the experiment was proposed. Moreover, three types of failure modes including elastic buckling, plastic yield and elastic-plastic combination for HSS under out-of-plane compression were studied. Good agreements were found in the theoretical prediction, the FEA calculation and the test result.

Keywords: honeycombs, equivalent mechanical properties, failure mode, FEM

1. Introduction

Due to its light density, high bending stiffness, easy to shape, thermal and acoustic insulation, energy absorption and other excellent properties compared with traditional materials, Honeycomb sandwich structures are widely developed and used in aerospace and aeronautic devices, such as, critical parts of spacecraft cabin, aircraft wings, and thermal-protection system [1-2].

Complex geometry and anisotropic mechanical properties of honeycomb core has brought many inconveniences to analysis of sandwich structures. If honeycomb cores can be equivalently treated as a continuous orthotropic entity by using equivalent material model, general plate can be used to simulate sandwich panel, which will significantly reduce the difficulty in modeling and effectively improve the computational efficiency. On macroscopic mechanical behavior characterization of the cellular structure, Gibson has made outstanding contributions, but further improvement is still required. In terms of equivalent elastic modulus, compressive test results of non-uniform cell-wall thickness honeycomb materials are often far less than the predicted value. On the hand, there are few studies on equivalent strength of the HSS, and general practice is to determine the constraint factor in the classic strength prediction formula based on the test results and experience. According to the cell-wall thickness, hexagonal honeycomb core can be divided into single-walled and double-walled ones. Based on hexagonal double-walled honeycomb, shown in Fig. 1, the elastic mechanical parameters and the equivalent strength of HSS under out-of-plane compression were investigated in this study.

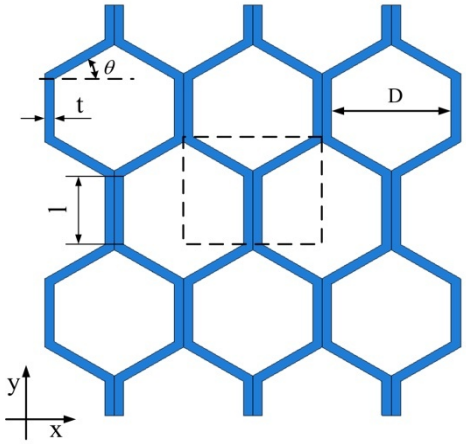


Fig. 1: Honeycomb core cross section diagram

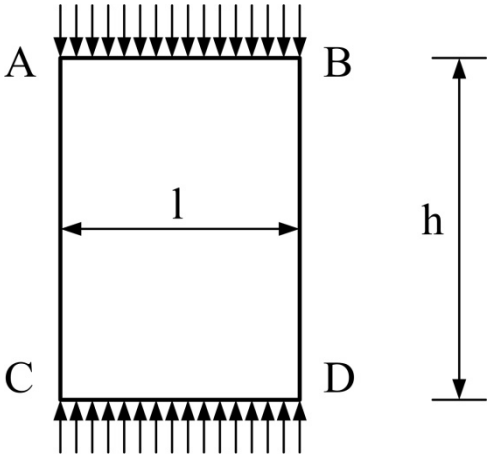


Fig. 2: Cell-wall pressure diagram

2. Theoretical analysis

2.1 Analysis of rectangular plates under in-plane compression

For a rectangular plate clamped at top and bottom sides (AB, CD) and simply supported or clamped along the other two sides (AC, BD), analytical method for uniformly compressed in the direction of the AC side have been given in reference [3]. Buckling load of the cell-wall is determined by the flexural rigidity, geometry dimensions (height, width, thickness) and the boundary conditions. The critical value of the compressive load is given by the equation:

$$P_{cr} = \frac{k\pi^2 E_s}{12(1-v_s^2)}\frac{t^3}{l} \tag{1}$$

Where l is width, t is thickness of plate and k is a numerical factor depending on the ratio h/l of the sides of plate. Several values of this factor are given in Table 1, when AC and BD sides are simply supported. When the two sides are clamped, several values are shown in Table 2.

Table 1: Values of the factor k in Eq. (1) when the AC and BD sides are simply supported

h/l	0.6	0.8	1.0	1.2	1.4	1.6	1.7	1.73	1.8	2.0	2.5	2.83	3.0
k	13.38	8.73	6.74	5.84	5.45	5.34	5.33	5.33	5.18	4.85	4.52	4.50	4.41

Table 2: Values of the factor k in Eq. (1) when the AC and BD sides are clamped

h/l	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50	3.75	4.00
k	10.07	9.25	8.33	8.11	7.88	7.63	7.57	7.44	7.37	7.35	7.27	7.24	7.23

2.2 Equivalent modulus of out-of-plane compression

Ignored the solder, the elastic modulus of HSS in z direction, E_3^c , is determined by the cross sectional area and the young's modulus of solids. Therefore, E_3^c is proportional to the density of the core.

$$\rho_c = \frac{2\rho_s tlh}{l\cos\theta(l+l\sin\theta)h} = \frac{8\sqrt{3}}{9}(\frac{t}{l})\rho_s \tag{2}$$

$$\frac{E_3^c}{E_s} = \frac{A_c}{A_{unit}} = \frac{\rho_c}{\rho_s} = \frac{8\sqrt{3}}{9}(\frac{t}{l}) \tag{3}$$

Test results of the single-walled core are agreed with the theoretical prediction, however, the double-walled ones are often far less. The buckling load of the double wall is as 8 times as the single

wall, which is shown in Formula 1, but the load borne by the double wall is just 2 times in the elastic stage. Therefore, when the single wall becomes buckling, the double wall is still in elastic stage. The assumption proposed by Zhang Dajun is that the carrying capacity of the buckling wall remains the same until the double wall instability [4]. Until the single-walled becomes buckling, honeycomb core is in elastic phase and the modulus value is predicted in Formula 3. From this to the double wall instability, due to the intact bearing capacity of single walls and the linear increase bearing of double walls, honeycomb core is also in macroscopic elastic phase, but the modulus value is only 50% of the forecast. So the elastic stage can be divided into the pre-elastic, which shown is 40% of the elastic stage, and the post-elastic stage. However, double elastic stages never happened in the experiment. According to the defects in specimens, the single wall could be buckling relative easily and the actual distribution of the load is more complex. Based on these reasons, the actual elastic stage is conjectured to be the post-elastic stage.

2.3 Equivalent strength of out-of-plane compression

There are two failure modes, the elastic and plastic instability, when a sheet under pressure. Therefore, honeycomb under out-of-plane compression also have elastic buckling and plastic collapse failure modes, while non-uniform cell-wall thickness honeycomb may have the combination mode.

a) Elastic buckling failure mode

For the cell wall, the AB and CD side are completely clamped and the boundary condition of the AC and BD side is between simply supported and clamped. The elastic buckling load is sum of the various cell walls.

$$\frac{(\sigma_{el}^{s})_3}{E_s} \approx \frac{5\sqrt{3}k\pi^2}{27(1-v_s^2)}\left(\frac{t}{l}\right)^3 \tag{4}$$

The condition for the elastic buckling failure mode is that the double-wall elastic buckling stress is less than the material yield stress.

$$\frac{P_{cr}}{2tl} = \frac{k\pi^2 E_s}{3(1-v_s^2)}\left(\frac{t}{l}\right)^2 < \sigma_{ys} \tag{5}$$

$$\frac{t}{l} < \sqrt{\frac{3(1-v_s^2)\sigma_{ys}}{k\pi^2 E_s}} \tag{6}$$

b) Plastic yield failure mode

If the honeycomb cell wall stress reaches the yield strength of the materials, cell walls will yield. Therefore the yield materials strength determines the compressive strength of honeycomb core.

$$(\sigma_{pl}^{s})_3 = \frac{8\sqrt{3}}{9}\frac{t}{l}\sigma_{ys} \tag{7}$$

The condition for the plastic collapse failure mode is that the single-cell wall elastic buckling stress is greater than the material yield stress.

$$\frac{P_{cr}}{tl} = \frac{k\pi^2 E_s}{12(1-v_s^2)}\left(\frac{t}{l}\right)^2 > \sigma_{ys} \tag{8}$$

$$\frac{t}{l} > 2\sqrt{\frac{3(1-v_s^2)\sigma_{ys}}{k\pi^2 E_s}} \tag{9}$$

c) Elastic-plastic combination failure mode

For the double-walled honeycomb core, the elastic-plastic combination failure mode could be happened, when the single cell-wall elastic buckling and the double cell-wall plastic yield.

$$(\sigma_{ep}^{s})_3 = \frac{\sqrt{3}k\pi^2 E_s}{27(1-v_s^2)}\left(\frac{t}{l}\right)^3 + \frac{4\sqrt{3}}{9}\sigma_{ys}\left(\frac{t}{l}\right) \tag{10}$$

The condition for this mode is that the elastic buckling stress of single cell-wall is less than the material yield stress, but the elastic buckling stress of the double cell-wall is greater than the material yield stress.

The single wall: $$\frac{t}{l} < 2\sqrt{\frac{3(1-v_s^2)\sigma_{ys}}{k\pi^2 E_s}} \tag{11}$$

The double wall: $$\frac{t}{l} > \sqrt{\frac{3(1-v_s^2)\sigma_{ys}}{k\pi^2 E_s}} \tag{12}$$

Failure condition: $$\sqrt{\frac{3(1-v_s^2)\sigma_{ys}}{k\pi^2 E_s}} < \frac{t}{l} < 2\sqrt{\frac{3(1-v_s^2)\sigma_{ys}}{k\pi^2 E_s}} \tag{13}$$

3 Finite element analysis and experimental verification

In accordance with ASTM C 365-03 [5], the flat-wise compression test was carried out, while the dimensions of specimens are shown in Table 3. The modulus and strength were obtained from the compression load-displacement curve. To ensure the accuracy of the displacement measurement, an extensometer was installed on each side of the compression fixture [6]. The Collapsed specimen is shown in Fig. 3.

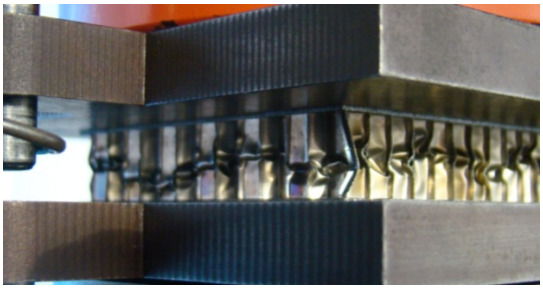


Fig. 3: Collapsed honeycomb specimen

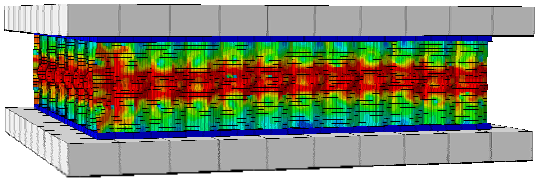


Fig. 4: Collapsed honeycomb core in FEA

Using commercial FEA software ABAQUS, the same dimensions honeycomb sandwich structure analysis model was established. The collapsed honeycomb core in FEA looks like the real one, shown in Fig. 4. According to the analysis of energy in the FEA calculation, the quality of calculation is assured [7, 8]. The load-displacement curve was extracted from the FEA calculation, and compared with the test curses. (Shown in Fig. 5)

Table 3: Titanium honeycomb core specimens for flatwise compression test

Length(mm)	Width(mm)	Height(mm)	Diameter (mm)	Thickness(mm)
75	75	15	6.4	0.05

From load-displacement curve, the FEA process can be divided into four stages: pre-elastic, post-elastic, buckling and collapse (Shown in Fig. 6). Elasticity modulus and the collapse strength of FEA results are compared with test statistical results and theoretical values, which is shown in table 4. The elastic modulus of theoretical prediction and numerical results are agreed well, but test result has 30% deviation. Due to the large discrete of the test data, further tests are needed. Collapse strength values are in good agreement among test results, theoretical prediction and FEA calculation, while the constraint factor k is 7.24, known in table 2.

Table 4: Results comparison

	Elasticity Modulus(GPa)		Collapse strength (MPa)
	First Modulus	Second Modulus	
Test Results	0.88		7.4
Theoretical Values	2.51	1.25	7.85
FEA Calculation	2.65	1.46	7.5

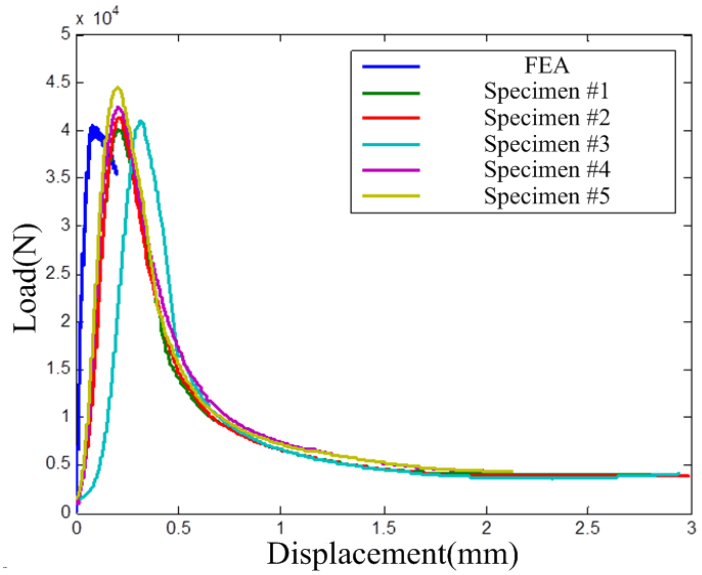


Fig. 5: Load-displacement curve of the FEA and test results

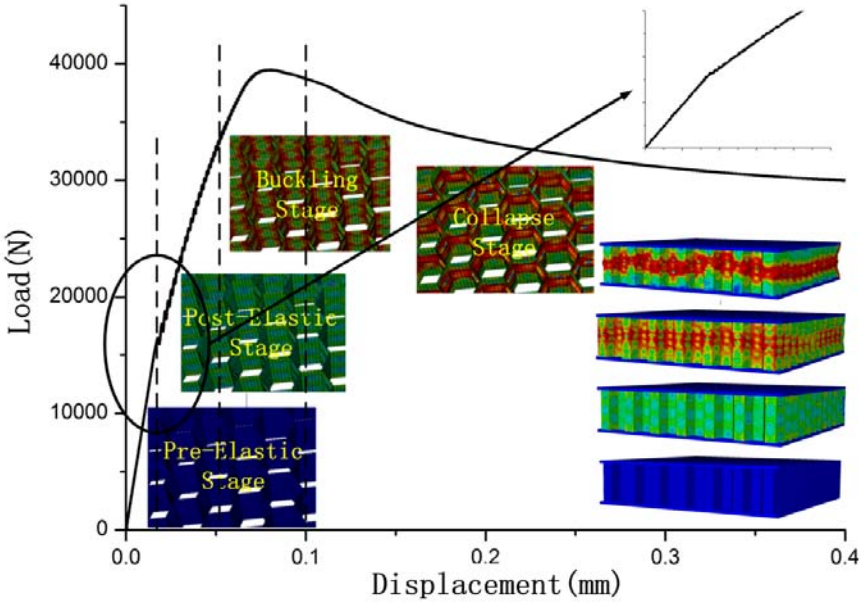


Fig. 6: FEA process of the HSS under flat-wise compression

4 Summery

The elastic mechanical parameters and the equivalent strength of HSS under out-of-plane compression were investigated in this study. The elastic modulus of theoretical prediction and numerical results are agreed well, but test result has 30% deviation. Due to the large discrete of the test data, further tests are needed. On some extent, the conjecture that the elastic stage of unequal cell-wall thickness HSS should include the pre-elastic stage and the post-elastic stage shown in the experiment was validated. Collapse strength values are in good agreement among test results, theoretical prediction and FEA calculation, which enriching the field analysis method for engineering applications and laying the Foundation for further study.

References

[1] Gibson L J, Ashby M F. Cellular solids: structure and properties[M], Cambridge University Press, Cambridge, 1997.

[2] Foo C C, Chai G B, Seah L K. Mechanical properties of Nomex material and Nomex honeycomb structure[J]. Composite structures, 2007.

[3] Stephen P. Timoshenko, James M. Gere. Theory of Elastic Stability(2nd Revised edition) [M], Dover Publications, 2009.

[4] Zhang Dajun, Yu Tongxi. The Calculation of the Initial Buckling Load of Honeycombs under Out-of-Plane Compression[J]. JOURNAL OF SHANGHAI JIAOTONG UNIVERSITY , Vol.24 NO 5,6(1990)47-55.

[5] ASTM C365-03, Standard Test Method for Flatwise Compressive Properties of Sandwich Cores.

[6] Xie Zonghong, Liu Xiaoyu. “Out-of-plane mechanical property test on titanium honeycomb cores”, Adv. Mater. Res, Vol. 718-720, 2013.

[7] Zhuang Zhuo, You Xiaochuan, Liao Jianhui. Finite Element Analysis and Application Based on ABAQUS [M]. Beijing: Tsinghua University Press, 2009. 1.

[8] Hibbitt, Karlsson. ABAQUS Example Problems Manaual [M]. Pawtucket, USA: HKS, 2002.

Geomagnetism Sensing & Calibration Systems of LING QIAO Satellite

CHEN Xi* WANG Jiabo WANG Menglu

（*Space Center, Tsinghua University）
Room 209, Floor 12, Main Building, Tsinghua University, 100084
chenxiee@tsinghua.edu.cn, wangjiabo13 @mails.tsinghua.edu.cn

Abstract: Geomagnetism-based attitude determinations have been widely used in modern small/macro satellites, which is attribute to the availability of small form factor, low power magnetometers and their accuracy meeting the attitude precision requirements of the space mission. In this work, we report the design of the geomagnetism sensing & calibration systems of LING QIAO experimental mobile communication satellite, which include a main and secondary magnetometer and a complex mathematical system for calibration. The LING QIAO satellite has worked in orbit steadily for less than a year, the in orbit results show that output of the main and secondary magnetometers corroborate each other and short term attitude variation given by the MEMS gyroscope is less than 0.2°.
Keywords: Satellite, Geomagnetism sensing, Magnetometer, LING QIAO

1. Introduction

The Geomagnetism sensing system block diagram Magnetometers have been widely used aboard spacecraft and satellites either for scientific missions and attitude determination. For scientific missions, magnetometers are used as instruments in measuring magnetic fields of planets. Until now, scientific explorer missions have detailed the magnetic fields of the Earth, Moon, Sun, Mars, Venus. When it comes to attitude sensing, 3-axial magnetometers have become essential components of modern small/macro satellites as geomagnetism sensor in geomagnetism-based attitude determinations and rotation cancellations.

During the past decades, designing better magnetometers have always been an active research topic in both the academic and industries. Considering that the band pass filter for filtering out the 2nd order harmonics is a performance limiting factor, an orthogonal fluxgate employing digital selective band pass sampling is developed in [1], which allows the fluxgate output to be sampled only once at a single time instance during a number N of excitation cycles. This provides reconstruction of a measured magnetic field with a 40% reduction of noise. In [2], a three-axial search coil magnetometer had been designed and built, which is optimized for 20mHz to 7Hz frequencies and for ultra-low(252uW) power consumption. In [3-4], digital fluxgate magnetometer was developed which substituted part of the analog circuit by digital processors. The performance of digital-to-analog converter in the feedback circuit in the digital fluxgate is the critical point which limits the resolution performance [5]. While most of the existing fluxgate magnetometers harvests the 2nd order harmonics, Yan-Ting Chen, etc., proposed a design and implementation of driving circuits for micro fluxgate sensors employing multiple-harmonic characteristics [6]. To reduce form factor, P.M.Drljaca [7] fabricated a planar fluxgate on a single chip. Heinrich.G [8] created an array micro-fluxgate used for nondestructive testing.

In this work, we report the design of the geomagnetism sensing & calibration systems of LING QIAO experimental mobile communication satellite, which includes a main and secondary magnetometer and a complex mathematical system for calibration. The LING QIAO satellite has worked in orbit steadily for less than a year, the in orbit results show that the output of the main and secondary magnetometers corroborate each other and short term attitude variation given by the MEMS gyroscope is less than 0.2°.

2. The Geomagnetism Sensing System of LING QIAO

The Geomagnetism Sensing System of LING QIAO logically includes two separate fluxgate magnetometers. The probe and analog part of the main magnetometer is a COTS product, which honors more than ten-times space flying experiences. The probe and analog part of the secondary magnetometer was a self-developed product, which had no space flying experiences. The digital circuitries of the two magnetometers are designed on a single PCB board. The digital circuity of each magnetometer consist of a MCU (Micro-Controller Unit), an ADC (Analog to digital converter) and 1/4 multiplexer. The MCU controls the multiplexer to select the outstanding channel to convert from analog to digital. The MCU reads out the ADC results for processing, which conducts all the work that acts as the so called mathematical platform.

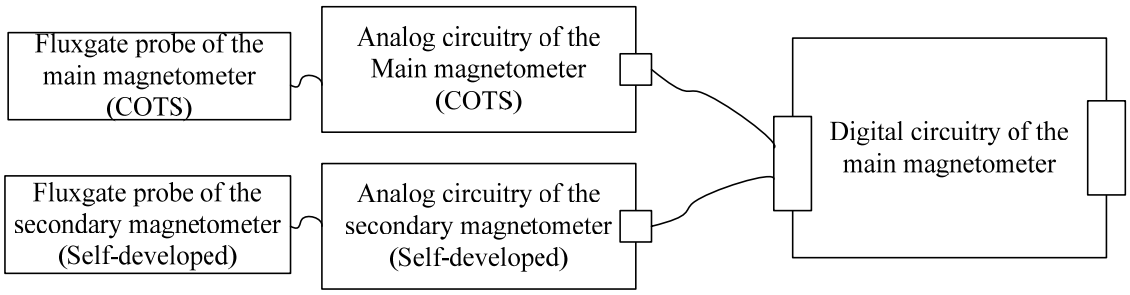


Fig.1 Block diagram of the Geomagnetism sensing system

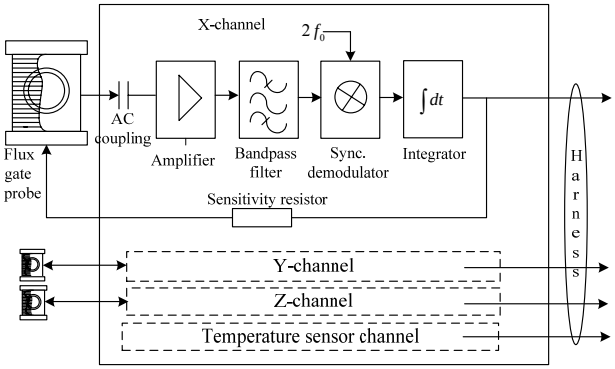


Fig.1 Analog circuitry of the self-developed fluxgate magnetometer

Structure of the self-developed fluxgate magnetometer is given by Fig.1. The analog part of the self-developed magnetometer has four channels: X, Y and Z channel are 3-axial fluxgate channel probing the environmental magnetic field for DS to hundreds of Hz; An additional channel, i.e., the temperature sensor channel is added to sense the temperature of the 3 fluxgate probes. The X,Y and Z-channel are the same with X-channel given in detail in upper part of Fig.1. The output of a fluxgate probe is AC coupled to an analog circuit. The first stage is the low noise amplifier which amplifies all input signals, the second state is a band-pass filter for filtering out signals besides the 2rd harmonic. A synchronization demodulator acts as the third stage to detect the envelope of the

2rd harmonic. The output of the synchronization demodulator is fed to an integrator, which acts as the final stage of the analog circuits. The output of the integrator is proportional to the amplitude of the environmental magnetic field whose direction is parallel to the normal of the fluxgate coil. The feedback path makes the fluxgate coil work around zero regions. The sensitivity resistor in the feedback path controls how much of the output is feedback to the feedback coil.

3. The Calibration

There exist zero point offset, calibration factor error, and orthogonal error due to the magnetometer's manufacturing process, the performance of the instruments and so on. Digital method is employed to cancel these errors.

3.1 The calibration of temperature drift

A reference Voltage channel is designed to help eliminate the temperature drift. Assume that the magnification of the operational amplifier is related only to the operational temperature, as contributes to the temperature drift. The reference channel has the same temperature drift as well as the measurement channels.

Assume that Ref is the analog to digital converter's output of the reference channel, and Mref is the reference channel's AD output in normal temperature. B_{xAD} B_{yAD} B_{zAD} is the AD output of axis x, y and z. The calibration formula is

$$\begin{aligned} B_{x0} &= B_{xAD} - B_{xAD}(\mathrm{Re}\ f - Mref)/Mref \\ B_{y0} &= B_{zAD} - B_{yAD}(\mathrm{Re}\ f - Mref)/Mref \\ B_{z0} &= B_{zAD} - B_{zAD}(\mathrm{Re}\ f - Mref)/Mref \end{aligned} \tag{1}$$

The digital circuit is put into a temperature chamber, while the magnetic field is 55000nT and the temperature declines from 55℃ to -25℃. The experiment results are shown in Fig. 1. The green line in the graph is AD the output of the reference channel, and the green line represents the AD output of axis x. The blue line is the final measurement values after calibration by the method in equation (1).

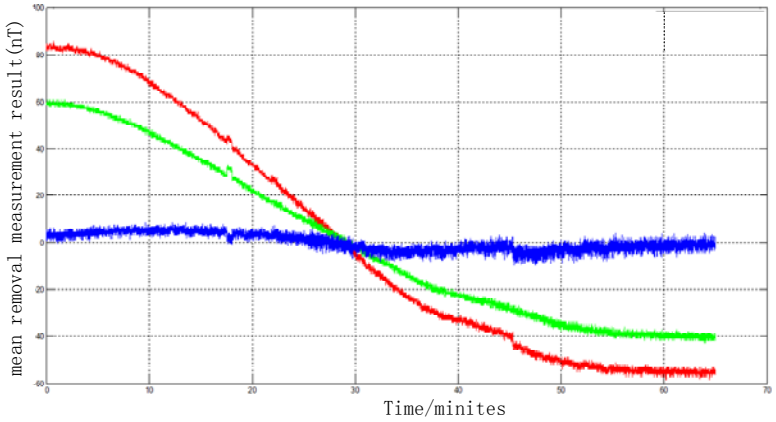


Fig. 1. The mean removal measurement result before and after calibration in the magnetic field of 55000nT

3.2 The zero point offset calibration

Assume that k_x, k_y, k_z are the calibration factors of axis x, y and z respectively, and $\Delta x, \Delta y, \Delta z$ are the zero point offset. The calibration formula is

$$\begin{cases} B_x = k_x(B_{x0} + \Delta x) \\ B_y = k_y(B_{y0} + \Delta y) \\ B_z = k_z(B_{z0} + \Delta z) \end{cases} \tag{2}$$

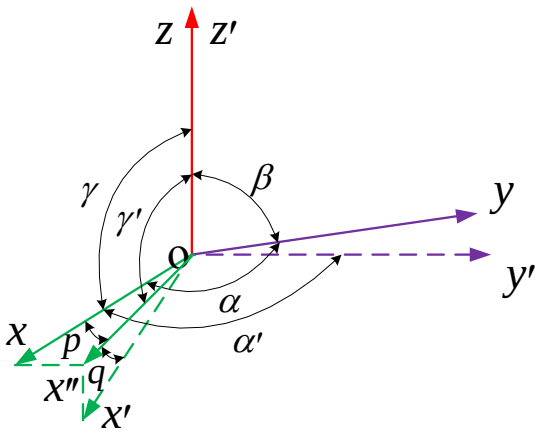
3.3 The orthogonal error calibration

Given a reference coordinate $0\text{-}x'y'z'$, and $0\text{-}xyz$ is the actual polar coordinate. Z-axis and Z'-axis coincide, Y-axis lies in the plane $z'0y'$, the projection line of X on plane $z'0x'$ is x''.

The relationship of $0\text{-}x'y'z'$ and $0\text{-}xyz$ can be formulated as

$$\begin{cases} B_x = B_{x'}\cos q\cos p + B_{y'}\cos\alpha' + B_{z'}\cos\gamma' \\ B_y = B_{y'}\sin\beta + B_{z'}\cos\beta \\ B_z = B_{z'} \end{cases} \tag{3}$$

where $B_{x'}, B_{y'}, B_{z'}$ are the projection of magnetic field intensity B in $0\text{-}x'y'z'$ (see Fig.3.).



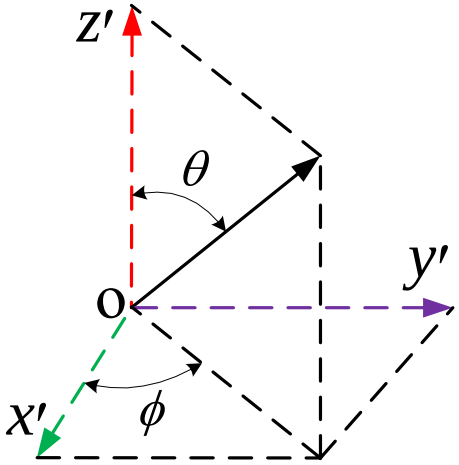


Fig. 2. the reference coordinate and the actual coordinate Fig. 3. The projection of magnetic field in $0\text{-}x'y'z'$

$$\begin{cases} B_{x1} = B\sin\theta\cos\phi \\ B_{y1} = B\sin\theta\sin\phi \\ B_{z1} = B\cos\theta \end{cases} \tag{4}$$

The orthogonal degree of three axis is quite small, i.e. $\left|\frac{\pi}{2}-\alpha\right|, \left|\frac{\pi}{2}-\beta\right|, \left|\frac{\pi}{2}-\gamma\right|$ are not more than 1°. Hence we have

$$\gamma-\gamma' \le 0.00006^\circ,\quad \alpha-\alpha' \le 0.02^\circ, \left|\frac{\pi}{2}-\alpha\right| - p \le 0.02^\circ,\quad q = \left|\frac{\pi}{2}-\gamma'\right|.$$

What is more

$$\cos q=\cos\left|\frac{\pi}{2}-\gamma'\right|=\sin\gamma'\approx\sin\gamma$$

$$\cos p\approx\cos\left|\frac{\pi}{2}-\alpha\right|=\sin\alpha$$

$$\cos\alpha'\approx\cos\alpha$$

$$\cos\gamma'\approx\cos\gamma$$

Thence equation (2) can also be formulated as

$$\begin{cases} B_x=B_{x'}\sin\alpha\sin\gamma+B_{y'}\cos\alpha+B_{z'}\cos\gamma \\ B_y=B_{y'}\sin\beta+B_{z'}\cos\beta \\ B_z=B_{z'} \end{cases} \tag{5}$$

Finally, according to equations (2) and (5), we can have the calibration formulation as

$$\begin{cases} B_{x'}=\dfrac{(\mathrm{k}_x(B_{x0}+\Box x)-(\mathrm{k}_y(B_{y0}+\Box y)-\mathrm{k}_z(B_{z0}+\Box z)\cos\beta)\cos\alpha\big/\sin\beta-\mathrm{k}_z(B_{z0}+\Box z)\cos\gamma)}{\sin\alpha\sin\gamma} \\ B_{y'}=(\mathrm{k}_y(B_{y0}+\Box y)-\mathrm{k}_z(B_{z0}+\Box z)\cos\beta)\big/\sin\beta \\ B_{z'}=\mathrm{k}_z(B_{z0}+\Box z) \end{cases} \tag{6}$$

4. Ground Test Results

The ground test performance of the self-developed magnetometer of LING QIAO is listed in Table 1. All the items can meet the requirement of the performance indexes.

Table 1: The performance Requirement of LING QIAO

Item	Requirement	Actual values
Power Consumption	1.2W(mean value)	1.02W
Pull off voltage	±10%	>10%
Operating temperature	[-25,55]℃	[-25,55]℃
Measuring range	±50000 nT	±60000 nT
Zero point noise	±10 nT	±7.3 nT
Linearity	0.1%	0.05%
Zero drift vs temperature	<0.2nT/℃	<0.2nT/℃
Temperature drift of circuit(Kc)	<50ppm/℃	<30ppm/℃
Temperature drift of probe(Ks)	<200ppm/℃	<149ppm/℃
3-axis orthogonal degree	<0.1°	<0.02°

5. In-orbit Results

A 24-hour in-orbit trace of the secondary magnetometer on Octo 12-13, 2014 is shown in Fig.1. The trace shows that the magnetometer had been working well on the satellite and had been proving a reliable geomagnetic measurement values for the attitude determination and control system. The maximum geomagnetism measured in LING QIAO orbit is no more than 45000nT.

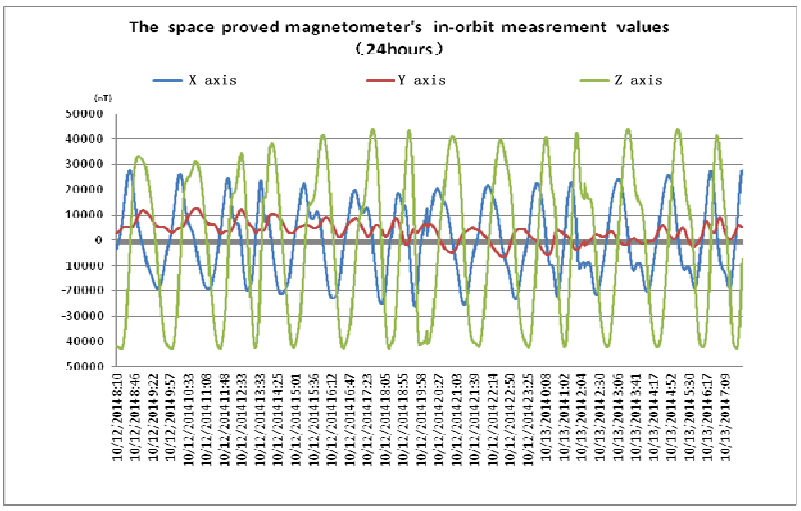
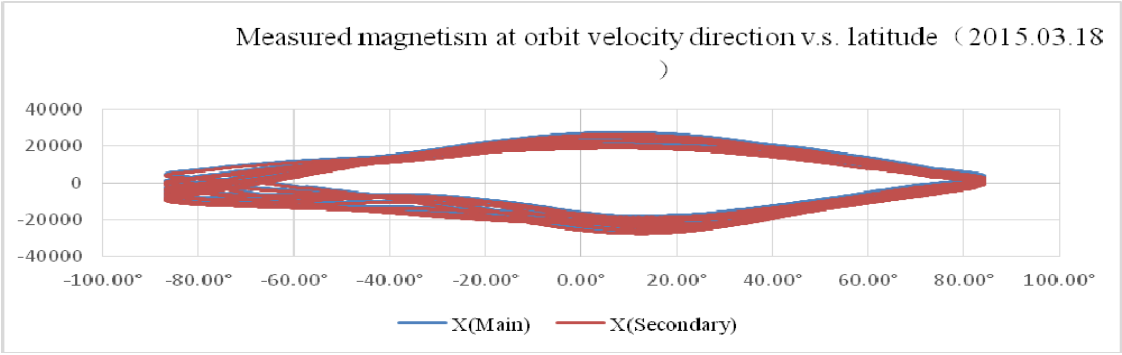
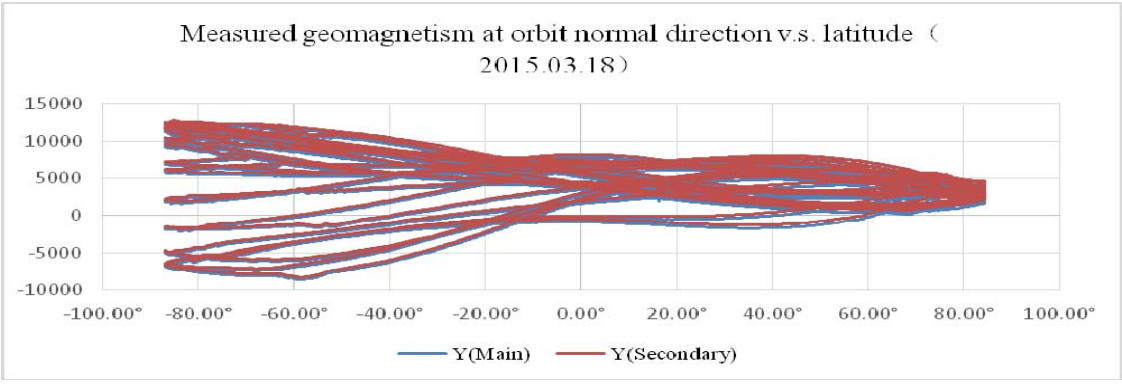


Fig. 1 24-hour in-orbit log Octo 12-13, 2014

In Fig.2, we compare the output of the main and secondary magnetometers, which corroborate each other





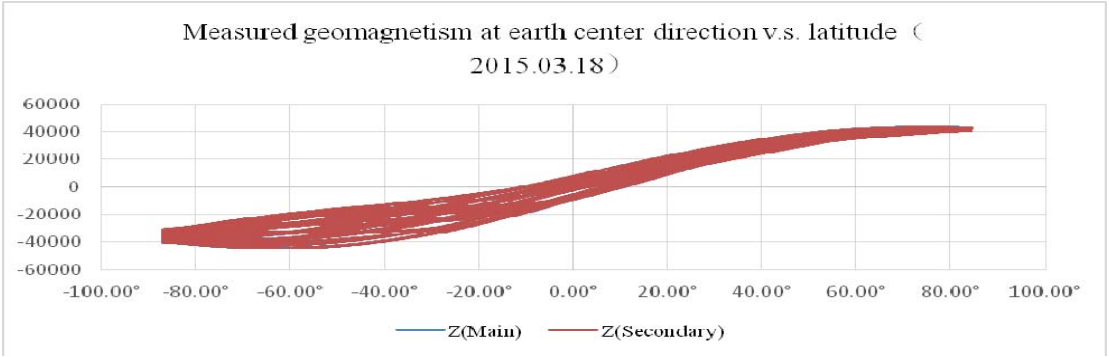


Fig.2 Measured geomagnetism of the main and secondary magnetometers v.s. latitude

In Fig.3, we give the 24-hour trace of the MEMS gyroscope of ADCS system of LING QIAO satellite. According to this result, short term attitude variation of LING QIAO is less than 0.2°, which is good enough for the requirement of mobile communications.

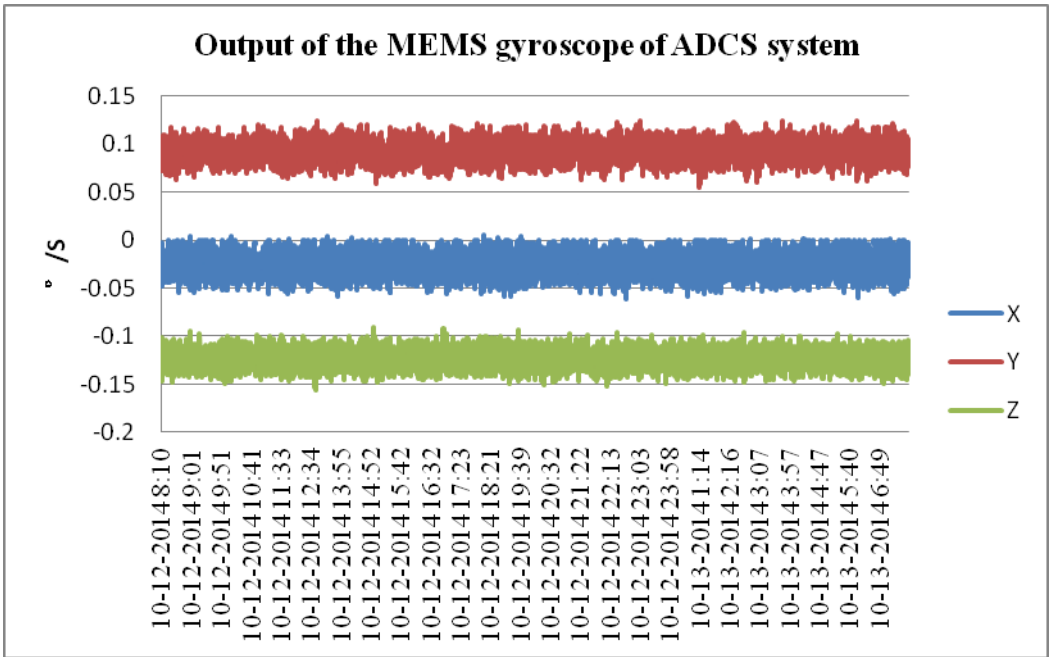


Fig.3 The 24-hour trace of the MEMS gyroscope of ADCS system of LING QIAO satellite Octo 12-13, 2014

6. Conclusions

Geomagnetism-based attitude determinations have been widely used in modern small/macro satellites, which is attribute to the availability of small form factor, low power magnetometers and their accuracy meeting the attitude precision requirements of the space mission. In this work, we report the design of the geomagnetism sensing & calibration systems of LING QIAO experimental mobile communication satellite, which includes a main and secondary magnetometer and a complex mathematical system for calibration. The LING QIAO satellite has worked in orbit steadily for less than a year, the in orbit results show that output of the main and secondary magnetometers corroborate each other and short term attitude variation given by the MEMS gyroscope is less than 0.2°. In future LING QIAO space missions, the self-developed magnetometer will replace the COTS analogy magnetometer and the design is expected to be more compact and light.

Acknowledgement

This work was supported in part by National Natural Science Foundation of China(91338101), National Basic Research Program of China(2014CB340206).

References

[1]Weiss, E.; Grosz, A.; Amrusi, S.; et al, "Orthogonal Fluxgate Employing Digital Selective Bandpass Sampling", IEEE Transactions on Magnetics, Vol.48, No.11, Pp.4089,4091, Nov. 2012

[2]Asaf Grosz; Eugene Paperno; Shai Amrusi; Boris Zadov, "A Three-Axial Search Coil Magnetometer Optimized for Small Size, Low Power, and Low Frequencies", IEEE Sensors Journal, Vol.11, No.4, Pp.1088-1094, April, 2011.

[3]Auster, H.; et al. "Concept and First Results of a Digital fluxgate Magnetometer", Meas. Sci. Technol., Vol.6,Pp.477-481, 1995.
[4]Zhou Xingping, "Design and Achievement of a Fully Digital Fluxgate Magnetometer". Master degree. Zhejiang University, Jiangsu, 2007,5.
[5]Kyosuke Iguchi; Ayako Matsuoka, "A digital-type fluxgate magnetometer using a sigma-delta digital-to-analog converter for a sounding rocket experiment". Meas.Sci. Technol., Vol.25, No.7, Pp.828-838, 2014.
[6]YanTing Chen; ChiCheng Lu; GuoMing Sung; et al, "Design and Implementation of Driving Circuits for Micro Fluxgate Sensors Employing Multiple-harmonic Characteristics", In Proc. of IEEE International Conference on Nano/Micro Engineered and Molecular Systems (NEMS), Pp.1071-1074, 20-23 Feb. 2011.
[7]P.M.Drljaca; Pavel Kejik; Franck Vincent; et al, "Single Core Fully Integrated CMOS Miro-fluxgate Magnetometer", Sensor and Actuators A,Vol.110,Pp236-241, 2004.
[8]Heinrich G; et al, "Array of Miniaturized Fluxgate Sensors for Nondestructive Testing Applications", Sensors and Actuators A,Vol.106, Pp.326-328, 2003.

A Study of Intelligent Power Supply for CubeSat

Yu-Ying Yao, Wei Gong, Long-fei Li, Hui-sheng Liang, Yuan-Mo Liu

（DFH Satellite Co.,Ltd）

Mail-Box 5616, Beijing, PRC, 100094

Email: yaoyyliang@163.com

Abstract: An advanced solar array-battery combined power supply system is proposed for CubeSat in this paper. The system consists of four standard PC/104 boards: Power Regulation Board (PRB), Power Interface Board (PIB), Power Management Board (PMB) and Power Distribute Board (PDB). These boards are stacked up without cable or with a few cables. A primary 12V power bus and multiple secondary centralized power buses are provided for a compact CubeSat. A half-regulated topology is applied for the stable primary bus voltage and charging battery safely by means of constant-current and limited-voltage. A combined power distribution principle is implemented for high efficiency. Load priority dynamic power management and fault diagnosis methods are implanted in power system software so that multi-satellite networking and management are processed handily and rapidly.

Keywords: CubeSat, power supply, power grade, fault diagnosis

1. INTRODUCTION

The launching amount of small satellites especially CubeSat is increasing rapidly in recent years and show great promise for space science breakthrough and plentiful innovation experiments. The typical advanced CubeSat power supply productions are NanoPower P series of CubeSatShop and EPS series of Clyde Space. The former is suitable for micro-satellite power less 30W and the latter can meet 1U, 1.5U, 2U and 3U CubeSat with power less 50W(1U means one CubeSat unit, size of 10cm×10cm×10cm，weight of 1kg[1]). The NSF Firefly CubeSat and MicroMAS are powered by EPS (both 3U) [2] [3]. The common design and novelty includes: standard PCB size (10cm×10cm) , uniform stack-up connector(PC/104 interface), centralized DC/DC, MPPT(Maximum Power Point Tracking)control, portable battery. The plane battery of Clyde Space is more suitable for capacity enlarges and board assembly.

2. SYSTEM SCHEME

An advanced solar array-battery combined power supply system is designed for CubeSat, as shown in Fig.1. The four physical standard PC/104 boards are named as Power Regulation Board (PRB), Power Interface Board (PIB), Power Management Board (PMB) and Power Distribute Board (PDB), and stacked up according to power process sequence as regulation, management, converter and distribution. The constant current phase of I-V curve is used to charging battery and the terminal charging voltage is limited by shunting circuit. Since the 12V primary bus is half-regulated and the battery discharges without BDR (Battery Discharge Regulator), a light PRB is implemented with only one cable. Two DC-DC converters with backups are configured for providing centralized secondary buses efficiently. Because of the utmost decrease of cables and dynamic load adjustment, high power density and power utilization are obtained simultaneously.

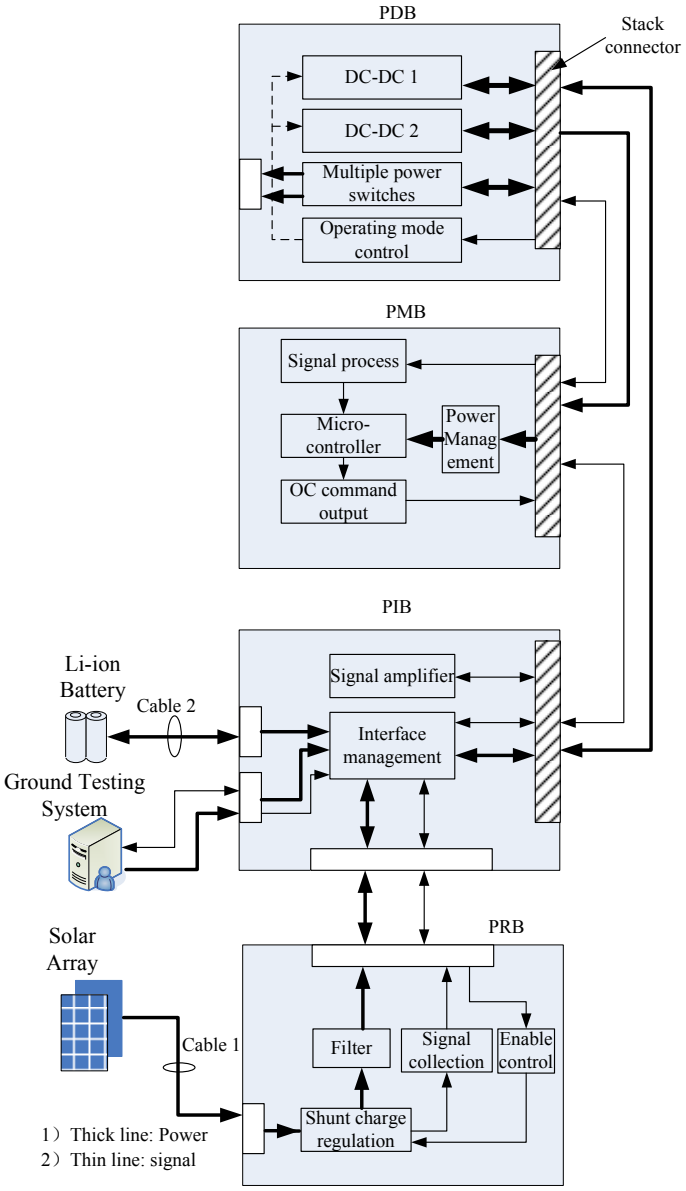


Fig.1 Power system block diagram of CubeSat

3. HARDWARE DESIGN

The CubeSat power system hardware is developed based on Fig.1 and is characterized of efficient DET topology, simple control, integrative power system and centralized power management. High quality DC-DC devices are fitted for high converter efficiency and low cost. 18650 Li-ion batteries are also equipped as light, compact and inexpensive power storage device.

3.1 Power Preassigned Regulator

Every solar array board is preassigned as power array and shunt array for decreasing power dissipation. The former is connected to primary bus only through isolated diode and is utilized for providing energy to loads directly. The latter is connected to PWM shunt circuit and is utilized for power regulation. PRB is implemented with a compact half-regulating topology which can charge the Li-ion battery safely by the constant current characteristic of solar array, as shown in Fig.2. The three blocks of circuits with dotted line are furnished on PRB and the other circuits including signal amplifiers are implemented on PIB (except battery, solar array and isolation diodes). The power

balance among the solar array, Li-ion battery and load is also realized by PRB. PIB receives the energy form battery and ground testing system and implements signal sampling and process of important power parameters, such as bus voltage, battery voltage, charging and discharging current.

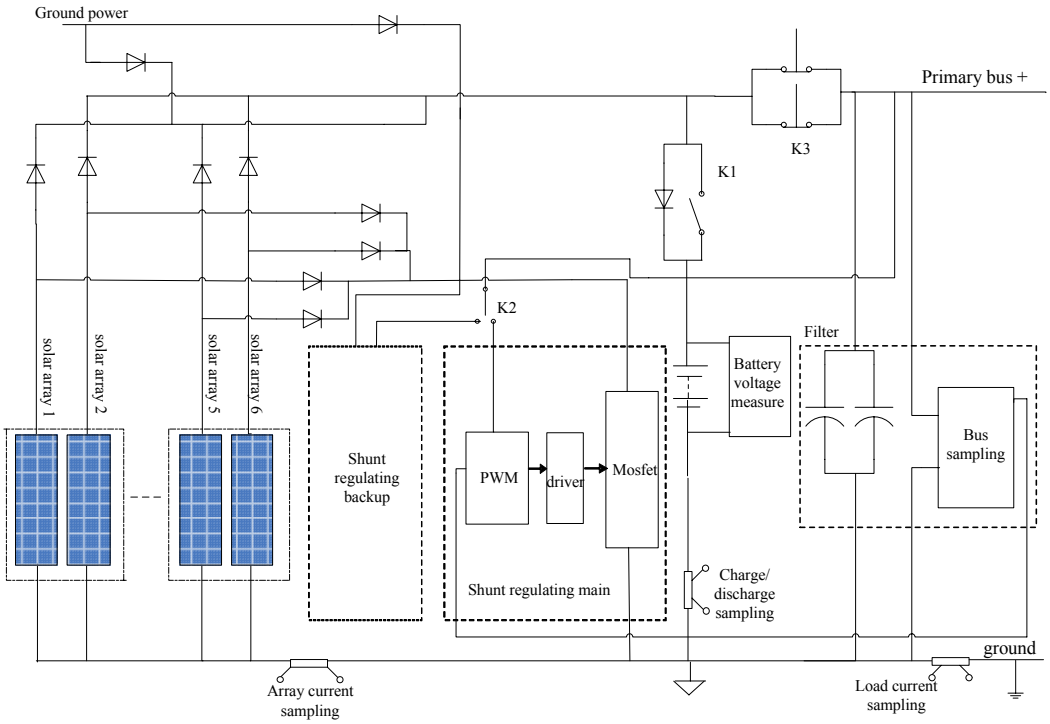


Fig.2 Preassigned power topology scheme

The Li-ion battery is connected to primary bus and discharge to loads directly during eclipse period. The redundant energy of solar array is charged up the battery to the limited voltage during insolation period for power balance. K1 is the battery charge/discharge switch which will be used to over-discharge protection on orbit and ground test. Charging current will flow through the parallel diode if K1 is turned off exceptionally. K1 is switched on before emission. A standby PWM shunt circuit is equipped and will take up the ball through shifting switch K2. K3 is a travel switch for separating indication and starting up the CubeSat.

3.2 High Efficient PDB

Due to he low average power in CubeSat, centralized DC-DC converter and power distribution are better than dispersive power supply for high efficiency. Two groups of DC-DC converters and multiple switches for short-term loads are provided on PDB. The principle of power distribution is:

(1) Low power load (≤5W) is supplied concentrated with DC-DC secondary bus from PC/104 stack-up interface;

(2) High power load (>5W) is supplied concentrated with primary bus from PC/104 stack-up interface;

(3) One-time instantaneous high power load (such as igniter) is supplied from individual connector not avoiding the generality of the stack-up interface connectors.

Two standby DC-DC converters are equipped as warm backups to secondary buses for work reliability. The DC-DC converters capable of +5V@3A and -12V@0.5A perform excellently with voltage volatility less than 1% and efficiency higher than 90%.

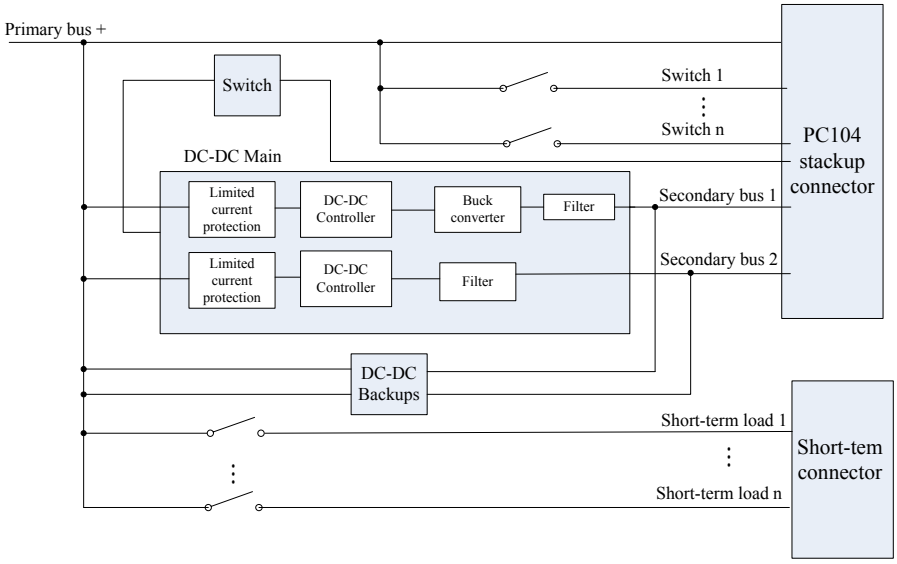


Fig.3 Block diagram of PDB

3.3 Power Management Design

PMB is the power management center and implements the analog signal sampling, command transmitting, bus communication and all software functions including charge/discharge control, fault diagnosis and safeguard procedures. For a compact design, high-speed integrated microprocessor (C8051F040) including ADC, analog switch, CAN controller and FLASH memory is furnished and the peripheral circuits are very concise. The power microprocessor communicates with the main computer through CAN bus preferentially and through I^2C bus complementarily. Power parameters such as charge/discharge ratio, battery under-voltage threshold and temperature range can be modified on-line.

4. SOFTWARE DESIGN

A more important task of power supply system software of CubeSat is realizing systematic dynamic power management except tele-measure, tele-command and communication. Solar panels of CubeSat are typically body mounted and the shining area is limited. This results in power budget insufficient and load dynamic adjustment is necessary.

4.1 Power Grade Management

Load dynamic adjustment is based on power grade management. According to the default load priorities, when the power supply output energy is not enough to meet consumption, some loads will be turned off from bottom grade to top grade. However when the output power is sufficient and battery is full, some loads will be turned on from top grade to bottom grade to match the balance. Load priorities are set based on the mission and actual power consumption in different working-modes. The power grades are divides by safe thresholds and optimized with hysteretic limits to avoid boundary oscillation. The power grades management flow chart is shown is Fig.4 for example. Primary bus voltage is divided to seven grade thresholds:
Th1>Th2>Th3>Th4>Th5>Th6>Th7, and Th7 is the battery over-discharge shut off voltage.

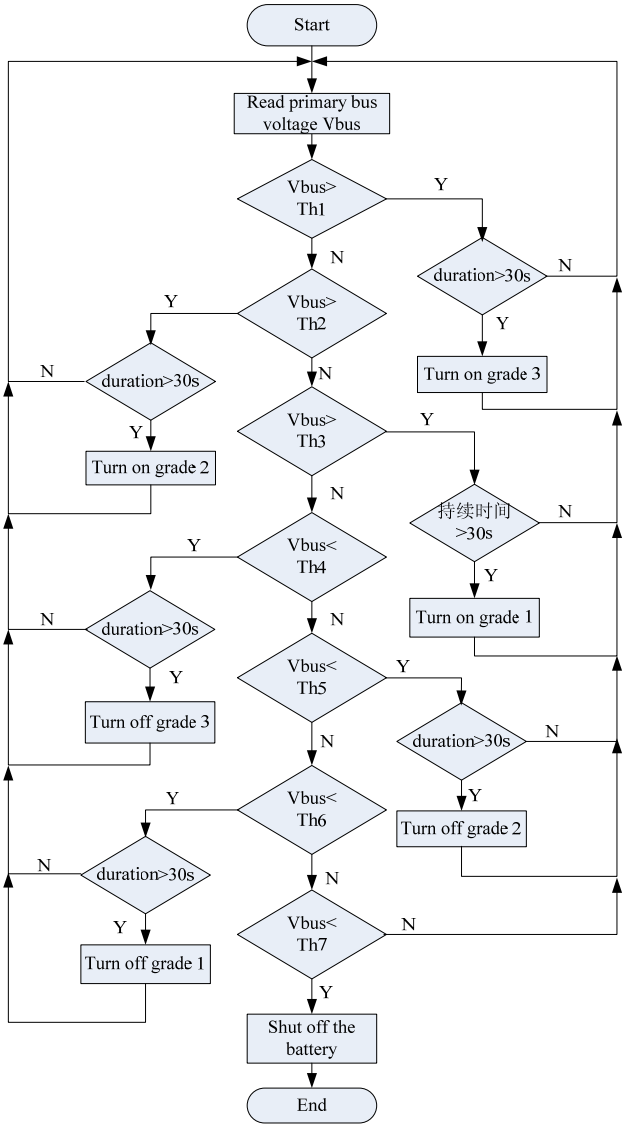


Fig.4 Flow chart of power grade management

4.2 Autonomous Fault Diagnosis

Generally, tens of CubeSats run in constellation networking mode. Mass data cannot be estimated in time, fault diagnosis is a plentiful work and reliability cannot be guaranteed. Autonomous fault diagnosis is a necessary function for CubeSats and other modern spacecrafts. The most common method is tele-measure parameter warning including limit warning, relative value warning and related parameter warning [4]. In the paper, all power safe modes and fault plans based on limit warning and related parameter warning are implanted in software of CubeSat power supply system. Primary bus over-voltage, battery over-discharge, shunt fault and DC-DC failure can be detected and handled or informed to ground autonomously. The power safe mode based on primary bus and battery voltage is processed in Fig.5 and Fig.6. Fig.5 is the software running flow and Fig.6 is the ground database warning system.

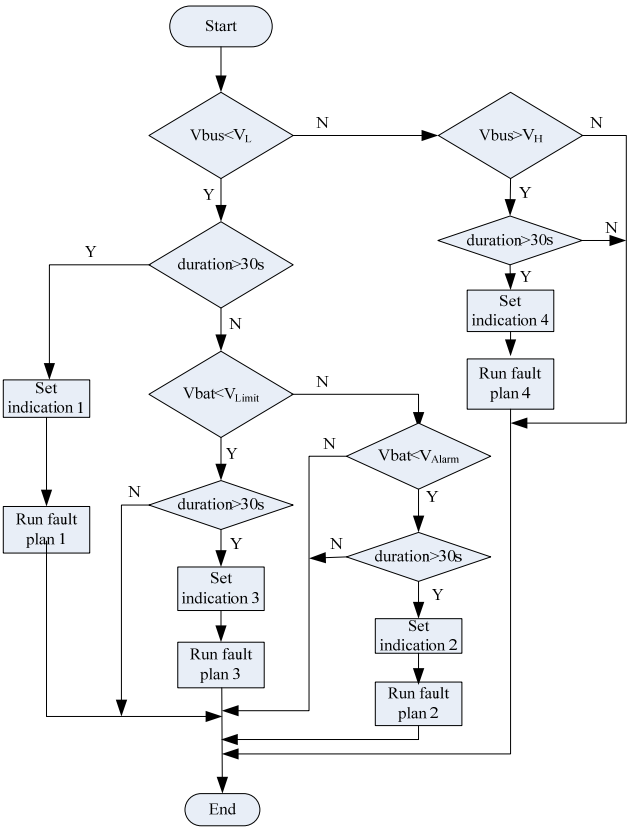


Fig.5 Flow chart of power safe autonomous fault diagnosis

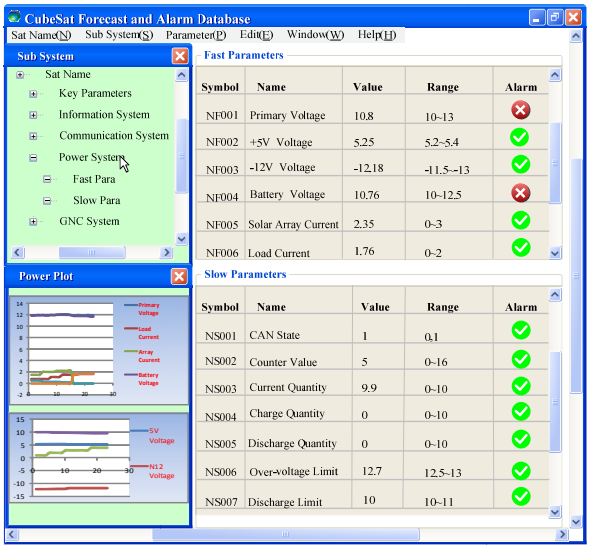


Fig.6 Ground database warning system

5. PROSPECT

One developing approach of CubeSat is to enhance systematic power density and compactness by adopting new devices such as single chip DC-DC converter module, micro rectangle connector and pocket batteries. On the other hand, innovative technology (new three-port power topology) and control method (MPPT) are more important which can improve the performance with less cost. Fuzzy inference based on component and hierarchical diagnosis for satellite power system based on diagraph are both effective ways and means of fault diagnosis to raise the efficiency and veracity on

orbit [5][6] . CubSats management based on networking terminal units is a rising research for the future.

ACKNOWLEDGMENT

I wish to express my deep gratitude to my team of CubeSat and my partners for their help and hard work in the whole system testing.

REFERENCE

[1] LIN Laixing. Technology Development and Application Prospects of CubeSat[J]. Spacecraft Engineering. 2013, 22(3): 90-96

[2] W. Blackwell, G. Allen, C. Galbraith,ect. Nanosatellites for Earth Environmental Monitoring:The MicroMAS Project[C]. Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International：206-209.

[3] Douglas E. Rowland, Joanne Hill, Paulo Uribe,etc. The NSF Firefly CubeSat Mission: Rideshare Mission to Study Energetic Electrons Produced by Lightning[C]. Aerospace IEEE Conference. 2011: 1-12.

[4] SHAN Changshen, LI Yuheng, WANG Libin. Study on Methods of Abnormai Alarm and Fault Diagnosis for In-Orbit Satellites[J]. Journal of Spacecraft TT&C Technology. 2011, 30((3): 6-10.

[5] CHEN Hongbo, SONG Zhengji, JIANG Xingwei. The Application of Fuzzy Inference Based on Component in the Satellite Fault Diagnosis[J]. Chinese Space Science and Technology. 2004, (1): 56-60.

[6] DONG Lei, MA Guifang, LI Qingdong. Hierarchical diagnosis for satellite power system based on diagraph[J]. Journal of Computer Applications. 2012, 32(S2): 38-40,84

STEP M: Space to Effectively Prepare for Migration

Xiaofeng Zhao[1], Xia Wang[2], Shihai Li[3]

1. Beijing Institute of Electronic System Engineering
2. Beijing Aerospace Control Center
3. Aerospace System Engineering Shanghai

Abstract: Human activities have affected the climate, and Climate change causes migration trends. The United Nations Framework Convention on Climate Change (UNFCCC) provides support to least developed countries (LDCs) for the implementation of National Adaptation Plans (NAPs) which aimed to deal with climate change impacts. In order to provide a solution for UNFCCC and LDCs inclimate change driven migration (CCDM) issues, this paper proposes STEM-M method,which is to create a framework for the LDCs identified by UNFCCC on Climate Change that introduces space-based resources to facilitate the creation of efficient migration strategies for adapting to climate change.The application of STEP M to the NAP process will allow LDCs to understand how space can play a role in NAPs dealing with migration.
Keywords: Climate Change, Migration, Space systems, National Adaptation Plans

Acronyms:
CCDM: Climate Change Driven Migration
NAP: National Adaptation Plan
LDC:Least Developed Country
UNFCCC: United Nations Framework Convention on Climate Change
GEO: Group on Earth Observation
GIS: Geographic Information System
IOM: International Organization for Migration
STEP M: Space to Effectively Prepare for Migration

1. INTRODUCTION

Climate change will increasingly impact the Earth's natural environment and well-beings of all humankind in many ways, such as frequent natural disasters, melting polar ice, and rising sea levels. One of the anticipated impacts of climate change is the displacement of millions of people who will be forced to extricate themselves from their homes and migrate to locations that are better suited to human habitation.CCDM results from complex interactions between several possible push and pull factors, but with climate change acting as a primary motivator for migration. CCDM is a highly complex issue that needs to be understood as part of the global migration trend. The International Organization for Migration (IOM) estimates that the number of climate change-driven migrants worldwide will be between 200 million and 1 billion by the year 2050 (IOM, 2008).

Although climate change and migration have been thoroughly investigated individually, little academic consideration has been given to the intersection of these two phenomena. Even less research has been conducted to investigate the effective use of space technologies, such as remote sensing and satellite communications, to address issues arising from CCDM.Although the utility of remote sensing for predicting and monitoring climate change is well known, no comprehensive

study exists on its utility as a tool for planning adaptation measures to migration caused by climate change.

The current space-based capabilities can provide new innovative solutions to the challenges of CCDM, which have only recently moved up the policy agenda.One organization that is aimed at tackling these problems is UNFCCC, which encourages and aids strategic adaptation to climate change throughNAPs for theLDCs(UNFCCC, 2011).STEP Mis an interdisciplinary approach for integrating space-based data into each step of the UNFCCC NAP process when it is applied to CCDM.A high-level representation of STEP M is shown in Figure 1. In this study, the use of space technology as a tool for assessing and planning CCDM is demonstrated from an inter-disciplinary perspective, which a LDC may follow if it chooses to adopt migration as an adaptation strategy to climate change. Overall, the aim is to increase the efficiency and effectiveness with which LDCs can manage CCDM through the use of space-based resources.

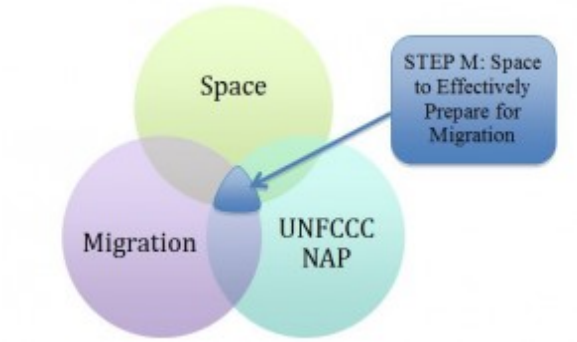


Figure 57 STEP M's role in UNFCCC NAP process

2. CLIMATE CHANGE AND MIGRATION

Migration is an adaptation strategy adopted by humans in the face of natural disasters and climate change. The inter-relation between migration and climate change has been largely ignored by experts and policy-makers until recently. Research on how climate change model outputs can be used as drivers for migration models is still at an early stage. But it is agreed by most experts in the field that climate change is a strong driver of migration andfits into existing migration theories. Table 1 shows the potential placement of environmental factors in classic migration theories.

Table 12 Placing Environment Factors within Migration Theories (IOM, 2009)

General typology of migration (Petersen, 1958)
Potential placement of environmental hazards: as ecological ‘’push’’ factor yielding migration as an ‘’innovative’’ response
Stress–threshold model (Wolpert, 1966)
Potential placement of environmental hazards: as ‘’stressors’’
Mobility transition hypothesis (Zelinsky, 1971)
Potential placement of environmental hazards: as related to ‘’personal preferences”
Residential mobility decision-making model (Speare, 1974)
Potential placement of environmental hazards: as ‘’locational characteristics’’
Neo-classical migration models (various contributors)
Potential placement of environmental hazards: as a “location-specific disamenity”

Generallythere are four different ways in which climate change can affect the movement of people (IOM, 2009):

- The intensification of natural disasters;

- Increased warming and drought that affects agricultural production and access to clean water;
- Rising sea levels making coastal areas uninhabitable and increasing the number of sinking island states (44% of the world's population lives within 150 kilometers of the coast);
- Competition over natural resources leading to conflict, and in turn, displacement.

Table2 lists somedetailed issues that drive migration, and associates them with related climate variables. Figure 2 depicts how climate change and other environmental factors can lead to forced and planned migration.The number of natural disasters has more than doubled over the last two decades. Further changes in the global climate will have a major impact on migration within the next century.

Table 13 Linking Migration Issues to Climate Variables (Ansdell, 2009)

Migration-related Issues	Climate Variables	
Sea level rise	Atmospheric	Air temperature, upper-air temperature, Carbon dioxide, methane, ozone, GHGs
	Oceanic	Sea-surface temperature, sea level, sea state, sea ice, temperature
	Terrestrial	Snow cover, glaciers and ice caps, albedo
Flooding	Atmospheric	Precipitation
	Oceanic	Sea level, sea state, sea ice, current
	Terrestrial	Lake levels, snow cover, glaciers and ice caps
Erosion	Atmospheric	-----
	Oceanic	-----
	Terrestrial	Land cover, biomass, soil moisture
Cyclone	Atmospheric	Wind speed, wind direction, cloud properties
	Oceanic	Current, surface temperature
	Terrestrial	Costal gradient,
Drought	Atmospheric	Precipitation, water vapor, cloud properties
	Oceanic	Soil moisture
	Terrestrial	Precipitation
Food (Crop, Fish, etc.)	Atmospheric	Salinity, carbon dioxide partial pressure, nutrients, carbon, ocean tracers/phytoplankton
	Oceanic	Permafrost and seasonally-frozen ground, albedo, land cover, fraction of absorbed photo synthetically active radiation, soil moisture
	Terrestrial	Precipitation, water vapor
Water	Atmospheric	
	Oceanic	-----
	Terrestrial	River discharge, water use, ground water

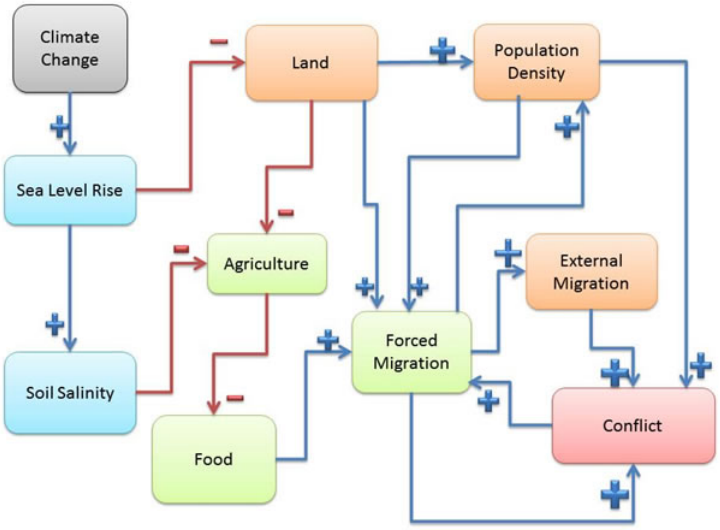


Figure 58 Climate Change Driven Migration Diagram(Litchfield, 2010)

3. UNFCCC NATIONAL ADAPTATION PLAN

During the 1992 United Nations (UN) Conference on Environment and Development, the assembled members negotiated the UNFCCC – an international treaty with the objective of stabilizing greenhouse gas emissions in the atmosphere. There are currently 195 Parties, observer states and observer organizations to the UNFCCC. Of the countries that submit NAPs to the UNFCCC, countries that are designated as the LDCs can apply for support in implementing NAPs. Support comes in the form of:

- Funding provided through a variety of mechanisms laid out in the UNFCCC;
- Technology transfer focused on providing climate-related capabilities and on the provision of green technologies;
- Capacity building directed toward enabling individuals and institutions to mitigate and adapt to climate change.

The UNFCCC has provided detailed guidance on how an LDC can develop a NAP. They established a process by which countries can address medium- to long-termclimate-change-related needs. This process is generalized to help LDCs generate adaptation plans to climate change. A simplified version of this process is represented graphically in Figure 3.

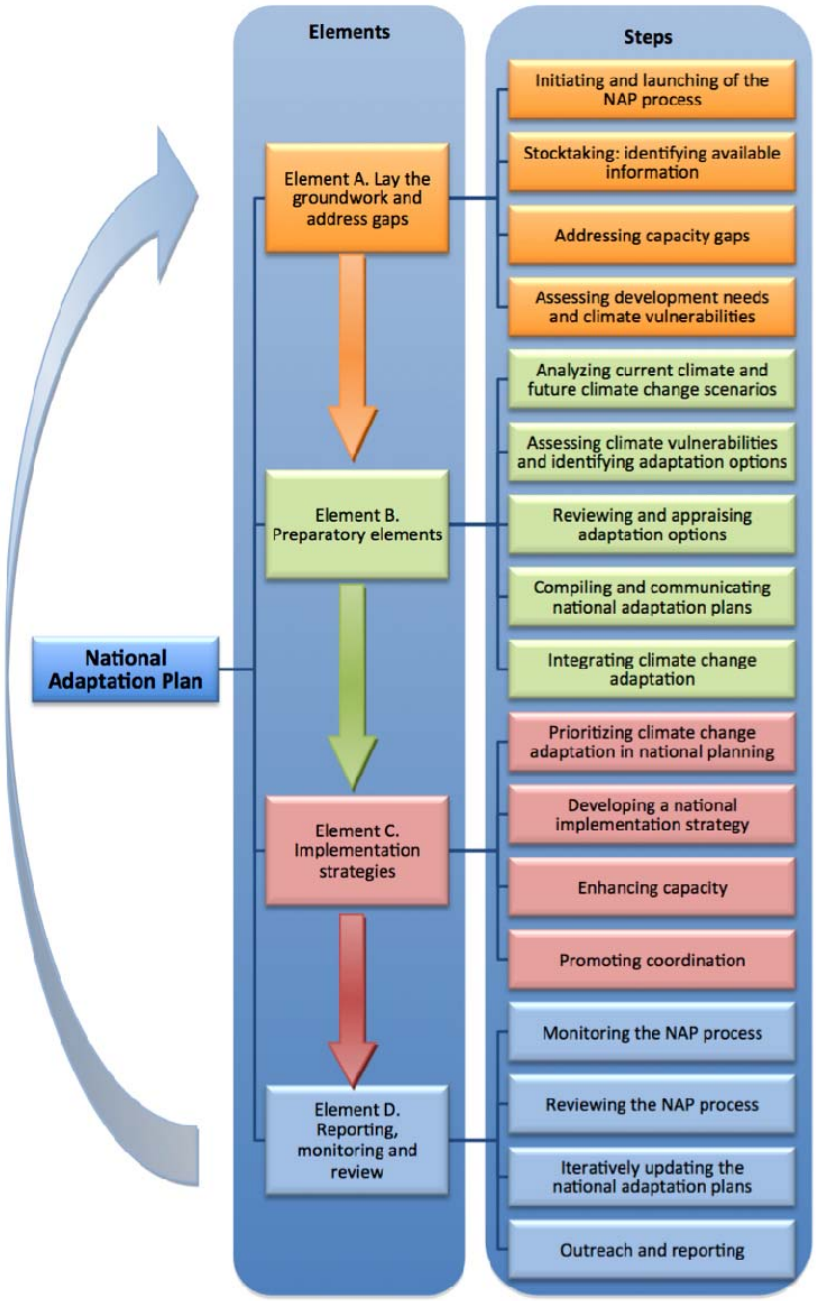


Figure 59Composition of the NAP Process

IOM noted in 2008 that the LDCs that were working with the UNFCCC had not yet included migration in their long-term strategies (IOM, 2008). Therefore, When an LDC is creating a NAP that addresses migration due to climate change, they can follow STEP M in parallel, which will provide advice on how to obtain, integrate and operate space data within their adaptation plan. Because most LDCs don't currently own the requisite satellites, STEP M also details how the LDCs can use the UNFCCC aid mechanisms of capacity building, technology transfer and funding to acquire data, processing capacity and the development of technical expertise for analysis of the acquired data.

4. ROLE OF SPACE

Since 1960, more than 240 environmental satellite missions developed by the National

Aeronautics and Space Administration (NASA), European Space Agency (ESA) and other space agencies have been launched. This includes over 160 meteorological satellites and more than 50 satellites operated as part of an ocean, land, or disaster monitoring series.More than 30 satellite missions speciallyaimed at monitoring climate components and supporting climateprocess management have been designed and developed by various space agencies (Dowell et al., 2013)

The global scale of CCDM requires information on an equally large scale. The only tools that can gather data on such a scale are satellites, which can be used to plan migration effectively and facilitate adaptation to CCDM. The key information provided by space technology is crucial to understanding both climate change and migration, such as the climate change variables listed in Table 2. Furthermore, space data can drive analysis methods, such as predictive climate models, that can better inform policymakers and local populations.More specifically, CCDM can be prevented, predicted, and managed more efficiently with space assets than compared to terrestrial means alone (Figure 4).

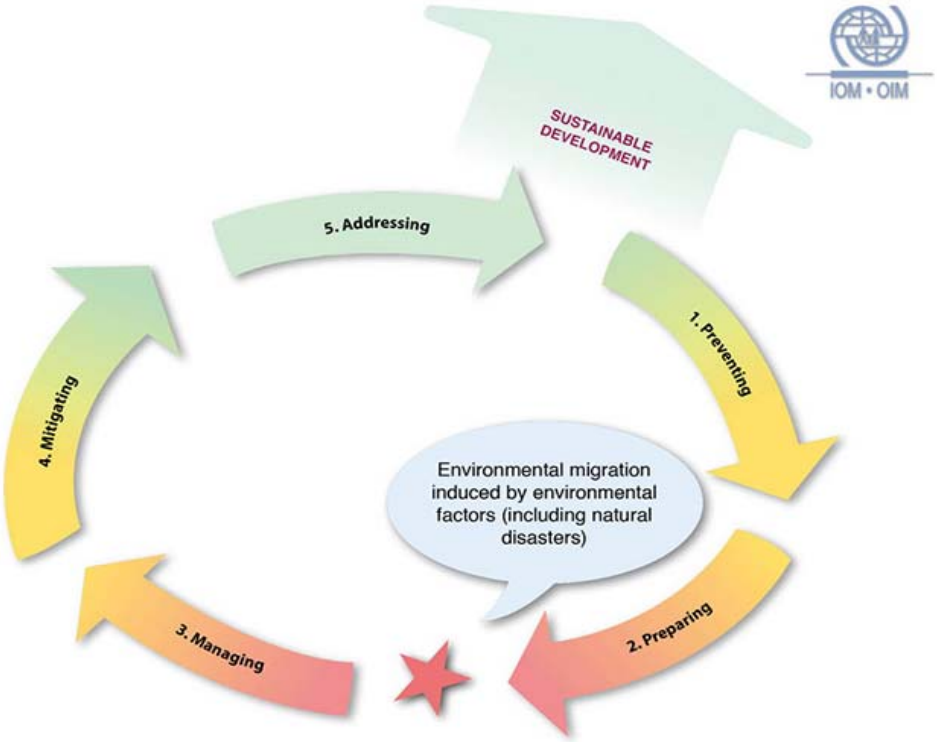


Figure 60Sustainability diagram (International Organization for Migration, n.d)

The process of how space assets can be utilized to address CCDM is illustrated in Figure 5. While GNSS and communication satellites are important tools, their utility is general and there are no detailed descriptionsin defining their role here.

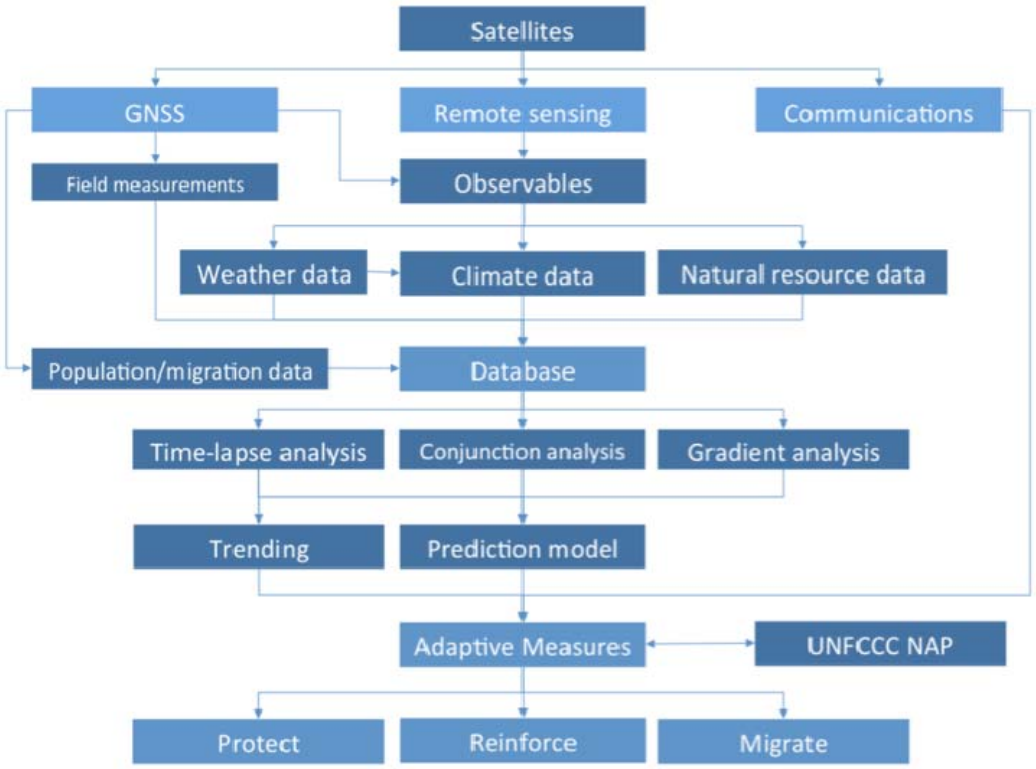


Figure 61 Proposed Data Analysis Process

Data remote sensing satellites act as inputs into:

- Time-lapse analysis: analysis of how key climate variables mapped over a spatially distributed area change over time;
- Conjunction analysis: analysis of interactions between, and aggregation of, key climate variables over a spatially distributed area;
- Gradient analysis: analysis of key climate variables for changes in the "dynamics, structure, and function of ecosystems" (Gosz, 1992).

In addition to these analysis methods, satellite data can support predictive climate and migration models. Outputs of these models subsequently support decision making and adaptation measures.The expectation is that LDCs could use these tools to achieve greater efficiency in climate change evaluation, as well as provide more accurate information to policy-makers. Furthermore, through investment in advanced technologies, LDCs can choose to turn the threat of CCDM into an opportunity to improve quality of life, advance development, and adapt to long-term environmental change.STEP M directs LDCs that are developing UNFCCC NAPs toward potential space-based tools that can be used to achieve the Steps in the NAP process.

5. THE CASE STUDY OF BANGLADESH

A case study was performed as a preliminary validation case for STEP M to illustrate how it operates in parallel with each step of the NAP development process. Bangladesh was chosen based on GDP per capita, population, vulnerability to climate change, and the probability of CCDM.Bangladesh is a country with a lot of biodiversity, a huge population, and a rich history. Its economy is based on its agriculture, which is highly dependent upon its water resources and arable land. The large population of the country will be a huge source of problems in terms of CCDM. Sea

level rise, flooding, erosion, and droughts make up some of the more serious climate change induced problems in the region, and Bangladesh will have to deal with them sooner rather than later.

Table 3 highlights the frequency of various types of disasters and their impact, measured in loss of life or total affected, and Table 4Table 15 provides details on two major cyclones that devastated Bangladesh. Future climatic scenarios of Bangladesh were used to prepare the NAPA by the Ministry of Environment and Forest (MOEF) in the year 2005, which are tabulated in Table 5.

Table 14Bangladesh Natural Disaster Profile (Columbia University, 2014)

EM-DAT Information (1907-2004): Disaster	# of Events	Total Killed	Avg. # Killed	Total Affected	Avg. # Affected
Cyclone	137	614,112	4,483	63,817,281	465,820
Drought	5	18	4	25,002,000	5,000,400
Earthquake	6	34	6	19,125	3,188
Flood	64	50,310	786	369,678,156	5,776,221

Table 15Major Cyclones (Shamsuddoha et al., 2007)

Date	Wind Speed (km/h)	Tidal Height (meters)	Human deaths (millions)	Financial Loss (USD)
12 Nov 1970	196	4.5-6	0.3 Govt. 1.2 Media	86.4m
29 April 1991	225	6-10	0.13 Govt.	1,780m

Table 16Future climate scenarios used for preparation of NAPA for Bangladesh (MOEF, 2005)

Year	Temperature change (°C) per year Mean (standard deviation)	Precipitation change (% per year) Mean (standard deviation)	Sea level Rise (cm)
2030	1.0	5	14
2050	1.4	6	32
2100	2.4	10	88

Every part of Bangladesh is vulnerable to climate change; however, the coastal region has been identified as most at risk.Sea level rise and increased salinity are two climatic events in the coastal region that have begun to reduce the amount of available land for cultivation and will be a strong driver of future migration. One-quarter of Bangladesh is inundated with excess water every year, but rising sea levels could eliminate up to an estimated 15% of land mass (UNEP/GRID, 1989) and an estimated 40% of arable land by the end of the century due to increased water-logging and salinity (CIA, 2013) (Figure626).

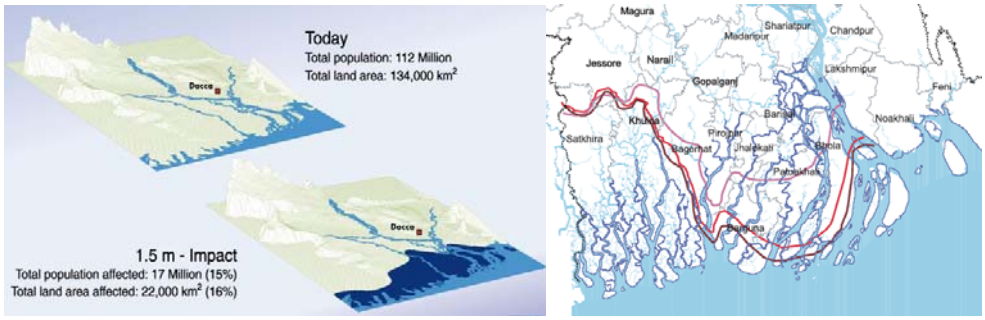


Figure62 Impact of sea level rise (Left: UNEP/GRID, 1989; Right: Siddique and Adrika, 2011) Left: Potential impact of a 1.5m sea-level rise. Modified rates of sea-level rise estimate land area affected will be ~15%. Right: Estimated impact of different levels of salinity ingress

Climate change has increased to flow rates of Bangladesh's rivers through upstream melting and runoff from the Himalayas (higher overall temperature) as well as increased rainfall during the year. This has led to riverbank erosion and a constantly changing flow pattern in the country. Flooding and increased flow rates rapidly change the riverbanks and meanders in the river system. Deforestation in the region has also led to weakening of the ground soil, which makes weathering and riverbank erosion effects much more efficient (Azim Uddin and Basak, n.d.). A map of the flooding areas in Bangladesh is shown below in Figure 7.

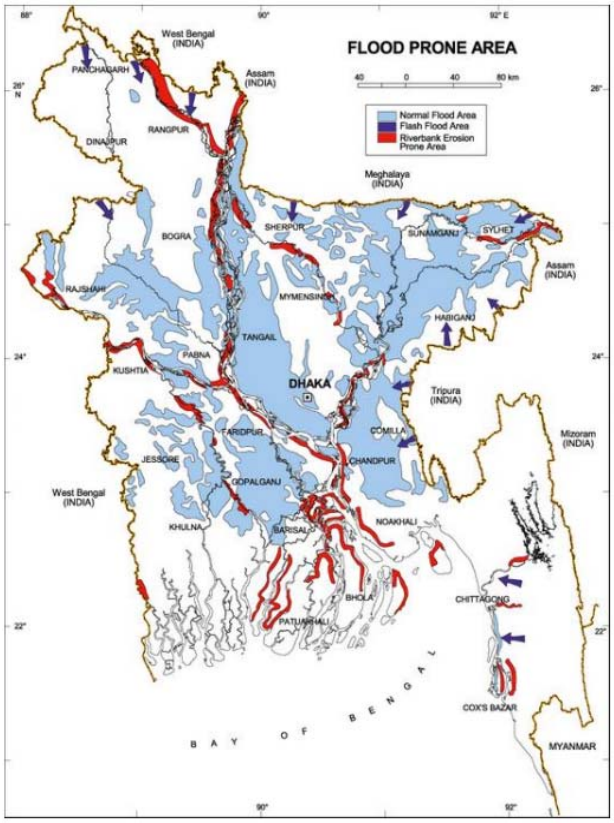


Figure 63Map showing Flooding Areas in Bangladesh (Mahmood, 2012)

Bangladesh experiences other climate change-related problems such asfloods, drought, cyclones, volcanic eruptions, etc. which have certain precursors. An essential climate variable is a geophysical variable that is associated with climate variation and change as well as the impact of climate change onto Earth. The following figures illustrate how Bangladesh has used some of the available remote sensing satellites for prediction and monitoring of natural disasters like floods, erosion, cyclone and drought, which impact Bangladesh frequently.The following figures (Figure 8, Figure 9, Figure 10, and Figure 11)illustrateEarth observation images used by for monitoring these disasters.

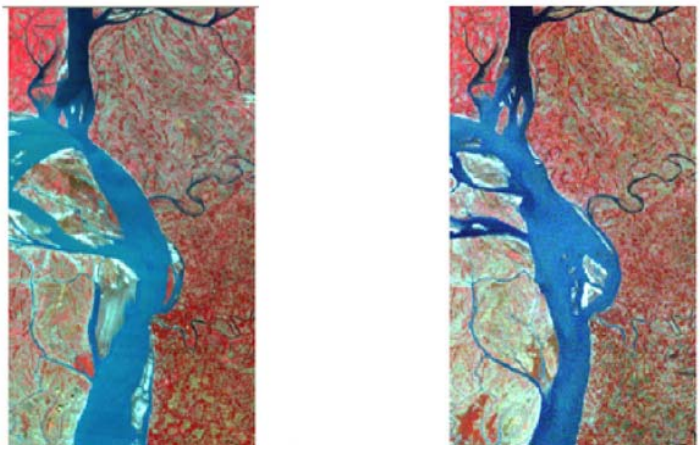


Figure64 Landsat images of erosion in Bangladesh (Nath et al., 2013)

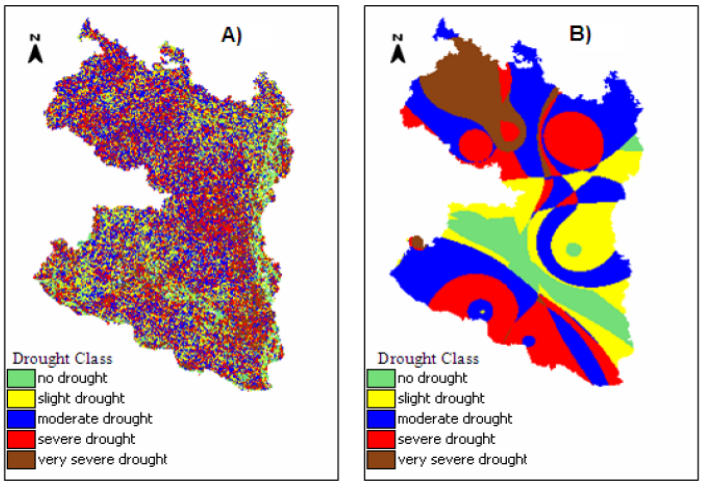


Figure 65 MODIS helping to monitor drought in Bangladesh (Murad et al., 2011)

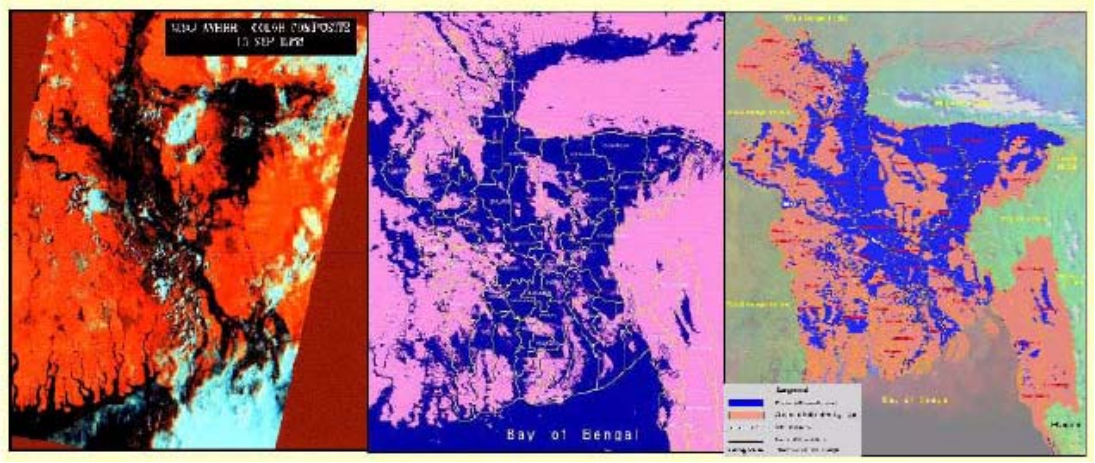


Figure 66 NOAA AVHRR and RADARSAT images of flooding in Bangladesh (Dhar, n.d)

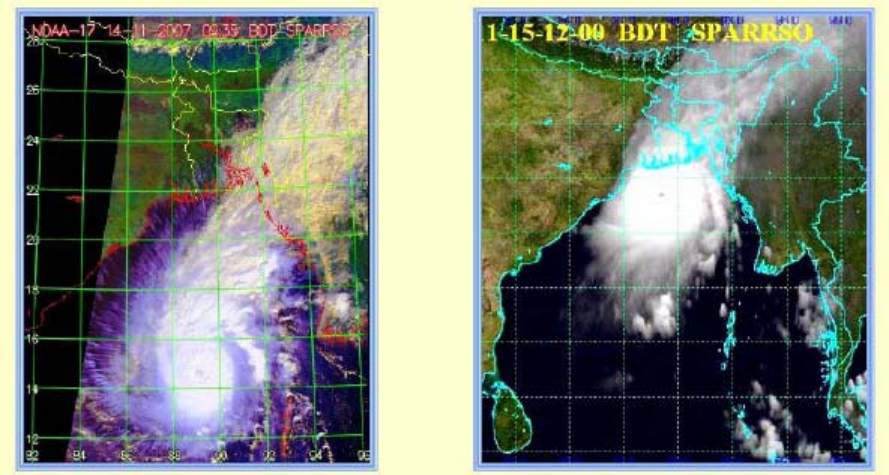


Figure 67 NOAA and FY images of cyclones in Bangladesh (Dhar, n.d)

Bangladesh is a member of GEO (Group on Earth Observations), which means it has access to GEO data coveringnine areas: Disasters, Health, Energy, Climate, Agriculture, Ecosystem, Biodiversity, Water, and Weather (GEO, 2014). The following tablesoutline the 26 Essential Climate Variables (ECV) that are measurable by satellite, as well as what satellites they are made from and what agency owns the satellite which is crucial for accessing data related to CCDM. The following three tables are organized into Terrestrial, Atmospheric and Oceanic variables (Ansdell, 2009).

Table 17 Terrestrial ECVs monitored by satellites

Variable (Terrestrial)	Satellite (agency, orbit, status)	Payload that is able to measure the variable (accuracy)
Lake observation (surface level)	Jason-1 (NASA/CNES, MEO, iO)	Poseidon-2 Altimeter (<4.2cm)
Lake observation (surface level)	Jason-2 (NASA/CNES, MEO, iO)	Poseidon-3 Altimeter (<3.3cm)
Lake observation (surface temperature)	Aqua (NASA, LEO, iO)	AIRS [Atmospheric Infrared Sounder]
Lake observation (surface level)	ENVISAT (ESA, LEO, iO)	RA-2 [Radar Altimeter 2]
Lake observation	Sentinel-3 (ESA, LEO, by 2012)	SRAL [Sentinel-3 Ku/C Radar Altimeter]
Glaciers/Ice caps/Ice sheets (area)	Landsat-5 (NASA, LEO, iO/restricted)	TM [Thematic mapper] (30m)
Glaciers/Ice caps/Ice sheets (area)	Landsat-7 (NASA, LEO, iO/restricted)	ETM+ [Enhanced Thematic mapper plus] (15/30m)

Table 18Atmospheric ECVs monitored by satellite

Variable (Atmospheric)	Satellite (agency, orbit, status)	Payload that is able to measure the variable (accuracy, if applicable)
Surface wind speed/direction	METOP-A/B/C (ESA, Polar LEO, iO/by 2012/by 2016)	ASCAT [Advanced scatterometer] (+/- 2m/s, +/- 20deg)

Variable (Atmospheric)	Satellite (agency, orbit, status)	Payload that is able to measure the variable (accuracy, if applicable)
Surface wind speed	ERS-2 (ESA, Near-Polar LEO, iO/restricted)	AMI/WS [Active Microwave Instrument/Wind Scatterometer] (+/- 3m/s)
Surface wind speed	Aqua (NASA, LEO, iO)	AMSR-E [Advanced Microwave Scanning Radiometer - Earth Observing System]
Surface wind speed	ADM-Aeolus (ESA, LEO, by 2010)	ALADIN [Atmospheric Laser Doppler Instrument]
Surface wind speed	QuikSCAT (NASA, LEO, iO)	SeaWinds?
Surface wind speed	OceanSat-1 (ISRO, LEO, iO)	MSMR [Multi-channel scanning microwave radiometer]

Table 19Oceanic ECVs monitored by satellite

Variable (Oceanic)	Satellite (agency, orbit, status)	Payload that is able to measure the variable (accuracy)
Sea Surface Temperature	Aqua (NASA, LEO, iO)	AIRS [Atmospheric Infrared Sounder]
Sea Surface Temperature	Aqua (NASA, LEO, iO)	AMSR-E [Advanced Microwave Scanning Radiometer - Earth Observing System]
Sea Surface Temperature	METOP-A/B/C (ESA, Polar LEO, iO/by 2012/by 2016)	AVHHR/3 [Advanced Very High Resolution Radiometer]
Sea Surface Temperature	METOP-A/B/C (ESA, Polar LEO, iO/by 2012/by 2016)	HIRS/4 [High resolution Infrared Radiation Sounder]
Sea Surface Temperature	METOP-A/B/C (ESA, Polar LEO, iO/by 2012/by 2016) (ESA, Polar LEO, iO)	IASI [Infrared Atmospheric Sounding Interferometer]

The STEP M process includes merging remote-sensing information with population data to allow the government to determine the links between climate change and migration, to predict CCDM, and strategically plan for it. Figure 13shows climate-change related events overlaid the migration pattern of Bangladesh for 2010; space assets relevant to measuring causes of CCDM also are mentioned.

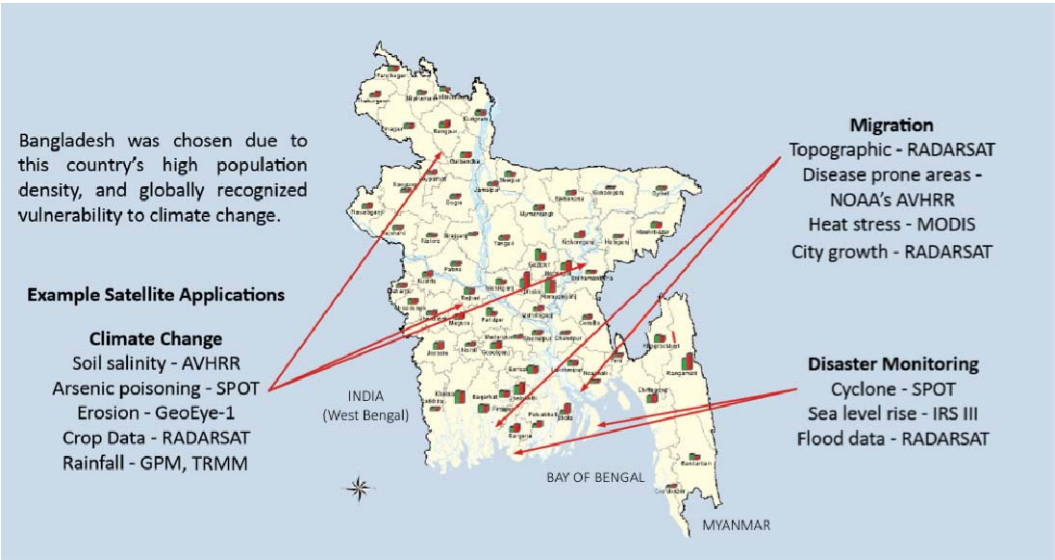


Figure 68 Examples of CCDM Issues in Bangladesh and the 2010 Migration Pattern

6. CONCLUSION

Climate change is expected to cause several global impacts, one of the most critical areas to be addressed is CCDM. STEP M will be useful to LDCs that sees potential in using space systems to provide data on migration and climate change variables. These can be distilled into the important information that policy maker require to properly confront the impacts of climate change.

While the STEP M process details where space resources can apply to a UNFCC NAP dealing with CCDM, this may seem intangible without a concrete application; therefore, STEP M is applied to Bangladesh as a validation case study. For Bangladesh, STEP M identified climate change variables with strong linkages to migration and identified relevant measurements for these variables that can be made using satellite remote sensing systems. The information acquired by these satellite systems can have several applications, such as climate change models, informing national or regional decision makers, and aiding the management of CCDM. As a country that is representative of other LDCs, the recommendations provided can be transferred to other LDCs.

Space-based capabilities are highly versatile and any LDC should not feel constricted to use space assets only for migration, but also should consider space as a tool to provide insights into other national issues. In this way, development of expertise with space capabilities can apply to several aspects of an LDC's development and potentially have positive social and economic impacts.

7. ACKNOWLEDGEMENTS

In addition to the listed authors, this article reflects the work of the 15 other members of STEP M: Alix Dudley, Andrew Alexander, Patricia Randazzo, Michio Hirai, PulkitKanwar, Raul Hernandez, Isaac Llorens, CristelDevrieze, Prateep Basu, Dawoon Jung and Jeremy Milne, Paul Kelly, LawalDanzangi, VatsalaKhetawat and Jan Stastna. Furthermore, the authors are grateful for the assistance of faculty advisers: Mr. J. Nakahara and Professor Emeritus John Farrow, as well as other associated members of the ISU faculty. We also would like to extend our gratitude to Dr. B. Ryan (GEO), Dr. V. Singhroy (NRCAN), Dr. J. Arnould (CNES), and Mr. A. Ismail (IPX) for their suggestions and help with this project.

References

[1] International Organization for Migration (IOM), 2008.Migration and Climate Change.doi: ISSN 1607-338X. Available from:

[2] <http://www.iom.int/jahia/webdav/site/myjahiasite/shared/shared/mainsite/published_docs/serial_publications/MRS35_updated.pdf> [Accessed 21 February 2014]

[3] UNFCCC, 2011. Report of the Conference of the Parties on its Sixteenth Session, 29-10 December 2010 Cancun. Available from: <http://unfccc.int/resource/docs/2010/cop16/eng/07a01.pdf#page=2> [Accessed 6 July 2014].

[4] Ansdell, M., et al.,2009. Climate Links: A Terrestrial Data Collection Network Complementing Satellite Observations. The International Space University.

[5] Dowell M., et al.,2013. Strategy Towards an Architecture for Climate Monitoring from Space. Available from: < http://www.cgms-info.org> [Accessed 6 July 2014].

[6] Gosz, J. R., 1992. Gradient Analysis of Ecological Change in Time and Space: Implications for Forest Management. Ecological Applications 2:248–261. Available at <http://dx.doi.org/10.2307/1941859>.Accessed on 17 March 2014.

[7] Litchfield, W. A., 2010. Climate Change Induced Extreme Weather Events & Sea Level Rise in Bangladesh leading to Migration and Conflict. ICE Case Studies. Available from: <http://www1.american.edu/ted/ice/Bangladesh.html> [Accessed 20 February 2014]

[8] Columbia University, 2014.Bangladesh Natural Disaster Profile. Available from:

[9] <https://www.ldeo.columbia.edu/chrr/research/profiles/pdfs/bangladesh_profile1.pdf> [Accessed 17 February 2014]

[10] Shamsuddoha, M., 2007. Climate change impact and disaster vulnerabilities in the coastal areas of Bangladesh. Dhaka: COAST Trust, Equity and Justice Working Group.

[11] UNEP/GRID. Vital Graphics: Potential Impact of sea-level rise on Bangladesh [online]. Available from: <http://grida.no/publications/vg/> [Accessed 11 March 2014]

[12] CIA, 2013. The World Factbook 2013-2014. Washington, DC. Available from: <https://www.cia.gov/library/publications/the-world-factbook/index.html> [Accessed 15 March 2014]

[13] Siddique, E. M. K. and Adrika, A., 2011. Sustainable Coastal Zone Management of Bangladesh: A Scoping Report for Mangroves for the Future. Available from: <http://www.mangrovesforthefuture.org/assets/Repository/Documents/Sustainable-Coastal-Zone-Management-of-BangladeshMFF.pdf> [Accessed 26 March 2014].

[14] Mahmood, S. A. I., 2012. Impact of Climate Change in Bangladesh: The Role of Public Administration and Government's Integrity. Journal of Ecology and the Natural Environment, 4(8), 223–240. Available from<http://www.gci.org.uk/Documents/Mahmood.pdf> [Accessed 15 March 2014]

[15] Azim Uddin, A. F. M., and Basak, J. K., n.d. Effects of Riverbank Erosion on Livelihood.Dhaka. Available from <http://www.unnayan.org/reports/climate/Effect_of_Riverbank_Erosion_on_Livelihood.pdf> [Accessed 15 March 2014]

[16] Nath, B., Naznin, S.N., Alak, P., 2013. Trends Analysis of River Bank Erosion at Chandpur, Bangladesh: A Remote Sensing and GIS approach, International Journal of Geomatics and Geosciences, 3 (3), pp. 454-463

[17] Murad, H.and Islam, A.K.M.S., 2011. Drought Assessment using Remote Sensing and GIS in Northwest Region of Bangladesh, 3rd International Conference on Water & Flood Management (ICWFM-2011)

[18] Dhar, S, n.d. Bangladesh Space Research and Remote Sensing Organization (SPARRSO), North South University, Available from: <http://shaurov.weebly.com/uploads/2/3/8/7/2387470/presentation_on_sparrso.pdf> [Accessed 3 March 2014]

[19] GEO, 2014.Available GEO satellite resources illustration. Available from:

[20] <http://www.earthzine.org/geo-and-geoss-the-group-on-earth-observations-and-the-global-earth-observations-system-of-systems/> [Accessed 3 March 2014]

Data Transmission Scheme for Micro-satellite Formation and Network

Li Peng, Zhong Lingling

Tianjin Jinhang Computing Technology Research Institute
Tianjin 300308, China

Abstract: The micro-satellite formation and network is one of the focuses of the current space-based platforms research field. In this paper, the data transmission scheme for micro-satellite formation and network is achieved based on the next generation of tactical networking research. The overall scheme, research objectives, technical specifications and main contents are described, and the space borne terminal equipment is mainly introduced. Furthermore, the key technology and solution of the scheme is particularized. The data transmission scheme for formation and network is realized by distributed, no center and IP-based networks, enabling multiple platforms dynamic network, which has a high-speed, broad bandwidth, low time delay characteristics. And what's more, the number of the network members can be more than 20 and data transmission network has high interference immunity, anti-interception capability. The data transmission scheme for micro-satellite formation and network can achieve high transfer rates between micro-satellite platforms. As a result it has a good prospect in the space-based platform data link communication field.

Keywords: Micro-satellite; Formation and network; Data transmission

I. INTRODUCTION

Space-based platform is one of the major projects in the current national focus on the strategic development, and the micro-satellite formation and network technology is an important research direction of space-based platforms. Micro-satellite has light weight, small size, low cost, short development cycle, and many other advantages, and thus subject to a high degree of importance to national space sector. Development of micro-satellite technology is of great significance to the construction of the national economy and national defense systems. In the military field, micro-satellite is indispensable for the future information warfare. In civilian areas, micro-satellite can be used for technology demonstration, scientific research, space exploration, satellite communications and earth observation. Since the 20th century, the mid-1980s, more than 300 modern micro-satellites have been launched, and the development of the industry has made great contributions.

Functions of single micro-satellite cannot be compared with the large satellites. But with the increasingly complex functions, the development of increasingly high cost of large satellites, the risk is growing. The micro-satellite constellation consists of multiple satellites can not only complete the functions of a single large spacecraft, but also significantly reduces costs and risks.

Compared with big single satellite, the advantage of micro-satellite constellation is as follows:

(1) Single micro-satellite is light and simple, and it can effectively reduce the cost of manufacturing and launching satellites, while simplifying daily operation and maintenance, thereby reducing the life cycle cost of satellites.

(2) The micro-satellite constellation is capable of higher performance tasks by intercommunicate of several satellites.

(3) Compared with the single satellite, distributed architecture is more tolerant of a single point of failure. When a virtual satellite fails, it can be excluded from the reconstruction of the system, to maximize the elimination of faults [1] [2].

With the emergence of micro-satellite constellation, smaller, faster, cheaper concept of user space will soon be tested. If three or three more satellites fly in formation as a group, a variety of tasks can be completed. For high target, to strengthen and enrich the capacity of micro-satellite in orbit, we need a comprehensive breakthrough in the new formation and network technology to solve the overall design of the satellite constellation network. The proposed scheme is focused on solving micro-satellite data transmission problems between each other.

According to the development trend of micro-satellite constellation, the research of micro-satellite data transmission is absolutely necessary and timely. Integrated Defense Advanced Research Projects Agency and information demonstration test of Rockwell Collins Company disclosed that TTNT (Tactical Targeting Network Technology) core technology uses new technology focused on highly dynamic Internet technology and large capacity, high dynamic, anti-interception, and interference of information processing technology. Micro-satellite formation and network technology use TTNT technology for reference. The IP-based network architecture is introduced for satellite network. The architecture has been widely used in the civilian Internet, and the technology is very mature. The micro-satellite constellation formation and network needs can be realized by making appropriate improvements of IP-based network architecture. As a result, it can greatly shorten the development cycle, while reducing development costs.

II. Research Status and Technology Trends

In recent years, many countries use globally distributed constellation of micro-satellites to achieve a single large satellites function.

On the military side, modern micro-satellite shows unique advantage in local emergency war by its fast and flexible launch manner and very short development cycle. In addition, the ground resolving power of modern small tactical imaging is better than 1m, and the covering width is up to hundreds of kilometers which is equal to the previous large satellites, but the quality is only 200 ~ 300kg and the life is up to five years. Reconnaissance satellites using multiple micro-satellites in low orbit satellite network can monitor the sensitive areas in real-time, which play a vital role in local or surrounding war. As in the Kosovo war, the US military use satellites imaging satellites system including KH12, lacrosse, Ra l and other micro-satellite get a lot of information.

In 2025 the Air Force's combat envisaged, micro-satellite constellation system will assume the information service provides real-time continuous task to combatants. US Department of Defense takes the small communications satellites, reconnaissance and early warning satellites as the focus of micro-satellite development. Among them, the most representative is SBIRS-Low includes 20 to 30 small low-orbit satellites and Starlite system includes at least 20 micro-satellites.

British Ministry of Defense also have a plan for using hundreds of satellites to form low-orbit satellite network, to meet the needs of combat forces tactical communications. It consists of 50 to 100 satellites, terrestrial users UHF (UHF) sends information to the satellite, and the information can be fast transferred to the destination.

On the civil side, a communication satellite network constituted by a plurality of micro-satellite not only enables global direct communication, the phone has become a truly global communication, but also has low transmit power, low delay and no dead ends, etc., the market prospect is very broad. Remote sensing satellite constellation of micro-satellite, which have a high temporal and spatial resolution, in obtaining information related to environmental protection and other aspects of human existence, has a very distinct advantage. The US Globalstar constellation is composed of 48 low-orbit satellites distributed in 8 orbital planes of composition, which can provide global seamless coverage of voice, fax, data, SMS and other mobile communications services. When the number of satellites in orbit business failure, the company can pass the orbit and other means to restructure satellites in orbit to form a new constellation of 40 satellites distributed in 8 orbital planes. India has the world's largest global remote sensing satellite constellation at present. It is composed of 66 satellites, currently in orbit are IRS-1C, IRS-1D, IRS-P3, IRS-P4 and TES satellites[3] [4].

III. Technical Solution to Achieve Data Transmission Scheme for Micro-satellite Formation and Network

A.Overall Scheme

For the realization of micro-satellite formation and network data transmission, the traditional coding and data processor is difficult to meet the multi-satellite formation flying networking needs, and the traditional communications equipment is not conducive to multi-satellite flight management. The establishment of inter-satellite communication links needs to process a lot of real-time communication data and control instructions.

The micro-satellite formation and network data transmission referred to herein are based on an IP-based Internet network which is high-speed, dynamic, and without centers, enabling multiple platforms (satellites) network architecture. The software architecture, interference, signals processing, integrated miniaturized key technologies are researched. The self-organizing network core terminal device is researched by the experience of successful product technology, and the development project plan is completed according to the engineering method.

B.Research Objectives

In this paper, the data transmission scheme for micro-satellite formation and network is in-depth studied based on the next generation of tactical networking research. The specific research objectives are: network topology, miniaturized network terminal, special routing protocols, and the distributed, no center, IP-based networks.

C.Technical Specifications

(1) Internet Protocol network, high-speed, dynamic and non-center-based, enabling dynamic network across multiple platforms.

(2) Including text conversation, streaming video and still images, including various types of IP applications.

(3) High-speed, broadband, low latency.

(4) The number of the support t users is large. Network can support more than 20 members.

(5) Have high anti-jamming, anti-interception capability.

D.The Main Contents

(1) System control and media access

The main task is to research the users of the network how to network, how to manage the transmit power, how to build / dismantle logical link and so on. For power control and management,

the need is to reduce energy consumption and emission power in the premise of ensuring network connectivity. For channel access control, the rational decision-channel methods can overcome the "Send conflict" and "Hidden terminal."

(2) Quality of service

The main task is to research how to guarantee a set of services to meet predetermined constraints such as end to end delay, jitter and packet loss rate.

(3) IP-based routing protocol

Satellite networks is a special wireless network, compared with other wireless terrestrial network, it has several notable features: 1) The distance between networks nodes is fast; 2) It need not only transfer data between the two nodes, but also for measuring distance; 3) Poor network maintainability. The features determines that the inter-satellite networks need to adopt a flexible network protocol in order to meet the demand for high reliability and flexible networking space. The IP-based routing protocol can meet the demand.

(4) Frequency bands selection

Space radio technology currently being developed to a higher band, there are two main reasons: 1) The higher frequencies and more particularly Ka-band has not been large-scale development, with less interference and easy to apply; 2) High-frequency itself has many advantages such as the devices has small size and light weight.

(5) Joint design of satellite antenna and satellite system

The overall design of the satellite link between the satellite constellation are inextricably linked, especially for narrow beam system, the joint design of the antenna and the satellite system will be the key project implementation. The main study included: 1) Joint design of satellite antenna and constellation network topology; 2) Iterative design of satellite antenna and assembly; 3) Joint control of satellite antenna and the satellite attitude.

(6) Anti-jamming technology

The main task is to research hybrid anti-jamming technology. The hybrid anti-jamming technology of frequency hopping + time hopping method is adopted, which has three anti-jamming capabilities, including the whole band of anti-jamming capability, frequency targeting anti-jamming capability and tracking anti-jamming capability.

E. Spaceborne Terminal Equipment

Spaceborne terminal equipment is critical core hardware for the scheme, which is mainly composed of spaceborne equipment terminals, spaceborne RF front-end, RF filter and receiving and transmitting antenna. Taking into account the limited space on a micro-satellite, the data link hardware resources should be fully integrated. Thus the image acquisition, image compression, baseband coding, data modulation, signal conversion, power amplification, low-noise receiver, baseband demodulation and other functions should be integrated into a spaceborne terminal equipment to achieve miniaturization and integration.

(1) Spaceborne equipment terminals: Receiving the image information, to complete the image source coding, channel coding and decoding, band pass signal processing, jump spread spectrum processing, the IF signal modulation and demodulation, frequency-changing the received IF signal to RF output, frequency-changing the received RF signal to IF output, and to achieve filtering and amplification in the transformation process, thus realize multi-channel control, while achieving the address and ID number stapling and other information.

(2) Spaceborne RF front-end: The RF signal is multi-stage amplified by the transmit channel in accordance with the system requirements. The signal takes enough gain provided by the final stage amplifier, and then output to the antenna after the filtered to realize wireless transmitting. The receiving channel is switched by the switching, and receives the microwave signal from the antenna. Then the signal is filtered, amplified and output to the RF terminal, as the whole of half-duplex operation is realized.

(3) RF filter: To ensure good electromagnetic compatibility of spaceborne equipment, the RF filter is set up between RF front-end and antenna. The filter uses the form of cavity, and the resonator structure is composed entirely of a mechanical cavity body filter with a very high Q value. It's very suitable for the present system require low insertion loss, high power transmission applications.

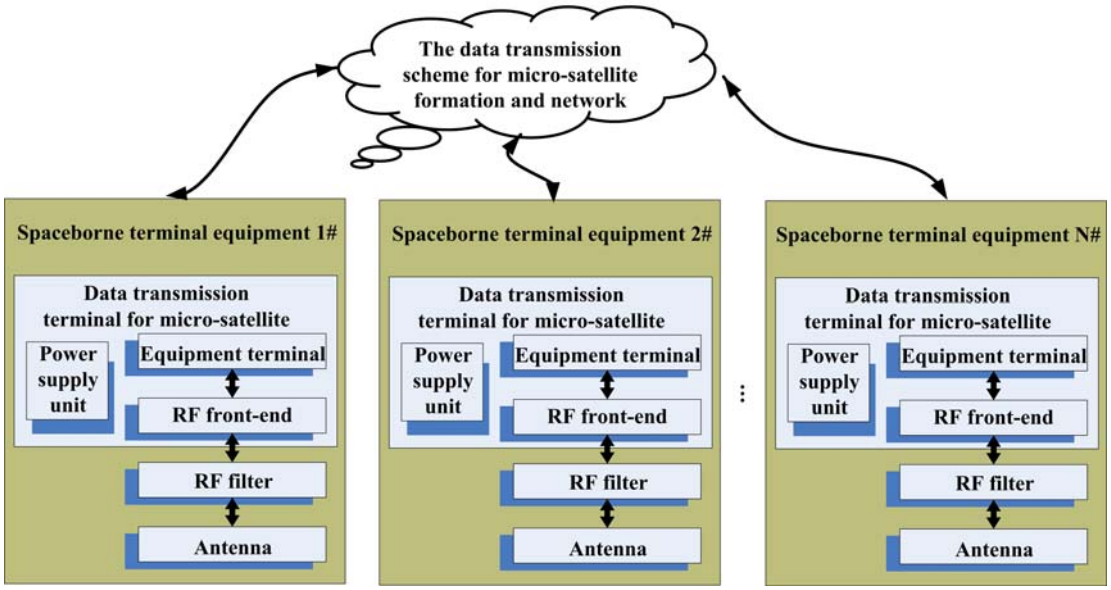


Figure 1. Spaceborne terminal equipment

IV. Key Technology and Solutions of the Data Transmission Scheme for Micro-satellite Formation and Networking

A. Measures to Improve the Quality of Service: Statistical Priority Access Control (SPMA) Technology

In order to ensure quality of service, the type of service is mapped to different service levels in accordance with the importance. The busy state of channel is monitored using the channel statistical methods in real-time and packets are controlled using SPMA protocol. The lower priority traffic is controlled when the channel overload, so the high priority, low-latency communication quality of service is assured. Various priority services are able to send on time when the channel is not busy. Under SPMA control the network load is maintained at a relatively stable level, so it provide ongoing robust network performance in a dynamic environment, to ensure the system's first access success rate.

B. Measures to Improve the Carrying Capacity of the Network: the Single-transmitter Multiple-receiver Parallel Communication Working Mode

The specific mode is: Each user controls pulse sending time in accordance with the duty ratio. The send pulses are randomized in the time domain and frequency domain. A single user can send messages to other users within the network, and can receive a plurality of users' information.

Compared with the traditional single-transmitter single-receiver communication system, single-transmitter single-receiver system network capacity = information rate of single user; single-transmitter multiple -receiver system network capacity = information rate of single user × number of users of the network. We can see that single-transmitter multiple -receiver system network capacity is more than single-transmitter single-receiver system network capacity.

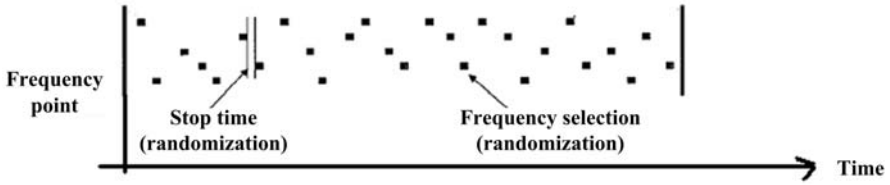


Figure 2. Single-transmitter multiple-receiver parallel communication working mode

C. Measures to Support the High-speed Information Transmission Capability: Constant Envelope Efficient Joint Coding Modulation System

The system includes multi-band continuous phase modulation and coding modulation techniques. Multi-band continuous phase modulation technique has the following advantages: Compared with the binary system, the higher rates of information transmitted in the same frequency band; Constant envelope for nonlinear channel; Out-band radiation is small; Side lobes decay quickly. Coding modulation techniques has the following advantages: The coding and modulation is optimized as a whole design; The signal Euclidean distance between the points can be expanded without increasing signal transmission power, so that the receiving end signal to noise ratio is maximum.

D. Measures to Enhanced Anti-jamming Capability: Hybrid Anti- jamming Technology of Frequency Hopping + Time Hopping Method

It has three anti-jamming capabilities, including the whole band of anti-jamming capability, frequency targeting anti-jamming capability, tracking anti-jamming capability. The whole band of anti-jamming: By hopping bandwidth of up to several megabytes of design, effective against full-band interference. Frequency targeting anti-jamming: By shortening the dwell time for each frequency, only up to tens of microseconds, and the continuous frequency hopping two points spaced frequency hopping to reduce the adjacent point frequency dependency, improve the ability of the system aimed at anti-jamming frequency. Tracking anti-jamming capability: By using of single-transmitter multiple-receiver communication mode, ensure that there is more than one user channel parallel transmission of business information, thus tracking for single user interference becomes more difficult.

V. CONCLUSIONS

Nowadays space technology is developed from a single satellite to the distributed satellites. Formation flying is the latest application of distributed satellites. As new applications continue to emerge, it’s not only to be a huge economic, defense and social benefits, but also made many new requirements to accelerate development. This paper discusses the implementation of data transmission scheme for micro-satellite formation and networking. It is realized by distributed and no center IP-based networks, enabling multiple satellite platforms dynamic network. As the combat platforms will be extended to the field of space-based, the warfare will be shifted from the "platform-centric warfare" to information technology-based "network-centric warfare". The achievement of the scheme will significantly improve the communication and information

equipment level of the military. It will greatly improve long-range precision strike capability and rapid response capability, and the military benefits are obvious. For the civil application, the data transmission scheme for formation and network also has a good prospect in mobile communications, terrestrial/coastal monitoring and agriculture management.

References

[1] Lin Laixing. Technological development and application prospects of distributed micro-satellite system [J]. Spacecraft Engineering, 2010(1), 60-66

[2] Deininger W D, Noecker M C, Wiemer D J, et al. Space technology-3: Mission overview and spacecraft concept description [J]. Acta Astro nautical, 2003(52), 455-465

[3] Schar f D P. Future technology directions-precision formation flying mission and technologies [R]. NASA JPL Tech Paper, 2008

[4] Leitner J. Spacecraft formation flying-An overview of mission and technology challenges [C]. AAS07-031, 2007

The Opportunities and Challenges of China's Commercial Communication satellite in global landscape

Liu Yue

Beijing Space Science & Technology Institute

1. Introduction

With the trend of Globalization in the Space Industry, China Space is actively seeking international development opportunities. By the virtue of increasingly mature space technology and improving infrastructure, Communication Satellite has become a pioneer in China's Space globalization development. With DFH-3 and DFH-4 communication satellite platform being mature and more powerful products under development, China's communication satellite has been in several competitions in international commercial communication satellite tenders, and has been awarded more than 10 commercial contracts [1].

In order to meet the international development requirements, China Space is also transforming itself. China Space is accustomed to complete scientific research tasks in the long term, but now they need "doing business" in the international market. Technology development is no longer the only factor that they need to consider, but they have to learn more about the market and think more about the profit. Therefore, an in-depth analysis of the strength and weakness in the market, a clear judgment of the competitive position are necessary. These issues will be important for future development of China's communication satellite in the global market.

This article reviews and summarizes the advantages & opportunities of China' communication satellite in global commercial competition, and analyzes the difficulties and challenges faced at the present stage, so as to provide advices and references for the future development of China' communication satellite.

2. Strengths and Opportunities

2.1 Take "One Belt and One Road" project as an opportunity, China's communication satellite have huge development potential.

At present, China is vigorously promoting the "Silk Road economic belt" & the "21st Century Maritime Silk Road" construction, where passing through areas with mountains, oceans and deserts. With complicated topography area, such areas have communication and transportation problems. Satellite coverage in these areas is also poor. Most current satellite communications is narrow band, which cannot meet the demand of big data transportation, commercial & entertainment needs spurt. China Space is now studying using Ka-band broadband satellite to set up an information highway in such region, so as to provide information support for Chinese enterprises and staff. "One Belt and One Road" project could be a tremendous opportunity for the globalization of China's communication satellite.[2]

2.2 China's communication satellite export benefits by Space diplomacy, which has achieved fruitful results.

Compared with other space-faring nations, China's Space Globalization has different meaning. China takes satellite manufacturing & launching as part of overall diplomacy measures with other

developing countries, which brings income and resources while improving these countries' capabilities in key areas of communications and remote sensing, reflecting the Chinese "win-win" diplomacy philosophy. China is attracting a growing number of emerging space-faring countries to cooperate with China and purchase Chinese communication satellite. Even China national leaders have promoted the export of communication satellite. For example, in May, 2015, during the Prime Minister's visiting four countries in South America, China has signed a space cooperation agreement with Brazil, Mexico and Peru, including the communication satellite and relevant applications agreement. [3] China Space is also attempting to embed the satellite communication applications into high-speed rail export, which will bring more opportunities to China's communication satellite.

2.3 Developing country users prefer China's "Package" Service and financing models.

As the target users are mainly developing countries, China adopts special marketing competition strategy. Due to less experience participating in space activities of these countries, China has adopted bundled service strategy, namely bundling satellite manufacturing, satellite launching, ground segment, concessional loans and customer training services together to provide customers with one-stop in-orbit delivery. These countries are usually lack of funds but rich in natural resources. China use satellite in-orbit delivery to exchange required raw materials, such as some earlier Chinese satellite export project with Venezuela & Nigeria.[4] Compared to some western satellite manufactures insisting on profits first, these measures are more attractive for developing countries.

2.4 International Human Resources provide adequate power for China Space enter the world market

From the aspects of human resources, China seems to have more advantages compared with western countries and other emerging space-faring nations. For example, the president of satellite manufacturer Loral Space System has publicly expressed such concerns. A growing number of Chinese students studying engineering and other space relevant majors graduate from the western top universities each year, return to China and join the China Space teams, as the United States refused to granted these students a green card.[5] This made China has a large number of high-quality young talent space engineers. Compared with aging problem in the US and Europe, the average age of Chinese Space engineers are much younger. Not to say in Brazil & other emerging space nations. [6] Their high-quality talent usually chose to go to the US and Europe after graduation for better research environment and pay, resulting a huge loss of talent for the country. Perhaps at this time short-term advantage of the talent is not obvious, but it will surely show the effect in a decade or longer.

3. The difficulties and Challenges

3.1 Emerging Internet enterprise flooded the market, traditional space enterprises facing fierce competition

With more and more US Silicon Valley internet enterprises entering into traditional communication satellite market, new industrial landscapes are accelerated forming. The new trends are to promote the more efficient combination of technology, industry and capital. Compared with ever-changing ground communication markets, the development of traditional communication satellite is slow. There is no essential change compared with five years before. Due to large satellites in geostationary orbit is capital intensive and risky, traditional satellite manufacturers fearing of risk are not willing to innovate and upgrade the technology. In recent years, with the mature of emerging

satellite manufactures, the competition becomes even more intense and the opportunities are even less than previous years. Emerging Satellite Interment companies such as OneWeb & SpaceX plan to bring the capital, technology and experiences accumulated in other fields to communication satellite market. [7][8]These companies would put the customer needs first, adjust satellite designs and reduce the costs to get more profit. Undoubtedly, the traditional satellite manufacturers are under pressure to change.

3.2 Lack of innovative products may squeezed Chinese communication satellite market space again

US space has problems in the past 20 years, especially the high satellite manufacturing & launch costs. However, with the development of all-electric propulsion satellite platform and Falcon series rockets, the United States is transforming its technology strength into business competitiveness, so that satellite development costs will be further reduced. China should closely follow and study the new technologies which can reduce the cost of the satellite. If China cannot catch up, its profit margins will be less and its market place will be squeezed again.

3.3 Local Protection made Chinese satellite manufacturers difficult to enter global commercial communication satellite market.

Local Protection is quite common in communication satellite industry. Satellite platforms are mainly purchased by the users of local country or of the region. Government's strong support is key to satellite manufacturer's competition in global market. Through the past years, world commercial satellite market has established its orders. The user and provider relationship are quite fixed. It would be very hard to break the existing partnership and to gain more market share. For example, Loral Space System, Inc.(SS/L) and Telesat are used to belong to the same parent company. Therefore, although Telesat has the right to select providers freely, most of its orders were awarded to SS/L. Many satellite manufacturers have shares of large satellite operators. Chinese satellite manufacturer currently does not take advantage of breaking these alliances and long-term relationship.

3.4 US trading policy block China participate in International cooperation

For a long time, the United States use export license, International Arms Trading Regulations (ITAR) to prohibit China to provide launch services for US satellites. [9]Other than that, the US satellite export restrictions also covers commercial communication satellites using US technologies Export or Re-Export to China. These initiatives have increased the difficulty of China's participation in international cooperation. However, Russian satellite manufacturer ISS-Reshetnev has developed capability through cooperation with European Company Thales Alenia Space (TAS). They have extended their existence to other countries in Central Asia region. For China, export control to some extent encourages the development of national space industry, but increased the difficulty of international cooperation and global purchasing satellite parts and components. In the short term, these measures have negative effect on Chinese communication satellite competitiveness in international market.

4. Implications and Recommendations

4.1 With more intense competition in the market, China Space must not solely depend on the government investment. China Space should learn from the emerging companies, making great efforts to encourage innovation & gain more competitiveness.

With more & more companies entering into satellite industry, the existing resources, technology, channel monopoly will gradually be broken. China space companies could not solely rely on

government's investment any more. Marketization requires traditional satellite manufactures explore how to convert current advantages (including technology & experiences) to new competitiveness. China's satellite manufacturers must learn as the start-up entrepreneurs and keep the same passions for innovative ideas, management manner and commercial investment, so that they could keep its positions in competition. Future competition will not be easy. Innovation requires a lot of capital investment, manpower, time and tolerance of failure. Chinese Space must recognize that although many attempts will fail, but many of them would become the cornerstone of future satellite designs. For companies which make great efforts to encourage innovation in business, the rewards will be great and the market will eventually prove it.

4.2 Taking "Internet Plus" and "One Belt and One Road" project as an opportunity, developing satellites in different orbits to provide internet connection services, Chinese satellite industry will expand rapidly.

For a long time, an important reason for China's communication satellite industry falling behind is Chinese policy restriction, which resulting domestic market demand cannot be released, thereby affecting the capabilities of ground equipment manufacturing, system integration and network operating. However, in western countries, satellite TV has formed a large scale industry. China not only has missed a hundred billion satellite broadcast television opportunity, but also has widened the gap with US & European countries. Now, Chinese government is vigorously developing "Internet Plus" and "One Belt and One Road" Policy. Communication satellite with large coverage and less likely influenced by geological conditions are the most suitable for solving communication problems in these regions. Under the influence of favorable national policies, the market demand will be released and new huge market traction will be formed. China Space should seize the historical opportunity and promote the prosperity of satellite industry.

4.3 Making full use of China Space image brand, China Space could draw from Chinese ground communication equipment providers' past experiences, and co-develop oversea foreign market.

China Space represents China National Image, on behalf of the highest technical level and high-end manufacturing capabilities. Communication satellite as the most important space infrastructure, on one hand, it could be bundled export with other high-end equipment like high-speed railway, nuclear power and so on; on the other hand, it could work closely with ground communication equipment providers such as HUAWEI and ZTE to expand the world market. China Space not only could use their resources and experiences, but also carry out in-depth integration with their products. China's Space must carefully consider ground system in future satellite system design, while improving its openness and comprehensive abilities to achieve complementary to each other.

Reference

[1] China Great Wall Industry Corporation [EB/OL].
http://cn.cgwic.com/In-OrbitDelivery/CommunicationsSatellite/Program/VeneSat-1.html
[2] Satellite Makers See Chinese Competitive Threat Differently [EB\OL].
http://spacenews.com/satellite-makers-see-chinese-competitive-threat-differently/,March 19, 2013
[3] The Space Report 2014, Space Foundation, 2014
[4] OneWeb Taps Airbus To Build 900 Internet Smallsats [EB\OL].
http://spacenews.com/airbus-wins-oneweb-contract/, June 15 ,2015
[5] SpaceX Confirms Google Investment [EB\OL].
http://spacenews.com/spacex-confirms-google-investment.January 20, 2015

[6] Loral Urges U.S. To Block Thales Alenia Use of Chinese Launcher [EB\OL].
http://spacenews.com/loral-urges-us-block-thales-alenia-use-chinese-launcher/, June 29, 2004

Trajectory Reconstruction Based on GPS Measure Data for Rocket Projectiles

BIAN Wei-wei, Ge Xiao-fei, XIN Zhen-fang, ZHAO Jin-xin

Beijing Machine and Equipment Institute, Beijing 100854, China

Abstract: In order to estimate the flight trajectory for rocket projectiles accurately so as to provide navigation information for control and guide system effectively by using GPS measure data, the Kalman filtering trajectory model was established in trajectory reconstruction by using Kalman filtering method and particle trajectory model. The parameters of a section flight trajectory detected by the projectile-borne GPS were filtered, and the trajectory falling point was obtained by the trajectory reconstruction. The simulation results show that, after Kalman filtering processing, the accuracy of trajectory reconstruction will decrease with the increase of range. Increasing the GPS tracking time can meet the trajectory accuracy requirements.

Keywords: GPS; trajectory reconstruction; Kalman filtering; numerical simulation

1 Introduction

Accurately measuring of the trajectory and estimating the falling point/the trajectory deviation are very important to the commander that it not only provides a basis to make correct firing decisions, but also make shooting correction rapidly and accurately. It mainly used artillery positioning radar to measure trajectory [1]. With the development of science and technology, especially the development of GPS, the accuracy and efficiency of the external trajectory measure are improved continuously. Usually, it mainly uses pseudo range measurement and carrier Doppler measurement to measure the position and velocity of projectile for trajectory measurement [2]. It can be used to predict the falling point, provide a basis for projectile control, and post processing analysis.

As projectile inevitably affected by some disturbances in the flight process, a large deviation exists between the actual flight state and the flight dynamics model. The application of GPS in trajectory measurement for guided rocket projectiles was studied in the subject literatures [3, 4], but the errors caused by disturbances or noises were not considered. The signals of radar measurement were used in the application of Kalman filtering in calculation of trajectory falling point for trajectory correction projectiles [5]. The signals of GPS measurement were used as state variables in the application of Kalman filtering in GPS-guided rocket projectiles[6]. This report studies the problem of trajectory reconstruction for a rocket projectile by using Kalman filtering based on GPS measure data.

2 Trajectory reconstruction principles

The Fig.1 illustrates the principles of trajectory reconstruction: it corrects the parameters of trajectory model and then estimates the falling point based on high reliability and availability GPS measure data.

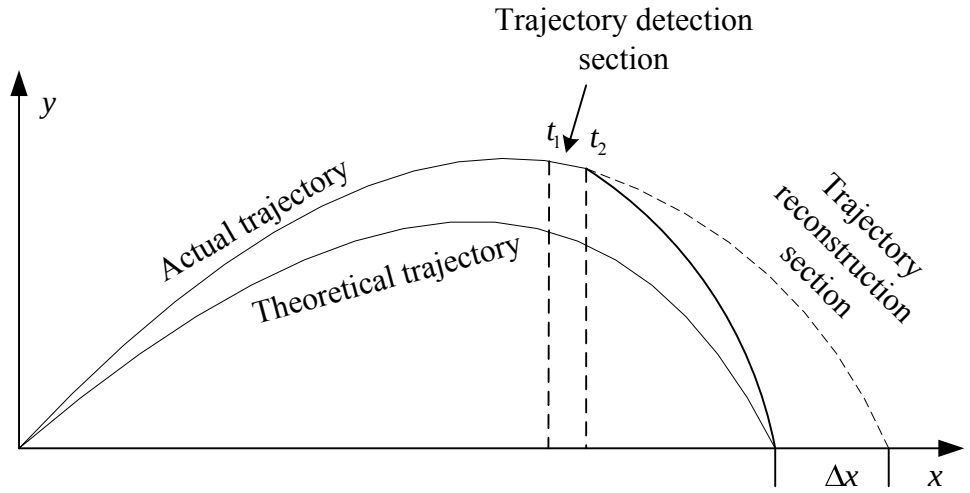


Fig.1 Illustration of trajectory reconstruction

GPS tracks the whole trajectory, collects measure data from t_1 to t_2, and adopts specific method to filter the trajectory. While the accuracy meets the requirement at t_2, then it takes the results of trajectory parameters filtering as the initial values to reconstruct the trajectory and predict the falling point. It will be closer to the actual trajectory by establishing trajectory model based on high accuracy data. Obviously, trajectory reconstruction will be an important technology to improve the control accuracy when trajectory control gets more and more attention.

3 Trajectory filtering models

3.1 System state equation

Trajectory control procedure can be described as follows: tracking and detecting the flight projectile by the projectile-borne GPS, processing the measure data, calculating the position of falling point, getting the trajectory deviation by compared with the target position, sending instruction to the projectile-borne control device after simple logic processing, and finally implementing the correction at a certain suitable moment by the projectile-borne actuator. It needs trajectory models in the process of GPS measure data filtering and trajectory reconstruction. Considering the rapidity/real-time of trajectory calculation, and GPS measuring parameters could be provided, the particle trajectory model can be used as the system state equation for trajectory filtering [7, 8 , 9]:

$$\begin{cases} \dot{x} = v_x \\ \dot{y} = v_y \\ \dot{z} = v_z \\ \dot{v}_x = -cH_\tau(y)G(v_\tau)v_x \\ \dot{v}_y = -cH_\tau(y)G(v_\tau)v_y - g \\ \dot{v}_z = -cH_\tau(y)G(v_\tau)v_z \end{cases} \tag{1}$$

Where:

x, y, z are the position components and v_x, v_y, v_z the velocity components of projectile in the ground coordinate system, and $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$;

c is the trajectory coefficient;

$H_\tau(y)$ is the air density function;

$G(v_\tau)$ is the drag function;

g is the gravity acceleration;

τ is the virtual temperature;

v_τ is the virtual velocity, and $v_\tau = v\sqrt{\tau_{0n}/\tau}$ in which τ_{0n} is the standard virtual temperature of ground surface and τ is the virtual temperature at height of y.

Define state variable:

$$\boldsymbol{X} = [x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad x_7]^{\mathrm{T}} = [x \quad y \quad z \quad v_x \quad v_y \quad v_z \quad c]^{\mathrm{T}},$$

Define $\dot{c} = 0$, then equation (1) can be written as:

$$\dot{\boldsymbol{X}} = \boldsymbol{f}(\boldsymbol{X}) = \begin{bmatrix} x_4 \\ x_5 \\ x_6 \\ -x_7 H_\tau(x_2) G(v_\tau) x_4 \\ -x_7 H_\tau(x_2) G(v_\tau) x_5 - g \\ -x_7 H_\tau(x_2) G(v_\tau) x_6 \\ 0 \end{bmatrix} \tag{2}$$

Some error always exists because of nonlinear equation (2) is a approximate description of the particle motion for rocket projectiles. $\boldsymbol{V}$ which is a zero mean Gauss white noise is introduced to compensate part of former error, and $\boldsymbol{V} \sim N(0, \boldsymbol{R})$. Therefore, equation (2) can be written:

$$\dot{\boldsymbol{X}} = \boldsymbol{f}(\boldsymbol{X}) + \boldsymbol{V} \tag{3}$$

3.2 System measurement equation

GPS works in inner earth coordinate system. It measures the velocity components v_{xs}, v_{ys}, v_{zs} and position components x_s, y_s, z_s of projectile (Fig.2).

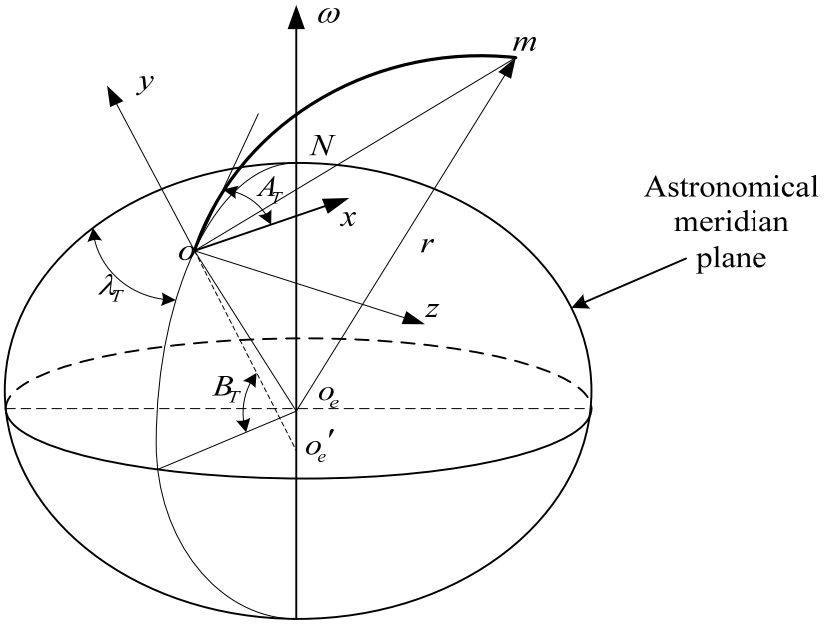


Fig.2 Inner earth coordinate system and ground coordinate system

Figure 2 shows the transform between inner earth coordinate system and ground coordinate system. Coupled with the conversion of velocity, the transform can be given:

$$\begin{cases} x_s = f_{11}(R_x + x) + f_{12}(R_y + y) + f_{13}(R_z + z) \\ y_s = f_{21}(R_x + x) + f_{22}(R_y + y) + f_{23}(R_z + z) \\ z_s = f_{31}(R_x + x) + f_{32}(R_y + y) + f_{33}(R_z + z) \\ v_{xs} = f_{11}(R_x + v_x) + f_{12}(R_y + v_y) + f_{13}(R_z + v_z) \\ v_{ys} = f_{21}(R_x + v_x) + f_{22}(R_y + v_y) + f_{23}(R_z + v_z) \\ v_{zs} = f_{31}(R_x + v_x) + f_{32}(R_y + v_y) + f_{33}(R_z + v_z) \end{cases} \tag{4}$$

With R_x, R_y and R_z representing the position coordinates of the inner earth in the ground coordinate system are known variables.

$$\begin{cases} f_{11} = -\sin A_T \sin \lambda_T - \cos A_T \sin B_T \cos \lambda_T \\ f_{12} = \cos B_T \cos \lambda_T \\ f_{13} = -\cos A_T \sin \lambda_T + \sin A_T \sin B_T \cos \lambda_T \\ f_{21} = \sin A_T \cos \lambda_T - \cos A_T \sin B_T \sin \lambda_T \\ f_{22} = \cos B_T \sin \lambda_T \\ f_{23} = \cos A_T \cos \lambda_T + \sin A_T \sin B_T \sin \lambda_T \\ f_{31} = \cos A_T \cos B_T \\ f_{32} = \sin B_T \\ f_{33} = \sin A_T \cos B_T \end{cases}$$

With B_T is astronomical latitude; λ_T is astronomical longitude; A_T is azimuth angle of astronomy.

Define measurement variable:

$$\boldsymbol{Z} = [x_s \quad y_s \quad z_s \quad v_{xs} \quad v_{ys} \quad v_{zs}]^{\mathrm{T}}$$

Then, system measurement equation can be written:

$$\boldsymbol{Z} = \boldsymbol{h}(\boldsymbol{X}) + \boldsymbol{d} = \begin{bmatrix} f_{11}(R_x + x_1) + f_{12}(R_y + x_2) + f_{13}(R_z + x_3) \\ f_{21}(R_x + x_1) + f_{22}(R_y + x_2) + f_{23}(R_z + x_3) \\ f_{31}(R_x + x_1) + f_{32}(R_y + x_2) + f_{33}(R_z + x_3) \\ f_{11}(R_x + x_4) + f_{12}(R_y + x_5) + f_{13}(R_z + x_6) \\ f_{21}(R_x + x_4) + f_{22}(R_y + x_5) + f_{23}(R_z + x_6) \\ f_{31}(R_x + x_4) + f_{32}(R_y + x_5) + f_{33}(R_z + x_6) \end{bmatrix} + \boldsymbol{d} \tag{5}$$

Where:

$\boldsymbol{d}$ is the GPS measure noise, and $\boldsymbol{d} \sim N(0, \boldsymbol{Q})$.

3.3 Kalman filtering trajectory equations

Equation (2) and equation (5) are nonlinear, but Kalman filtering is only applied to linear system. Consequently, Kalman filtering trajectory equations can be derived only when equation (2) and equation (5) are linearzed and discretized.

Define $E[\boldsymbol{d}_k] = 0$ and $D[\boldsymbol{d}_k] = \boldsymbol{Q}_k$ are the mean and variance of $\boldsymbol{d}_k$, respectively, which is the discrete measurement noise variable. $E[\boldsymbol{V}_k] = 0$ and $D[\boldsymbol{V}_k] = \boldsymbol{R}_k$ are the mean and variance of $\boldsymbol{V}_k$, respectively, which is the discrete state disturbance variable. Predictive equation can be written:

$$\hat{\boldsymbol{X}}_{k+1/k} = \hat{\boldsymbol{X}}_k + \boldsymbol{f}(\hat{\boldsymbol{X}}_k)\Delta t \tag{6}$$

Predictive variance can be written:

$$\boldsymbol{P}_{k+1/k} = \boldsymbol{\varphi}_{k+1/k} \boldsymbol{P}_k \boldsymbol{\varphi}_{k+1/k}^{\mathrm{T}} + \boldsymbol{R}_k \Delta t \tag{7}$$

Where:

$\boldsymbol{\varphi}_{k+1/k} = \mathbf{I} + \left.\frac{\partial \mathbf{f}(\hat{\mathbf{X}})}{\partial \hat{\mathbf{X}}}\right|_k \Delta t$, with $\boldsymbol{I}$ is an unit matrix.

Predictive equation of measurement can be written:

$$\hat{\boldsymbol{Z}}_{k+1/k} = \boldsymbol{h}(\hat{\boldsymbol{X}}_{k+1/k}) \tag{8}$$

Filtering equation can be written:

$$\hat{\boldsymbol{X}}_{k+1} = \hat{\boldsymbol{X}}_{k+1/k} + \boldsymbol{K}_{k+1}(\boldsymbol{Z}_{k+1} - \hat{\boldsymbol{Z}}_{k+1/k}) \tag{9}$$

Gain equation can be written:

$$\boldsymbol{K}_{k+1} = \boldsymbol{P}_{k+1/k}\boldsymbol{H}_{k+1}^{\mathrm{T}}(\boldsymbol{H}_{k+1}\boldsymbol{P}_{k+1/k}\boldsymbol{H}_{k+1}^{\mathrm{T}} + \boldsymbol{Q}_{k+1})^{-1} \tag{10}$$

Where:

$$\boldsymbol{H}_{k+1} = \left.\frac{\partial \boldsymbol{h}(\boldsymbol{X})}{\partial \boldsymbol{X}^{\mathrm{T}}}\right|_{\boldsymbol{X}=\hat{\boldsymbol{X}}_{k+1/k}} .$$

Filtering variance equation can be written:

$$\boldsymbol{P}_{k+1} = (\boldsymbol{I} - \boldsymbol{K}_{k+1}\boldsymbol{H}_{k+1})\boldsymbol{P}_{k+1/k} \tag{11}$$

Filtering initial valves can be written:

$$\hat{\boldsymbol{X}}_0 = E(\boldsymbol{X}_0)\ , \quad \boldsymbol{P}_0 = \mathrm{Var}(\boldsymbol{X}_0) \tag{12}$$

Thus, 7 state variables of x , y , z , v_x , v_y , v_z and c output from Kalman filtering can be taken as the initial conditions for the trajectory reconstruction adopting equation (1). And then the trajectory falling point can be estimated.

4 Example and analysis

In this report, a rocket projectile is taken as an example. Trajectory reconstruction is based on GPS measure data of a flight test. The relationship between range and flight time is put forward in table 1.

Table 1 Relationship between range and flight time

Angle of fire (θ_0 /°)	Range (x /km)	Flight time (t /s)
45	16	43
	19	69
	23	77

Trajectory parameters are calculated by particle trajectory model. GPS measure noise is introduced by random generator to obtain measure data. Errors range of x , y , z , v_x , v_y and v_z are [-20,80], [-30,120], [-10,40], [-2,8], [-3,12] and [-1,4] ,respectively.

The initial values of system state are:

$$\boldsymbol{X}_0 = [x_0 \quad y_0 \quad z_0 \quad v_{x0} \quad v_{y0} \quad v_{z0} \quad c]^{\mathrm{T}} = [7013.2 \quad 2375.2 \quad 168.3 \quad 364.3 \quad 54.3 \quad 8.7 \quad 0.8]^{\mathrm{T}} ,$$

$$\boldsymbol{P}_0 = \begin{bmatrix} 1637 & & & & & & \\ & 6550 & & & & & \\ & & 1343 & & & & \\ & & & 17.5 & & & \\ & & & & 70 & & \\ & & & & & 14 & \\ & & & & & & 0.04 \end{bmatrix}.$$

Assuming there is no error in system state model, the covariance matrix of $\boldsymbol{R}$ can be written: $\boldsymbol{I} \times \boldsymbol{0}_{7\times 7}$.

The error of measurement equation is given according to the maximum of the errors of parameters, and variance matrix can be written:

$$\boldsymbol{Q} = \begin{bmatrix} 100^2 & & & & & & \\ & 150^2 & & & & & \\ & & 50^2 & & & & \\ & & & 10^2 & & & \\ & & & & 15^2 & & \\ & & & & & 5^2 & \end{bmatrix}.$$

For the convenience of comparison, both of Directly Trajectory Reconstruction (DTR) and Kalman Filter Trajectory Reconstruction (KFTR) are taken to do numerical simulation.

According to the range shown in table 1, the relationship between falling point deviation and range is put forward in Table 2.

Table 2 Relationship between falling point deviation and range

Range x (km)	Deviation(m)	Reconstruction method	
		DTR	KFTR
16	Δx	54.323	16.456
	Δz	26.174	7.553
19	Δx	61.228	19.395
	Δz	29.239	8.451
23	Δx	70.848	21.942
	Δz	31.441	8.955

With Kalman filtering, trajectory reconstruction can reduce the falling point deviation and improve trajectory accuracy effectively (Table 2). In addition, the accuracy of trajectory reconstruction is affected by range. The relationship between falling point and range in x direction and z direction are described respectively in Fig.3 and Fig.4.

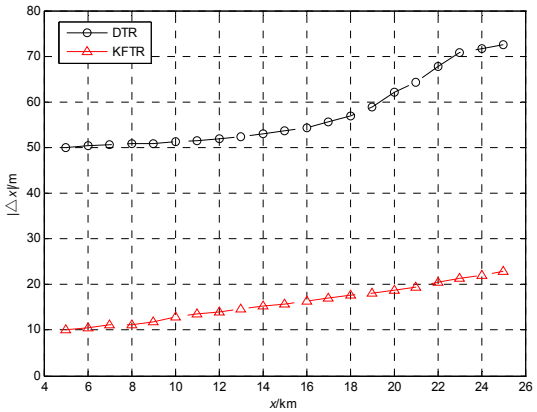


Fig.3 Range deviation

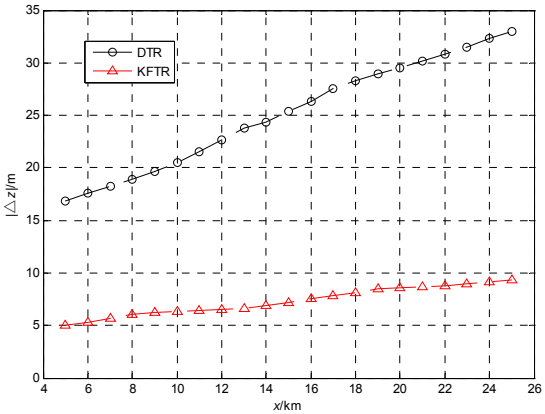


Fig.4 Lateral deviation

Falling point deviations in both x direction and z direction are increased with the increase of range. This is due to the increase of range in reconstruction process, which means that the projectile flying in the air is fast, the flight trajectory will become longer, and the error will accumulate. Secondly, the velocity measure error of GPS will be increased along with the increasing of the velocity (Fig.3 and Fig.4).

In addition, with a large amount of trajectory calculations, the accuracy of trajectory reconstruction will be improved rapidly with the increasing of the detection points in the flying process of projectiles. And the trend will tend to flatten after a period of time.

The fallowing parameters are used as the initial conditions: sampling interval is 0.1s, tracking time of GPS is 5s, and tracking points are 10, 20, 30, 40, and 50 corresponded to 16s, 17s, 18s, 19s and 20s. The simulation results are given in Figure5 and Figure6. It indicates that falling point deviations both in x direction and z direction are tend to flatten, and tracking points (tracking time) of GPS are 50 can meet trajectory accuracy requirements (Fig.5 and Fig.6).

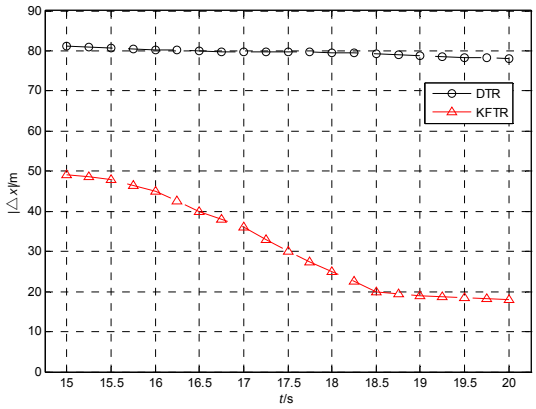


Fig.5 Range deviation

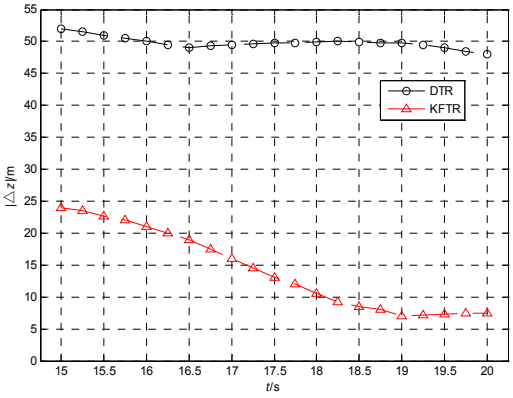


Fig.6 Lateral deviation

5 Conclusions

In this paper, trajectory reconstruction principle based on GPS measure data was addressed. Kalman filtering trajectory equations were established with system nonlinear particle trajectory state equation and system measurement equation from the transform of GPS measure data between inner earth coordinate system and ground coordinate system. An example of numerical simulation and analysis for rocket projectile were carried out. The simulation results show that the estimation of range and lateral can be reduced greatly by using Kalman filtering algorithm. The precision of trajectory reconstruction will decrease with the increase of range. Increasing the GPS tracking time can meet the ballistic accuracy requirements.

References

[1] ZHANG Xianmeng, WANG Jianxin. The applications of Kalman filter in fire-finder radar[J]. Journal of East China Institute of Technology, 1987, (3): 64-72. (in chinese)

[2] GUO Xifu, ZHAO Zihua. Theory and Application for Exterior Ballistics Modelling in Fire Control System[M]. Beijing: National Defend Industry Press, 1997. (in chinese)

[3] NIU Zhongxing, WU Jianfeng, ZHAO Yuqin. The Positioning Principles of GPS to MDI Measurement of Missile[J]. Tactical Missile Technology, 2006, (2): 66-70. (in chinese)

[4] ZHU Limei, LIANG Yan, ZHANG Xuchun. Study on the Methods for Exterior Ballistic Measurement Based on GPS[J]. AEROSPACE SHANGHAI, 2004, (3): 11-14. (in chinese)

[5] SHI Jin-guang, XU Ming-you, WANG Zhong-yuan. Application of Kalman filtering in calculation of trajectory falling point of trajectory correction projectiles[J]. Journal of Ballistics, 2008, 20(3):41-43. (in chinese)
[6] DING Chuan-bing, WANG Liang-ming, CHANG Si-jiang, HAO Ting, YANG Rong-jun. Application of Kalman Filtering in GPS-guided Rocket Projectiles[J]. Journal of Nanjing University of Science and Technology, 2010, 34(2):157-160. (in chinese)
[7] XU Ming-you. Higher Exterior Ballistics[M]. Beijing: Higher education press, 2003. (in chinese)
[8] Changey S, Fleck V, Beauvo is D. Projectile attitude and position determination using magnetometer sensor only[J].Proceedings of SPIE, 2005, 58(3): 49-58. (in chinese)
[9] WU Chunling, HAN Chongzhao. Finite-difference ex tended Kalman filtering algorithm for ballistic target tracking [J]. Journal of Xian Jiaotong University, 2008, (2): 143- 146. (in chinese)

Digraph-model-based Dynamic Command Scheduling: An Emerging Paradigm for Smart Satellite Operation

Tian Zhi-xin, Cui Xiao-ting, Li Zhen, Qiao Yi-shi, Li Xiao-juan

Beijing Institute of Spacecraft System Engineering, China

Abstract: The operation of space missions through sending of goals, instead of command sequences, is an emerging paradigm, with advantages in reduced operation-costs, increased responsiveness and simplified self-monitoring. Therefore, user-oriented on-board autonomous command scheduling is a critical issue. In this paper, the satellite control flows are mapped to a digraph model, in which the vertices and edges represent mission commands and minimum time intervals allowed between command pairs respectively. Thereafter, a novel digraph-model-based dynamic command scheduling algorithm is proposed to generate executable and conflict-free command sequences. It is applicable for satellite on-board computer with affordable computation/storage complexity. Experiment on an Earth-Observing-Satellite shows that, with 7 sensors combination, 6 payload data processing methods, 3 data transmission strategies and 2 attitude maneuvers, 3 mission-oriented service interfaces can achieves 43 kinds of command sequences, user control interface complexity can be reduced 93% and mission operation up-linking efficiency improved over 500%. The proposed dynamic command scheduling algorithm has great potential in Earth-Observing-Satellite in-orbit operation.

Keyword: smart satellite, directed-graph-model, user-oriented operation, dynamic command scheduling

1 Introduction

Autonomous management technology provides all the benefits for satellites of reduced operation costs, increased emergency responsiveness and simplified self-monitoring[1]. As a research and industry hot topic, autonomous management technology gets progressively more attention from worldwide nations. The NASA Jet Propulsion Laboratory and the Ames Laboratory made a series of research on autonomous management technology, such as intelligent mission planning technology, science data analysis, distributed autonomous system, autonomous safe and accuracy landing, fault detection and mitigation [2]. In the “New millennium project” funded by NASA, the on board autonomous management technology was first successfully demonstrated by DS-1[3]. European Space Agency developed a highly autonomous on board operation system to accomplish spacecraft operations in the PROBA project [4].

Time-based command sequences are the traditional paradigm for control of remote sensing satellites. With a unique operation model for each satellite, operation personnel in the ground control center choose a certain mission command sequences, set parameters and uplink to the satellite by TT&C system. The satellite has to strictly follow the commands sequences by invariable time interval [5]. Through the method of commands sequence is straightforward and practical， it restricts the ability of on board autonomous operation by incapability to optimize the mission performance according to the changes in the environment. Therefore, the operation of space missions through the sending of goals, instead of sequences of commands, becomes an emerging

paradigm. With the application of the autonomous management technology, the Earth Obseving-1 has exhibit the emergency response ability by observing the floodwater, volcano, etc[3]. The Mars Exploration Rovers improved the autonomous survival ability under complicated space environment[6]. Kucinskis F N [7] proposed a dedicated design scheme for payloads modeling, on-board autonomous software architecture and satellite-ground interface. Wu B F [8] integrates different command sequences into a series of on-board control procedures, which simplified the satellite-ground interface. However, the satellite's payload working modes in references [2-8] involves only recording and playback working modes, therefore, none of the abovementioned research can be applied to complicated remote sensing satellite with multi-sensors and multi-working modes. As the development of nation's high spatial resolution earth observation system, satellite's ability of high resolution imaging and agile attitude maneuver leads to more complexity of working mode and on-board mission planning[9-10]. Therefore, a generic on-board mission-oriented autonomous command sequences generation algorithm, applicable to high performance and complicated satellites, is critical to improve satellite's autonomous mission planning ability.

In this paper, a novel digraph-based algorithm is proposed to generate a sequence of executable and conflict-free commands. Firstly, the satellite mission control flow analysis is introduced and the digraph-model is established. Then, the mission command sequences generation algorithm based on colouring of the digraph-model is brought forward and by the example of a specific satellite application, the effectiveness of the algorithm is fully proved and executed.

2 Satellite mission control flow analysis

The satellite mission command sequence, i.e., mission control flow or on-board control procedures, is designed to coordinate the on-board equipment, such as the payloads data generation, data processing and transmitting equipments, antennae and attitude control equipments to accomplish the payloads date generation, processing, recording and/or transmitting under each specific working mode. In general, most remote sensing satellites support the following working modes:

1. Multi-payload imaging and real-time transmitting mode under various attitudes maneuver strategy (WM1);
2. Multi-payload imaging and recording mode under various attitudes maneuver strategy (WM2);
3. Satellite-earth data playback mode (WM3);
4. Satellite-relay-earth data playback mode (WM4);
5. Multi-payload imaging data fast recording and semi-real-time satellite-earth playback mode (WM5) ;
6. Multi-payload imaging data fast recording and semi-real-time satellite-relay-earth playback mode (WM6).

Figure 1 illustrates a representation of a typical remote sensing satellite mission control flowchart. All the command sequences under the six working modes mentioned above is reorganized into four basic control flows as data processing, payloads data generation, antennae control and attitude maneuver. Each basic control flow, orderly executed under required time intervals, is redefined as a basic command sequences vector (in brief BASIC_VEC) and delegated to fulfil certain independent mission. Constraint matrix among the four BASIC_VECs is defined to

coordinate the independent missions of payload data generation, processing, recording, transmitting and attitude maneuver.

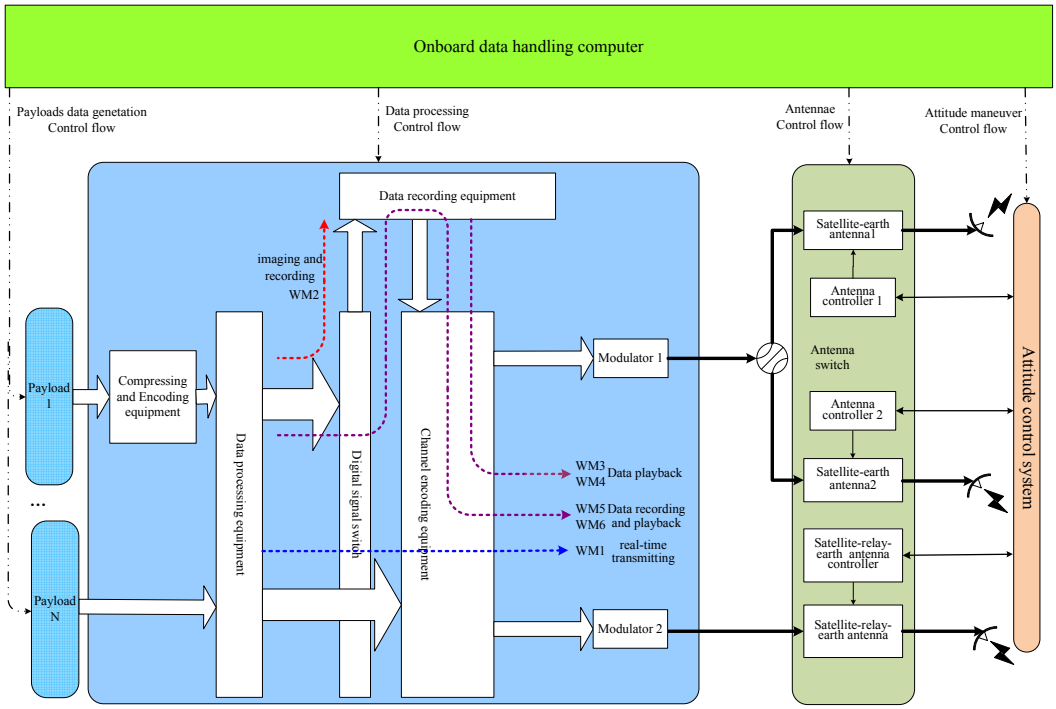


Figure.1 Satellite mission control flowchart

Table 1 lists all the basic command sequences for mission control flows involving six working modes:

1. DP_1~ DP_6 represent the six data processing BASIC_VEC under each working mode.

2. PL_1~PL_N represent all the N kinds of payloads, where PL_Mode={i, j, ……, k} represent any of the i, j, ……k working payloads.

3. AT_a, AT_b, AT_{relay} represent the antennae control BASIC_VEC, assuming there are two sets of satellite-earth antennae and one set of satellite-relay-earth antenna.

4. $AOCS_1$~$AOCS_M$ represent all the M kinds of target imaging attitude maneuver BASIC_VEC.

Table.1 Basic command sequences involving in typical working modes

Working mode	Data processing BASIC_VEC	payloads data generation BASIC_VEC	antennae control BASIC_VEC	attitude maneuver BASIC_VEC
WM1	DP_1	$\cup PL_j$（j∈PL_Mode）	AT_a/AT_b	$AOCS_l$（l∈[1,M]）
WM2	DP_2	$\cup PL_j$（j∈PL_Mode）	/	$AOCS_l$（l∈[1,M]）
WM3	DP_3	/	AT_a/AT_b	/
WM4	DP_4	/	AT_{relay}	/
WM5	DP_5	$\cup PL_j$（j∈PL_Mode）	AT_a/AT_b	$AOCS_1$
WM6	DP_6	$\cup PL_j$（j∈PL_Mode）	AT_{relay}	$AOCS_1$

Table 2 shows the constraint matrix of four mission control flows under six working modes:

1. C_{DPi_PLj}（i=1~6，j=1~N) represent the constraint between data processing BASIC_VEC and

payloads data generation BASIC_VEC, which defines the sequence between the two mission control flows as follows: Data processing should commence work after the payloads output the valid data, and the payloads should not stop the valid data outputting before data processing ends.

2. C_{DPi_ATk}（i=1~6，k=a，b，relay）represent the constraint between data processing BASIC_VEC and antennae control BASIC_VEC, which defines the sequence between the two mission control flows as follows: Data processing should commence work after the antennae finish the tracking of the ground station or the relay satellite and the antennae should stop the tracking after the data processing mission;

3. C_{DPi_AOCSl}（i=1~6，l=1~M）represent the constraint between data processing BASIC_VEC and attitude maneuver BASIC_VEC, which defines the sequence between the two mission control flows as follows: Data processing should commence work after the required attitude maneuver for target imaging and the next attitude maneuver shouldn't start until the data processing mission is over.

Table.2 Constraint matrix involving in typical working modes

Working Mode	C_{DPi_PLj} （i=1~6，j=1~N)	C_{DPi_ATk} （i=1~6，k=a，b，relay）	C_{DPi_AOCSl} （i=1~6，l=1~M）
WM1	$\cup C_{DP1\_PLj}$（j∈PL_Mode）	C_{DP1_ATk}（k=a/b）	C_{DP1_AOCSl} （l=1~M）
WM2	$\cup C_{DP2\_PLj}$（j∈PL_Mode）	/	C_{DP2_AOCSl} （l=1~M）
WM3	/	C_{DP3_ATk}（k=a/b）	/
WM4	/	$C_{DP4\_ATrelay}$	/
WM5	$\cup C_{DP5\_PLj}$（j∈PL_Mode）	C_{DP5_ATk}（k=a/b）	C_{DP5_AOCS1}
WM6	$\cup C_{DP6\_PLj}$（j∈PL_Mode）	$C_{DP6\_ATrelay}$	C_{DP6_AOCS1}

3 Digraph-model of satellite control flows

Digraph-model is an effective method to solve the scheduling optimization problems under timing constraint, which has been broadly applied to satellite mission scheduling, integrated circuit logic optimization, etc[11]. In this paper, a mapping of mission-oriented command sequences and digraph-model is proposed to indicate the constraint relations between the basic command sequences and the mutual constraint matrix. The digraph-model G(V,E) of the satellite control flows is constructed as follows (Table 3 is the notations used in the graph-model construction process),

1. Initialization. Add Source and Sink to the digraph-model G(V,E) to represent the start and end of the mission respectively. Initialize the executing time of Source and Sink by zero.

2. Digraph-model construction. For each command of the BASIC_VEC, insert a vertex into G(V,E). For each constraint between a pair of commands in the BASIC_VECs, insert an edge into G(V,E), and the edge length symbolizes the minimum time interval allowed between the two commands. Insert an edge between Source and each first command of the BASIC_VEC, initializing the edge length by zero. Insert an edge between Sink and each last command of the BASIC_VEC, initializing the edge length by zero. Add edge Cons($V_p \rightarrow V_q$) to G(V,E), representing the constraint among the basic command sequences. The direction of the edge represents the executing order of the vertices V_p and V_q and the edge length symbolizes the minimum time interval allowed between

the two commands.

3. Digraph G(V,E) slicing. Define Source as level 0 vertex. The vertices whose forward vertices only relevant to Source are defined as level 1 vertices, the vertices whose forward vertices only relevant to Source and level 1 vertices but not belong to Source, level 1 vertices are defined as level 2 vertices, and so on, the vertices whose forward vertices only relevant to Source, level 1 …, and level n-1 vertices but not belong to Source, level 1…, and level n-1 vertices are defined as level n vertices.

Table.3 Notations for digraph-model of satellite control flows

Notations	Definition
V	The vertices of the overall basic command sequences
E	Timing constraints derived from basic command sequences and constraint matrix
v_i,v_j	Vertex v_i and v_j
$e(v_i \rightarrow v_j)$	The directional edge between v_i and v_j
$pred_l(v_i)$	The l-th predecessor vertex of v_i
S_r	Sub-digraph of basic command sequences r
C_{rt}	Edge set representing the constraint relation between basic command sequence r and t
N	The number of the entire digraph vertices
M	The number of the entire digraph edges
Source	The start of the mission command sequence
Sink	The end of the mission command sequence
T_0	The imaging time
$TIME(v_j)$	The command executing time of v_j
$t(v_i \rightarrow v_j)$	Time interval between v_i and v_j
$X = \{S, X_l, X_m, \ldots\}$	S representing Source executing time, X_l, X_m,… representing time intervals between v_l, v_m,…and Source

4 Command sequence generation based on digraph-model coloring

4.1 The introduction of the algorithm

With digraph-model of the satellite control flows, the autonomous mission command generation could be regarded as an optimum problem of the digraph's vertices and edges. Since the digraph-model is the mapping set of the overall time constraints under six typical working modes, the constraints under each particular working mode could be regarded as a dedicated sub-digraph under this working mode. And the mission command generation problem turns into the process of sub-graph construction and vertices optimum under the dedicated sub-graph constraints.

The basic command sequences are the minimum function unit of the satellite missions. Two distinct features of BASIC_VEC are as follows:

1. Commands correlation. With the atomic property of the BASIC_VEC, the commands either

all participate in the mission or not a single command takes part in it.

2. Constraint accumulative effect. If a BASIC_VEC participates in one particular mission, the mission commands sequence generated should be subject to the constraint within the BASIC_VEC. And if two BASIC_VECs participate in one particular mission, the mission commands sequence generated should be subject to the constraint of both BASIC_VECs.

Based on the two distinct features mentioned above, we define the coloring status of each sub-graph's vertices and edges as follows:

1. Vertices coloring. Define all the commands of some BASIC_VEC as a sub-graph. If the BASIC_VEC takes part in certain working mode, then all the vertices of the sub-graph is in certain coloring status.

2. Edges coloring. If a pair of vertices in the sub-graph is coloring status, then the edge between the pair of vertices is in coloring status. Otherwise, either of the two vertices isn't in coloring status, then the edge is not in coloring status.

The sub-graph based satellite mission commands generation process is showed in Figure.2. Assign each digit of the "working mode word" to every dedicated BASIC_VEC under each working mode. If the BASIC_VEC takes part in the mission, the value of the corresponding digit is 1. Otherwise, if the BASIC_VEC doesn't take part in the mission, the value of the corresponding digit is 0. Then define the sub-graph vertices set related to a certain mission and the sub-graph edges related to the constraint within the dedicated working modes. The mission command sequences generation algorithm based on digraph-model coloring is as follows:

1. Mission parameter settings. Assign the mission related parameters, usually inputted by the users, such as imaging or data transmitting time to the corresponding edge length, and assign the imaging command timing property with T_0.

2. Define the mission-related sub-graph. Define the coloring status of the sub-graph vertices and edges based on the "working mode word" value.

3. Sub-graph vertices timing property calculation. Compute the G(V,E) vertex timing property by a recursive algorithm from level 0 to level n. For every vertex v_j, the timing property equals to the maximum of all the $pred_l(v_i)+ t(pred_l(v_i)\rightarrow v_j)$.

4. Mission time window calculation. The executing time delay between Sink and Source, in brief EST(Sink), is defined as the mission time window. The executing time of Source is the beginning of the mission executing time.

5. Mission confliction detection. Check the confliction between the mission executing time span [TIME(Source), TIME(Source)+EST(Sink)] and the existing missions executing time span. If there is a time window confliction, rearrange all the missions based on the mission priority and delete the mission with the lowest priority. If there is no confliction, add the new mission command sequence to the satellite command buffer. When satellite time reaches TIME(Source), execute the mission command sequence autonomous generated based on the algorithm.

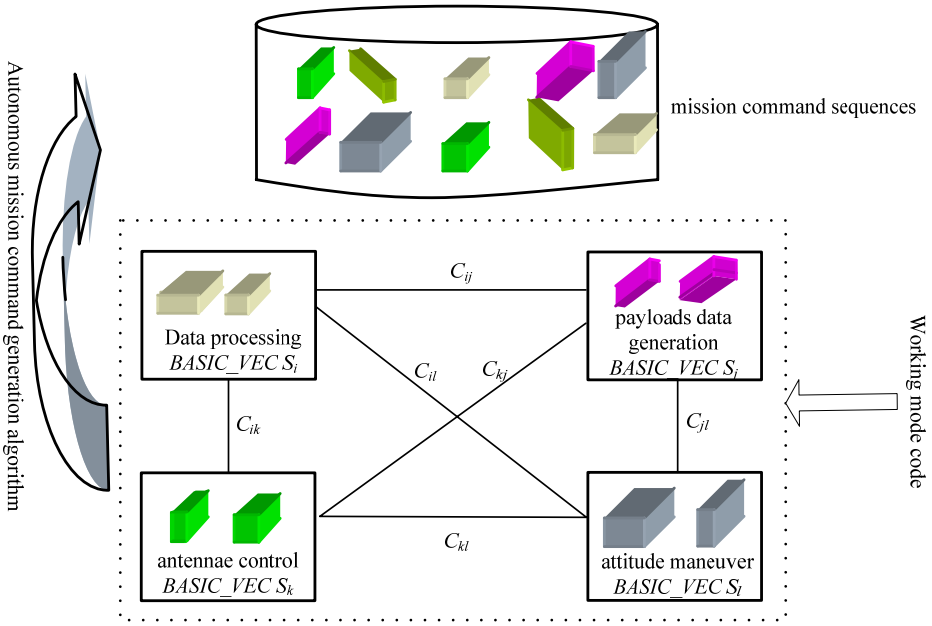


Figure.2 Command sequences generation based on sub-graph colouring

4.2 Algorithm analysis

The algorithm proposed in this paper is applicable to on-board computers with affordable storage and computing ability for its linear storage complexity and computing complexity.

Since the satellite digraph-model is the mapping of the basic command sequences and the mutual constraints, the sub-digraph under a certain working mode could be attained by the coloring of the satellite digraph-model. The mission command sequences under different typical working modes could be generated by the coloring of the digraph's vertices. The on-board computer only needs store N vertices and M edges.

In the processing of mission command sequences generation, the computing complexity in step1 "Mission parameter settings" is in direct proportion to the number of the mission related parameters. In the processing of the digraph-model coloring in step2 "Define the mission-related sub-graph", the coloring time of digraph vertices is no more than N, and the coloring time of digraph's edges is no more than M. In step3 "Sub-graph vertices timing property calculation", only basic mathematical algorithm is applied to cover the digraph coloring, and the computing time is equal to digraph coloring vertices and no more than N. Step4 "Mission time window calculation" and step5 "Mission confliction detection" involve only basic mathematical calculations.

5 Application effect

The algorithm has been applied to a Chinese earth observing satellite. There are three kinds of payloads on the satellite, supporting 7 payload working combination modes with single payload imaging, two payloads combination or all three payloads combination. There are 2 kinds of attitude maneuver strategies, one target imaging attitude maneuver or consecutive target imaging attitude maneuvers. With 2 satellite-earth antennae and 1 satellite-relay antennae, there are 3 kinds of payload data transmission methods. Meanwhile, the satellite supports 6 data processing modes. For the imaging mission within China territory, there are 3 kinds of operation modes can be applied, (1) real-time imaging and transmitting, (2) imaging data fast recording and semi-real-time satellite-earth playback, and (3) imaging data recording and semi-real-time satellite-relay-earth playback. For the imaging mission abroad, there are 3 kinds of operation modes can be applied, (1)

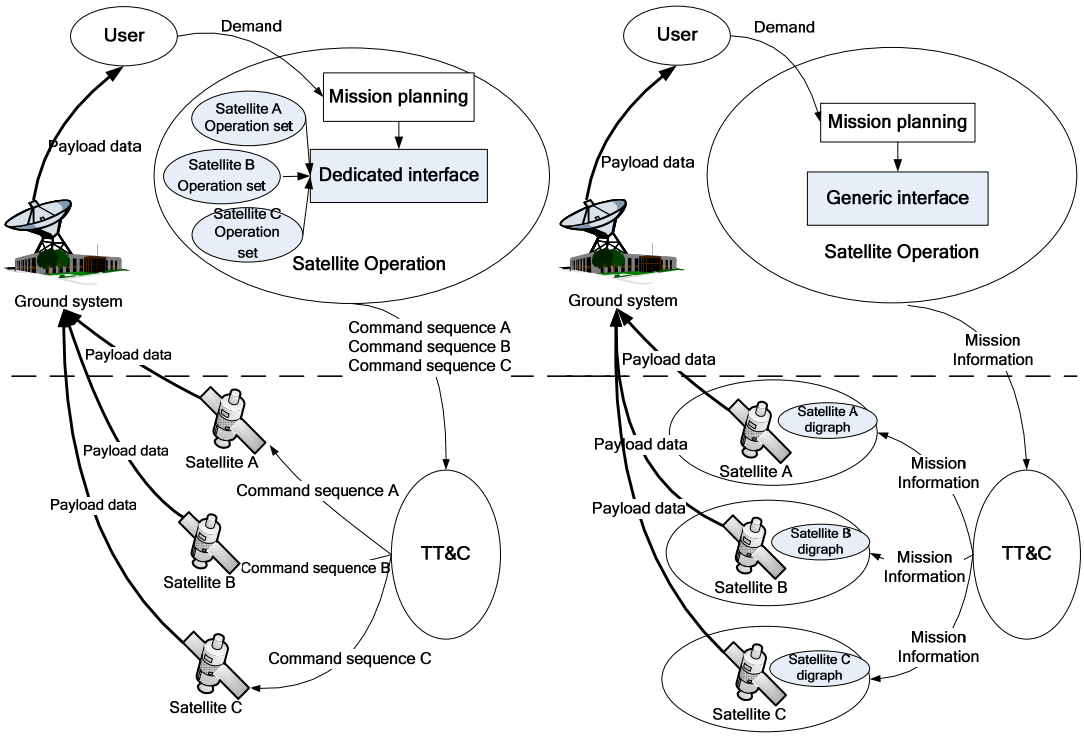
imaging data recording and satellite-earth data playback, (2) imaging data recording and satellite-relay-earth data playback, and (3) imaging data fast recording and semi-real-time satellite-relay-earth playback.

Therefore, the satellite supports totally 52 kinds of working modes, including 14 kinds of imaging data recording modes with 7 payloads combinations and 2 attitude maneuvers, 14 kinds of imaging and real-time transmitting modes with 7 payloads combinations and 2 data processing strategies, 21 kinds of imaging data recording and playback modes with 7 payloads combinations and 3 data processing strategies, and 3 kinds of data playback modes.

The satellite on-board computer is configured with TSC695F CPU, 2MB SRAM and 128KB PROM, and the schedule time slice interval of the real-time operation system is 125ms. According to the on-board verification, the control flow digraph-model takes up 42.6KB SRAM, the command sequences generation algorithm takes up 29.4KB PROM. To reduce the computation time interval in each slice, the digraph-model coloring takes up one time slice, the recursive algorithm takes up one time slice, and the mission command vertices sorting and storage take up one time slice. According to the result of the on-board experiment, the computing time from the satellite receiving the customer's mission information to the accomplishment of the command sequences generation is no more than 0.5 seconds.

The digraph-based algorithm proposed in this paper has achieved prominent application benefit. As shown in Figure.3 (a), for the traditional satellite operation paradigm, the users have to set a dedicated operation set corresponding to each satellite. There are an ever-increasing number of satellites, the ground operation and control management system is overburdened. To accomplish the user's mission, a large quantity of command sequences of payloads, data processing, antennae, storage, attitude control are programmed and the uplink ability of the TT&C is quite demanding. Furthermore, most of the low orbit remote sensing satellite's revisiting period is about 3~5 days, once on-board failure occurs, the satellite couldn't realize mission reconstruction.

As shown in Figure.3 (b), there are three advantages of our proposed operation mode. (1) A generic mission-oriented high-level autonomous command operation interface, independent of satellite properties, is adopted to realize the on-board operation of all kinds of satellites. (2) The up-linked mission-oriented commands are no more than 40 bytes. Each uplink data block, in general 256 bytes, can contain 6 missions. Thus, the mission uplink efficiency is increased about 500%. (3) The satellite can realize mission reconstruction once on-board failure occurs. For example, the traditional satellites have to follow the commands sequences strictly to accomplish a certain mission, once on-board failure occurs, the imaging schedule would be missed. However, if the algorithm proposed in this paper is adopted, the on-board computer can reschedule the imaging time and recalculating the imaging attitude based on the object longitude/latitude and orbit parameters. Under the new imaging schedule, the mission-oriented command sequence is regenerated and the imaging mission is accomplished.



(a) The traditional operation paradigm (b) The proposed operation paradigm

Figure.3 Mission operation schema evolution

6 Conclusions

In this paper, the satellite mission control flow under different working modes is reorganized into four basic command sequences of payloads data generation, data processing, antennae control and attitude maneuver. The digraph-model and the mission command sequences generation algorithm based on the digraph-model colouring are proposed to simplify the remote sensing satellite operation ability and to increase the mission uplink efficiency. There are two distinct advantages for proposed operation paradigm. (1) Since the computing complexity and storage complexity are in direct proportion to the number of the mission commands and the constraint relations, the algorithm is applicable for storage/computation rare on-board computers. (2) The algorithm is generic and thus applicable for different satellite through the setting of the digraph-model parameters.

The method proposed in this paper increases the on-board autonomous operation and control ability. And it establishes the technical foundation for satellite mission operation schema evolution from dedicated operation to generic operation, from static on-ground scheduling to dynamic on-board scheduling, from open-loop satellite-ground management to close-loop satellite-ground management.

References

[1] Jeremie P, Sylvain J, Patxi O. Autonomous mission planning in space: mission benefit and real-time performances[C]// Proceedings of the Embedded Real Time Software and Systems, 2014

[2] Pell B, Bernard D E, Chien S A, et al. An autonomous spacecraft agent prototype[J]. Autonomous Robots, 1998, 5(1): 29-52

[3] Thompson D R, Bornstein B J, Chien S A, et al. Autonomous spectral discovery and mapping onboard the EO-1 spacecraft[J]. IEEE Trans. on Geoscience and Remote Sensing, 2013, 51(6):3567-3579
[4] Santandrea S, Gantois K, Strauch K, et al. PROBA2: Mission and Spacecraft Overview[J]. Solar Physic, 2013, 286(1):5–19
[5] He R J, Li J F, Yao F. Earth Observing Satellite Mission Planning[M]. Beijing: Science Press, 2011（in Chinese）
[6] Grotzinger J P, Crisp J, Vasavada A R, et al. Mars science laboratory mission and science investigation[J]. Space Science Reviews, 2012, 170(1-4):5-56
[7] Kucinskis F N, Ferreira M G V. Planning on-board satellites for the goal-based operations for space missions[J]. IEEE Latin America Trans., 2013, 11(4):1110-1120.
[8] Wu B F, Li Z G, Li J Y, et al. Design of mission-oriented autonomous commands for small satellite[J]. Spacecraft Engineering, 2013, 22(4) : 68-71（in Chinese）
[9] G. Beaumet, G. Verfaillie, and M.C. Charmeau. Feasibility of autonomous decision making on board an agile earth-observing satellite. Computational Intelligence, 27(1):123–139, 2011.
[10] Huang Qun-dong, Yang Fang, Zhao Jiang. Calculation of image motion velocity for agile satellite dynamic imaging to changed continuously attitude point[J].Optics and Precision Engineering, 2012, 20(12): 2812-2820.
[11] Wang S H. Graph Theory[M]. Beijing: Science Press, 2004（in Chinese）

Innovation Potential by Small Satellites for Spacecraft Design

Klaus Schilling

University Würzburg, schi@informatik.uni-wuerzburg.de

Abstract: Small satellite technologies reached a maturity level such that building blocks for the design of specific missions are available. This offers opportunities to test small components in space at relatively low costs, increasing technology readiness levels (TRL) significantly in very short time. This offers the potential to accelerate innovation cycles in spacecraft design dramatically. Thus advantages in faster access to commercial market result.

This contribution reports about the UWE Pico-satellite program, where since 2005 advanced miniature satellite components were tested in orbit, including new solar cells, standardization of electrical interfaces, attitude determination systems and robust miniature on-board data handling systems.

1. Introduction

One of the most innovative fields the space sector relate to very small satellites. Modern miniaturization techniques allow to implement complete, fully functional satellites at a mass of just a few kg. The pico-satellites (~ 1 kg) and nano-satellites (~10 kg) are the fastest growing segment in space with 113 launches in 2014 (cf. Fig. 1). Here currently the transition from the educational and research area [6], [10], [11] to commercial applications has started [5], [7]. Since the conceptual design of CubeSats in the year 2000 [6], the maturity of technologies rapidly increased and exhibits today impressive capabilities and robustness. In consequence from universities and research institutes several commercial companies spun off in USA, Europe and Japan.

Current emphasis in development with respect to applications concern:

- implementation of advanced fault detection, identification and recovery (FDIR) methods to increase robustness and in-orbit lifetime,
- improved attitude and orbit determination and control, for precise pointing of payload instruments,
- standardized electrical interfaces, to encourage exchangeability of payloads and satellite subsystems and support cooperation in international teams with many partners.

In the following section, related approaches and solutions developed in the UWE-program will be described in more detail. This offers interesting perspectives for the future realization of high performance sensor networks composed of pico-satellites.

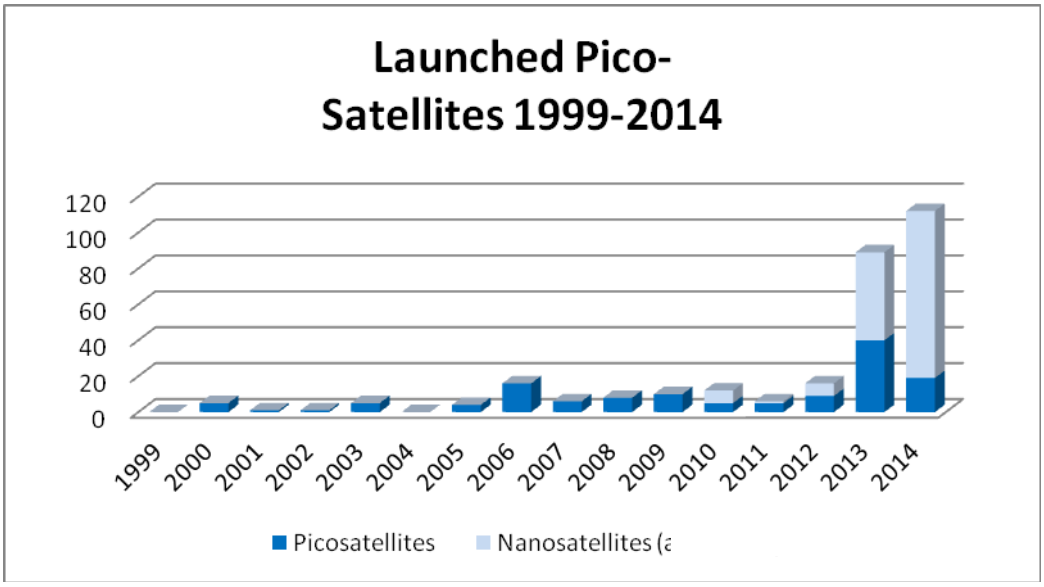


Fig. 1: Global amount of launched small satellites

2. Technology Development in the UWE Program

In 2003 the UWE-Program (**U**niversity **W**ürzburg's **E**xperimental satellites) was initiated with the objective to educate Aerospace students in system design, by developing relevant small satellite technologies and demonstrating them in orbit [10], [11]. In order to compensate negative effects of miniaturization, such as higher susceptibility to radiation, main emphasis was placed on electronics and software methods. Those are considered most promising for innovation of modern advanced spacecraft, too.

Typically a pico-satellite is implemented within a period of about one year and advantage is taken from the previous design, such that extensive documentation to transfer between different generations of students can be avoided.

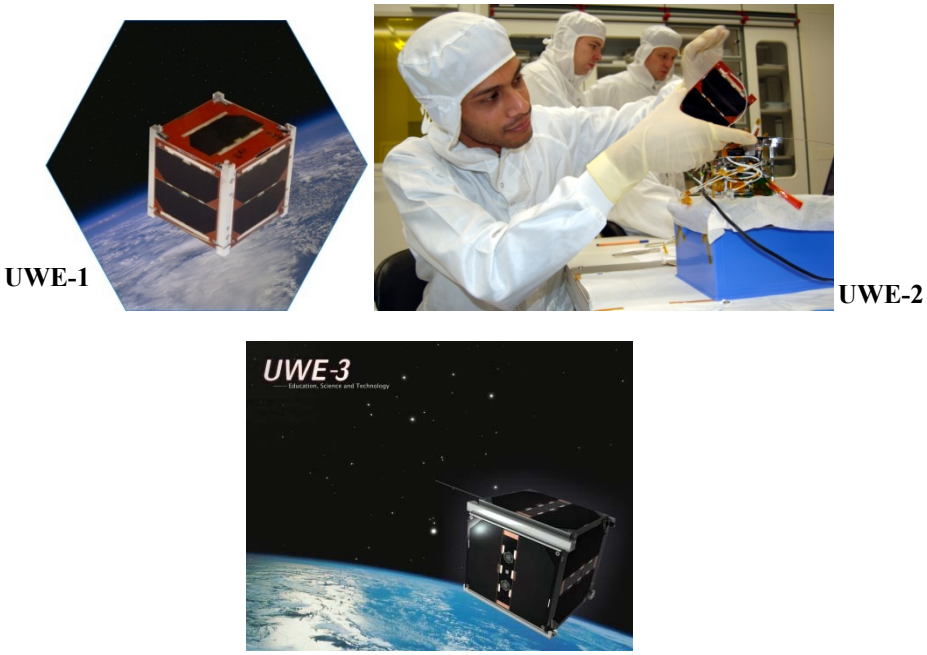


Fig. 2: The University Würzburg's Experimental (UWE) satellites placed so far in orbit

2.1 Mission UWE-1

The first german pico-satellit UWE-1 [2] investigated 2005 telecommunication aspects related to adaptions of parameters in Internet Protocols to the space environment, characterized by significantly increased signal propagation delays and noise levels. A position dependent tuning of these parameters during passage of a ground station was analyzed. In parallel, UWE-1 tested also new triple-junction GaAs solar cells, with an high conversion efficiency of about 30 %. from solar radiation to electrical energy. Having been tested by the pico-satellite, it became a space proven component and quickly progressed in technology readiness levels.

2.2 Mission UWE-2

UWE-2 analyzed determination of attitude and position of the satellite within its very limited mass budget [12]. A sensor data fusion with a Kalman filter of measurements from Sun sensors, gyros, magnetometers and a GPS sensor, as well as complementing software of the Earth's magnetic field and orbit propagators were implemented. This provided the basis to perform targeted orientation changes as basis for future remote sensing observations.

2.3 Mission UWE-3

The main scientific emphasis of UWE-3 related to attitude control in order to achieve an orientation towards an observation object [1]. UWE-3 initiated a new generation of pico-satellites, as here core properties of future commercial systems were realized and tested, such as

- reliable operations in space during extended periods,
- modular architecture for flexible exchange of subsystem- or payload-boards [3].

3. Crucial Fields for Technology Progress

The maturity of small satellites progressed by recent technology breakthroughs addressing crucial satellite components, which are addressed in this chapter.

3.1 Reliable On Board Data Handling System

The miniaturization necessary to realize a complete satellite at the mass of 1 kg causes increased noise susceptibility, in particular when only commercial, not radiation hardened components are used For UWE-3 the implemented advanced software on Fault Detection, Identification and Recovery (FDIR) for the redundant components of the OBDH [4] enabled a continuous uninterrupted operation since launch in November 2013, despite all problematic effects from space radiation were detected (e.g. Single Event Upsets, Latch-ups). Without interruption of service, malfunctions were immediately detected by an intelligent watchdog and a transfer of responsibility to the hot redundant unit was performed. This continued the running processes as well as the initiated failure analyses and correction of the disconnected faulty parts in order to return it after rehabilitation as soon as possible. Similar approaches are of interest on Earth for highly reliable systems, too.

Fig 3: The On-Board-Data-Handling (OBDH) system of UWE-3 with redundant micro -processors and storage chips.

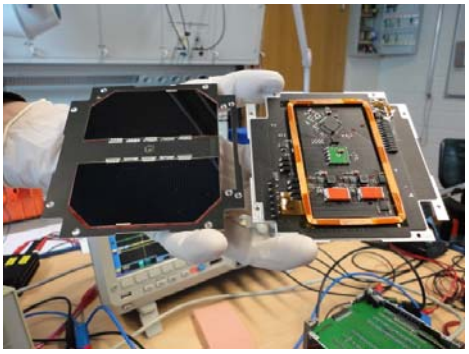


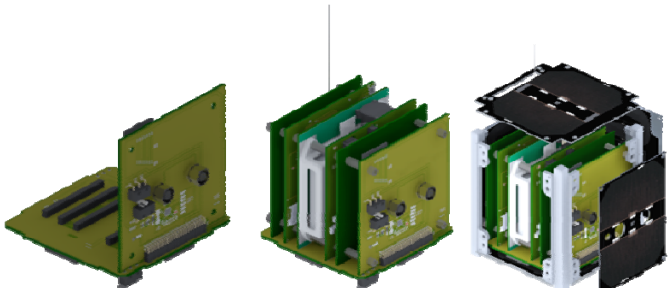
Fig. 4: Side panels of UWE-3 with solar cells for energy generation on the outside, and magnetorquers for attitude control and the related electronics on the inner side.

3.2 Attitude Control System

For Earth observation applications as well as in telecommunication applications appropriate pointing of instruments and of antennas are of importance for efficient operations. Therefore capabilities for attitude determination and control are crucial [1]. The orientation of the satellite is derived from data acquired by gyros, magnetometers and Sun sensors with an accuracy of a few degrees (cf. UWE-2). Subsequently UWE-3 changes its attitude very energy efficient with magetorquers in interaction with the Earth's magnetic field gradients. Along the magnetic field lines a single small reaction wheel is used to complement this to a 3-axis control system. Various algorithms for attitude control were uploaded to test and fine-tune in many experiments the related performance.

3.3 Modular Design and Standardized Electrical Interfaces

So far the CubeSat-Standard concerned only the outer dimensions and the mass of the satellite, such that launcher adapters could be shared on different rockets. Recent efforts focus on standardization of electrical interfaces in order to enable integration of subsystem and payload components from different manufacturers. A flexible, modular design was implemented at UWE-3 by a baseplate accommodating all power and data lines [3]. Via connectors all boards can be plugged in. By removing this way cables and harness, a very flexible and quick exchange of boards can be realized. A detailed description can be found at the UNISEC Europe webpages at http://unisec-europe.eu/standards/bus/



(a) (b) (c)

Fig. 5: The UWE architecture uses a base plate (a) in which all sub-system board are plug in by standardized connectors (b). The guiderails in the edges of the cube and the side panels form the

This modularization supports production of many units for storage, and for subsequent fast and flexible integration, whenever needed. Thus the total time for a satellite implementation can be significantly further reduced.

3.4 Orbit Control

Control of orbit and position is another important function for self-organizing pico-satellite formations. In preparation to the forthcoming UWE-4 mission a very compact electric propulsion system [8] is in preparation in laboratory tests. Two good candidates are Micro Vacuum Arc Thrusters (μVAT) and Nano Field-Emission Electric Propulsion thrusters (NanoFEEPs), which are currently evaluated in comparative laboratory tests. Due to the limited fuel consumption electric pro-pulsion is very suitable for small satellites. But due to the limited mass of the ejected fuel particles, accelerated to high velocities by a magnetic field, the resulting specific impulse is also low. Therefore orbit changes require some patience. For the objective of keeping the topology of a satellite formation despite the noise effects of the space environment, such small forces are sufficient. Current estimates indicate that a fuel consumption of ca. 3 gramm/year might be sufficient for this task. The characterization of these electric propulsion systems in orbit after extensive laboratory tests is planned with UWE-4 in 2016. This enables also to place the satellite in a suitable graveyard orbit after end of mission.

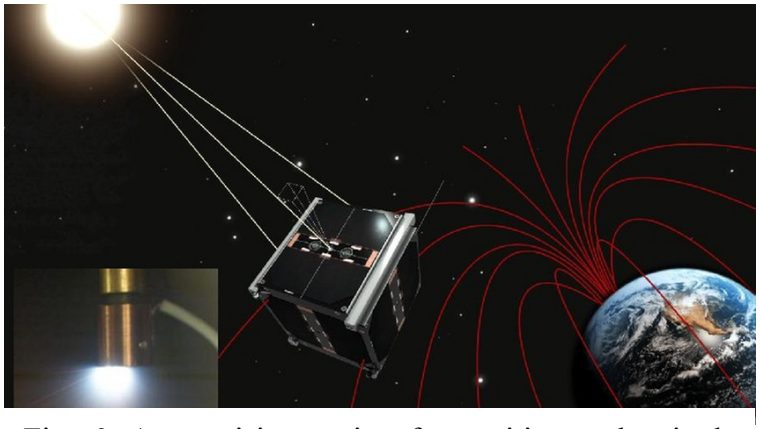


Fig. 6: A promising option for position and attitude control at small satellites concern electric propulsion systems integrated in the 4 guiderails of the satellite. The image in the lower left corner displays the electric propulsion system during tests in the laboratory.

4. Conclusions

Rapid technology evolution for very small satellites provides now a suitable base for commercial applications. Specific points addressed in this contribution concern system robustness, improved attitude and orbit control, as well as modularization and standardization. First commercial fields for commercialization address Earth observation and telecommunications.

The limited costs of placing small satellites in orbit and the shorter lifecycles due to low Earth orbits encourage “try and improve” design approaches instead of the traditional careful test procedures before launch. Many parameters can still be adapted in orbit and new on-board software can be uploaded to improve performance. Thus a significant increase of flexibility can be observed. Innovation cycles are in the field of small satellites very fast as

- there are reliable basic components with appropriate performance available which can be adapted to specific tasks,
- the time from conceptual design to a launch is much shorter than for traditional satellites,
- more frequent launch opportunities are available,
- in-orbit verification costs with small satellites are often lower compared to careful testing on ground.

Thus future satellite design is expected to benefit from this new emerging field as new approaches and techniques are expected to be transferred from this highly innovative sector.

Acknowledgements

The contributions of all the engaged collaborators in the UWE-projects are acknowledged, as well as the financial support from the German Ministry of Economics via the German Space Agency DLR, the Space Research program of the Bavarian Ministry for Economics and the European Research Council.

References

[1] Bangert, P.; Busch, S.; & Schilling, K.; Performance Characteristics of the UWE-3 Miniature Attitude Determination and Control system, Proceedings 2nd IAA Conference on Dynamics and Control of Space Systems (DYCOSS), Rome 2014, IAA-AAS-DyCoSS2

[2] Barza, R., Y. Aoki, K. Schilling; Cubesat UWE-1 – Technology Tests and in Orbit Results. In: Proceedings 57th International Astronautical Congress 2006, IAC-06-B5.3.

[3] Busch, S.; Schilling, K.; UWE-3: A Modular System Design for the Next Generation of Very Small Satellites. Proceedings of Small Satellites Systems and Services - The 4S Symposium, Slovenia. 2012

[4] Busch, S., Schilling, K.; Robust and Efficient OBDH Core Module for the Flexible Picosatellite Bus UWE-3, 19th IFAC Symposium on Automatic Control in Aerospace, Würzburg 2013, p. 218 – 223

[5] Hand, Eric; The rise of the CubeSat , Science vol. 346 issue 6216 (19.12.2014), p. 1449 http://www.sciencemag.org/content/346/6216/1444.full.pdf

[6] Heidt, H., J. Puig-Suari, A. S. Moore, S. Nakasuka, R. J. Twiggs; CubeSat: A new Generation of Picosatellite for Education and Industry Low-Cost Space Experimentation, 14th Annual USU Conference on Small Satellites, paper SSC00-V-5 http://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=2069&context=smallsat

[7] Jones, N.; Mini satellites prove their scientific power, Nature 508 (17 April 2014),p. 300 – 301 www.nature.com/news/mini-satellites-prove-their-scientific-power-1.15051

[8] Kronhaus, I., Pietzka, M., Schilling, K., 2013. Pico-Satellite Formation Control by Electric Propulsion. Proceedings 19th IFAC Symposium Automatic Control in Aerospace, Würzburg 2013

[9] National Science Foundation (NSF) / NASA ; CubeSat-Based Science Missions for Geospace and Atmospheric Research, Annual Report 2013 www.nsf.gov/geo/ags/uars/cubesat/nsf-nasa-annual-report-cubesat-2013.pdf

[10] Schilling, K., Design of Pico-Satellites for Education in System Engineering, IEEE Aerospace and Electronic Systems Magazine 21 (2006), pp. 9-14.

[11] Schilling, K. ; SpaceMaster: Master in Space Science and Technology As an International Interdisciplinary Educational Effort . Proceedings Advances in Control Education, Nizhny Novgorod. 2012, p. 300 - 305
[12] Schmidt, M., K. Ravandoor, O. Kurz, S. Busch, K. Schilling; Attitude Determination for the Pico-Satellite UWE-2. Space Technology 28 (2009), pp. 67-74

Small Satellite Formations for Earth Observation and Telecommunication

Klaus Schilling

University Würzburg and Center for Telematics
klaus.schilling@telematik-zentrum.de

Abstract: Cooperating small satellites offer innovative distributed system approaches as an alternative to traditional satellite design, with obvious similarities to the computer evolution from large mainframes to today´s networked smart phones. There was significant progress of small satellite capabilities related to formation flying: attitude and orbit determination and control systems were developed for pico-satellites at the level of a few kg of mass, as well as reliable miniature communication systems. On this basis the NetSat formation mission, composed of 4 pico-satellites, is now in preparation stage.

In addition to the review of recent technology developments, this contribution also outlines the application potential of future small satellite formations. With respect to Earth observation, a distributed multi-satellite systems offer higher temporal and spatial resolution, as well as a higher robustness against defects, compared to a traditional system at the same cost level. For telecommunication applications, low Earth orbiting small satellites can provide cost-efficient approaches for low-bandwidth communications.

1. Introduction

Formations of small satellites were analyzed for many years due to their advantages based on provision of a large baseline distance for observations [1], [2]. Advantageous applications include interferometry, gravimetry, distributed telescopes, and virtual apertures in observation scenarios. There are several satellite constellations in orbit (by example in telecommunications: Iridium, Globalstar, TDRSS, Orbcomm; in navigation: GPS, Glonass, Galileo, BeiDou; in Earth observation; Rapid Eye, Dove, Skybox; and in science: Cluster, Swarm), where each satellite is individually controlled from ground. In contrast, *formations* use a networked closed-loop control in-orbit in order to preserve the topology in the multi-satellite system on basis of determination and control of relative distances. Here so far only three missions based on only two spacecraft provide examples in orbit: GRACE, PRISMA, and Tandem-X [6]. As cost implications constrained traditional satellites to only these few missions, nowadays the dramatic technology progress for very small satellites [7], [9], [16] enables formations for realizing sensor networks of small satellites at the same cost level as one traditional large satellite [11]. A trade-off between both, the traditional and the cooperating small satellite approach identifies the benefits for specific application fields, in particular in telecommunications and Earth observation.

This contribution addresses at the example of the UWE-program the technology development of the essential functionalities for formations at the pico-satellite level. This includes in particular attitude and orbit control capabilities, production efficiency by modularization and standardization, as well as networked control aspects. Section 3 explains the NetSat mission, which prepares on this basis the placement of a formation of four pico-satellites in orbit The benefits of formations in Earth

observation and telecommunication scenarios outline the huge application potential for future formations.

2. The UWE Pico-Satellites for Technology Preparation of Formation Flying

The UWE program (***U***niversity ***W***ürzburg's ***E***xperimantal satellites) established a roadmap (cf. Fig.1) to provide the basis of crucial technologies for formation flying by pico-satellites at the mass of just 1 kg [12], [13], [15]. The approach is to develop step by step the related crucial technologies and to demonstrate them in orbit before the next functionality will be added in the next mission.

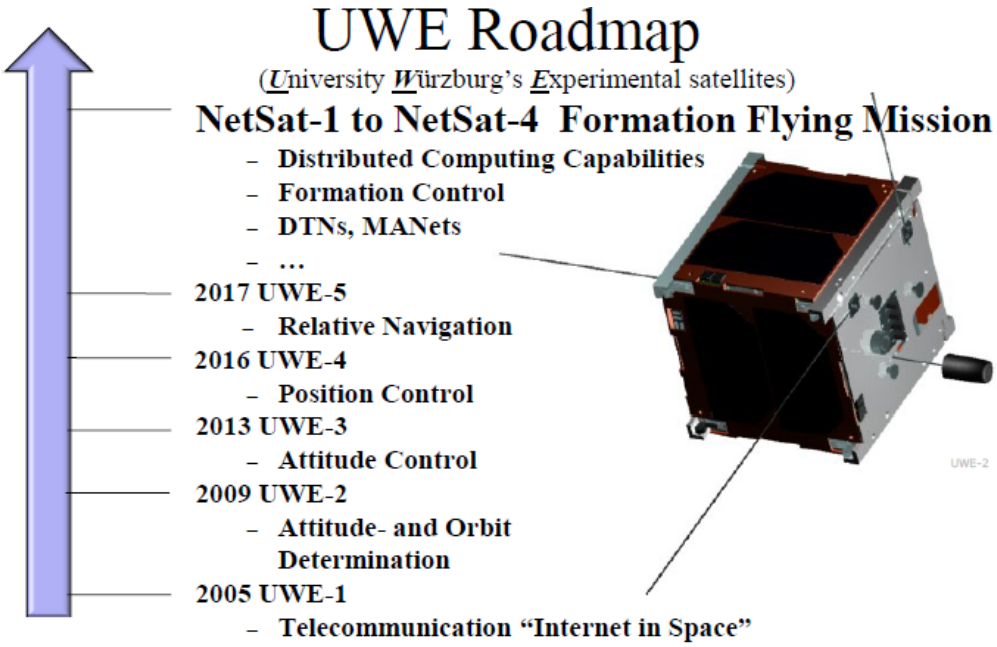


Fig.1: The roadmap to pico-satellite formation flying of University Würzburg and Center for Telematics

UWE-1 (launched 2005) had the scientific objective to analyze optimization of parameters in Internet Protocols in order to adapt to the space environment with signal propagation delays due to large distances and much higher noise levels [4].

UWE-2 (launched 2009) addressed the topic of attitude and orbit determination at the very constrained framework of a pico-satellite [14]. By using GPS for position acquisition and a combination of MEMS inertial sensors, magnetic field sensors and sun sensors for attitude determination, in combination with magnetic field models and an orbit propagator a Kalman filter performs the related sensor data fusion to derive attitude and position.

UWE-3 (launched 2013) realizes on that basis related attitude control actions. For this purpose magnetorquers are placed on the inside of each of the 6 side panels and one miniature reaction wheel in the center. By coordinated activity of these actuators a very energy-efficient 3-axis control system had been implemented [3].

UWE-4 (with a planned launch in 2016) is complementing the relevant technologies by orbit control capabilities and is in the implementation and testing stage [10]. An electric propulsion system (on basis of nano-FEEPs and on mikro-Vakuum Arc Thrusters) placed in the four guiderails in the edges of the 10 cm cube will provide position and attitude control capabilities.

3. The NetSat Mission

These UWE missions prepared the scientific and engineering basis to envisage now formation flying. The European Research Council (ERC) selected „NetSat“ for an „ERC Advanced Grant“ to analyze first time a formation of 4 satellites in orbit, enabling and controlling thus a 3-dimensional configuration [13]. This opens new perspectives in photogrammetry to provide in Earth observation three-dimensional surface images, as well as in telecommunication networks for testing of redundancy concepts in routing. Thus ambitious objectives are to be realized in the limited volume, power and mass budgets of pico-satellites. The different mission phases include initialization of the formation after deployment of the 4 satellites from the launcher moving to the planned relative distances, as well as the subsequent formation keeping, despite the encountered disturbances of the space environment.

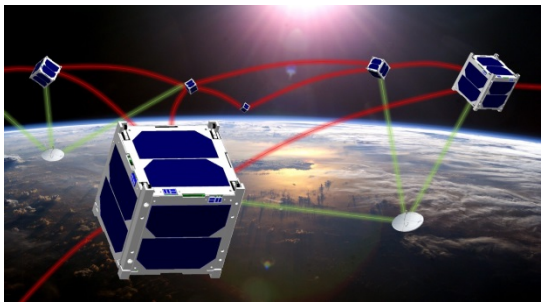


Fig.2: Future network composed of ground control stations and self-organizing, distributed, small satellite formations.

The challenges relate to

- attitude and position control based on relative navigation inputs,
- networked control techniques, closing control loops via telecommunication links, as well as autonomous reaction capabilities.

The combination of digital, packet-oriented communication techniques with control theory, requiring either continuous or at least by fixed sampling intervals generated inputs, leads to innovative challenging research tasks. Current research addresses the interrelation between different communication protocols and appropriate control approaches. Another relevant control challenges concerns the robust integration of autonomous reaction of the formation in orbit with remote operations commands from the ground control center in order to avoid conflicts and contradictions.

4. Applications in Earth Observation and in Telecommunications

While the advantages of sensor networks are already realized in our standard terrestrial environments, in space we are still in a pioneering phase to exploit related advantages. Distributed networks of multiple satellites support

- multi-point measurement capabilities
- higher temporal / spatial resolution in observation data,
- higher availability,
- graceful degradation capabilities in case of failures on-board.

In a distributed multi-satellite system data from various perspectives can be fused to provide added value in resolution and 3-dimensional data generation in remote sensing from 2-dimensional raw-data images. In comparison with constellations the formations don’t have to wait for ground control contact to configure in a suitable way for appropriate observations, but continuously optimize their positions for the planned mission objectives in a self-organized manner by exchanging the relevant information in the formation.

An interesting example of a small satellite approach to access atmospheric layers, where due to limited satellite lifetime only few data are available, is provided by the mission QB50 [17], supported by the SPACE-Program of the European Union. Here a satellite constellation composed of 50 CubeSats is employed for exploration of the lower Thermosphere by in-situ multi-point-measurements in the altitude range between 200 km and 380 km.

Interesting commercial perspectives are pointed out in use of many small satellites in Earth observation ([7], [9]), where Planet Labs (www.planet.com) plans for continuously receiving images from the Earth's surface by nominally 125 „Dove" satellites (CubeSats with the dimensions 30 cm x 10 cm x 10 cm) in orbit [8]. In the field of telecommunication the companies/consortia OneWeb and SpaceX plan a global provision of Internet with many (about 700 respectively 4000) low Earth orbiting small satellites. All these missions are currently planned as constellations, but it is expected that the operational complexity and costs will lead in the next generation to implementations as formations.

5. Conclusions

Pico-satellite formations offer interesting application perspectives for distributed multi-satellite missions within a limited cost frame. This offers attractive opportunities in Earth and Space weather observations by sensor networks, distributed in orbits to cover a wide area. Cooperating distributed satellite networks thus offer a high temporal and spatial resolution for the payloads, as well as graceful degradation properties in case of malfunctions. Current constellations in Earth observation and telecommunications might in future be replaced by formations in order to increase real-time reaction capabilities and to reduce operational complexity in ground control centers. Also combinations of traditional large satellites as energy-demanding emittors and networks of small satellites as detectors collect the reflected signals over a large area. This approach promises significant, cost-efficient performance increases in Earth observation scenarios

Acknowledgements

The contributions of all the engaged collaborators in the UWE-projects are acknowledged, as well as the financial support from the German Ministry of Economics via the German Space Agency DLR, the Space Research program of the Bavarian Ministry for Economics and the European Research Council.

References

[1] Alfriend, K. T. , S. R. Vadali, P. Gurfil, J. P. How, L. S. Breger,. *Spacecraft Formation Flying. Dynamics, Control and Navigation , Elsevier Astrodynamics.* 2010.

[2] Ankersen, F. (ed.), Proceedings *3rd International Symposium on Formation Flying, Missions and Technologies*. *ESA SP-654.* 2008

[3] Bangert, P.; Busch, S.; & Schilling, K.; Performance Characteristics of the UWE-3 Miniature Attitude Determination and Control system, Proceedings *2nd IAA Conference on Dynamics and Control of Space Systems (DYCOSS), Rome* 2014, IAA-AAS-DyCoSS2

[4] Barza, R., Y. Aoki, K. Schilling; *Cubesat UWE-1 – Technology Tests and in Orbit Results*. In: Proceedings *57th International Astronautical Congress* 2006, IAC-06-B5.3.

[5] Busch, S.; Schilling, K.; *UWE-3: A Modular System Design for the Next Generation of Very Small Satellites*. Proceedings of *Small Satellites Systems and Services - The 4S Symposium*, Slovenia. 2012

[6] D'Errico, M. (ed.), *Distributed Space Missions for Earth System Monitoring*. Springer Verlag 2012

[7] Hand, Eric; The rise of the CubeSat , Science vol. 346 issue 6216 (19.12.2014), p. 1449 http://www.sciencemag.org/content/346/6216/1444.full.pdf

[8] Hand, Eric, Startup Liftoff, Science Vol. 348, Issue 6231 (10.April 2015), p. 172 – 177

[9] Jones, N.; *Mini satellites prove their scientific power*, Nature 508 (17 April 2014),p. 300 – 301 www.nature.com/news/mini-satellites-prove-their-scientific-power-1.15051

[10] Kronhaus, I., Pietzka, M., Schilling, K., 2013. *Pico-Satellite Formation Control by Electric Propulsion*. Proceedings 19th IFAC Symposium Automatic Control in Aerospace, Würzburg 2013

[11] Sandau, R., Nakasuka, S., Kawashima, R., Sellers, J., 2012. *Novel Ideas for Nanosatellite Constellation Missions*, IAA Book Series

[12] Schilling, K., *Design of Pico-Satellites for Education in System Engineering*, IEEE Aerospace and Electronic Systems Magazine 21 (2006), pp. 9-14.

[13] Schilling, K.; Networked Control of Cooperating Distributed Pico-Satellites, Proceedings IFAC World Congress Cape Town. 2014. ID: WeC19.6

[14] Schmidt, M., K. Ravandoor, O. Kurz, S. Busch, K. Schilling; *Attitude Determination for the Pico-Satellite UWE-2*. Space Technology 28 (2009), pp. 67-74

[15] Schmidt, M. & Schilling, K.; Formation Flying Techniques for Pico-Satellites *6th International Workshop on Satellite Constellation and Formation Flying*.2010

[16] Walker, R. ; *CubeSat Nano-Satellite Systems*, Technical Dossier ESA/ESTEC TEC-SY/145/2013/TNT/RW, 2013

[17] www.qb50.eu

Trajectory/General Parameters Optimization for Suborbital Launch Vehicle

Sun Shijie, Zhang Yechen, Min Changwan, Xu Chunling, Wang Guanyu

Science and Technology on Space Physics Laboratory, Beijing, 100076
China Academy of Launch Vehicle Technology, Beijing, 100076

Abstract: Suborbital Space Vehicle, running at the edge of the atmosphere, has many advantages of low cost, high reliability, rapid entry/reentry into space, which has raised full attention. In this paper, the issue of optimization design for Suborbital Launch Vehicle (SLV) is discussed. According to the characteristics of the Suborbital Space Vehicle, an idea of integrated optimization for trajectory/general parameters is proposed. The mathematical model of SLV trajectory/general parameters is established, including the mass model, dynamic model, aerodynamic model, flight procedure controlling model, and engine interior ballistic model. The variables and constraints of the trajectory/general parameters integration optimization are determined. The hierarchical optimization strategy, combined with genetic algorithm and pseudo-spectral method, is proposed. The handover velocity is set as objective function. Genetic algorithm is adopted in external layer to search the optimized solution of system parameters, the propellant fuel and thrust of each stage. Pseudo-spectral method is adopted in interior layer for trajectory optimization to ensure each group of system parameters is corresponding to an optimal flight procedure. Simulation results illustrate, the handover velocity raised effectively, with fixed fuel mass. The model and the method can also be applied to other launch vehicle design, and have high value of engineering application and prospects.

Keywords: Sub-orbital, Launch Vehicle, Integrated Optimization Design, Pseudo-Spectral Method, Genetic Algorithm

1. Introduction

Suborbital flight refers to a status that the vehicle run at the edge of atmosphere, but it's speed is not enough to complete the circumduction of the earth orbit[1].The suborbital flight's height is approximately between 80km to 120km, which has wide application and great military value. Compared with the traditional spacecraft, suborbital vehicle has lower orbit altitude, and the flight path of the vehicle is all within the atmosphere. That's requires the launch vehicle turn quickly in the atmosphere, in order to achieve a small height and path angle. Traditional launch vehicle was designed for space orbit launch, which leads to increased costs and decreased reliability if utilized for suborbital launch. As the exploitation of suborbital space by many countries, it is necessary to optimize the design of existing launch vehicle to better fit the suborbital launch mission.

According to the characteristics of the suborbital space vehicle, a launch vehicle model is established in this paper .A hierarchical optimization strategy combined with genetic algorithm and pseudo-spectral method is utilized for the integrated optimization of suborbital launch vehicle. The result shows that after the optimization, the first stage fuel mass is reduced, the second and the third stage fuel mass are increased; the thrust of the first stage and the third stage are increased, the thrust of the second stage is reduced, enhancing the burnout point speed of the launch vehicle.

2. Mathematics model of launch vehicle

2.1 Mass model

The launch vehicle is a three-stage solid rocket. The mass of the rocket consists of the mass of each stage's engine M_{Ei}, the mass of interstage M_{12}, M_{23} sections and the mass of the payload M_{load}.

$$M_{total} = \sum_{i=1}^{3} M_{Ei} + M_{12} + M_{23} + M_{load} \tag{1}$$

The mass of each stage's engine M_{Ei} can be obtained by the fuel mass M_{Fi} of the engine and the mass ratio r_i of the engine. At the general design phase, the mass ratio r_i is considered as constant, and mass of the engine M_{Ti} can be calculated from the fuel mass.

$$M_{Ei} = M_{Fi} / r_i \tag{2}$$

2.2 Dynamics model, aerodynamic force model and propulsion model

Dynamic Model is built in a longitudinal movement plane of the rocket. In order to simplify the model, we ignore terms related with attitude angular velocity and angular acceleration in the rocket dynamic equation; ignore the rotation of the earth and treat the earth as a homogeneous sphere; ignore the control force. Therefore, the 3-DOF rocket dynamic equation under the trajectory coordinate system can be expressed as:

$$\begin{cases} \dfrac{dV}{dt} = \dfrac{P\cos\alpha}{m} - \dfrac{D}{m} - \left(\dfrac{R_0}{R_0+H}\right)^2 g_0 \sin\theta \\ \dfrac{d\theta}{dt} = \dfrac{P\sin\alpha}{mV} + \dfrac{L}{mV} - \left(\dfrac{R_0}{R_0+H}\right)^2 \dfrac{g_0\cos\theta}{V} \\ \dfrac{dH}{dt} = V\sin\theta \\ \dfrac{d\beta}{dt} = \dfrac{V\cos\theta}{H+R_0} \\ \dfrac{dm}{dt} = -\dot{m}_s \end{cases} \tag{3}$$

Where $\dot{m}_s$ represents fuel flow rate; θ represents local path angle; β represents geocentric angle; L represents lift coefficient, D represents drag coefficient, and S represents the reference area of the vehicle.

$$\begin{aligned} D &= \frac{1}{2} C_D \rho S V^2 \\ L &= \frac{1}{2} C_L \rho S V^2 \end{aligned} \tag{4}$$

Lift coefficient and drag coefficient can be calculated by the existing rocket fitting formula.

$$C_L = f(\alpha, Ma), \quad C_D = f(\alpha, H, Ma) \tag{5}$$

Propulsion characteristics of the engine can be affected by many factors. At the general design phase, in order to simplify the model, we consider the parameters of the engine are constants.

$$F_i = C_i \tag{6}$$

2.3 Flight procedure control model

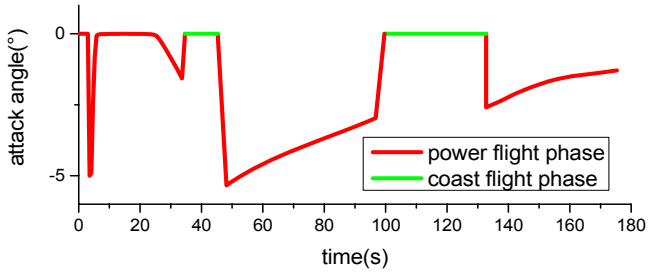


Figure 69 flight procedure control model

The design of the attack angle of the launch vehicle is divided into five phases, including the first stage phase, the second stage phase, the third stage phase, and two interstage unpowered glide phases. The first stage phase can be divided into the vertical ascending section, the turning section, and the transonic section. In the vertical ascending section, the transonic section, and the interstage unpowered glide phases, the attack angle need to be zero. Attack angle of other phases/section, including the first stage turning section, the second stage phase, and the third stage phase, should be design under the constraint of maximum negative attack angle and maximum change rate of the attack angle.

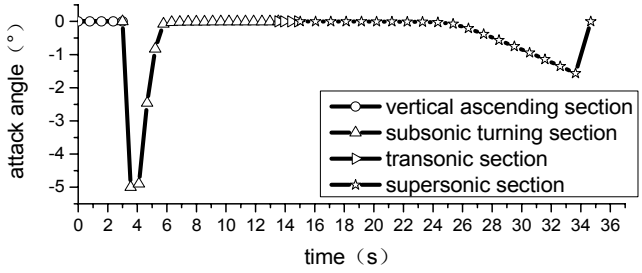


Figure 70 flight procedure control model

3. Optimization model

3.1 Optimization target and optimization variables

The handover velocity of suborbital vehicle is very important to its range, therefore, the handover velocity of suborbital vehicle is taken as the optimization target, which is equivalent to the following formula:

$$\min\ J = -v_f \tag{7}$$

Variables in the mass model including engine mass, interstage mass and payload mass. The mass of interstage is mainly determined by the load environment and the stage separation scheme during the flight. In order to simplify the optimization model, the interstage mass and the payload mass are considered as constants. The engine mass of all stages account for more than 90% of the total rocket mass, and the engine mass is determined by the ratio of the fuel mass and the engine mass. At the general design phase, the mass ratio is fixed, and the engine mass is only determined by the fuel mass. So the fuel mass of the first to third stage of the engine m_{F1}, m_{F2}, m_{F3} are chosen as the optimization variables from mass model, and we assume that the change of fuel mass only affect the engine's length without affecting the engine's diameter.

Select the thrust of each stage's engine T_1 T_2 T_3 as optimization variables from the interior ballistics model.

In the flight procedure controlling model, flight time of each stage is determined by fuel mass and thrust. Attack angle of each stage flight and unpowered glide time are chosen as optimization variables.

Generally, select the fuel mass M_{Fi} and thrust Ti of each stage, coast time between first-stage and second-stage, coast time between second-stage and third-stage t_{12}, and the attack angle of the rocket flight procedure $\alpha(t)$ as optimization variables.

3.2 Constraints

Take the total propellant fuel mass of the rocket as constant, take the handover altitude and path angle as constraint.

Restrictions the maximum negative attack angle of each phase、attack angle change rate of each phase. Attack angle of slide phase、vertical ascent phase and transonic phase is limited to 0. Restrictions maximum negative attack angle of other phase no more than 12°.

$$H(f)=H_f,\theta(f)=\theta_f$$

$$\sum_{i=1}^{3} M_{Fi}=M_F$$

$$|\alpha|\le 12^\circ,|\dot{\alpha}|\le 4^\circ/\mathrm{s}$$

$$n<n_{\max},Q<Q_{\max},Q_{12}<Q_{12\max} \tag{8}$$

3.3 Hierarchical Optimization Strategy

Optimization variables can be divided into two types: static variables, including engine fuel mass $[m_{F1},m_{F2},m_{F3}]$, thrust $[m_{F1},m_{F2},m_{F3}]$, unpowered glide time t_{12} and t_{23};The other is dynamic variables which is attack angle sequence $\alpha(t)$.

For this optimization problem which mixed with static variables and dynamic variables, a hierarchical optimization strategy combined Genetic Algorithm and pseudospectral method is used in this paper. Optimization strategy is divided into inner and outer layer. Genetic algorithm of the external layer is used to optimize the fuel mass, thrust and unpowered glide time. Pseudospectral method of the interior layer is used to optimize the interior flight procedures. Pseudospectral method is a rapid development and widely applied method which can transform the trajectory optimal control problem into a nonlinear program problem. The basic steps pseudo-spectral method is as follows

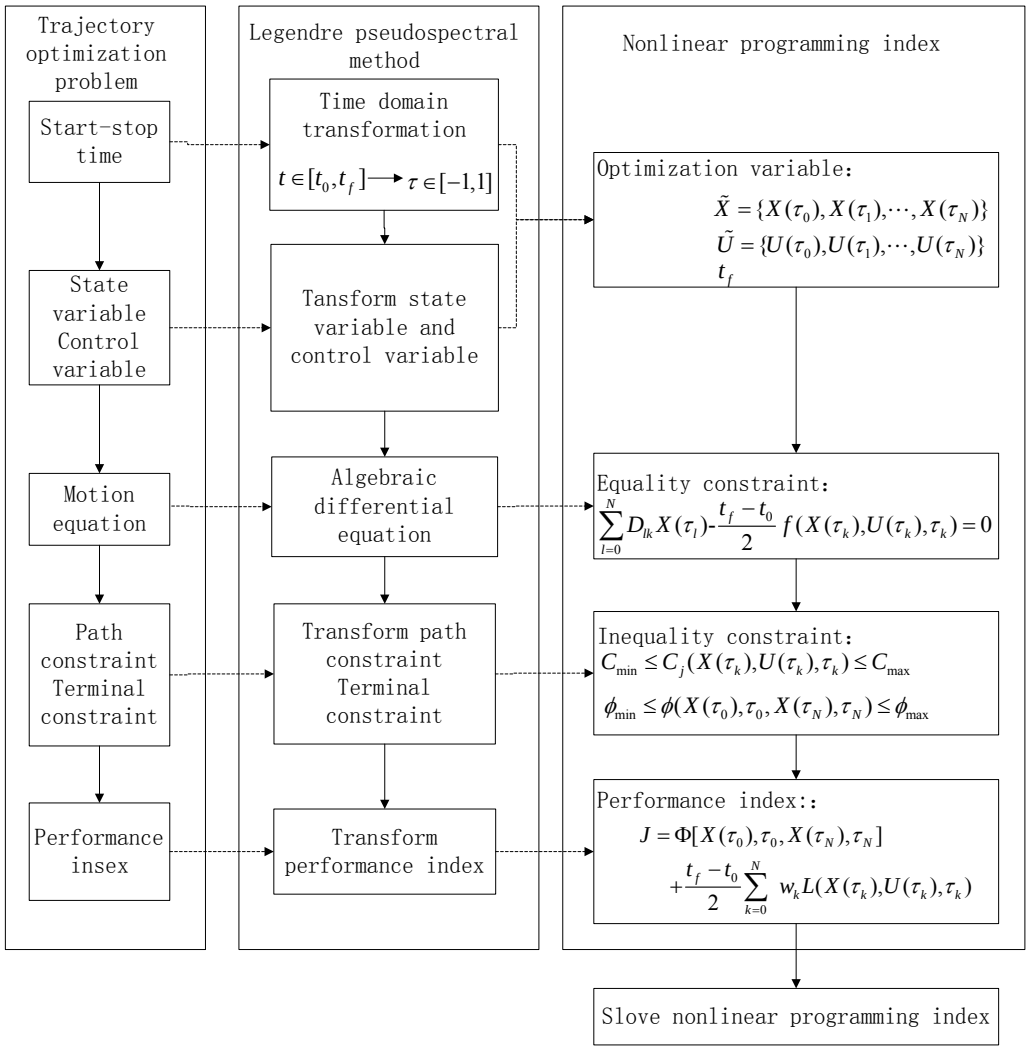


Figure 71 The basic steps of pseudo-spectral method

The basic steps of hierarchical optimization strategy:

1） The genetic algorithm of the outer layer gives a number of parameters set which contains fuel mass，thrust and unpowered glide time as the initial population.

2）The pseudospectral method of the inner layer takes parameters set as input condition, take the handover velocity as the objective function for solving the optimal flight attack angle procedures

3）The pseudo spectrum method is used to compute the handover velocity of each parameters set, then take the handover velocity as the fitness of individuals.

4）Check whether the termination condition is satisfied: if the iteration termination condition is satisfied, select the best optimum individual from the last generation as the optimal results of optimization; otherwise continue the iteration.

5）Generate a number of new individuals with selection, crossover, mutation and other genetic operations and keep the best individual; return step 2) continue the iteration.

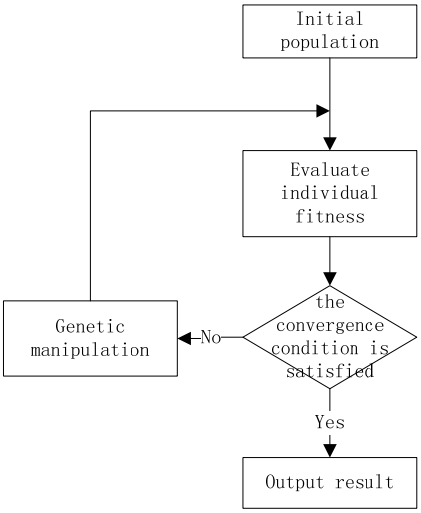


Figure 72 hierarchical optimization strategy

4. Optimization results analysis

Optimize the fuel mass, trust, coasting time and flight procedures of the rocket under the constraints 80km handover altitude and 2 degree handover path angle. Population number is 50. Convergence in the twentieth generation. The optimization results are compared with the performance index of the original rocket as follows:

Table 20 performance index contrast

	handover velocity (m/s)	ε_{12}	ε_{23}	t_{12} (s)	t_{23} (s)
before optimization	6837.3	1.939	2.875	0.0	0.0
after optimization	6962.4	1.310	2.803	10.6	33.1
	thrust-weight ratio of first stage	thrust-weight ratio of second stage	thrust-weight ratio of third stage		
before optimization	1.91	3.37	4.11		
after optimization	3.22	3.38	4.95		

In table 1, ε_{12} is the ratio of the first stage fuel mass and the second stage fuel mass, and ε_{23} is the ratio of the second stage fuel mass and the third stage fuel mass. In the case of same total fuel mass, after the optimization, the fuel mass of the first stage is reduced, and the fuel mass of the second and third stage is increased, at the same time, the thrust of the first and third stage is increased, and the thrust of the second stage is reduced. Besides, two coasting phases are added to the flight procedures, these factors make the optimized handover velocity increased by 125m/s.

Table 21 velocity loss contrast

	handover velocity（m/s）	ideal velocity（m/s）	gravity loss（m/s）	drag loss（m/s）	nozzle pressure loss（m/s）	attack angle loss（m/s）	Total loss（m/s）
before optimization	6837.3	7665.2	724.3	174.7	-102.7	31.4	827.7
after optimization	6962.4	7746.9	587.5	210.0	-28.2	15.2	784.5
increment	125.1	81.7	-136.8	35.3	74.5	-16.2	-43.2

In this table, the nozzle pressure loss is a negative loss, the rest of the losses are positive losses. Nozzle pressure loss depends on the impact of the atmosphere on the rocket thrust. In this paper, when calculate the velocity loss, the thrust of the first stage is ground thrust, with the increase of altitude, the air density decreases, and the actual thrust will be larger than the ground thrust; the thrust of the second and the third stage is vacuum thrust, the actual thrust will be smaller than the vacuum thrust. When the nozzle pressure loss is calculated, the former has a large proportion, combine the two factors, the former leads to a negative loss of the nozzle pressure loss. So the nozzle pressure loss will increase the velocity of the rocket in this case. As we can see in this table, after optimization, the ideal velocity increased, the gravity loss and the attack angle loss decreased the loss of resistance, drag loss, nozzle pressure loss increased. The change amplitude of drag loss and attack angle loss is relatively small.

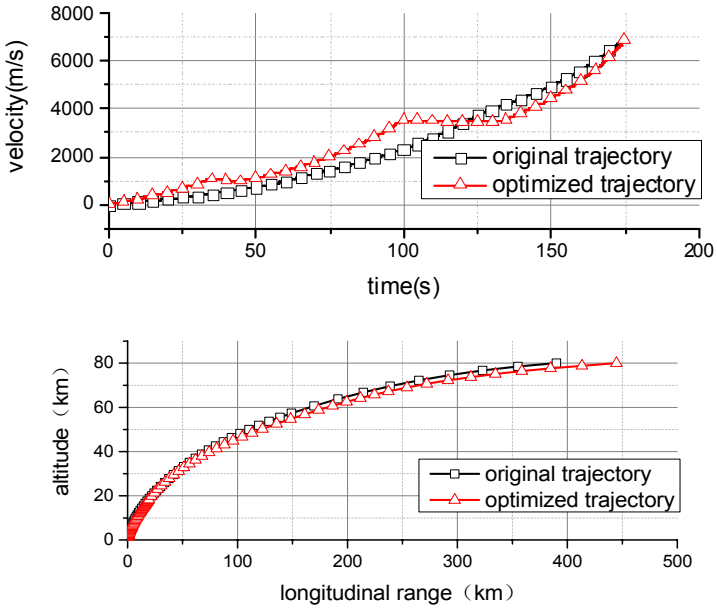


Figure 73 trajectory contrast

After optimization, there are several differences in the general parameters and flight procedures of the rocket: interstage ratio configuration changes, the optimized interstage ratio configuration makes the ideal velocity increase; the optimized flight procedures has a large negative attack angle at the early flight, so that the rocket can quickly turn to achieve smaller handover altitude and path angle, besides, a large negative attack angle at the early flight will extend the flight time at lower altitude, and increase the nozzle pressure loss and drag loss, but the trajectory of the rocket will be more gentle, and the reduced gravity loss is sufficient to compensate for the increased

nozzle pressure and drag loss. The two coast flight time will also make the trajectory more gentler, and because of the altitude of the coast phase between the second stage and the third stage is higher than the coast phase between the first stage and the second stage, the air density and path angle is smaller, the velocity loss is smaller, so the coast phase between the second stage and the third stage is more suitable for sliding, and also has a long coast flight time.

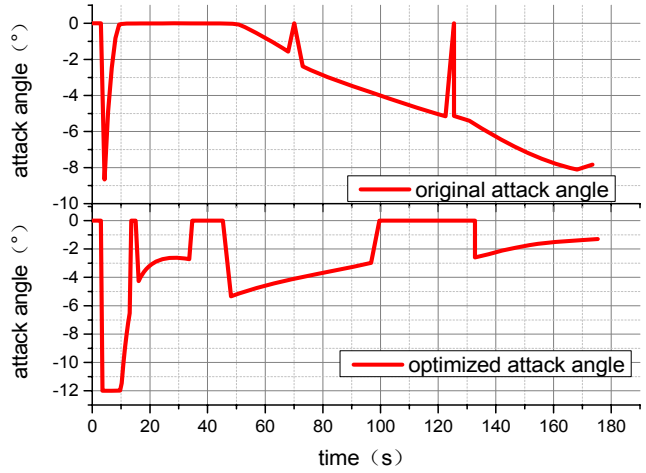


Figure 74 attack angle contrast

5. Conclusion

In this paper, genetic algorithm and pseudo-spectral method are combined for the integration optimization of the general parameters and flight procedures of suborbital Launch Vehicle. The optimization results show the algorithm is effective. In the case of the same total propellant fuel mass, Optimization of the fuel mass distribution, the thrust of each stage and the flight procedures of the rocket, the handover velocity increased by 125m/s. Better distribution of the fuel mass will increase the ideal velocity. The optimized flight program showed that the launch vehicle should have a large negative attack angle at the early flight, and add coast flight phase to the flight program, make the trajectory as gentle as possible, reduce the loss of velocity.

reference：

[1] Mao Zhihui; Yang Huabo; Zhang Qiang; Zhang Shifeng. The Study on Dynamic Characteristic of Suborbital Launcher [J]. Journal of Projectiles,Rockets,Missiles and Guidance,2012,03:152-156.

[2] Zong Qun; Tian Bai-ling; Dou Li-qian. Ascent Phase Trajectory Optimization for Near Space Vehicle Based on Gauss Pseudospectral Method [J]. Journal of Astronautics,2010,07:1775-1781.

[3] Sha Jianke; Shi Yuyang; Wan Ziming; XU Min. Integrated Optimization Design of Trajectory and Overall Designing Parameters for Ramjet Powered Missile [J]. Modern Defence Technology,,2014,03:37-42.

[4] Zhang Ning-ning; Min Chang-wan; Liu Hui; ZHANG Ye-chen. Integrated optimization design for mass distribution / trajectory of near space directly injecting carrier rocket [J]. Flight Dynamics, 2014, 03.

[5] Jia Peiran, Chen Kejun, He Li. Long-range Rocket Ballistics［M］. Changsha: National University of Defence Technology Press，1993:230.

Innovational Design and In-orbit Test of Small SAR Satellite Electrical Power System

JIA Xiaodong, YAN Wanjuan, WANG Wentao, ZHANG Running

（DFH Satellite Limited Corporation Beijing China）

Abstract： At present, the Synthetic Aperture Radar (SAR) is developed to be applied in the field of small satellite. This paper presents the Electrical Power System (EPS) design, with particular attention on specific requirements of the peak power requirement by the SAR antenna and the volume and mass restriction of the whole small satellite. Then, the paper presents test and flight data demonstrating the effectiveness of the EPS design, corresponding to the peak power.
Key words: Small satellite; High power; EPS design

1. INTRODUCTION

The Synthetic Aperture Radar（SAR）is kind of earth observing payload of spacecraft， which can take high resolution microwave image of ground objects. The Synthetic Aperture Radar can observe all-weather and all day. Since 1970's, the US and EU countries started to launch SAR satellites. For now, the SAR satellite is developing to the direction of high resolution, wide width, multi-polarization, light miniaturization and constellation network.

This paper presents the small satellite Electrical Power System (EPS) design on regulated primary power bus managed by power control unit (PCU) and there are two separated power buses with voltage in the range from 27.5 to 29.5V and 33V to 35V. The higher power bus supplies electrical power to solid state amplifier (SSA) of SAR, and the lower power bus supplies electrical power to platform equipments and other payload electronic equipment.

2. TOPOLOGY OF EPS

By analysis the global SAR satellites, such as TerraSAR, TecSAR, Cosmos-skymed, there are three kinds of topology of EPS to satisfy the peak power requirement of SAR: unregulated primary bus, combined primary bus and dual-primary bus.

The unregulated primary bus has low output impedance which can mostly satisfy the requirements short peak power and high repeat power. And the combined primary bus can use the SA-BAT union system effectively, at the same time, satisfy the different requirements of platform and payloads. The dual-primary bus can supply the high quality output power to both of the platform and payloads.

Table 1 Summarize of different EPS topology

Unregulated primary bus	Combined primary bus	Dual-primary bus
One primary bus supply for both the payloads and platform. Power bus voltage changes a lot within the battery voltage	One primary bus(unregulated bus) supplies for the payloads The other one primary bus (regulated bus) supplies for the payloads platform. One set of EPS（SA×1，battery×1）	Two regulated primary bus supply for the payloads and platform independently Two sets of EPS（SA×2，battery×2）

According to the characteristic of SAR and requirements of whole satellite, the EPS satisfy the following requirements:

- Because of the restriction of volume and mass (100kg level), EPS adopts SA-BAT union system with high efficiency charge/discharge regulator, in order to use the limited power of the satellite more efficiently.
- EPS should avoid the influence of payload reflection ripple to other equipment.
- EPS should supply the primary power bus voltage to SSA in range of 33.5V±0.5V

EPS made a tradeoff table in the design phase as Tab 1 to select the topology of EPS.

Table 2 Tradeoff table of EPS topology

Item	Unregulated primary bus	Combined primary bus	Dual-primary bus
Mass at 100kg level	√	√	√
SA-BAT union system	√	√	√
Avoid the influence of payload	×	√	√
High stability of bus voltage for SSA	×	×	√

So we chose the dual-primary bus for The EPS block diagram is shown in Fig.1.

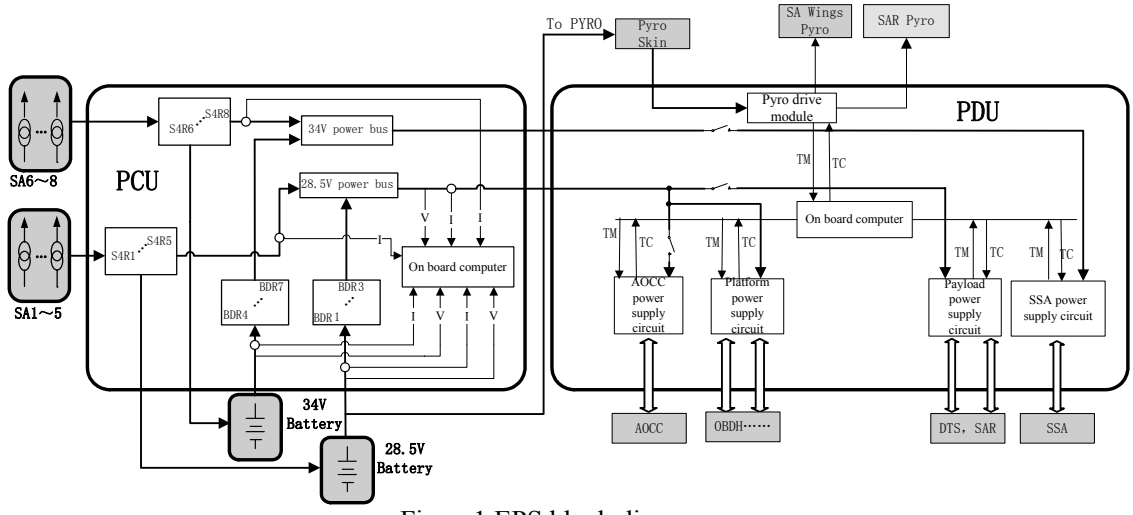


Figure1 EPS block diagram

The constitution of EPS is shown in Tab.3:.

Table 3 Constitution of EPS

Item		34V power bus	28.5Vpower bus
Power storage		70Ah Cd-Ni battery	50Ah Cd-Ni battery
Power generation		1.3kW@BOL/AM0/25℃ 0.8kW@EOL/AM0/80℃	
Power management	Regulator and charge	S4R/8levels/efficiency≥92%	
	Discharge	Rated power≥400W×4	Rated power≥240W×3
	Power distribution	Switch for 34V power bus×1 Closed-loop surge elimination circuit	Directly for platform Switch for AOCS×1 Switch for payload×2

Based on the topology of dual-primary bus, EPS focus on the characteristics of SAR and do some innovational design such as ducal- primary bus isolation, closed-loop surge elimination circuit and high stability of power bus voltage.

3. INNOVATIONAL DESIGN OF EPS

3.1 DUAL-PRIMARY POWER BUS ISOLATION DESIGN

EPS uses the dual-primary power bus solute the EMC problem caused by high power pulse payload. The 28V power bus supplies power for platform and parts of payload except SSA. This design is similar to the cast2000 satellite as before, which is good inheritance and proven technology. So EPS paid more attentions on the power supply for high power SSA. EPS adopt 34V power bus and SA-BAT union system for SSA. When the SAR payload doesn't work, SA supplies power to all equipment and the surplus power to charge the battery. When the SAR payload is working, the SA-BAT union system supplies power to payload together.

Because there are two sets of SA-BAT, only one PCU for both power buses, EPS focus on the isolation in PCU. The power bus and grounding lines should be isolated in PCU, and come together in PDU. The grounding scheme is depicted in Fig.2.

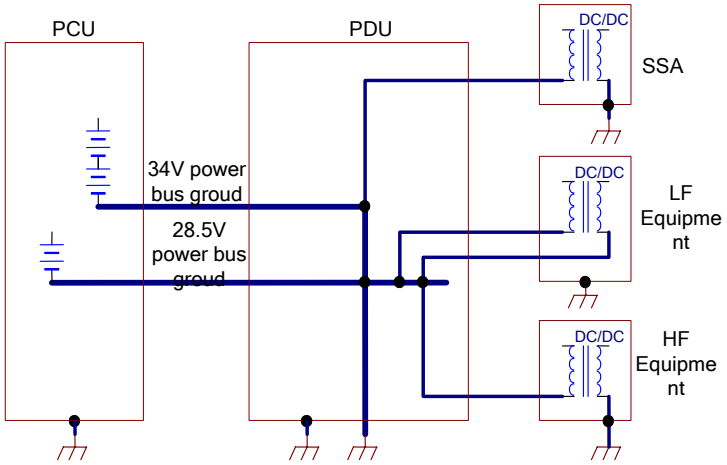


Figure2.EPS grounding scheme

In design phase, EPS test the voltage ripple of dual-primary bus as follow in maximum power test:

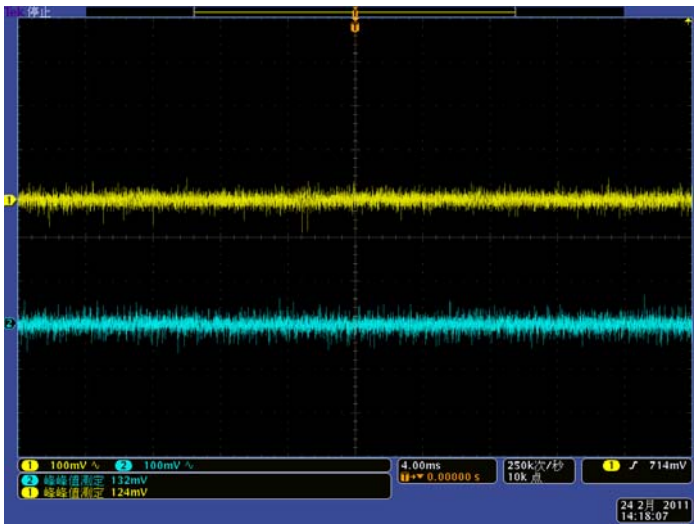


Figure3 Dual-primary bus voltage ripple in maximum power test

In maximum power test, 34V power bus supplied power both from SA and battery and the SAR worked for 15minutes. The oscilloscope showed the voltage ripple of 28V and 34V power bus. The peak-to-peak value is less than 150mV which can satisfy the EMC requirement.

3.2 SURGEELIMINATION CIRCUIT

The filter capacitor at the input of SSA which is more than 20000μF could cause a 200A surge current when SSA is power on. EPS designed closed-loop surge elimination circuit in PDU to limit the surge current to keep the switches safe, and reduce the impact for primary bus.

The principle diagram of the closed-loop surge elimination circuit is shown in fig.5. The circuit contains pre-charge circuit, voltage detecting circuit, CPU and control circuit. When the CPU receives the power on tele-command, the primary bus will charge the filter capacitor C through the pre-charge circuit containing K1, R1. And the Vo will increase to an preset value detecting by the voltage detecting circuit containing R2, R3. Then CPU could control the switch K2 power on.

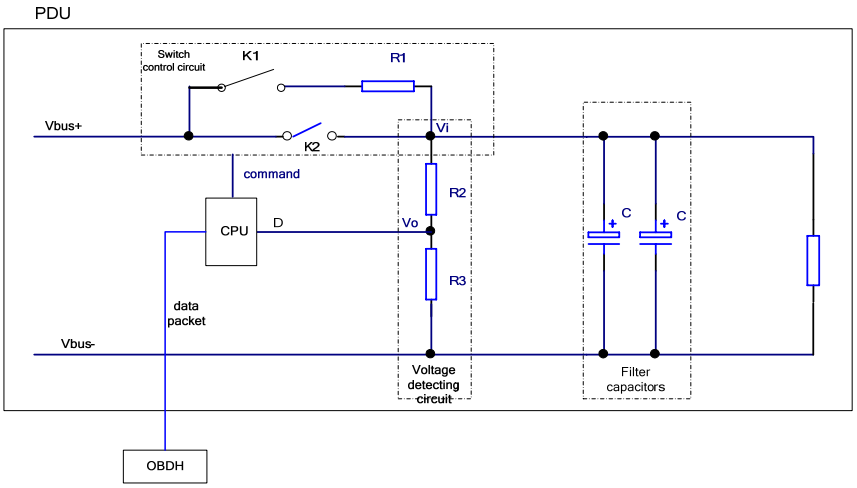


Figure4 Principle diagram of the closed-loop surge elimination circuit

In design phase, the circuit simulation result shows the circuit could limit the surge current to 1.6A/0.7s. The simulation result is the shown as follow. The upper figure shows the surge current without the closed-loop surge elimination circuit. The other one shows the result after using the circuit.

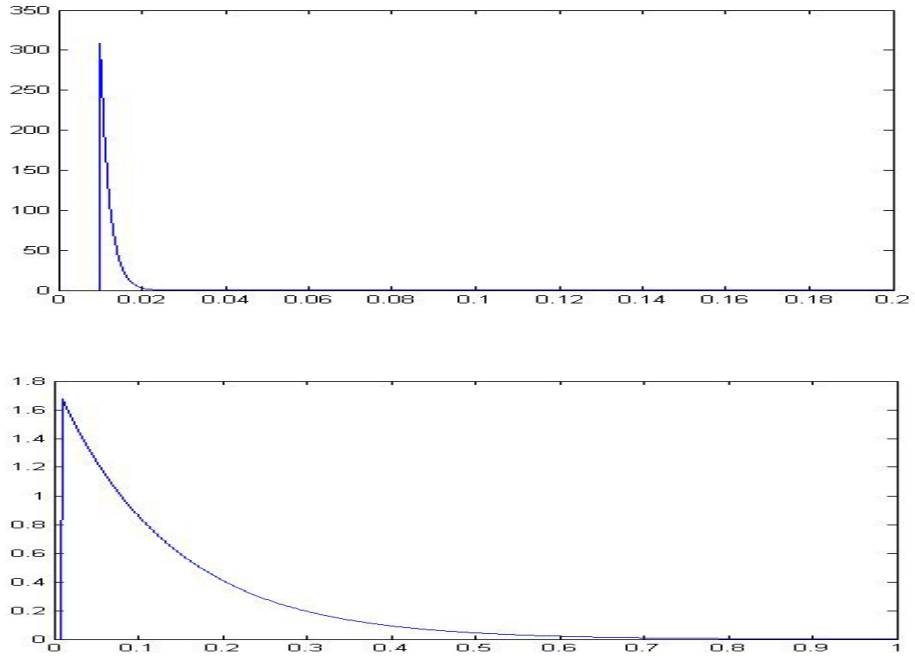


Figure5 The simulation result of the closed-loop surge elimination circuit

3.3 HIGH STABILITY OF HIGH VOLTAGE POWER BUS

The SSA of SAR required that the power bus should be in the range of 33.0V~31.0V when SSA is powered on. EPS adopts 2-region control by using main error amplifier of power bus, This design ensures that the power bus voltage in high stability in different modes.

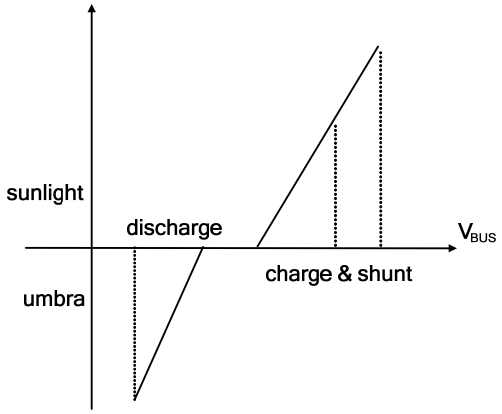


Figure6 Schematic diagram of 2-region control

EPS using strategies as follow:

● SAR powered off in sunlight: high voltage power bus works in the charge and shunt region and the voltage is above 34V.

● SAR powered on in sunlight or powered off in umbra: high voltage power bus works in the discharge region and the voltage is in range of 33V~33.8V.

The strategies ensure the power bus voltage can satisfy the SSA requirement. Fig.7 shows the actual power bus voltage in orbit in different modes which can prove the result of EPS design.

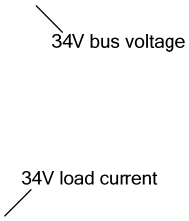


Figure7 Power bus voltage in different modes

From fig.7, the data shows before the SAR powered on (the load current at low level), the power bus voltage was stable above 34.6V. Then the load current increased step by step, the power bus voltage descent to 34V.When the load current increased to the peak value and lasted 3 minutes, the power bus voltage was in range of 33.68V～33.83V. The curve proves the result of EPS design could satisfy the SSA requirement.

4. IN-ORBIT PERFORMANCE

The Ni-Cd Battery and the PCU performances during the flight operations are detailed in this chapter. In Tab.4 the nominal battery characteristics and performances that have been taken into account for the EPS design are shown.

Table 4 Battery nominal characteristics

Items	28.5V	34V
Nameplate capacity	50Ah@0.2C	70Ah@0.2C
Depth Of Discharge	13.78%（0.32C）	9.66%（0.71C）
Max discharge current	0.32C	0.71C
Max charge current	12.5A	11.57A
BAT discharge voltage	19.8V	23.1V
BAT charge voltage	27.18V	31.71V

In ground test of PCU, the S4R charge efficiency is 95.3％ and BDR efficiency is 92.1%@28.5V, 93.4%@34V.

4.1 BATTRY CHARGE FUNCTION

The battery charge control is realized by the on board computer through a Ampere-Hour algorithm based on battery discharge/charge current telemetries received by the PCU. Fig.8 and Fig.9 show the behavior of this charge control system. There also is a voltage-temperature curve control method as a redundant method via a reliable circuit in PCU.

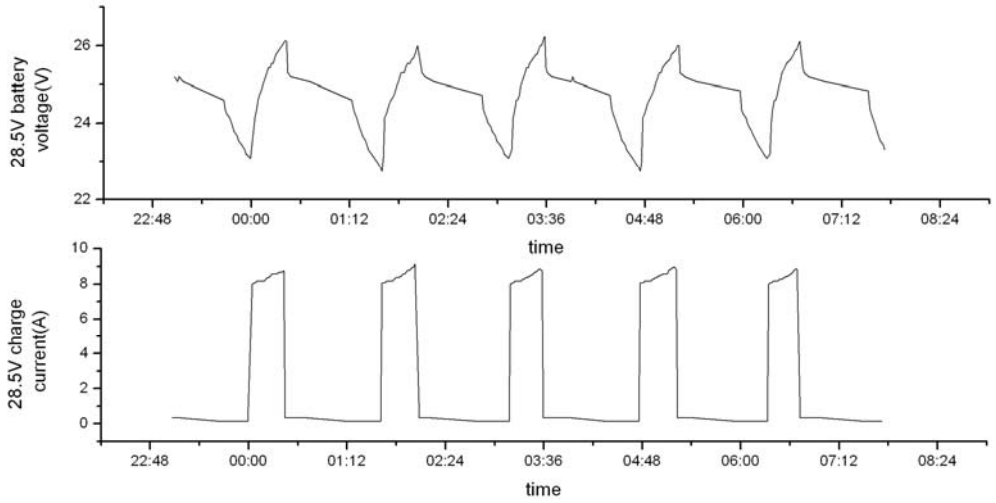


Figure8 28.5V battery charging (end of Dec 2013)

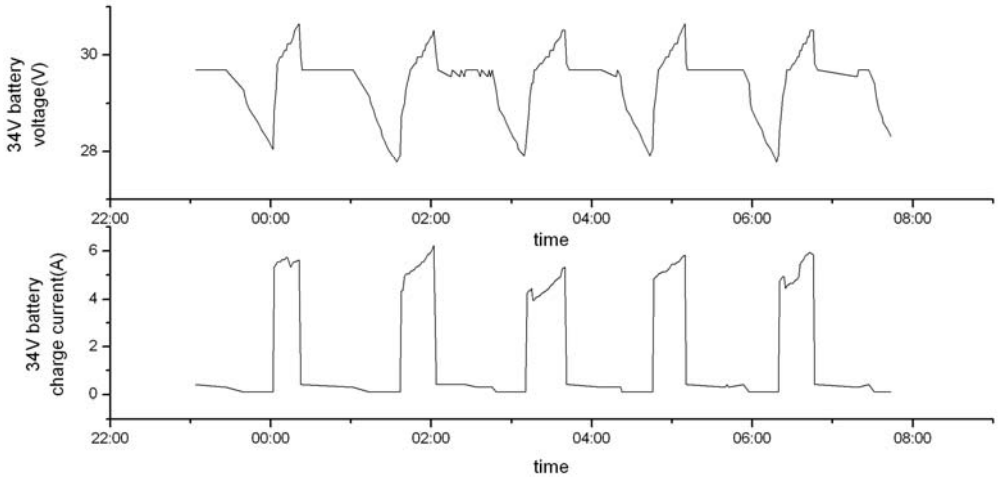


Figure9 34V battery charging (end of June 2013)

4.2 BATTRY DISCHARGE FUNCTION

The BDR efficiency is also an important evidence of the PCU performance. The TMs show the BDR efficiency during the orbits with the most demanding power request.

Table 5 34V BDR in orbit efficiency

date	34V bus voltage(V)	battery voltage(V)	discharge current(A)	BDR efficiency
2012-12-09	33.68	28.32	26.29	92.41%
2013-10-1	33.68	28.45	30.48	93.08%
2014-1-20	33.68	28.73	24.56	93.92%

4.3 SUPPLY POWER TO HIGH POWER LOAD

From in orbit data, it is possible to verify the effectiveness of the PCU design in terms of power delivered to the bus during SAR working. Because the high power demand of SAR is on the 34V power bus, we could draw a figure of 34V power bus by the telemetries involving SA current and battery discharge current. The function of PCU is to maintain the bus voltage stable and extract the power from battery to compensate the SA output power during SAR working. Fig10 shows the PCU performance during SAR working, the bus voltage is 33.68V which has fulfilled the requirement voltage of SSA as 33.5V±0.5V.

The high power request from the SAR payload for the imaging is reflected on the battery discharge current telemetries. The battery performances during SAR working also have been

demonstrated and validated by TM (see Fig.10). In working period, the max battery discharge current is about 36.4A which is under the designed max battery discharge current as 0.71C.

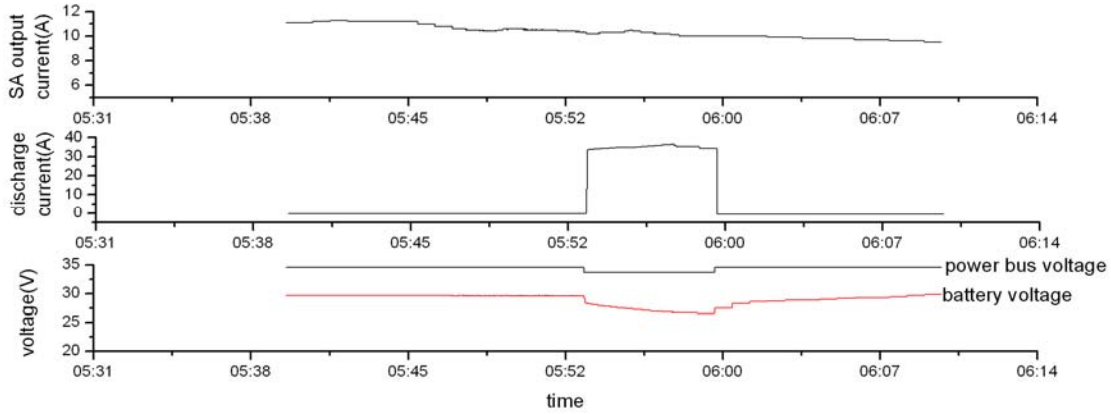


Figure10 34V performance during SAR working (end of Dec 2013)

5. SUMMARY OF THE EPS PERFORMANCE

The EPS mentioned in this paper has been test both in ground and orbits. The data which has been processed has shown an excellent performance of EPS, in terms of power delivered by the two solar array wings and by the two sets of battery. The dual-primary bus and the SA-battery union system can fulfill the high power payload perfectly.

REFERENCES

[1] Marco Manfreda, Stefano Costantini, et al., COSMO-SKYMED Electrical Power System in Flight Performances. Proceedings of the 8th European Space Power Conference, Constance, Germany, Sep, 2008.

[2] JIA Xiaodong, YAN Wanjuan, WANG Wentao, Electrical Power System Design of HJ-1-C Satellite, JOURNAL OF RADARS, 2014,Vol.3 Sup: 14-19

[3] ZHANG Xiaofeng, ZHANG Wenjia, GUO Weifeng, LIN Wenli, LIU Zhigang, Analysis and Enlightenment of Electrical Power System for SAR Satellite, SPACECRAFT ENGINEERING, 2015, Vol.24, No.3: 107-113

[4] ZHAO Zhiwei, JIN Lihua, Inspiration From Development of Oversea SAR Satellite System Technologies, SPACECRAFT ENGINEERING, 2010, Vol.19,No.4：86~91

Study on Application of Wireless Ad Hoc Network for Moonlet

WANG Li-sheng, WANG Di, ZHU Bo, TU Zhi-jun, Lu Yi-bin, Lai An-xue

(Shanghai Key Laboratory of Spacecraft Mechanism, Shanghai Institute of Aerospace System Engineering, Shanghai,201109)

Abstract: With the development of moonlet, the network and corresponding technique is needed. Then the wireless ad hoc network is introduced in this article, including network characteristic, structure and route. It is Analyzed that the wireless ad hoc network technique is applied for moonlet to form a network. And it is pointed out that the scheme can be perfected in subsequently study.
Keywords: Ad hoc network; Structure; Route; Moonlet; Framework

1. Introduction

Wireless mobile self-organizing network (Ad Hoc network) is a network of temporary, self-organizing, multi-jump and no center, which is composed of mobile terminals with wireless transceiver system. Because of the Ad Hoc network transmits data to the destination node through other nodes, and each node is both terminals and routers, therefore, Ad Hoc network is also known as multi-jump wireless network.

Ad Hoc Network technology research began in the late 70s, formerly of the us military Packet Radio Network, and its production and later development have the close military application background. As the result of the development of packet radio network, Ad Hoc network is especially suitable for the battlefield communications command and control because of its flexible network mode, rapid expansion and the characteristics of high anti-destroying ability. In recent years, Ad Hoc network also got rapid development in the field of civil. When no or can't take advantage of the existing fixed communication facilities, Ad Hoc network can rapidly complete the configuration and provide communication function, broadening the application environment of mobile communication network. With the continuous development of wireless communication and application, Ad Hoc networks will eventually become a good part of the whole communication environments for human beings.

2. Wireless self-organizing network system

2.1. The network characteristics

Ad Hoc network can not rely on any existing network infrastructure, and can at any time and any place quickly build up a multi-jump temporary autonomous mobile communication network without central access. When communication both in each other's transmission range, they can communicate directly between the two communication equipment; When communication is beyond the scope of each other's transmission, the communication between the two nodes need jump through a number of intermediate nodes. The characteristics of the Ad Hoc network can be summarized as follows:

a. no center and self-organizing: In the network there is no absolute control center, and it is a peer-to-peer network. Each node coordinates each other's behavior by a layered protocol and

distributed algorithm. Any node's fault will not affect the operation of the entire network, so the network has a strong anti-destroying ability.

b. the network topology can dynamically change: The mobile terminal can move independently, in any possible speed and movement pattern. Combined with different types of wireless transmitting antenna, the network topology of mobile terminals formatting through the wireless channel may change at any time.

c. multi-hops routing: the multi-hop routing of network is done jointly by ordinary nodes, without using special routing equipment;

d. wireless transmission bandwidth limited: Ad Hoc network uses wireless transmission technology. Because of the characteristics of wireless channel itself, the network bandwidth is relatively limited, much lower than cable channel.

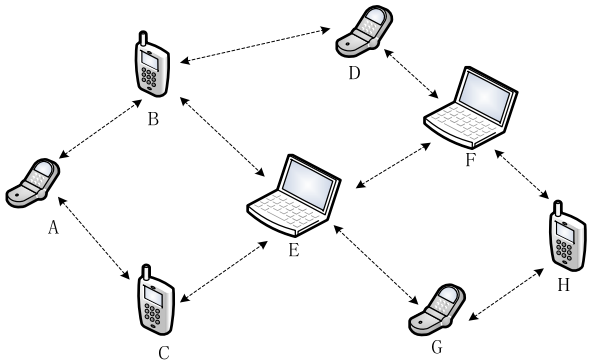


Fig.1 Wireless mobile ad hoc network

As shown in figure 1, the node A need to communicate with node F, but transmission range of node A can not covered node F, so this two nodes can't directly communicate with each other. With the aid of the other nodes, the path of A - > B - > D - > F, A- > C - > E - > F, and so on, can be chosen to transmit data.

2.2. The network structure

There are generally two kinds of structure of the Ad Hoc network, surface structure and hierarchical structure. All nodes have equal status in the flat structure, therefore this structure is also named equation structure. In hierarchical structure, the network is divided into clusters, and each cluster is composed of cluster head nodes and cluster member nodes. All these cluster head nodes formed a high level network, and the new network can be again divided into clusters, forming a higher level network, finally the highest.

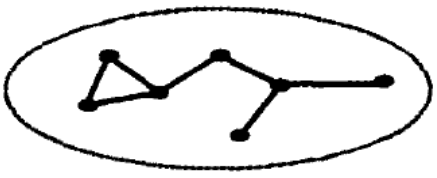


Fig.2 Wireless mobile ad hoc network

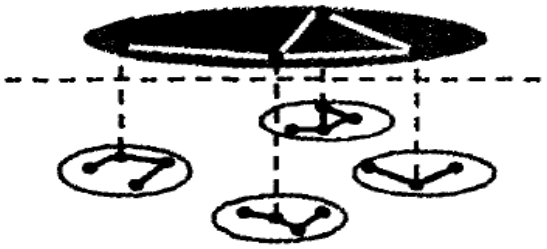


Fig.3 layered topology of ad hoc network

The network of surface structure is simpler, without any structure maintenance process. Generally there exist multiple hops between source and destination, so you can use multiple paths to realize the load sharing, and you can also choose the appropriate path for different business types. All nodes in the network are equivalent, and in principle there is no bottleneck, so the network is more robust. In this structure the coverage of nodes is small, and the network is relatively safe. The biggest drawback of the flat structure is that the network scale is restricted. In flat structure, each node needs to know the route to all other nodes. Because of node mobility, it needs a lot of control information to maintain the dynamic changes of routing information. The more is the network size, the greater the routing maintenance overhead is. When the network size increases to a certain extent, all of the bandwidth will be routing protocol consume, so the network of surface structure has poor scalability.

Hierarchical structure has the biggest advantages of good scalability. In this structure the network scale is relatively not limited, and when necessary, it can be improved by increasing the number of clusters or levels. In hierarchical structure, the function of the cluster members is simpler, and the cluster members basically do not need to maintain routing, all of this greatly reduces the amount of network control information. Cluster head nodes is complex, because it need to maintain the routing reaching to other cluster heads, and to know the ownership of all the nodes and clusters. On the whole, under the condition of the same scale of network, the routing overhead is smaller than that of flat structure. When the proportion of communication in a single cluster is bigger, each cluster can not interfere with each other, so the system throughput obviously is higher than that of the flat structure. However hierarchical structure also has its disadvantages. First of all, the maintenance of hierarchy needs more complex cluster head selection algorithm, thus this algorithm requires to be carefully designed. Secondly, because all information of cluster need the cluster head to find the transmission path, the route may not be the best one.

2.3. The network routing

The network performance greatly depends on the design of routing protocols. Because the network bandwidth is limited and topology changes frequently, the traditional routing protocol used in fixed network is not suitable for Ad Hoc networks. In Ad Hoc networks, because of the differences between the wireless transceiver and the impact of wireless channel, there may be a one-way channel. In addition, the periodic routing updates of conventional routing protocols can bring great burden to the network, and consume a lot of energy of the node at the same time. Therefore, the impact of these factors must be considered in routing protocols of Ad Hoc network. Ideal Ad Hoc network routing protocol should have the following features:

a. using distributed routing algorithm;

b. adaptive ability to the rapid changes in the network topology structure;

c. without loop;

d. small control overhead;
e. scalable, suitable for large-scale network;

3. Small satellite development and the network structure

3.1. Current situation of the small satellites

With the increasingly complex of spacecraft function, the cost of the single large satellites is higher and higher, and the risk is becoming more and more big. This gradually become a bottleneck of the development of the aerospace. And at the same time, small satellites and micro-satellite, for its low cost, light quality, small volume, short development cycle, high technical content, launching and using flexibly, etc, has been developed in various application fields. Many small satellites form a network by constellation or formation, and work together to undertake the task of signal processing, communications and payload test. This can realize the function of the single large satellite, reduce the production cost and the risk of failure, increase the ground coverage, shorten the return cycle, simplify the task to upgrade and maintain satellite components, and complete the complex space missions better. In addition, the network of small satellites will occupy the important position in the tactical integration environment of the future.

3.2. Small satellite network structure

The network of small satellites usually has two kinds: polling type and master-slave type. It happens to have the same view with flat structure and hierarchical structure of mobile self-organizing networks. Integrated thinking of the factors of node mobility, survivability and link anti-destroying ability of small satellite group, the hierarchical distributed architecture is more suitable for small satellite network. Hierarchical structure can not only guarantee the normal communication within the cluster member without many resources, and can meet the communication requirements of above cross-layer and collaborative work. in special circumstances. In addition, this architecture can largely reduce then control information and improve the utilization rate of channel. It can also manage the network routing hierarchically, which has good scalability, has no limit on the network scale, and meets the practical requirement of space missions.

4. Routing technology of small satellite network

4.1. The traditional interstellar routing technology

Traditional interstellar routing technology is divided into two kinds: based on interstellar link(ISL) and not based on interstellar. Routing technique based on interstellar routed without ground facilities. And because of independence on earth station, more reliable and safety, it is the ideal development direction to the future spacecraft communications. Routing calculation based on ISL completely rely on periodic and foreseeability of satellite operations, and most or all of the routing calculation are in offline calculated in advance. In some special cases, some satellite can leave the original orbit under the control of ground station, and some may be also failure because of malfunction or outside interference and attacks on them. Then the satellite network topology will appear unpredictable changes. At this point, if the routing algorithm depends entirely on the operation law of satellite set in advance, without probing the network topology, the fault link will not be bypassed, leading to a large amount of data connection interruption and loss of data, a significant reduction in the network performance and even crash. Although by monitoring of ground station the above problems can be found and solved, the delay time is too much to be suitable for

some system where the real-time demand is strict. In addition, this way increases the satellite network dependence on ground station, reduces the autonomous ability of the satellite system.

Domestic and foreign scholars put forward many routing algorithm based on interstellar, such as the satellite network routing strategy based on ATM, the satellite network topology design and routing algorithm, the connection-oriented satellite network routing algorithm based on the simple label switching, etc. These algorithms simplify the function of satellite by the way of off-line calculated in advance reducing the on-board real-time computation. Compared with traditional shortest path routing algorithm, offline routing algorithm greatly reduces the consumption of dynamic routing computing and signaling overhead, adapt to the characteristics of satellite network. However, offline routing algorithm also has its limitation unable to overcome, reflected in the following aspects:

a. It cannot dynamically adjust the route scheme based on statistical properties of the current network business, and cannot achieve optimal allocation of entire network resources, so there will be a certain link overload while others that can share the load was empty inequality, unable to guarantee the quality of service.

b. The network adaptability is bad. Because the offline routing table was calculated in advance, it can not be dynamically adjusted according to the current state of the network when part of link fails. Thus the system performance will be dropped significantly, resulting in a large number of existing connection interruption. Unless human intervention, symptoms can't ease.

4.2. Small satellite network routing technology

There is many similarities between small satellite network and the self-organizing network, such as mobile nodes, to transmit data by wireless channel, multiple hop, etc.. In addition, the satellite network is more stable than self-organizing network. In self-organizing network a mobile node could leave the network at any time, and new mobile node could join the network at any time. While in the satellite network, although the failure of some satellite may lead to a drop in the node, new satellite usually do not join the network randomly. In addition, the interstellar link is more reliable than wireless channel of self-organizing networks, and the structure is more stable. We separated the satellite network based on interstellar link from the satellite communication system including ground equipments such as terminals and gateway, and made the following assumptions: the data sent to the satellite by ground equipment is produced by the satellite itself and transmitted in the satellite network, that is, the satellite is the source of the data. So satellite network can be considered as a kind of wireless multi-jump network.

Satellite network, however, are not wholly self-organizing network, and there are obvious differences between them, mainly reflects in the followings:

a. Satellite network topology can change frequently, but the vast majority of change is predictable. It is because that the movement of the satellite is cyclical and predictable, and mobile node can move at any time in self-organizing network. The change of the satellite network topology can be accurately predicted, unless the satellite or interstellar link line is failure or the satellite moving orbit is artificially adjusted.

b. Different from the omni-directional antenna using in self-organizing network mobile node, beam antenna is used in satellite communication. This leads that, the mobile node in self-organizing network broadcasts data for all its neighbors and satellite communication based on interstellar link is point to point.

c. The distance between the satellite is greater than the distance between the mobile node in the

self-organizing network, thus the time delay of interstellar link transmission is greater than that of data transmission between mobile nodes through wireless channels in the self-organizing network. So the transmission delay, usually negligible in self-organizing network, cannot be ignored in the satellite network, in some cases, even bigger than the node processing delay.

In order to improve the adaptive ability of the network under the premise of not reducing the autonomy ability, a new small satellite network routing mechanism should be introduced. Taking the satellite network as a special case of the wireless multiple hops network, drawing lessons from the path algorithm of self-organizing network design ideas, on the basis of the existing various protocols to do the following:

a. in part of the interstellar link failure, the algorithm can automatically avoid the fault link, without human intervention;

b. having smaller routing signaling overhead;

c. by making use of satellite operation rule, the algorithm can optimize routing and further reduce routing signaling overhead;

d. algorithm supports multiple routing, can balance load between different link, can transmit the information of QoS and guarantee the QoS when routing;

e. using hierarchical network structure, having strong adaptability, and to ensure that there will be no routing loops.

Through the above measures, a feasible routing algorithm of satellite network in space environment, that is, a new dynamic routing algorithm based on interstellar link using self-organizing network design ideas, can be developed.

5. Conclusion

In this paper, it was analyzed that the wireless self-organized network technology is applied in small satellite network. Based on the characteristics, structure and rout of the wireless ad-hoc network, combined with the present situation of small satellite and network requirements, from the traditional comparative analysis of the advantages and disadvantages of two kinds of interstellar routing technology, a improved new type of dynamic routing algorithm is put forward for small satellite network. In view of the particularity of small satellite self-organizing network, it should be constantly studied, by means of simulation and test, to improve the safe and flexible design of the small satellite network .

Reference

[1]Wang haitao, song lihua song, design of military self-organizing network architecture, communication world, February 28, 2003.

[2] Zhao Ning, military self-organizing network architecture research, China Academic Journal Electronic Publishing House, 1994-2014.

[3]Tang zu, research on small satellite constellation ad-hoc network MAC protocol, master thesis, graduate school of national university of defense technology, in November 2011.

[4] Yao Zhongbang, research on the MAC protocol and routing algorithm based on topology information among Ad Hoc network, engineering Ph.D. Thesis, tsinghua university, April 2004.

[5] Qian shengxiao, air high-speed self-organization network MAC protocol, xi 'an university of electronic science and technology, master thesis, December 2011.

[6] Wu Chenyang, research on wireless self-organized network attack methods and countermeasures, master thesis, Beijing university of posts and telecommunications, December 12, 2012.

Extract Feature Frequency of Rocket Vibration Signal Based on EEMD and Correlate Coefficient Test

Zhang Jianhai, Han Yingchun , Liu Jun

Abstract: In order to eliminate the transient high frequency oscillation signal which contained in the rocket vibration signal, and extract feature frequency of the vibration signal, the paper proposed a novel method which bases on ensemble empirical mode decomposition and correlation coefficient test. First, the vibration signal is decomposed by EEMD, and a series of intrinsic mode functions (IMF) are obtained. Then, the correlation coefficient between the IMF components and the original signal is calculated, and the effective IMF components are identified by the correlation coefficient significance test. In the end, Hilbert transform (Transform Hilbert, HT) method is used to analyze the effective IMF components, and the purpose of extracting the feature frequency is achieved. The extract results of simulation signal and rocket vibration signal show that this method is feasible and effective.

I. INTRODUCTION

The process of fault diagnosis for rocket can be classified into three steps: signal acquisition, feature extraction and diagnosis decision-making. One of the most important and difficult problems is feature extraction. In a sense, feature extraction is a bottleneck problem in the research of mechanical fault diagnosis. It is used to analyze and extract the fault feature information, which is a service for fault pattern recognition and diagnosis. Many signal analysis methods have been successfully applied in signal feature extraction, such as spectral analysis, wavelet analysis, correlation analysis, modern time frequency analysis, maximum principal component analysis, time series analysis and so on[1].

Sometimes, when the rocket is stationary, affected by internal or external factors, the rocket will appear the transient high frequency and the low frequency interference signal. The real vibration frequency can be affected by the interference signal, and the analysis results will not reflect the frequency characteristic of the vibration signal. Therefore, in order to extract the characteristic frequency accurately, effective removal of interfering signal is critical.

EEMD methods apply the noise analysis to the empirical mode decomposition, In order to promote the anti-decomposition and the mode mixing phenomenon is suppressed effectively[2]. The IMFs obtained by EEMD can reveal the physical meaning of the original signal, and the physical essence of each IMF is clearer[3]. Firstly, the correlation coefficient of the IMF components and the original signal is calculated after the EEMD decomposition of the vibration signal; Then, the method of correlation coefficient test is used to eliminate the high frequency transient interference; In the end, the effective IMFs are obtained, and the feature frequency of the signal is extracted.

II. Basic principle

A. Ensemble empirical mode decomposition principle

The EEMD method was proposed in 2009 by Huang et al. The signal is introduced Gauss white noise in the process of being decomposed. The additive white noise is distributed in the whole time

frequency space, which is divided into different scales. Although the white noise is added to the signal components, it is considered that the additive white noise is random, so that the mode mixing phenomenon can be eliminated by enough times.

The EEMD analysis method was used to process time series signal $x(t)$, and the steps are as follow[4]:

1) A new data sequence $x_i(t)$ is obtained by adding Gauss white noise $n_i(t)$ to target signal, whose length is certain and amplitude coefficient is K:

$$x_i(t) = x(t) + k * n_i(t) \tag{1}$$

In the above formula: $x(t)$ is a signal to be decomposed, $n_i(t)$ is white noise signal, and $x_i(t)$ is a signal to be decomposed with white noise.

The strength of $n_i(t)$ is determined by the ratio of the standard deviation Rstd (Ratio of Standard Deviation):

$$\mathrm{Rstd} = \mathrm{STD}_1 / \mathrm{STD}_2 \tag{2}$$

In the above formula: STD_1 and STD_2 are white noise standard deviation and target signal standard deviation.

2) All the local extreme value points of the signal $x_i(t)$ were found. Up-and-down envelope lines are formed by using cubic spline interpolation function. Take note the average value of the up-and-down envelope line as $m(t)$;

3) After removing the mean sequence $m(t)$ from the $x_i(t)$, the signal to be analyzed was got. According to the criterion of the intrinsic mode function, if the signal does not satisfy the criterion, step 2) was repeated until it does.

4) The residual signal was calculated

$$R_{i1}(t) = x_i(t) - \mathrm{c}_{i1}(t) \tag{3}$$

$R_{i1}(t)$ and $\mathrm{IMF}_{i1}(t)$ are the remaining signals and the intrinsic mode functions

5) The remaining signal is treated as the signal, repeat the step 2 and 3, IMF $\mathrm{c}_{i1}(t), \mathrm{c}_{i2}(t), \cdots \mathrm{c}_{ik}(t)$ are got in turn.

6) Another noise signal was add into the original signal, after the step 2 ~ 5, a series of IMF were obtained;

7) According to the principle that statistical mean value of non-correlation random sequence is zero, the IMFs from above steps were averaged. The influence that Gauss white noise is added to the

real IMF repeatedly was eliminated. The IMFs and the remainder after EEMD decomposition are as follow.

$$\mathrm{c}_j(t) = \frac{1}{N}\sum_{i=1}^{N}\mathrm{c}_{ij}(t) \tag{4}$$

$$r(t) = \frac{1}{N}\sum_{i=1}^{N} r_i(t) \tag{5}$$

In the above formula: N is the number of IMFs, and $\mathrm{c}_{ij}(t)$ is the first jth IMF of mixing first ith noise.

Finally, we get K IMF component and a remainder $r(t)$:

$$x(t) = \sum_{j=1}^{K}\mathrm{c}_j(t) + r(t) \tag{6}$$

B. Correlation coefficient significance principle

The correlate coefficient is a quota to measure correlative degree between investigated signals. Usually, the greater of the correlate coefficient shows, the higher correlative degree of the signal is; and vice versa. However, if the length of investigated signals increases, the error also increases and the correlate coefficient is often low. So, it is obviously unreasonable to use correlate coefficient to indicate the close degree of investigated data. Therefore, the significant test of correlate coefficient is used to judge how close the degree of investigated signals is [5].

According to statistics principles, t distribution is adopted to test the significant of correlate coefficient. The test procedure is as follows:

1)The simulation signal is decomposed into a series of IMF components $c_i(\mathrm{t})(i = 1, 2\cdots n)$. The correlate coefficient ρ_i for each component with simulation signal is defined as

$$\rho_i = \frac{\sum_{j=1}^{N}(c_{ij} - \overline{c_i})(x_j(t)' - \overline{x(t)'})}{\sqrt{\sum_{j=1}^{N}(c_{ij} - \overline{c_i})^2}\sqrt{\sum_{j=1}^{N}(x_j(t)' - \overline{x(t)'})^2}} \tag{7}$$

The t test value of the correlate coefficient ρ_i is defined as

$$t_i = \frac{\rho_i\sqrt{N-2}}{\sqrt{1-\rho^2{}_i}} \tag{8}$$

2)According to the given significance level α and freedom degrees (n-2), the critical values $t_{\alpha/2}$ is gained by checking the t distribution table.

3)Judgment. If $|t_i| \geq t_{\alpha/2}$, ρ_i is considered to be significant in statistics; otherwise, it is considered to be not significant, and the low frequency redundant components will be eliminated.

C. Feature frequency extraction algorithm

After decomposing the original signal, each component is calculated into Hilbert spectrum. The process is as follows:

1) For each component $c_i(t)$, its Hilbert transform is defined as

$$\overline{c_i}(t) = \frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{c_i(\tau)}{t-\tau}\mathrm{d}\tau \tag{9}$$

2) Analytic signal $z_i(t)$ is constructed as

$$z_i(t) = c_i(t) + j\overline{c}_i(t) = a_i(t)e^{j\phi_i(t)} \tag{10}$$

where

$$a_i(t) = \sqrt{c^2{}_i(t) + \overline{c}_i^2(t)} \quad \text{and} \quad \phi_i(t) = \arctan\frac{\overline{c}_i(t)}{c_i(t)} \ . \tag{11}$$

$a_i(t)$ is the instantaneous amplitude of $c_i(t)$, which can reflect the energy of the $c_i(t)$ varing with time. $\phi_i(t)$ is the instantaneous phase of $c_i(t)$.

The instantaneous frequency of $c_i(t)$ is showed as

$$f_i(t) = \frac{1}{2\pi}\omega_i(t) = \frac{1}{2\pi}\frac{\mathrm{d}\phi_i(t)}{dt} \tag{12}$$

3) The vibration signal is designated as the Hilbert spectrum $H(\omega,t)$.

$$H(\omega,t) = \mathrm{Re}\sum_{i=1}^{n} a_i(t)e^{j\int\omega_i(t)\mathrm{d}t} \tag{13}$$

Re means getting the real part of the above formula. In general, $r_n(t)$ is a monotonic function or constant, so it can be omitted as trend. The amplitude of the signal is expressed as a function of t and ω. the above formula is Hilbert three-dimensional spectrum of original signal

The distribution of frequency can be accurately analyzed in the three-dimensional spectrum. If the characteristic signal of the $m\,th$ component is concerned, the feature information of the $m\,th$ component can be extracted according to following formula:

$$H(\omega,t) = a_m(t)e^{j\int\omega_m(t)dt} \tag{14}$$

III. Simulation signal analysis

The simulation signal $x(t)$ consists of sinusoidal signal $x1$ whose frequency is 20Hz and high frequency transient oscillation signal $x2$. The sampling frequency f_s of simulation signal is 2 000, and the sampling points N is 500. The simulation signal are shown in Figure 1.

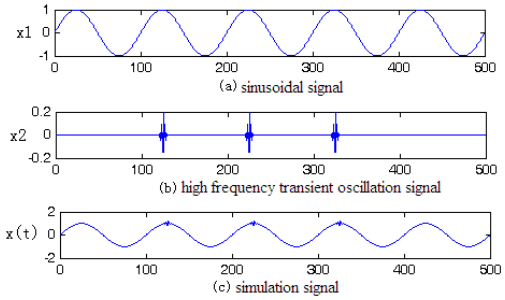


Figure 1. Time domain of simulation signal

$x(t)$ is decomposed by EEMD, while the coefficient k of Gauss white noise is0.25, and the EEMD total number M is 100.All the IMF component is shown in Figure 2.

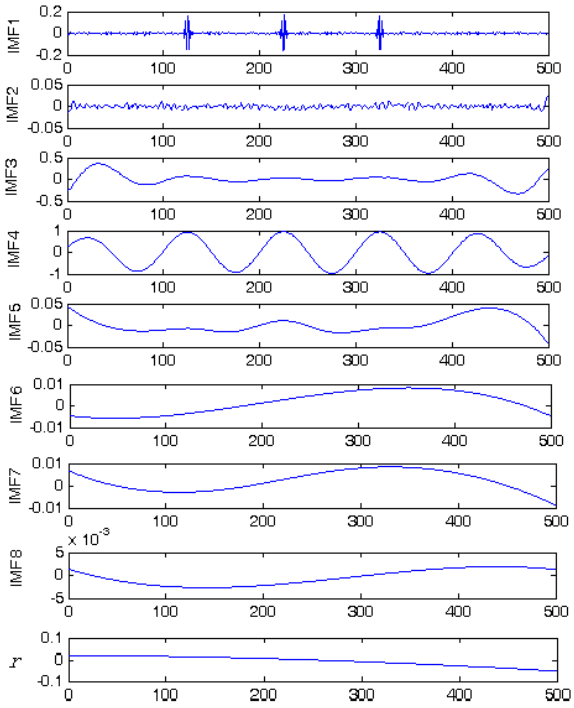


Figure 2 The IMF component of simulation signal

From Figure 2, the first component IMF1 is very similar to the transient high frequency signal $x2$, which indicates that EEMD can distinguish the abnormal events from the original signal. The very small amplitude of the IMF2 is Gauss white noise. IMF4 is close to the signal $x1$, but at the begin and end of the signal, the amplitude of IMF4 is slightly less than that of $x1$. The signal merging IMF3 and IMF4 are more close to the sine wave part in $x(t)$. So, EEMD can separate each part from simulation signal with high quality, and there is no aliasing, which reflects the signal' physical essence.

The correlate coefficients ρ_i and t test values are calculated between each IMF and original signal, and the data is shown in Table 1.

The significant level α =0.05, and the freedom degree n-2=998. The t distribution table shows $t_{\alpha/2}$=1.96. According to the calculate results of table 2, the test values of *IMF3*～*IMF7* satisfy the equation $|t| \geq t_{\alpha/2}$ which indicates that the correlate coefficients are significant. Contrarily, the other correlate coefficients are not, which will be eliminated. Therefore, the effective component is *IMF3*～*IMF7*

Table 1.Correlate coefficient and t - distribution test value

component	ρ	$\|t\|$	component	ρ	$\|t\|$
IMF1	0.036	0.804	IMF6	0.152	3.432
IMF2	0.044	0.983	IMF7	0.101	2.265
IMF3	0.596	16.564	IMF8	0.082	1.836
IMF4	0.914	42.199	r	0.031	0.692
IMF5	0.380	9.168			

The effective IMF components are synthesized and compared with the original single, and the result is shown in Figure 3.

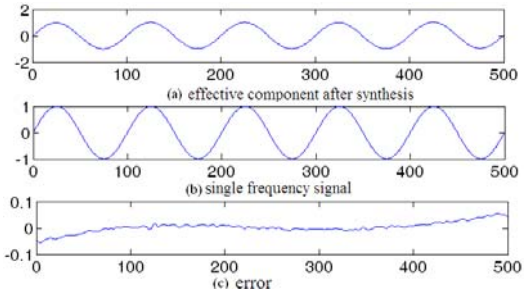


Figure 3. Synthetic signal, single frequency signal and error

The correlation coefficient of synthesized signal and original single is 0.991. The test value is 165.2, which is far greater than 1.96. Thus, the effective IMF components can be used as real signal.

The effective IMF components are calculated by Hilbert transform, and the time spectrum is shown in Figure 4.

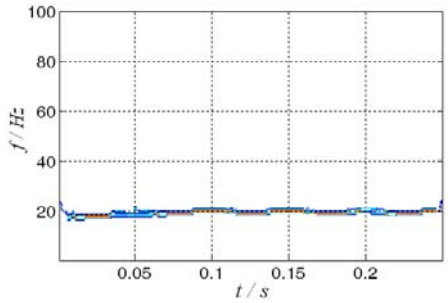


Figure 4. time spectrum of effective IMF components

Hilbert three-dimensional spectrum is obtained, as shown in Figure 5.

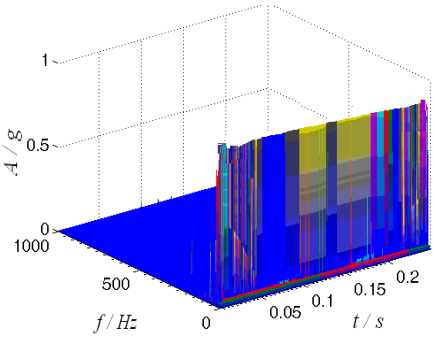


Figure 5. three-dimensional spectrum of effective IMF components

IV. Feature frequency extraction of rocket vibration signal

The vibration signal of a certain type of rocket in the stationary stage is analyzed, in order to extract feature frequency. The sampling frequency f_s is 640, and the sampling points N is 1024. The domain is shown in Figure 6.

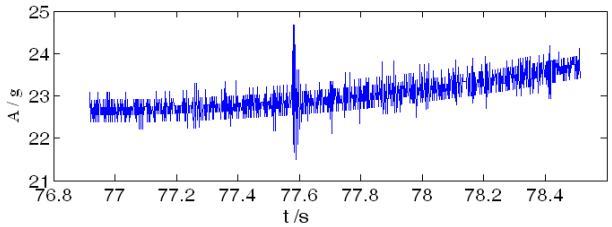


Figure 6 Vibration signal of a certain type of rocket

From the domain, we known that the stationary signal contains the transient high frequency vibration signal. After EEMD decomposition of the signal, the IMF component is obtained , and shown in Figure 7.

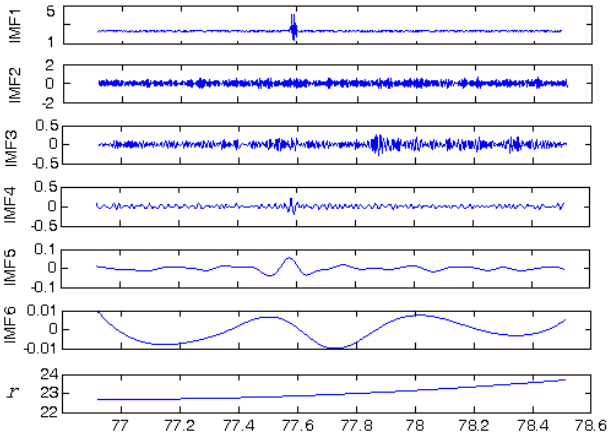


Figure 7. IMFs of the measured signal

From Figure 7, the first component is transient high frequency signal, and r is the trend. The correlate coefficients ρ_i and t test values are calculated between each IMF and original signal, and the data is shown in table 2.

The significant level α =0.05, and the freedom degree n-2=998. The t distribution table shows $t_{\alpha/2}$=1.96. According to the calculate results of table 2, the test values of *IMF2*□ *IMF5* satisfy the equation $|t| \geq t_{\alpha/2}$ which indicates that the correlate coefficients are significant. Contrarily, the other correlate coefficients are not, which will be eliminated.

Table 2.Correlate coefficient and t - distribution test value

component	ρ	$\|t\|$	component	ρ	$\|t\|$
IMF1	0.059	1.889447	IMF4	0.252	8.324785
IMF2	0.672	29.00946	IMF5	0.152	4.916374
IMF3	0.457	16.42525	IMF6	0.035	1.119592

After eliminating the transient high frequency and low frequency interference, the effective components are remixed. Hilbert spectrum is calculated, and is shown in Figure 8. The frequency signal is consistent with the actual vibration frequency of rocket.

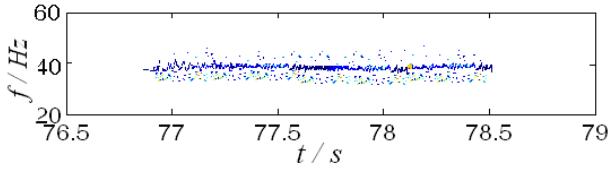


Figure 8. Hilbert spectrum of effective components

V. Conclusion

Aiming at the problem of transient high frequency and low frequency interference in stationary signal, a method based on the combination of EEMD and correlation coefficient

test is proposed. The novel method is used to analyze and process simulation signal and rocket vibration signal which contains transient high frequency and low frequency interference signal. The effective IMF components are obtained, the feature frequency is extracted, and the correctness of this method is demonstrated.

References

[1] WU Zhaotong, Yang Shixi. A new method for fault feature extraction of rotating machinery and pattern classification[M]. Science Press, 2012：111-13

[2] Wu Z H，Huang N E. Ensemble Empirical Mode Decomposition: a Noise -assisted Data Analysis Method[J]. Advances in Adaptive Data Analysis，2009，1(1) : 1- 41

[3] CHEN Juan, LI Xiang. The method Ensemble Empirical Mode Decomposition in fatigue signal denoising[J]. Journal of Vibration, Measurement & Diagnosis, 2011, 31(1) : 15-19.

[4] CHEN Renxiang, TANG Baoping, LV Zhongliang. Rotor vibration signal denoising method based on correlation coefficient and EEMD[J]. Journal of Vibration, Measurement & Diagnosis, 2012,32(4):542-546.

[5] Zhang Jianhai, Han Yingchun, Li Lianzhou . An Improved EMD Time-Frequency Analysis Method for Rocket Vibration Signal[C]. IEEE Chinese Guidance, Navigation and Control Conference, 2014: 1842-1846

Development and Testing of Nitrous Oxide/Propane Rocket Engine

Wang Dong[1], Liang Guozhu[2]

1, Xi'an Aerospace Propulsion Institute, Xi'an, China.
2, Beihang University, Beijing, China

Abstract: A new rocket test stand facility, built at Beihang University, equipped with palletized propellant feed systems, thrust stand, and data acquisition systems, was constructed to test a nitrous oxide/propane(N_2O/C_3H_8) rocket engine. A unique rocket ignition concept using N_2O/C_3H_8 torch igniter was explored and influence of various mass mixing ratio were summarized. The hot test study for N_2O/C_3H_8 principle engine is described. The N_2O/C_3H_8 engine can be ignited quickly by igniter and work stably when the mixing ratio between 2.63 and 6.81. Meanwhile, the combustor equilibrium pressure varies in the scales of 0.257~0.514MPa, and the maximum average thrust is 140N. The experimental specific impulse improves obviously with the rise of mixing ratio, which increases from 1300m/s (mixing ratio is 3.0) to 1895.7m/s (mixing ratio is 6.56). The hot test results show that the intermediate frequency combustion instability (900~1000Hz) occurs at starting phase and low frequency combustion instability (65~80Hz) occurs at stable working phase.
Keywords: nitrous oxide; propane; rocket engine

1. Introduction

The current rocket liquid propellant generally has some disadvantages, such as high toxicity, heavy corrosive and cryogenic. With the development of manned spaceflight technology, a cheap, non-toxic, non-pollution, high energy and maintainable propellant, has become the main direction for the development of liquid propellant. Using nitrous oxide(N_2O) and propane(C_3H_8) as the rocket propellant combinations can effectively satisfy the requirements for the new generation propellant. So far, the overseas studies on N_2O/C_3H_8 engine are still under test[1-3], but a comprehensive understanding about its performance is still not found. Therefore, this paper tries to analyze and study the N_2O/C_3H_8 engine.

2. Advantages of N2O/C3H8 Engine

N_2O has added benefit as a space propellant in that it stores as a liquid and injects as a gas. This is important for attitude control(AC), since liquid injection rockets can't provide the shorter pulse times required for an ACS mission. Liquid storage gives tremendous weight benefit since a liquid tank can hold multiple times its own weight in propellant, whereas the same is not true for gas storage systems[2].

Generally, most propellants commonly used today have relatively low vapor pressure (lower than the rocket chamber pressure) and consequently require a separate expulsion system. In contrast to these systems the N_2O/C_3H_8 propellants are self-pressurizing due to their relatively high vapor pressures (higher than the rocket chamber pressure). Consequently, they don't require separate expulsion systems and the entire tank volume can be used to store propellant. The vapor pressure of N_2O is approximately 5.17MPa and that of C_3H_8 is 0.76MPa at ambient temperature. The vapor

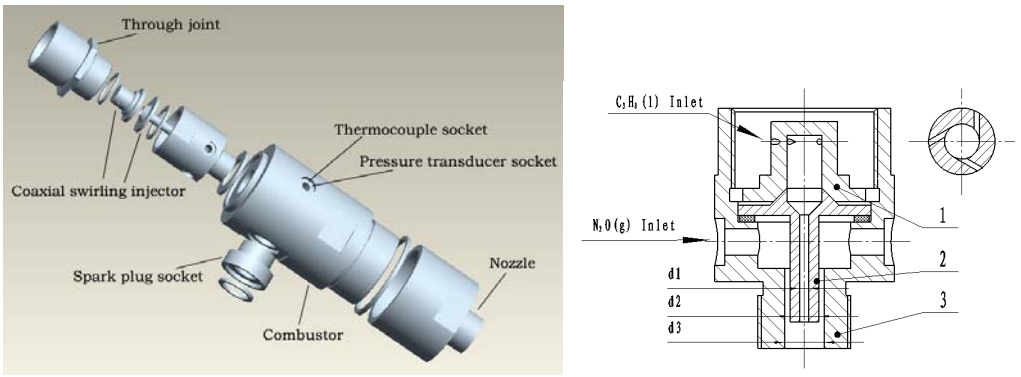
pressure of C_3H_8 is slightly low for it to be used as a true self-pressurizing propellant. For a N_2O/C_3H_8 rocket system, high-pressure N_2O vapor would be used as a pressuring gas for the propane.

The N_2O/C_3H_8 rocket offers a non-toxic, environmentally benign propellant combination that is storable in space over long periods of time and offers comparable specific impulse(2800m/s~3200m/s) to current systems. The non-toxic nature of the N_2O/C_3H_8 propellants will serve to reduce operating costs due to the handling issues associated with the hypergolic propellants currently in use for space applications. The N_2O/C_3H_8 propellants are benign and not highly reactive. They remain so until the N_2O is catalytically decomposed and combined with the fuel. Exhaust products consist mainly of nitrogen, water, and carbon dioxide. Another beneficial feature of the N_2O/C_3H_8 rocket propellants are that they are storable over long periods of time without degradation. For example, inadvertent decomposition is one of the main technical obstacles for hydrogen peroxide use in space propulsion systems.

3. Igniter Design

The ignition is the key of the rocket engine technology, which is very difficult that determined by physicochemical properties of N_2O/C_3H_8. There are about two ignition methods was used by some research institutions, such as Alabama University[2]. One is N_2O catalytic ignition, which structure is simple, but the catalyst is expensive, lost activity easily and must preheating before ignition. Another is precombustion ignition by using O_2/C_3H_8, which supply system is complicated due to the third propellant O_2 is introduced.

In order to ignite the N_2O/C_3H_8 rocket engine repeatedly, immediately and directly, the N_2O/C_3H_8 torch igniter was designed. Figure 1 shows the schematic of the igniter. A coaxial swirling injector was employed due to its excellent atomization and mixing behaviors. The liquid C_3H_8 flows through the inner injector and is swirled when the gas N_2O flows through the outer annulus injector, then N_2O/C_3H_8.mix rapid near the injector face. A spark plug socket located outside the combustion chamber. Table 1 shows the main parameters of igniter.



(a) Schematic (b) Coaxial swirling injector schematic

1. Swirling chamber; 2. Inner injector; 3. Outer injector

Fig. 1 N_2O/C_3H_8 Torch igniter schematic

Table 1 The parameters of N_2O/C_3H_8 igniter

Parameter	Value
L(mm)	109
dc(mm)	20
dt(mm)	3.3/2.0
d_1(mm)	0.53
d_2(mm)	5.9
d_3(mm)	6.0
spark plug	10Hz, 1J

4. N_2O/C_3H_8 Rocket Engine Design

The engine utilizes gas N_2O as the oxidizer and liquid C_3H_8 as the fuel. The chamber pressure is 2MPa for this 720N(ground thrust) engine, and a fuel-rich propellant combination is used to minimize combustion instabilities. According to common practice for space-based thrusters, a pressure drop between 20-30% of chamber pressure is taken across the injectors, requiring an injection pressure of approximately 2.5MPa. Since the vapor pressure of C_3H_8 at 20℃ is only 0.76MPa，the C_3H_8 needs to be pressurized for this application. This is accomplished using nitrogen pressurant for the ground-based testing, and can be accomplished using the N_2O as pressurant in space.

To measure rocket performance by specific impulse (I_{sp}) and characteristic velocity (c*), accurate measurements of thrust and propellant (fuel and oxidizer) flow rate are essential. Measuring the flow rate of N_2O (g) requires using a sonic nozzle and C_3H_8(l) requires using a venturi.

The rocket design involves: design of the injector for N_2O (g) and C_3H_8(l), design of the rocket combustor with an appropriate L* (characteristic combustor length) and design of a rocket nozzle for sea level static conditions and space conditions. Table 2 summarizes the key design parameters for N_2O/C_3H_8 rocket engine.

Table 2 The parameters of N_2O/C_3H_8 rocket engine

Parameter	Value
Ac/At	10
dt(mm)	18.12
L* (m)	1.5
Dc(mm)	57
ground thrust(N)	720
vacuum thrust(N)	1000

The engine injector design is probably the single most important component with regard to affecting combustor stability and performance. The injector design for N_2O/C_3H_8 rocket engine was based on experience and knowledge gained from the igniter injector design and numerical simulation. As shown in Figure 2, the engine injector include six coaxial swirling injectors and eighteen N_2O(g) film cooling orifices around the jet surface.

The current injectors are sized for two criteria: the liquid propane injector is sized for an injector pressure drop of 30%P_C, and the gaseous nitrous-oxide injector is sized for a pressure drop of 20%P_C. Approximately 25% of the N2O will flow through the orifices, providing for film cooling.

Fig. 2 N_2O/C_3H_8 Rocket engine injector

Figure 3 shows the designed N_2O/C_3H_8 rocket engine with vacuum nozzle(Ae/At=220). For sea level conditions, the nozzle is truncated at an area ratio, Ae/At of 3.42 for the supersonic region, to achieve perfect expansion to the ambient conditions.

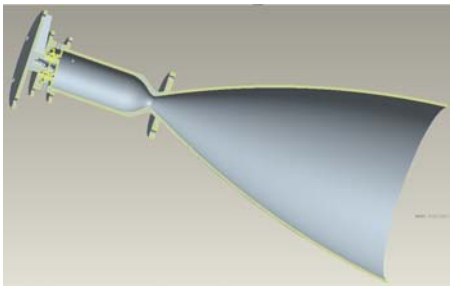


Fig. 3 N_2O/C_3H_8 Rocket engine(vacuum nozzle)

5. Rocket Engine Test Stand

In order to investigate the performance of N_2O/C_3H_8 engine, the hot test system of rocket engine is designed and installed, including ignition experiments system, engine performance experiment system and dynamic thrust measuring system. This experimental platform can satisfy the (dynamic) performance testing requirements for the liquid rocket engine, while regulating the supply pressure and flow rate of propellant flexibly for the N_2O/C_3H_8 igniter and engine, conducting data collection effectively for each testing point in propellant supply system, engine combustor temperature & pressure and thrust.

The basic idea behind the propellant system layout is to provide the user with a safe, modular and self-contained process for loading and self-pressurizing or pressurizing N_2O and C_3H_8 for use in the spark ignition circuit and rocket engine propellant feed-systems. The N_2O and C_3H_8 are first loaded into their respective run tanks, and a nitrogen pressurization tank is used to further pressurize the C_3H_8 run tanks in order to ensure that the C_3H_8 are in the liquid state at least through the venturi, thus assuring accurate mass flow rate measurements. A nitrogen purge circuit is also used to purge the lines before and after the rocket firing sequence is performed. The fuel and oxidizer systems are separated and are both separated from the rocket compartment. The rocket is mounted on a thrust stand, and is connected to the propellant system by stainless lines. Schematics and photographs of the propellant system, and thrust stand are shown in Figure 4~Figure5.

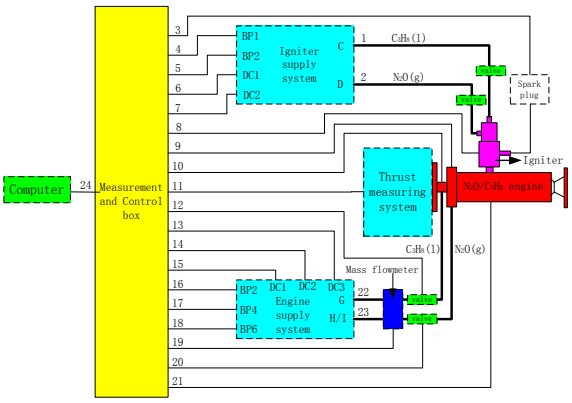


Fig. 4 Propellant supply and measuring control system

Fig. 5 Rocket engine test stand

6. Experimental Results

6.1 N_2O/C_3H_8 igniter experimental results

The typical igniter experimental result are shown in Figure 6. The C_3H_8 supply valve opened firstly, when the supply pressure P_{C3H8} trends to be stable, the N_2O supply valve opened at 45.6s, then spark plug worked at 50.5s and continued for 0.5s. It can be seen the combustor pressure Pc is 1.0MPa. The flame extinction immediately when the C_3H_8 supply valve closed at 56s, the N_2O supply valve closed about 5s later in order to cool the igniter.

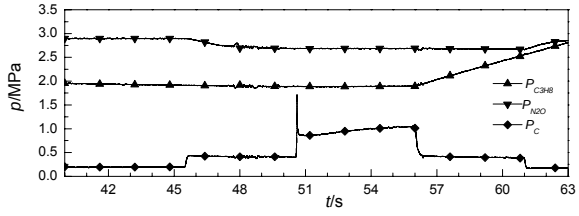
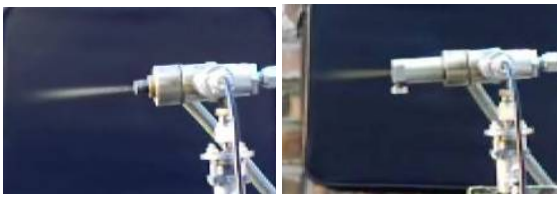


Fig. 6 Pressure sampling curve(dt=2.0mm)



a) Ignition torch(dt=3.3mm) b) Ignition torch(dt=2.0mm)

Fig.7 Ignition torch

The factors affecting the ignition characteristics are various. Due to the high decomposition activation energy of N_2O, the ignition energy input must be sufficient. If the N_2O content too high and exceed the energy input, then ignition failure. If the N_2O content too low, then the decomposed

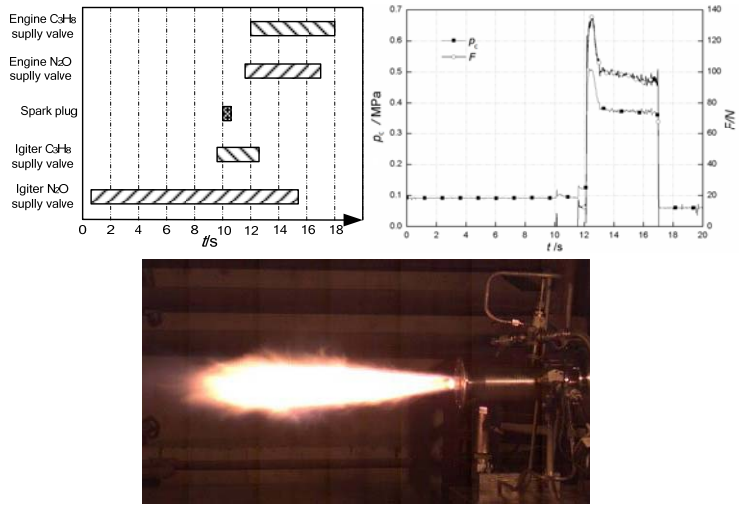
oxygen from N_2O too thin to sustain combust with C_3H_8. In additionly, the distribution of flow field of N_2O/C_3H_8 mixture, atomization and mixture effect, energy input position, and so on, will be affect the ignition result directly. Table 3 summarizes the experiment result of igniter.

Table 3 The experiment result of igniter

q_{N2O} (g/s)	4.07~6.18
q_{C3H8} (g/s)	1.24~1.41
Experimental Mr	2.99~4.54
Successful Mr	2.99~4.07
Pc (MPa)	0.50~0.65(dt=3.3mm) 1.0~1.3(dt=2.0mm)

6.2 N_2O/C_3H_8 engine experimental results

After the experimental system built and tested normally, the $N_2O(g)/C_3H_8(l)$ engine performance investigation was conducted. Figure 8 shows the typical hot test result. Table 4 presents a summary of the average performance data(q_{N2O}, q_{C3H8}, Mr, Pc, F, c^*, I_{sp}). Date shows that the N_2O/C_3H_8 engine can be ignited quickly by igniter and work stably when the mixing ratio between 2.63 and 6.81. Meanwhile, the combustor equilibrium pressure varies in the scales of 0.257~0.514MPa, and the maximum average thrust is 140N. The experimental specific impulse improves obviously with the rise of mixing ratio, which increases from 1300m/s (Mr is 3.0) to 1895.7m/s (Mr is 6.56).



(a) Thrust and chamber pressure (b) Test photograph

Fig. 8 N_2O/C_3H_8 rocket engine tests

Table 4 Summary of performance data for N_2O/C_3H_8 rocket tests

Avg.q_{N2O}, (g/s)	Avg.q_{C3H8}, (g/s)	Avg.Mr	Avg.Pc, (MPa)	Avg.F, (N)	Avg.c^*, (m/s)	Avg.I_{sp}, (m/s)
46.7	14.7	3.14	0.328	80.7	1373.0	1309.5
56.3	8.3	6.81	0.417	102.7	1665.0	1589.3
54.6	13.9	3.92	0.393	94.0	1478.8	1371.6
54.3	15.5	3.51	0.377	90.7	1392.6	1298.9
54.3	15.0	3.62	0.257	91.3	955.5	1317.3
69.4	16.9	4.11	0.354	98.0	1058.4	1136.0
67.9	19.4	3.51	0.370	96.0	1094.7	1100.9
67.3	20.3	3.32	0.366	93.6	1077.0	1068.5
58.4	21.0	2.78	0.360	89.6	1169.1	1128.5

Avg.q_{N2O}, (g/s)	Avg.q_{C3H8}, (g/s)	Avg.Mr	Avg.Pc, (MPa)	Avg.F, (N)	Avg. c^*, (m/s)	Avg.I_{sp}, (m/s)
68.4	21.0	3.26	0.448	117.2	1292.6	1311.0
66.8	19.6	3.41	0.444	110.8	1325.0	1281.8
65.4	22.0	2.97	0.424	109.6	1251.5	1254.0
58.9	19.4	3.03	0.424	108.4	1396.5	1384.1
58.9	20.6	2.85	0.403	103.3	1307.3	1299.6
58.2	21.9	2.66	0.383	101.5	1233.8	1268.2
57.9	22.0	2.63	0.382	98.8	1233.8	1237.4
57.9	20.8	2.78	0.276	65.4	905.3	830.6
57.9	20.8	2.78	0.286	66.0	938.1	838.5
59.5	22.0	2.70	0.309	72.3	977.6	887.0
65.1	22.2	2.94	0.337	84.2	995.0	964.9
74.5	21.8	3.41	0.440	105.6	1177.6	1096.2
64.1	9.8	6.56	0.514	140.0	1794.9	1895.7
61.0	14.3	4.27	0.376	111.4	1289.3	1480.5

Limited by the N_2O supply flow resistance in the experimental process, the maximum q_{N2O} is 74.5g/s, cause the thrust tests result(140N) less than the design value(720N). Meanwhile, the intermediate frequency combustion instability (900~1000Hz) occurs at starting phase and low frequency combustion instability (65~80Hz) occurs at stable working phase due to the low mixture ratio and low combustion pressure.

7. Conclusion

A new rocket test stand facility, equipped with palletized propellant feed systems, thrust stand, and data acquisition systems, was constructed to test a nitrous oxide/propane(N_2O/C_3H_8) rocket engine. The engine was tested over a range of mixture ratios(2.63~6.81). A unique rocket ignition concept using N_2O/C_3H_8 torch igniter was explored and ignited the engine successfully. Future work will focus on increasing the N_2O supply mass flow, improving the test thrust and restrain combustion instability.

References

[1] Yves Maisonneuve. Green propellant perspectives for future missions. AIAA-2008-5028, 2008.

[2] Nicholas Tiliakos, Jason S Tyll. Development and testing of a nitrous oxide/propane rocket. AIAA-2001-3258, 2001.

[3] Charles F Kopicz, Roger Herdy. Nitrous oxide/propane rocket test resilts. AIAA-2000-3222, 2000.

[4] Zakirov V A. Investigation into nitrous oxide propulsion option for small satellite applications[D]. Univ. of Surrey, 2001

[5] Zakirov V A. Li Luming. N2O research in Tsinghua[C]. 2th International Conference on Green Propellant for Space Propulsion, Sardinia, Italy, 2004.

[6] Dominique Valentian, Martin Sippel, et al. Green Propellants Options for Launchers, Manned Capsules and Interplanetary Missions.

[7] Dominque Valentian, Alain Souchier, et al. Green Propellant Implementation for Space Missions and Upper Stage Propulsion. AIAA 2007-5483.

[8] D Haeseler, V Bombelli, P Vuillermoz, et al. Green Propellant Propulsion Concepts for Space Transportation and Technology Development Needs. 2th International Conference on Green

Propellant for Space Propulsion, Italy. June 7-8, 2004.

[9] Wang Dong, Guo Hong-jie, et al. Preliminary experimental investigation on nitrous oxide/propane igniter[J]. Journal of Beijing University of Aeronautics and Astronautic. 2010,36（10）.

[10] Wang Dong, Guo Hong-jie, et al. Experimental investigation on nitrous oxide/propane ignition characteristics[J]. Journal of Propulsion Technology. 2010,31（6）.

Burnout Point Sensitivity Analysis for General Parameters of Nearspace Launch Vehicle

Zhang Yechen, Sun Shijie, Min Changwan, Chen Feng, Jiang Huan

Science and Technology on Space Physics Laboratory, Beijing, 100076
China Academy of Launch Vehicle Technology , Beijing, 100076
yexuan1008@qq.com

Abstract: The traditional launch vehicle, designed for space missions, exist high costs and risks when applied for nearspace missions. As depth development of the nearspace recently, it is imminent that designing launch vehicle for nearspace mission or altering traditional launch vehicle to adapt nearspace mission. In this paper, the burnout point sensitivity for general parameters of nearspace launch vehicle is analyzed. The mathematical model of trajectory/general parameters is established for nearspace launch vehicle, including the mass model, dynamic model, aerodynamic model, flight procedure controlling model. A complex variable sensitivity analysis method is adopted. The regular pattern sensitivity of burnout point parameters (velocity, height, flight path angle) about thrust, fuel, flight procedures of each stage is obtained. Results and method in this paper have reference value on nearspace launch vehicle optimal design.
Keywords: Nearspace; Launch Vehicle; Sensitivity Analysis; CVD Method

1. Introduction

Different from traditional Launch Vehicle, nearspace launch vehicle has a lower orbit altitude, the ascent phase is all in the atmosphere. That's requires the rocket rapidly turning in the atmosphere, in order to achieve low height and path angle. The subjects of the general parameters of vehicle, aerodynamic, trajectory, and power are interacting with each other closely. As depth development of the nearspace recently, it is imminent that designing launch vehicle for nearspace mission or altering traditional launch vehicle to adapt nearspace mission. Sensitivity analysis, used for analyze the sensitivity of the performance of the system with the changes of parameters or design variables, is a very important part in the optimization design. It can be used to determine the influence of the system design variables or parameters for the objective function or the constraint function, ascertain the coupling strength between the various subsystems, and make guidance for design, optimization process and decisions

In this paper, Complex Variable Differentiation (CVD) Method is applied to analysis the burnout Point sensitivity. The thought of using complex variables to calculate the sensitivity was first proposed by Moler [1] and Lyness [2] in 1967.Using complex variable method to calculate the sensitivity derivative has the advantages of high precision, good robustness and short development time. It has been widely used in the field of aircraft design. The sensitivity of burnout point velocity, altitude and path angle for the launch vehicle general parameters (thrust, fuel, flight procedures of each stage) is analyzed in this paper, and the variation of sensitivity has a certain reference value to the optimization design of the near space launch vehicle.

2. launch vehicle model

The launch vehicle is a three-stage solid rocket. The mass of the rocket consists of the mass engines of each stage, the mass of interstage sections and the mass of the payload.

In order to simplify the model, we ignore terms related with attitude angular velocity and angular acceleration in the rocket dynamic equation; ignore the rotation of the earth and treat the earth as a homogeneous sphere; ignore the control force. Therefore, the 3-DOF rocket dynamic equation under the trajectory coordinate system can be expressed as:

$$\begin{cases} \dfrac{dV}{dt} = \dfrac{P\cos\alpha}{m} - \dfrac{D}{m} - \left(\dfrac{R_0}{R_0+H}\right)^2 g_0 \sin\theta \\ \dfrac{d\theta}{dt} = \dfrac{P\sin\alpha}{mV} + \dfrac{L}{mV} - \left(\dfrac{R_0}{R_0+H}\right)^2 \dfrac{g_0\cos\theta}{V} \\ \dfrac{dH}{dt} = V\sin\theta \\ \dfrac{d\beta}{dt} = \dfrac{V\cos\theta}{H+R_0} \end{cases} \tag{42}$$

Where m represents the mass of the rocket; θ represents local path angle; β represents geocentric angle; L represents lift , D represents drag , S represents the reference area of the vehicle.

$$\begin{aligned} D &= \frac{1}{2} C_D \rho S V^2 \\ L &= \frac{1}{2} C_L \rho S V^2 \end{aligned} \tag{43}$$

Attack angle procedures of the launch vehicle are divided into five phases to design, including the phase of each stages and unpowered glide phase between each stages. The first stage flight phase is also divided into the vertical ascending section, the turning fight section, the transonic section and the supersonic section. The attack angle of flight procedures need to be zero in the vertical ascending section, the transonic section and the unpowered glide phase, considering the load and control capability of the rocket. Rectangular flight angle of attack procedures is used in the supersonic section, second and third stage phase. Maximum negative attack angle of each phase needs to be designed. Considering the influence of dynamic pressure of turning fight section, attack angle of this section should not be too large [3], trigonometric attack angle is used at this section.

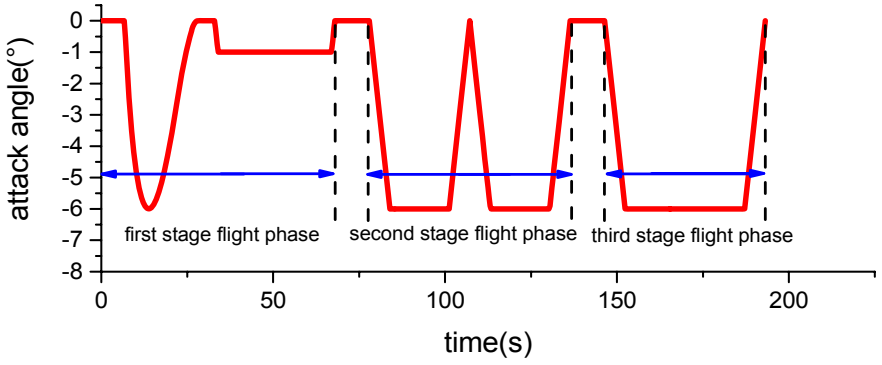


figure 75 attack angle

3. Sensitivity analysis based on complex variable method

3.1 General description of sensitivity

Sensitivity reflects the change of system response when the input parameters are slightly perturbed. Analytical method and finite difference method are usually used in sensitivity analysis. For model $U = U([T],\{X\})$, U is the output matrix, $U = [u_1, u_2, \cdots, u_m]$, $\{X\}$ is the parameter vector, $\{X\} = [x_1, x_2, \cdots, x_n]$, $[T]$ is the coefficient matrix, and the Sensitivity s_i of the i-th parameter x_i is defined as the first derivative of U with Respect to x_i :

$$s_i = \frac{\partial U}{\partial x_i} \tag{44}$$

Parameter sensitivity matrix is generally written as:

$$s = \begin{bmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} & \cdots \frac{\partial u_1}{\partial x_n} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} & \cdots \frac{\partial u_2}{\partial x_n} \\ \vdots & \vdots & \ddots \ \vdots \\ \frac{\partial u_m}{\partial x_1} & \frac{\partial u_m}{\partial x_2} & \cdots \frac{\partial u_m}{\partial x_n} \end{bmatrix} \tag{45}$$

The analytical method is rapid and simple, only need to calculate the derivative about all the analyzed variables, high accuracy can be obtained with the method. But the main difficulty of this method is the derivation. In practical engineering application, models always have too many variables, or the expression is too complex, and even the explicit expression can't be obtained. So the system sensitivity can't be obtained by analytical method generally.

Finite difference method is different, it using the numerical method to get the relevant derivative information instead of using analytical method. The basic principle is: slight disturbance the input parameter x_i , and calculate the new state of model, and then calculate the approximate derivative of output parameters with respect to the input parameters. Forward difference scheme is usually used as the difference format:

$$\frac{\partial U}{\partial x_i} \approx \frac{U(x_i + \Delta x_i) - U(x_i)}{\Delta x_i} \tag{46}$$

Central difference scheme has better accuracy:

$$\frac{\partial U}{\partial x_i} \approx \frac{U(x_i + \Delta x_i) - U(x_i - \Delta x_i)}{2\Delta x_i} \tag{47}$$

Generally, Δx_i is taken as

$$\Delta x_i = \lambda x_i \quad (10^{-5} < \lambda < 10^{-2}) \tag{48}$$

Difference method principle is simple, and it is convenient to implement on computer, but there are some problems:（1）The efficiency and accuracy of the calculation are closely related to the selection of Δx_i ,When Δx_i is not appropriate, the computational efficiency and speed will be reduced.（2）The State variables need to be calculated twice when the difference scheme is used to calculate the sensitivity.

3.2 Complex variable method

Complex variable method can avoid the above inadequate, sensitivity can be obtained with a high accuracy by calculate the output parameters only once. Specific principles and procedures are as follows:

For real function $u(x)$, construct complex variables $x+ih$, x is the real part, and h is a real tiny number (Usually $h=10^{-20}$), Taylor expansion of this complex functions is:

$$u(x+ih)=u(x)+ih\frac{du}{dx}-\frac{h^2}{2}\frac{d^2u}{dx^2}-i\frac{h^3}{3!}\frac{d^3u}{dx^3}+\cdots \tag{49}$$

Since h is a tiny real number, ignore the three times higher-order terms on the right-hand side. Imaginary part of the Left and right sides should be equal, so

$$\frac{du}{dx}=\frac{\mathrm{Im}(u(x+ih))}{h} \tag{50}$$

The derivative $\frac{du}{dx}$ can be obtained without the derivation operation, and derivative information obtained directly by calculating the function value.

4. Sensitivity analysis of the burnout point parameters

Sensitivity analysis of the launch vehicle is based on a certain launch vehicle. The general parameters of the launch vehicle are shown in the following table:

Table 22 General Parameters

	First stage	Second stage	Third stage
Fuel mass(kg)	20000	12000	4000
thrust(kN)	800	600	250

Select height, path angle and velocity of the burnout point as the research objects. Study on the influence of fuel mass, thrust and flight procedures.

The sensitivity calculation program is changed by the trajectory calculation program. There are no major changes to the original program, only alter the general parameters from real variable to complex variable.

4.1 The Influence of burnout parameters for fuel mass of each stage

Curves of burnout point parameters sensitivity for fuel mass as the Figure 2-4 show, the change range of each stage fuel mass is 50% to 100% of original fuel mass. The first stage phase includes two sections: the procedure of the turning section and the supersonic turning section. The fuel will be run out at procedure of the turning section, the supersonic phase will not exist in the flight stage phase, when the fuel is too few, and the derivative of the sensitivity curve is not continuous.

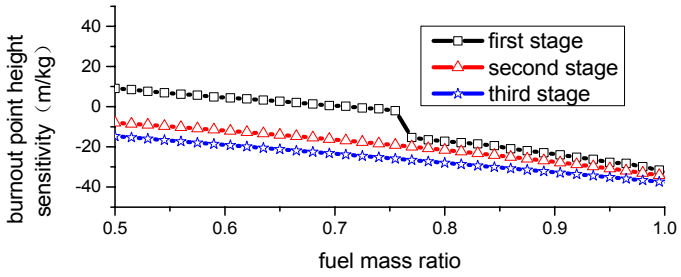


figure 76 burnout point height sensitivity for fuel mass

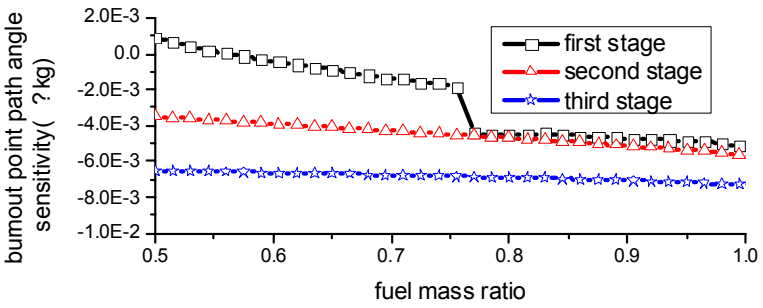


figure 77 burnout point path angle sensitivity for fuel mass

The increasing of fuel in first stage will lead the height and flight path angle increased, when the fuel is too few. With the increase of fuel in first stage, the time of the negative attack angle is also increases, and the increase of fuel will lead the height and flight path angle decrease. The increasing of fuel in second and third stage will lead the height and flight path angle decrease either. The mass of third stage rocket is smaller, so the height and flight path angle are more sensitive.

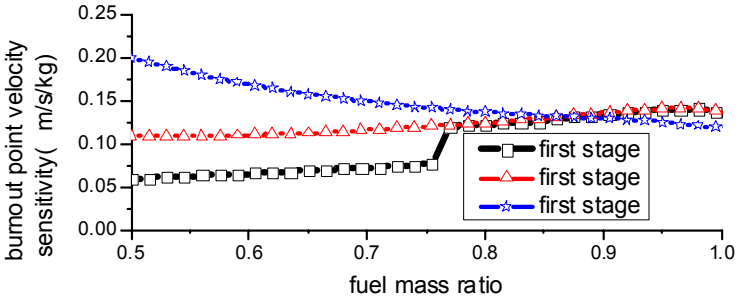


figure 78 burnout point velocity sensitivity for fuel mass

The increase of fuel in each stage will lead the speed increase. In first and second stage the speed will increase in the increasingly obvious, on the contrary, with the fuel increase of the third stage, the speed amplitude of increase decreases gradually.

4.2 The Influence of burnout parameters for thrust of each stage

Figure 5 to 7 give the sensitivity curve of burnout point parameters for the thrust of each stage, the vertical axis is the sensitivity of parameters, and the horizontal axis is the ratio of the current thrust compare to original thrust.

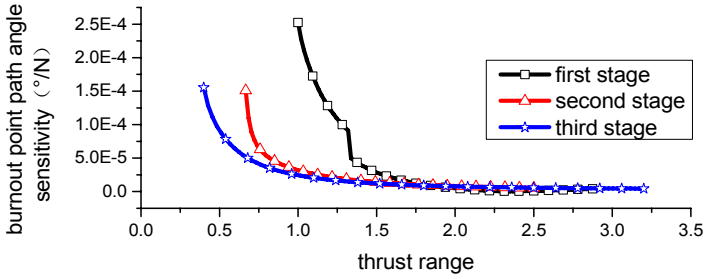


figure 79 burnout point path angle sensitivity for thrust

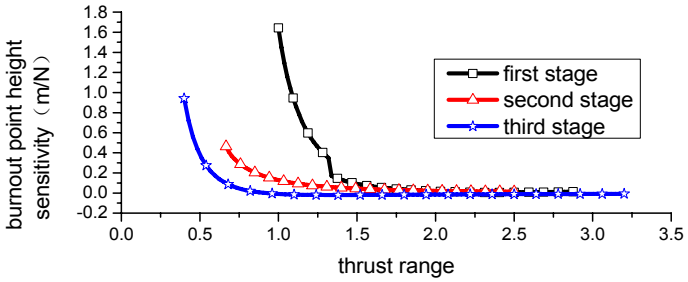


figure 80 burnout point height sensitivity for thrust

Effect levels of thrust on the height and flight path angle decrease gradually as the order: first second third. The increase of thrust can raise the height and the flight path angle, and the sensitivity decreased gradually with the increase of thrust.

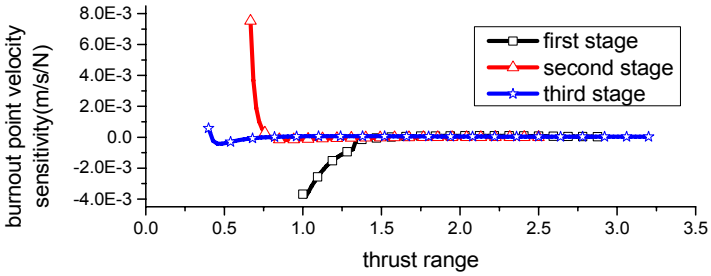


figure 81 burnout point velocity sensitivity for thrust

Thrusts of each stage have a great influence to the burnout point parameters in a certain range. With the increase of thrust, the impact of the thrust on the speed will be reduced rapidly.

4.3 The Influence of burnout parameters for flight procedures

The influences of the four maximum negative angles of attack are analyzed in: the procedures turning fight section, the supersonic turning section in first stage, the second stage, the third stage.

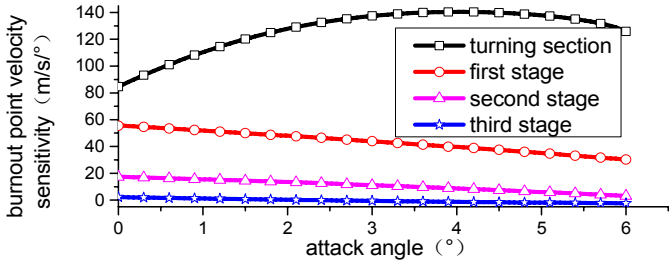


figure 82 burnout point velocity sensitivity for attack angle

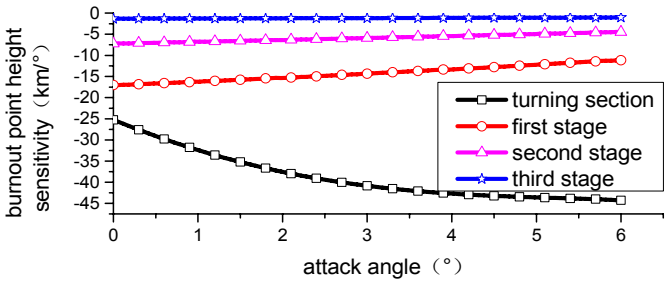


figure 83 burnout point height sensitivity for attack angle

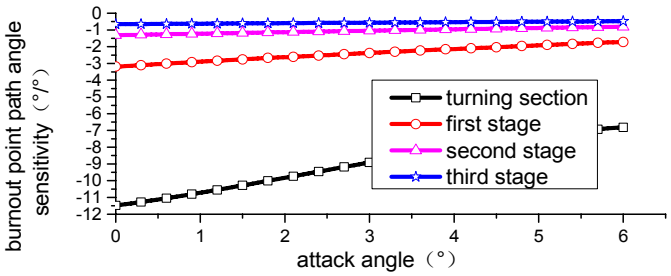


figure 84 burnout point path angle sensitivity for attack angle

In the four phase, the influence of procedures turning section is the largest, and the higher the angle of attack, the greater the influence on the height and speed. With the increase of altitude, the atmosphere becomes thin, and the influence of the angle in other three phases also reduced. In the ascent flight, when the negative angle of attack becomes larger, depressed trajectory, height and flight path angle decreases, negative work produced by gravity is reduced and speed is improved.

5. Conclusion

This paper analyzes the burnout point sensitivity of near space launch vehicle by CVD method. The influence of the parameters of the three sets of the masses of fuel, the thrusts and the flight program is analyzed. The results can make guidance for nearspace launch vehicles' design, optimization and decisions

Reference

[1]Lyness,J.N., and Moler, C.B.Numerical Differentiation of Analytic Functions[J]. SIAM Journal on Numerical Analysis, Vo.l 4, No. 2, 1967.

[2]Lyness, J.N.Numerical Algorithms Based on the Theory of Complex Variables[R]. Proceedings of the ACM 22nd National Conference, Thomas, Washington, DC, 1967.

[3] ZHANG Ning-ning; MIN Chang-wan; LIU Hui; ZHANG Ye-chen. Integrated optimization design for mass distribution / trajectory of near space directly injecting carrier rocket [J]. Flight Dynamics, 2014, 03.

[4] Yan Li; Chen Xiaoqian; Wang Zhenguo. Study of the Complex Variable Method for Sensitivities Calculation in MDO of Flying Vehicles [J]. Tactical Missile Technology, 2006, 03:5-8. DOI:doi:10.3969/j.issn.1009-1300.2006.03.002.

[5] Jia Peiran, Chen Kejun, He Li. Long-range Rocket Ballistics［M］. Changsha:National University of Defence Technology Press，1993:230.

[6] YANG Xi-xiang; JIANG Zhen-yu; ZHANG Wei-hua. Flight program design for small launch vehicle in atmosphere flight stage [J]. Flight Dynamics, 2010, 04:68-72.

Application of Model-based Systems Engineering in Small Satellite Conceptual Design – A SysML Approach

Muhammad Waseem*, Muhammad Usman Sadiq, and Muhammad Frooq

Satellite Research and Development Centre, SUPARCO, Lahore, Pakistan

*Contact: vasim98@gmail.com

Abstract: In this paper we have explored the application of model-based systems engineering (MBSE) in conceptual design of remote sensing satellites. The purpose of this study was to investigate the usefulness of the MBSE approach in order to assist the system engineering process and practices within the organization for future space missions. We have used System Modeling Language (SysML) to develop a model repository in which mission requirements are formulated as requirement diagrams. Physical aspects of the satellite design specifications are captured in the form of SysML block definition diagrams (BDD) at system, subsystem and unit level. The behavioral modeling is performed using SysML activity and state machine diagrams. Internal block diagrams (IBD) and parametric diagrams are used to establish parametric relationship among different design entities and perform mathematical analyses. Integration of existing analytical modeling tools with SysML model was also investigated in this effort.

INTRODUCTION

Satellite design industry has successfully accepted the Systems Engineering (SE) since its inception taking advantage from SE methodologies and techniques to harness underlying complexities. In traditional systems engineering approach it becomes difficult to assess the completeness and consistency of information spread across plenty of tools and documents. In particular, it is difficult to perform traceability and assess the change impacts of certain design parameters.

Model-based Systems Engineering (MBSE) is an emerging technology that is providing the new horizon in systems engineering domain. MBSE produces a "system model" contained in a model repository which includes more formal, complete and semantically rich system specifications and design for rigorous analysis, verification and validation at earlier phase of projects. MBSE provides a consistent and automatically traceable system design which enhances the quality of traceability and change impact assessment.

System Modeling Language (SysML) has been standardized by the INCOSE (International Council on Systems Engineering) as a graphical modeling language for system engineers [15]. SysML is a domain-specific modeling language used to specify, analyze, design, optimize, and verify systems and system-of-systems [13]. We have used SysML to develop a scalable model repository for remote sensing satellite.

Objectives

This research was conducted at first hand in order to investigate the usefulness of the MBSE approach to assist the existing system engineering process and practices for future space missions. The prime focus of the work presented in this paper is:

- To establish parametric relationships among various design elements with parametric modeling using SysML internal block diagrams and parametric diagrams
- To show how the SysML modeling language can be useful to complement the requirements engineering process, i.e. to better define, trace and validate requirements.
- To enhance understanding of SysML pertaining to satellite modeling for our system engineers
- Seamless interface of existing system level design tools (STK® , Matlab/Simulink® and Microsoft Excel®) with SysML Satellite Model to assist on-the-fly parameters tuning

To achieve these objectives, we have selected ongoing remote sensing student satellite (NSS-1) as case study and developed its SysML model repository. We have used MagicDraw 18.1 from No Magic Inc. [14] as SysML modeling tool.

PREVIOUS WORK

The system engineering division at Satellite Research and Development Center, SUPARCO has been involved in state-of-the-art R&D in systems engineering applied to spacecraft design and development. To assist the system engineering process and procedures several automated applications / tools have been devised. Development of System Integrated Design Environment (SIDE) was one of such efforts [10]. SIDE is a spreadsheet based integration tool which connects various satellite system engineering design and analysis tools in a local area network (LAN).

A significant work on use of SysML has been done by Spangelo et al. for CubeSat [1]. The paper reports on using MBSE and SysML to model a standard CubeSat modeling framework and applying that model to Radio Aurora Explorer (RAX) mission [7].

M. Sayanjali et al. [6] have used MBSE approach using SysML for remote sensing (RS) satellite conceptual design. The RS satellite captures image in two visual and IR spectrum. Block definition diagrams (BDD) are created to specify physical architectures for systems and subsystems. By defining parametric diagram (PD) for system sizing and efficiency measuring, trade-off study is explored.

G.R. Alessandro et al. [4] have used SysML to model Attitude and Orbit Control System (AOCS) software for a satellite using SysML (Systems Modeling Language). Constraints and parametric diagrams for AOCS subsystem are depicted and explained. Physical aspects are modeled using sequence diagrams. Finally, model is refined to become a PSM model from software viewpoint, allowing code generation.

In the literature, model based approach is compared with document based approach. This develops a sense that the MBSE is a paradigm shift which contradicts the document based systems engineering approach. This is one of the reasons that model based approach is currently less accepted by conventional system engineers. The document based approach may be inefficient in certain aspects, which can be improved by employing MBSE approach making the practice of systems engineering as effective as possible. We therefore advocate a hybrid MBSE methodology as described in subsequent sections which is assisting our conventional systems engineering process in most appropriate way.

SYSTEM MODELING LANGUAGE (SYSML)

Object Management Group's (OMG) SysML is a UML based standard graphical language for modeling systems. SysML is comprised of a graphical notation and an information model that

emphasizes the formal capture, description, and communication of systems specifications. Diagrams can be constructed to describe various systems based on the structure and behavior of the system.

A. SysML Diagrams

Systems Engineers are used to MS Visio diagrams in which precise meanings are not associated with each type of arrows and blocks, and in which the purpose of the diagram is supporting a paragraph, or a document. The information model distinguishes SysML from conventional drawing and simulation tools in the following specific ways:

- Formal information model for underlying diagrams.
- SysML model of a system can be tested to ensure that the models are complete and consistent
- The SysML diagrams have robust semantics and details captured for formal specifications.
- The SysML model can be used to interface and inter-operate with other tools and data sources. This also assists the integrated product development process [10]

B. Hybrid Approach

The promoters of MBSE advocate the benefits of using SysML and similar modeling paradigms. However, in practice we have found that aerospace industry is less familiar with the inherent advantages and usefulness of MBSE approach. The aerospace system engineers have mostly electrical and mechanical engineering background. They are inherently at weaker side in Software Engineering and Object Oriented Programming (OOP).

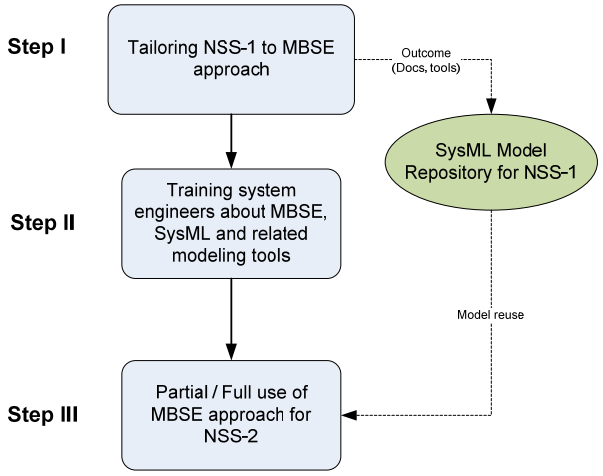


Fig. 13: MBSE hybrid approach

They usually prefer conventional ways of doing system engineering through the tools they are most familiar with, such as MS Office tools chain (Excel, Word, Visio), STK, Matlab etc. Enabling our system engineers to adopt SysML is the foremost challenge for us. In order to fill the gap between traditional document-based approach and model based approach, we have devised to follow a step-by-step hybrid approach as shown in Fig. 13.

C. Associated Challenges

Although model based systems engineering has a long history [12], it has not yet fully taken hold within the satellite systems engineering. One of the reason is that MBSE needs some other disciplines knowledge particularly from modeling and software engineering which is usually absent in conventional system engineers belonging to core engineering disciplines (electrical, mechanical).

They are inherently uncomfortable with UML/SysML diagrams and follow conventional ways of writing design documents.

Another important aspect of complexity is that MBSE approach not only requires a language such as SysML to represent the system, but also a method that defines the activities and artifacts, and a tool to implement the modeling language and method. The language, method, and tool each introduce their own concepts, and must be learned to master model-based systems engineering. This skill must then be applied to a particular domain, such as designing a spacecraft. Additional modeling challenges are associated with scaling the modeling effort to larger projects, and in the context of a diverse development environment.

There are also many different types of models involved in MBSE effort beyond the SysML model, such as a multitude of analysis models, hardware models, and software models. The integration among the different models and tools, and other engineering artifacts is another challenge associated with MBSE.

SysML Model of NSS-1 Satellite

NSS-1 Satellite and System Engineering

NSS-1 is a remote sensing student satellite. The design philosophy of NSS-1 is conceived in such a way that all the systems engineering related tasks are performed by system engineers at SUPARCO, while design and development of various satellite units is to be undertaken by university students. Throughout this process the system engineering team at SUPARCO has to be involved iteratively in the system level documents, design and analysis tools in order to tune and freeze the design parameters as early as possible.

At this moment system engineering tasks have been performed using conventional approach based on heritage knowledge based tools and documents. This approach inherently involves a lot of manual procedures which are error-prone due to parametric mismatch. We have therefore decided to benefit from model based approach to address these problems.

Model Repository for NSS-1

We have reused some model elements from FireSat SysML model which is used as a case study example in [9]. The NSS-1 modeling group consists of individuals from multi-disciplinary areas in SUPARCO and academia. The modeling framework uses SysML to capture common design patterns of typical nanosatellites class: managing values, describing scenarios, and describing functions, parts, and subsystems as well as the relationships between these design patterns.

The framework illuminates a path to an integrated model-based engineering environment, including interoperability with existing system engineering models ready in existing analysis tools such as MS Excel, Matlab, STK. This NSS-1 model repository demonstrates the possibility of a diverse set of analysis applications that are provided with information about the space system from the system model to accomplish analysis driven by a formal description of the mission, flight and ground systems.

Following major modeling elements have been introduced in NSS-1 model which were not found in the FireSAT model [9].

- Modeling elements for standard SI units have been identified and created in NSS-1 repository
- Mass, power and reliability budgets have been incorporated

- State machine diagrams for satellite's operational modes and their transitions at system and subsystem level
- Description of satellite bus (CAN) protocol via SysML sequence diagram

NSS-1 model is developed using diverse SysML diagrams. We have briefly described each type of diagram used in the NSS-1 context in following paragraphs.

Package Diagram

The SysML package diagram is used to organize the model elements contained in the model. Fig. 14 shows the package diagram for the NSS-1 satellite model and includes packages for Requirements, Behavior, Architecture and Parametrics. Package diagrams also assist to navigate through the model.

Requirement Diagram

Requirement Diagram (RD) defines requirements, traceability and test case that verify each requirement. A requirement diagram is shown in Fig. 15 which represents the requirements contained in the 'Requirements' package to specify the NSS-1 system. For demonstration purpose some of the requirements are suppressed in order to save the paper space.

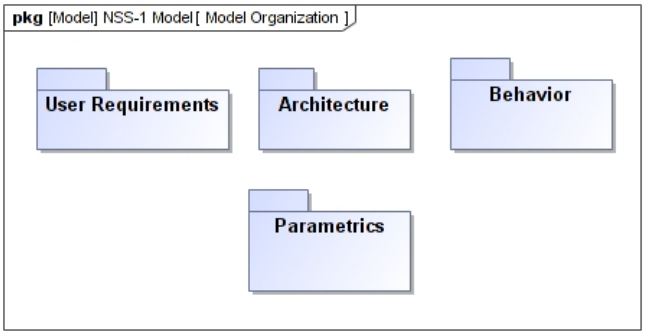


Fig. 14: NSS-1 Model Organization

The relationships among the requirements and other model elements are important aspects of requirement diagrams. These include relationships for defining a requirements hierarchy, deriving requirements, satisfying requirements, verifying requirements and refining requirements. Practically speaking, being able to ask a modeling tool to generate requirement traceability and verification matrices, and to perform an automated impact analysis when the requirements change, is a very important time-saving feature for system engineering. SysML requirement diagram in this context are quite helpful to assess NSS-1 requirement specifications and their verification and validation (V&V).

Block Definition Diagrams

Block Definition Diagram (BDD) defines physical architecture of the system. We have displayed various kinds of model elements and relationships between NSS-1 design elements on a BDD to express information about spacecraft architecture. We have also adopted design techniques for creating extensible system structures. This practice reduced the model development time. In future this will also reduce cost to change the design as the stakeholders' needs evolve.

The top level block diagram of the NSS-1 is depicted in Fig. 16. SysML diagrams have formal semantics for each of the modeling elements. The diamond-ended arrow indicates the "directed composition" relationship between the blocks. The diagram may be interpreted as "NSS-1 System is composed of a space segment, a ground segment and associated launch services".

Internal Block Diagrams

The internal block diagram (IBD) shows the interconnection and external interfaces between the system and its external user and environment. They formally describe the qualitative flows between model elements. The internal block diagrams are created for each block of a BDD which represents its internal parts in the form of instances connected with each other and the environment through ports.

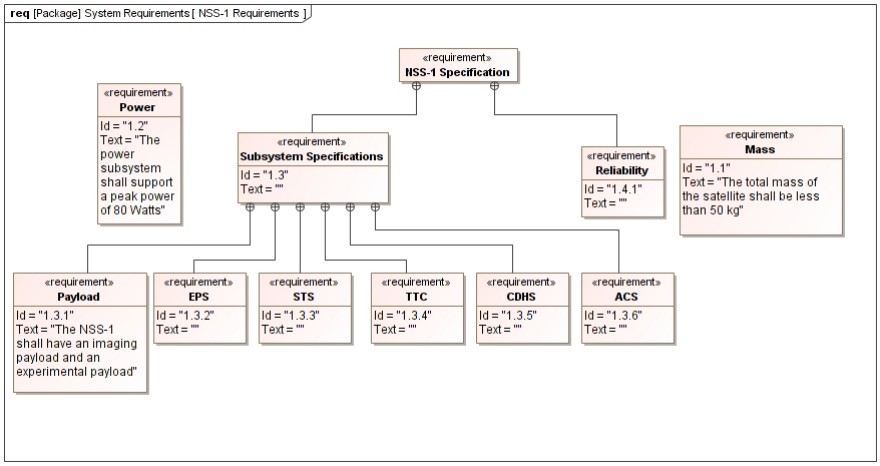


Fig. 15: NSS-1 Requirement Diagram

SysML Parametrics

SysML Parametric Diagrams provide a way to integrate engineering analysis models described in mathematical equations and constraints. For demonstration of SysML parametrics, we have modeled mass and power budget of NSS-1 in this paper. The procedure starts with requirement diagrams in which the constraint blocks have been created which calculate the mass and power margins in the budgets keeping in view the mass and power requirement by the user. This scenario is depicted in SysML requirement diagram with 'verify' relationship shown in Fig. 20.

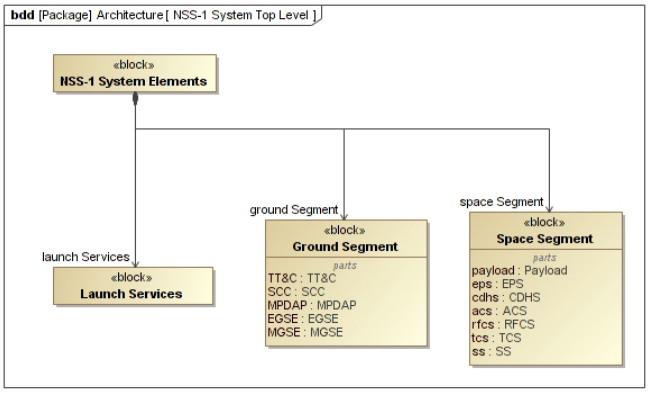


Fig. 16: NSS-1 Top Level Block Definition Diagram

A test case or other named element can be used in a similar fashion to represent any of the standard verification methods such as review, inspection, code walk through, test or analysis. NSS-1 is composed of seven subsystems. Each subsystem has a mass and there is a total mass budget for the satellite of 50 kilograms. The ACS, CDHS, Payload, TCS and RFCS subsystems draw electrical power from the EPS, which can supply a maximum of 80 watts. Given the masses of all subsystems and the power requirements of five subsystems, we calculated the total mass and power demand for the satellite and compare it against the requirement.

To meet the mass and power constraints the parametric diagram is created from the “NSS-1 Satellite” block as shown in Fig. 18. After connecting the parameter ports to respective inputs/outputs, we have validated the model using ParaMagic.

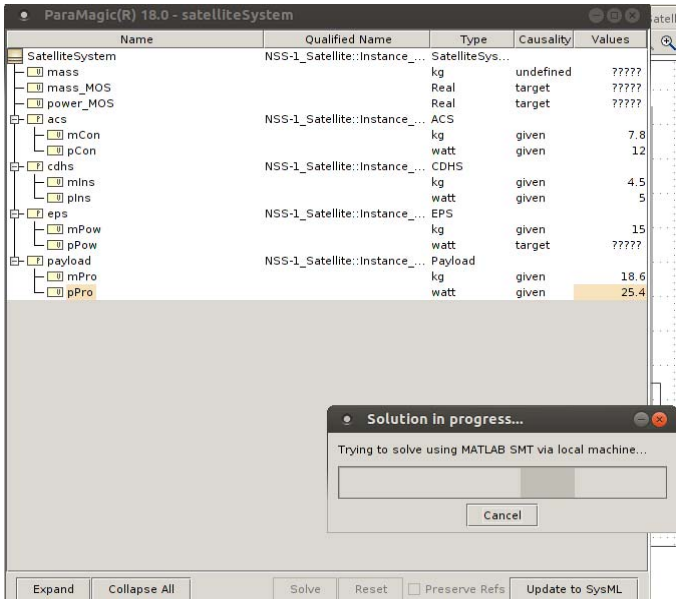


Fig. 17: Solving Budget Equations using SysML Parametrics

Following solvers are supported by ParaMagic for mathematical analysis: Mathematica, Modelica MATLAB SMT. For our analysis, we have used local MATLAB SMT solver to analyze and tune mass and power allocations to various subsystem. After assigning causalities and values to all the parameters, MATLAB SMT is invoked within the ParaMagic in order to solve for known / target parameters. In our case, margins of safety (MoS) for power and mass are target (output) parameters whereas satellite ‘mass’ is the unknown parameter to be computed through following trivial equation:

The power constraint on EPS subsystem is computed based on power demanded by electronic subsystems. Following equation is specified in the constraint block in order to compute the power budget.

A snapshot of the parameters settings and ongoing solution by the MATLAB SMT solver is depicted in Fig. 17 in which four parameters (mass, pPow, mass_MOS and power_MOS) are being calculated based on valued parameters listed in right most column named as “Values”.

The values of these parameters can either be given directly or may be imported from MS Excel worksheets.

The mass budget tool was earlier created and enhanced within MS Excel in SIDE platform [10]. In SysML model repository the tool has enhanced capabilities to trace down the requirements which in turn are very useful for change impact and V&V of the requirements.

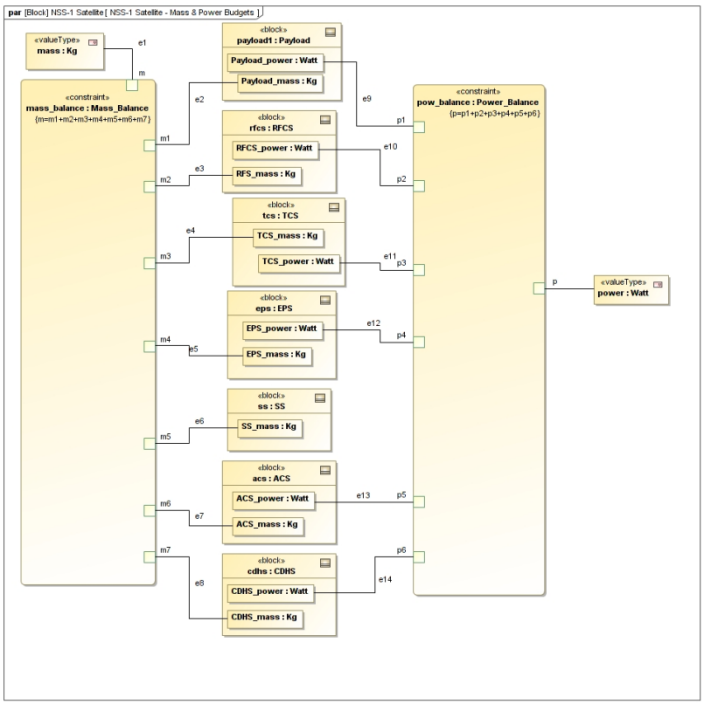


Fig. 18: Parametric Diagram for Mass and Power Budgets

The resulting values of the parameters can be seen on the same dialogue after the solution is complete. The values of all parameters after the calculation are updated to SysML instance which appear as in Fig. 19. For the sake of brevity, we have only shown the analysis for four subsystems.

We can see that mass and power budget margins are calculated to be positive (mass_MOS = 0.082, power_MOS = 0.47) which verify the mass and power requirements as given in requirement diagram (Fig. 20). We have demonstrated only mass and power budgets in this paper. The NSS-1 model repository contains following different analysis tools implemented in the form of constraint blocks and parametric diagrams:

- Sensors Field of View (FOV) analysis
- Reliability budget
- Orbital analysis
- Trade-Off analysis for ACS configurations

Interface with Existing SE Tools

One of the main objectives of this work was to establish seamless interfaces between our existing system engineering design and analysis tools and the SysML repository. With the help of MagicDraw plug-ins, we are able to communicate between common engineering analyses tools with NSS-1 model repository as follows.

Interface with MS Excel

MagicDraw provide an interface to import and export Data between SysML model and Microsoft Excel files using ParaMagic plugin. In our NSS-1 case study, we have computed some parameters using spreadsheets, read these values in ParaMagic to populate SysML instance slot values, solved the SysML model using ParaMagic, and exported solved values back to spreadsheets.

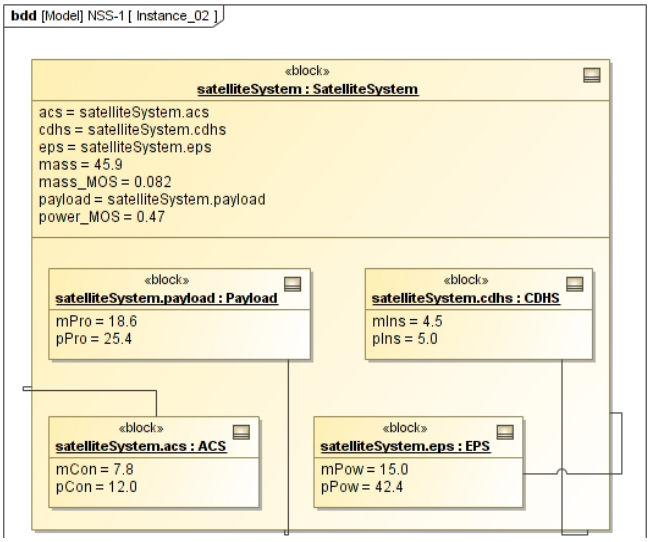


Fig. 19: Mass & Power Budget Instance after Solution

Interface with MATLAB

One of the main objectives of this work was to transfer mathematical expressions from MATLAB source code into the NSS-1 model repository. This was seamlessly achieved using InterCAX plug-ins. InterCAX offers parametric solver plug-ins for SysML modeling tools. The plug-ins treats MATLAB functions and scripts as constraints in SysML parametric diagrams. When the SysML product calls MATLAB, the subject MATLAB function/script runs locally on the user's machine and creates an output text file that is automatically read back into the SysML model. Simulink models can also be accessed through an intermediary MATLAB script.

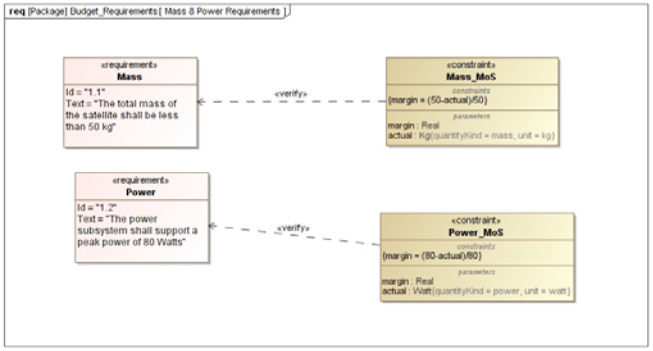


Fig. 20: Verification of Mass and Power Budgets Requirements for NSS-1

Automatic Report Generation

An important aspect of the MBSE is automatic documentation generation from the model. Specifically from SysML models, allocation table, requirement table and coverage analysis are generated automatically from model repository. Similar to UML, we can also generate corresponding code structure for SysML model in popular object oriented programming languages such as Java, C++ or C#. This code can be used to develop prototype models for simulation of the system. For software based models, the code structure can be directly used for implementation of operations/functions by software developers.

FUTURE WORK

The work presented in this paper is an initial study and modeling to evaluate the effectiveness of the MBSE approach for our satellite design projects. In future we envisage training our system

engineers about the SysML, its development environment and the underlying object oriented paradigm for system modeling. The outcomes would be utilized in the future mission using partial or full utilization of the MBSE approach. An indirect link between STK and SysML can be established using MS Excel spreadsheet as intermediate tool. A direct interface between a SysML model and AGI Components was demonstrated in INCOSE MBSE workshop [15]. The interface was developed using Systems Lifecycle Management (SLIM) methodology [3]. In future we wish to incorporate SLIM to interface STK and SysML model repository.

CONCLUSIONS

Model based approach is found to be promising for a reduced development risk, improved quality and increased productivity. The NSS-1 model repository developed during this phase can be used effectively for knowledge transfer. After this research, following key advantages of MBSE using SysML have been identified:

- Project documentation becomes more consistent
- Well structured models and clear separation between 'architecture' and 'behavior'
- SysML model and diagrams can point out missing design information and relevant information about a particular design element can easily be found
- Opens the way to automated systems engineering, through automatic document generation, traceability, assisted design verification and system behavioral simulation

Nevertheless, the following risks have also been foreseen:

-SysML learning curve is steep. It may require a considerable knowledge of software engineering and object oriented programming (OOP) by the satellite system engineers before its advantages come out

-Diagrams may be misinterpreted if SysML modeling is not learned beforehand

REFERENCES

[1] Spangelo, S.C. Kaslow, D. ; Delp, C. ; Cole, B. ; Anderson, L. ; Fosse, E. ; Gilbert, B.S. ; Hartman, L. ; Kahn, T. ; Cutler, J. "Applying Model Based Systems Engineering (MBSE) to a standard CubeSat", IEEE Aerospace Conference, PP 1 – 20, 2012.

[2] K.D. Bell and L.A. Hsu, "Balancing Performance and Cost for Cost-Effective Satellite Systems Design Using an Integrated Cost Engineering Model", Aerospace Corporation, Aerospace Report TOR-95(5409)-2, Air Force Material Command under Contract No. F0470 1 -93-C-0094, 18 December 1994.

[3] Bajaj, M., Zwemer, D., Peak, R., Phung, A., Scott, A. and Wilson, M. (2011). "Satellites to Supply Chains, Energy to Finance — SLIM for Model-Based Systems Engineering", Part 1: Motivation and Concept of SLIM. 21st Annual INCOSE International Symposium, Denver, CO, June 20-23, 2011.

[4] Alessandro Gerlinger Romero, Mauricio Gonçalves Vieira Ferreira, "Modeling An Attitude And Orbit Control System Using SysML", 2nd Workshop em Engenharia e Tecnologia Espaciais, 2011.

[5] C.E.G. Cardenas, "Modeling Embedded Systems Using SysML", Master Thesis, 2009.

[6] M. Sayanjali, O. Nabdel, "Remote Sensing Satellite Design using Model Based System Engineering", Journal of Science and Engineering, 2013.

[7] Bahcivan, H., and J. W. Cutler (2012), “Radio Aurora Explorer: Mission Science and Radar System”, Radio Sci., 47, doi:10.1029/2011RS004817, 2012.
[8] Weilkiens, T. (2006). Systems Engineering with SysML / UML, Modeling, Analysis, Design. London: Elsevier, 2006.
[9] Larson, W. J., & Wertz , J., (2005). “Space Mission Analysis and Design”. London: Kluwer Academic Publishers, 2005.
[10] Muhammad W., Sadiq M.U., Ahmed I., Muzammil F., “Development of an Integrated Design Environment to Assist Conceptual Satellite Design”, Proceedings of the Asia Pacific Council on Systems Engineering Conference (APCOSEC-2013), Yokohama, Japan, Sept-2013.
[11] Bruno B. F. Leonor,Walter A. Dos Santos and S. Stephany, “A Knowledge-Based and Model-Driven Requirements Engineering Approach to Conceptual Satellite Design”, Proceeding: ER '09 Proceedings of the 28th International Conference on Conceptual Modeling, Pages 487 – 500, 2009.
[12] Wymore, A. W., "Model Based Systems Engineering”, CRC Press, 1993
[13] OMG Website: http://www.omgsysml.org
[14] No Magic Inc. Website: http://www.nomagic.com
[15] INCOSE Website: http://www.incose.org/

Small Satellites and Non-State Actors: A New Challenge To State Sovereignty

George Anthony Long, Esq.

Managing Member, Legal Parallax, LLC

P. O. Box 18282, Fountain Hills, Arizona 85260

Tel: 480.201.1891/Fax: 866.779.4805

email: gal@spacejurist.com

I. Introduction

Space exploration is no longer the exclusive domain of governmental entities; instead, non-state actors utilization of outer space, especially Earth orbit activities, is the growing trend. The re-emergence of Small Satellites ("SmallSats") serves to accelerate non-state actors access to and use of outer space as it lowers the costs and overhead associated with non-crewed space objects. Additionally, the advent of new SmallSat related technologies will broaden the scope and nature of non-state actors Earth orbit activities. In fact, some envision that within the next generation SmallSats will transition into a consumer product like previous innovations associated with technological advances in the communication and informational industries such as radio, television, personal computers and mobile phones.

Small satellites encompass a spectrum of space objects with various designations but generally the category refers to space objects having a mass of less than 500 kg. This encompasses a range of space assets designated as Mini Satellites ("Minisats"), Micro Satellites ("Microsats"), Nano Satellites ("Nanosats"), Pico Satellites ("Picosats"), Femto Satellites ("Femtosats") and Spires. Although the definitions for these varying sub-categories are arbitrary, an object's mass generally serves as the point of differentiation among SmallSats. A SmallSat's mass for classification purposes is based on its in-orbit fully fueled mass. Minisats' mass range between 100 and 500 kg. Microsats have a mass between 10 and 100 kg. Nanosats have a mass of 1 to 10 kg. Picosats possess a mass of 10g to 1 kg while the mass of Femtosats is less than 10 g. Spires are about the size of a postage stamp and they contain all the essentials for a satellite such as a radio, aerials a solar cell and instruments. It is estimated that about 100 Spires can fit inside a Cube Satellite ("CubeSat"). While CubeSats ("CubeSats") come within the SmallSat classification, they are not truly a distinct sub-category as their name derives from their design and not their mass. Cubesats normally fall within the nanosat or picosat classification. The Cubesat design, however, is the standard most utilized by SmallSats. A newer design known as a Tube Satellite ("TubeSat") is emerging to compete with the CubeSat design. TubeSats are s a low cost alternative to CubeSats and have a maximum mass of 0.75 kg. Regardless of its designation, a SmallSat can be deployed and operated singularly or as part of a collective referred to as a constellation or swarm. SmallSats in a collective operate in unison to accomplish a common goal.

The re-emergence, proliferation and anticipated consumerization of SmallSats is viewed by some as unleashing "'the most powerful force in the universe - human creativity."' Human creativity, however, is not limited to ideas and exploits deemed acceptable by a State or the community of nations. Exercising human creativity can result in some non-state actors using

SmallSats as a portal to deliberately or indirectly infringe on a State's territorial sovereignty. This paper will focus on the potential use of SmallSats by non-state actors that can infringe on State sovereignty and the legal recourse and remedies, if any, available to the offended State under The Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies ("Outer Space Treaty").

II. Current Concept of Sovereignty

The current inclination of State sovereignty has served as the cornerstone of international law since the 1684 Peace of Westphalia which ended the armed European power struggle commonly referred to as the Thirty Years War. This conflict transformed the legitimacy of European State sovereignty from a religious basis to a secular foundation. Since the Peace of Westphalia, sovereignty has evolved to mean that a State has exclusive and absolute jurisdiction and control over its territory and the people that inhabit or are otherwise present in its territory. This Westphalian concept of sovereignty also views all States as equal and adhere to the principle of non-intervention in the domestic affairs of another State. Additionally, the current reach of sovereignty extends to the air space over a State's territory. It does not, however, extend to the area of outer space above a State's territory. Outer Space Treaty Article II expressly provides that outer space "is not subject to national appropriation by claim of sovereignty, by means of use or occupation or any other means."

Under contemporary international law, a State is deemed to exist and be a juridical person when it possesses (1) a permanent population, (2) defined territory, (3) a government and (4) the capacity to enter into relations with other States. Customary international law recognizes five tentacles for a State to exercise jurisdiction over non-state actors. The five basis are territorial jurisdiction, protective jurisdiction, national jurisdiction, passive personality jurisdiction, and universal jurisdiction. The territorial theory rests on a state possessing sovereignty over its territory which includes acts occurring within the territory and acts occurring outside of the territory but having an effect within the territory. The protective theory flows from a state protecting its vital national interests. The national theory flows from a state having jurisdiction over the conduct of its nationals while the passive personality theory is premised on a state having jurisdiction based on its nationals being a victim. Each of these four theories is rooted directly or indirectly on a State's territorial sovereignty or integrity. Universal jurisdiction, however, is unique as it alone does not require a nexus between an act and the State exerting jurisdiction.

Universal jurisdiction allows a State to prosecute certain offenses under customary international law which do not have any nexus to it and the prosecution can proceed "even over the objection of the defendants' and victims' home states." The mere fact that certain conduct is universally condemned or prohibited by States in their domestic law "is not necessarily significant or relevant" to classifying conduct as an international crime under customary international law and subject to universal jurisdiction. Murder, for example, is not a violation of customary international law or a crime subject to universal jurisdiction because the act is not of mutual concern to the international community. For conduct to qualify as an offense subject to universal jurisdiction it must be "'contrary to a peremptory norm of international law'" and "'be so serious and on such a scale that [it] can justly be regarded as an attack on the international legal order.'" Although the list is unsettled, there is no dispute that acts subject to universal jurisdiction under customary international law include piracy, slavery, genocide, crimes against humanity, war crimes, and torture.

International law grants a State broad jurisdiction to prosecute or otherwise target non-state actors who engage in activities that infringe on its territorial integrity or who commit an international crime subject to universal jurisdiction. However, prosecuting or targeting individuals or non-state entities does not necessarily equate with a cessation of the offending activity. This circumstance presents a challenging scenario, especially in the context of outer space.

III. Space Law, SmallSats and Infringement of State Sovereignty

A general principle of international law is that what is not prohibited is permitted. Pursuant to this principle, a use of outer space which is not expressly prohibited by the Outer Space Treaty is generally viewed as being legal under international law.

Outer Space Treaty Article 1 expressly provides that space "shall be free for exploitation and use by all States without discrimination of any kind, on a basis of equality and in accordance with international law." Accordingly, the treaty does not exclude ot restrict access to and use of space based on the ideology or other belief of a State or a non-state actor. It does, however, contain 6 express exclusions to the use of outer space. The Outer Space Treaty expressly precludes space activities (1)which entail any national claim of sovereignty or other appropriation (2) which are not in compliance with international law, (3) which involve placing weapons of mass destruction in Earth orbit, on a celestial boy, or in outer space in any manner, (4)which are not for a "peaceful purposes" on the Moon and other celestial bodies," (5) which cause harmful contamination to space, the Moon or other celestial bodies, and (6) and which cause "adverse changes in the environment of the Earth resulting from the introduction of extraterrestrial matter."

Space activities which comply with the broad usage afforded by the Outer Space Treaty do not violate international law. Space activities may, however, violate a State's domestic law enacted pursuant to its sovereign powers or otherwise infringe on a State's territorial sovereignty or national security interests. Non-state actors using space for such a purpose, while rare, is not without precedence.

It is not unusual for non-state actors to utilize satellites to surreptitiously broadcast, promote or advertise religious, political or ideological beliefs or propaganda. Such activity first occurred in the United States almost 30 years ago in 1986. In that incident, a person using the alias "Captain Midnight" hijacked the satellite broadcast signal for a cable television network known as HBO. The signal was hijacked for about four minutes and the HBO programming was replaced it with a printed message regarding HBO's fee. The message stated:

> "GOODEVENING HBO
> FROM CAPTAIN MIDNIGHT
> $12.95/MONTH ?
> NO WAY !
> [SHOWTIME/MOVIE CHANNEL BEWARE !]

Within six months after the incident, the United States enacted a domestic law criminalizing the interference with the operation of a satellite. That statute, 18 U.S.C. § 1367 reads as follows:

> (a) Whoever, without the authority of the satellite operator, intentionally or maliciously interferes with the authorized operation of a communications or weather satellite or obstructs or hinders any satellite transmission shall be fined in accordance with this title or imprisoned not more than ten years or both.

(b) This section does not prohibit any lawfully authorized investigative, protective, or intelligence activity of a law enforcement agency or of an intelligence agency of the United States.

There has been one known prosecution under the statute. That case involved a signal hijacking which temporarily interrupted the scheduled adult satellite programming on the Playboy Channel with a religious Christian message.

Asian satellite operators reportedly experienced such signal hijackings in 2002 perpetrated by a non-state actor. The hijacking is said to have involved 9 stations of the China Central Television Station (CCTV) and 10 provincial TV channels during a seven day period in June, 2002. This was accomplished by illegally procuring the satellite frequency and technical parameters used by China's television programs and using the stolen information to disrupt the transmissions of two satellites and briefly transmit propaganda material.

SmallSats potentially offer new and creative ways for non-state actors motivated by ideological, religious, or political beliefs to infringe on a State's sovereignty. While it is difficult to accurately forecast how such "creativity" will manifest itself, one potential avenue consists in advertising.

The Outer Space Treaty nor any other space treaty precludes the placement of advertising in outer space whether for commercial, ideological, religious, or any other purpose. One enterprising entrepreneurial outfit has offered its services to use a constellation of satellites to place corporate logos in Earth orbit for an astronomical amount of money. While this particular service offer is farcical, it is technologically feasible to use a constellation of satellites to place a symbol, logo or 'billboard" in outer space for advertising, promotional or propaganda purposes. Such an act would not violate international law or the law of any State except the United States.

The United States apparently considered the possibility of a non-state actor placing advertising in space so viable that it enacted legislation tailored to control such conduct within the scope of its jurisdiction. The statute, 51 U.S.C. § 50911 reads as follows:

a) Licensing.--Notwithstanding the provisions of this chapter or any other provision of law, the Secretary may not, for the launch of a payload containing any material to be used for the purposes of obtrusive space advertising--

(1) issue or transfer a license under this chapter; or

(2) waive the license requirements of this chapter. (b) Launching.--No holder of a license under this chapter may launch a payload containing any material to be used for purposes of obtrusive space advertising.

(c) Commercial space advertising.--Nothing in this section shall apply to nonobtrusive commercial space advertising, including advertising on--

(1) commercial space transportation vehicles; (2) space infrastructure payloads; (3) space launch facilities; and (4) launch support facilities.

By its plain language, the United States prohibition against space advertising is limited to "obstructive space advertising" by a United States licensee. The phrase "obstructive space advertising" means "advertising in outer space that is capable of being recognized by a human being on the surface of the Earth without the aid of a telescope or other technological device." A person violating the statute is subject to a civil penalty of not more than $100,000 a day for each day the violation continues.

While the statute does not define the term advertising, it is unclear whether it is limited to commercial advertisement or if it also reaches non-commercial advertising. If the prohibition extends beyond commercial speech, then it may run afoul of the First Amendment of the United States Constitution which, with certain judicially created exceptions, forbids Congress from enacting a law restraining the freedom of speech.

Nevertheless, if a non-state actor uses SmallSats to place a logo, symbol or other form of advertising, promotion or propaganda in Earth orbit that is deemed to be an ideological, political or religious adversity of a State, then the offended State will desire removal or elimination of the logo, symbol or advertising to prevent it from being seen or displayed over its sovereign territory. This goal will exist despite the length of the temporal period such a display is in orbit. Prosecuting the non-state actor or actors perpetrating such a space activity will, in all likelihood, simply be insufficient for sovereignty purposes as that does not ensure removal or elimination of the orbiting logo, symbol or advertising. A State will seek ways to neutralize or nullify such a SmallSat network. This circumstance present the core issue of how to quell the sovereignty infringing space activity without contravening the Outer Space Treaty.

IV. Space Law Remedies For SmallSats Infringement of Sovereignty

Preventing the deployment of an offensive SmallSat constellation is the first remedy for addressing SmallSats infringement of sovereignty. This necessitates implementing SmallSat licensing protocols and procedures at the international or State level similar to the legislation promulgated by the United States. As long as SmallSats are deployed as secondary payloads from launch vehicles or from the International Space Station ("ISS") such regulatory measures should prove viable. Anticipated technological advances, however, will eventually erode the effectiveness of any such licensing schemes, at least to the extent non-state actors may desire to circumvent such regulatory protocols.

The growth in the SmallSat industry has generated demands for developing launch and deployment systems specifically designed for SmallSats. One such system envisions using aircraft as a launch platform for SmallSats. Launching SmallSats from a dedicated launcher such as an aircraft, rather than as a secondary payload or from the ISS, vests SmallSat operators with the flexibility to specify the launch date, launch site and "the exact orbital parameters" the operator requires. Developing this type of mobile and dedicated launch platforms for SmallSats is a necessary step for the maturation of the SmallSat industry. Unfortunately, the progress also inures to the benefit of non-state actors who may possess the intent to surreptitiously place a SmallSat constellation in Earth orbit for purposes of infringing on the sovereignty of one or several States. This essentially relegates SmallSat licensing schemes to the status of a temporary measure which diminishes with technological advancements associated with launching SmallSats.

Since prevention presents a limited window of opportunity, it seems the alternatives for dealing with an SmallSat constellation which infringes on territorial sovereignty are physical removal of the SmallSats from orbit, destroying the constellation or disabling the SmallSats comprising the constellation. The problem is that there is no clear legal path for using any of the alternatives.

Removing, destroying or disabling a SmallSat constellation entails employing Anti-Satellite ("ASAT") weaponry. An ASAT is best described as a weapon or device, whether terrestrially based or stationed in outer space, which can destroy or disable a satellite in space or Earth orbit. It can take the form of a kinetic energy projectile or object, which destroys its target by physical impact or

explosion and in the process creates a debris field. It can also consist of a "directed energy" beam such a laser, and devices which can generate and aim an electromagnetic pulse or a high powered radio frequency. A directed energy ASAT disables the SmallSats comprising the constellation which in turn creates space junk in the form of inoperable and uncontrollable space objects. There is a dispute over whether an ASAT includes a space object which can physically "capture" and remove another space object from its orbit or from outer space. For purposes of this paper, a space object which can "capture" another space object will be viewed as an ASAT.

There is not any international agreement which expressly prohibits the use of an ASAT. Nevertheless, if a State desires to use a kinetic energy or directed energy ASAT, then the decision triggers a prior consultation requirement under the Outer Space Treaty Article IX since a successful result would produce either a debris field or inoperable space junk. Article IX provides, in part, that if a State plans an activity which it believes "would cause potentially harmful interference" with the peaceful outer space activities of other States, then that State "shall undertake appropriate international consultations before proceeding" with the planned activity. An activity intended to create either a debris field or space junk undeniably "causes potentially harmful interference" to other States peaceful use of outer space.

In the context of an offending SmallSat constellation, compliance with Article IX essentially requires a State to consult with other States prior to taking action to destroy or disable an offense to its territorial sovereignty. This prior consultation can possibly generate political tension in as much as some States may not consider the orbiting offensive symbol, logo or advertisement as a sufficient affront to sovereignty which justifies creating a new debris field or new space junk.

Pursuant to the first paragraph of Outer Space Treaty Article VIII, a State retains jurisdiction and control over a space object on its registry. The agreed upon construction of this provision leaves "little doubt" that a State may not destroy or remove the space object of another State or its component parts without permission. A SmallSat constellation which intentionally infringes on a State's territorial sovereignty will probably not be on the registry of any State. However, this does not necessary mean an unregistered SmallSat constellation is not entitled to any protection under space law.

Many satellites are not registered in accordance with Article VIII. In fact, it is believed that State's, as a matter of practice, do not register any military satellites,, spy satellites, or other secret space object. Nevertheless, space faring States have apparently treated such unregistered space objects the same as registered space objects. In other words, space faring States have traditionally refrained from destroying, incapacitating or removing from orbit space objects that it did not launch. An argument could be made that this practice has produced the customary law that a State does not destroy, incapacitate or remove from orbit a space object that it did not launch, especially an operational space object. If the treatment of an unregistered space object does not constitute customary international law, then physical removal from orbit seemingly represents the more practical and reasonable alternative.

One drawback to physically removing a space object from orbit is that it seems that the United States may be the only space faring State which currently possesses the space asset capable of accomplishing this feat. Although the space shuttle has retired, the United States does operate a space asset designated as X-37 which is a small robotic version of the space shuttle. This asset, like the Space Shuttle, has a cargo bay and apparently tthe potential to physically remove a space object from orbit. Another drawback in removing an unregistered space object owned by a non-state actor

exists if some or all of the SmallSats comprising the constellation are equipped with a small detonating device, effectively making the constellation a "defensive" ASAT. Such a tactic would discourage physical removal of an offending SmallSat constellation.

Lastly, regardless of which alternative a State elects to pursue, removal, destruction or disabling, it will entail the use of force against a space object not subject not registered to it. The United Nations Charter governs a Sate's use of force. Article 2(4) of the Charter prohibits the use of force except when (1) authorized by the United Nations Security Council, (2) in self-defense as authorized by U.N. Charter Article 51 and (3) as anticipatory self-defense provided the use of force is a necessity and is proportional to the response. Also, there are uses of force which do not come within the scope of U.N. Charter 2(4) and accordingly the use of such force would not be a violation of Article (2(4). This distinction developed as a State's use of force is said to take three forms which are "force as armed violence, force as coercion, and force as interference." Coercive force generally takes the form of economic or political sanctions while interference consists of impeding or impairing the rights of another State. Of the three forces, the predominate view is that U.N. Charter 2(4) applies only to armed force while coercive force and forceful interference do not rise to the level of Article 2(4). The legal parameters governing a State's use of force against another State also apply to a State's use of force on the international level against non-state actors.

A State employing an ASAT to destroy or disable a space object not subject to its jurisdiction reasonably constitutes a use of force prohibited by U.N. Charter 2(4). This means using a kinetic or directed energy weapon or device against another's space object is legal under international law only if it is approved by the U.N. Security Council, is in self-defense or constitutes a justifiable peremptory strike. Obtaining the U.N. Security Council's approval to destroy or disable an offending SmallSat constellation is impractical in light of a State's urgency to eliminate the offending constellation, the time it will take to have the matter heard and the uncertainty regarding whether the Security Council will approve the action. For the response to be deemed a legitimate self-defense action, the constellation would have to constitute an armed attack. An orbiting constellation alone, without any other compelling circumstances such as associated kinetic damage, should not be construed as an armed attack. Again, this renders physical removal as the viable alternative provided the SmallSat constellation is not registered pursuant to Outer Space Treaty VIII and the previously discussed drawbacks are absent.

CONCLUSION

SmallSats possess great potential for the advancement of non-state actors participating in outer space activities. In fact, they have been referred to as "democratizing" or "equalizing' the access to and use of space as they reduce the portal control by the traditional space faring States. With this advancement, however, comes the potential "creative" activity by non-state actors which infringe on the territorial sovereignty of a State for ideological, religious, political or criminal purposes. This poses a challenge to the international community as all States are subject to being confronted with this potential dilemma. Perhaps this circumstance can generate common ground for international collaboration and cooperation among States to reach a modern agreement establishing new parameters regulating Earth orbit activities.

FOOTNOTES

Hyperspectral satellite application in Asia-pacific regional space cooperation

Shendan[1], Gao Yufeng[2]

(1 Shanghai Institute of Spaceflight Control Technology Email: shendan_82@163.com)
Zhongchun Road 1555, Shanghai, China
(2 Shanghai Academy of Spaceflight Technology Email:hammergao@126.com)
Yuanjiang Road 3888, Shanghai, China

Abstract： Hyperspectral remote sensing is used in a wide range of application, originally developed for mining and geology (identifying various minerals which is ideal for mining and oil industry), it has now spread into fields as forest coverage survey, crop monitoring, vicious weather surveillance, etc.. Organizations like NASA have catalogues of various minerals and their spectral signatures and have posted them online to make them readily available for researchers.

Countries of the Asia-pacific region are most agricultural intensified and often suffer from various geological disasters, meanwhile, nonferrous metal resources in some of these countries are abundant, making a great supplementary to their national economy. However, most regional countries are short of critical technology and measures, as well as fiscal support to accomplish the disaster prevention and mitigation work. Consequently, a joint-developed project which can provide proper spectral data and serve for all the regional countries will be the best solution, in this sense a space-borne hyperspectral imaging system may be the best choice, and small satellite platform carrying the imager is the most economical and feasible scheme.

The hyperspectral small satellite is designed to operate in the sun-synchronous orbit , the visible and short wave infrared hyperspectral imager data carried by the hyperspectral small satellite play an important role in geological and mineral survey, ecological environmental monitoring, agricultural conditions monitoring, forestry survey, drought and ice and snow monitoring, ocean environment monitoring as well as disaster prevention and mitigation.

Keywords : hyperspectral imaging, small satellite, regional space cooperation

1. Natural and economic conditions in Asia-pacific region

Asia and the Pacific is the most disaster prone region in the world. A person living in the region is almost twice as likely to be affected by a disaster as a person living in Africa, almost six times as likely compared with Latin America and the Caribbean, and 30 times more likely than a person living in North America or Europe. In 2013 alone, natural disasters in Asia and the Pacific affected more than 57 million people and caused US $128 billion in damages. Rapid economic growth and population expansion over the coming decades, along with the impacts of climate change, will increase the exposure and vulnerability of the region to disasters. As disasters disrupt all sectors of the economy and destroy hard-earned development gains, it is crucial that effective disaster risk reduction measures are integrated into development plans and poverty reduction strategies.

At the same time, technological innovations provide unprecedented opportunities to build resilience and deepen connectivity. Experiences from the region and around the world have proven that disaster prevention and preparedness, enabled by space technologies, can be far more effective

and less costly than ever before. Space technologies have proven highly effective in disaster monitoring, early warning, and emergency response efforts. Despite the great progress in deepening regional connectivity through remote sensing technologies, Asia and the Pacific is still the most digitally divided region in the world, with less than eight per cent of the population connected to affordable and reliable high-speed Internet. As a result, millions of people are shut out from transformative digital opportunities in education, health and financial services.

Scientific and effective disaster monitoring and assessment is not only the basis of scientific disaster management, but also a popular and difficult topic for international scholars. There are varieties natural disasters in China, Bangladesh, Iran, Pakistan and other countries, such as sand storm, landslide, flood, drought, landslides, etc. There are several times of serious tsunami, forest fires and other disasters wreak havoc on Southeast Asian countries every year, such as Thailand and Indonesia. Disaster monitoring and assessment system can be used to warn these disasters in advance, conduct real-time monitoring to disaster situation and support post-disaster assessment.

Disaster monitoring and evaluation functions of small multi-mission satellite constellation system include:

- Image product of data integration, processing and production appreciation for multi-resource remote sensing image;
- Advanced image processing and analysis;
- Thematic map production for disasters;
- Flood monitoring and assessment;
- Drought monitoring and assessment;
- Sandstorm monitoring and assessment;
- Forest fire monitoring and assessment;
- Landslide and mountain collapse monitoring and assessment;
- Debris flow monitoring and assessment;

2. Taking Drought Monitoring from Space as an example

Drought is a creeping disaster and a silent killer. It rolls back development gains and exacerbates poverty, especially in least developed countries. Due to climate change, the frequency, severity and duration of droughts will likely be more serious in the future. Focusing on drought is proactive, not reactive, as taking action once drought has occurred is more costly and less effective than acting preemptively.

The socio-economic and environmental impacts of droughts have increased significantly, particularly among the most vulnerable groups in Asia and the Pacific. Between 1985 and 2013, 110 drought events occurred in the region, affecting 1.2 billion people and costing US$52 billion. The tragic consequences of drought include:

- Loss of human life and livelihoods;
- Reduced water and food security;
- Increase in debt among farmers;
- Deepening poverty with intergenerational consequences;
- Farmer suicides;
- Potential for unrest and violence;
- Land degradation and desertification.

In order to save lives and livelihoods, time is of the essence. Signs of drought can be observed from space long before they are visible to the human eye on the ground. Therefore, space-based data is a vital complement to ground-based information in combating drought. However, there is a lack of resources and capacity to perform such analysis in many drought-prone developing countries. It is critical and urgent for the Asia-Pacific region to place a stronger emphasis on preparedness by applying innovative hyperspectral remote sensing for effective drought monitoring and early warning, through a strengthened regional cooperative mechanism and South-South Cooperation.

3. Merits of hyperspectral remote sensing

The key element for hyperspectral remote sensing is the visible short wave infrared hyperspectral imager, which is a kind of optical remote sensing instrument with a combination of image and spectrum, it can obtain the image information and spectral information of surface features within the broadband spectrum of 400-2500nm with the swath of 60km, spatial resolution of 30m and spectral resolution of 5/10nm, and the number of spectral channels is up to 320.

To realize the technical requirements of high spectral resolution, high spatial resolution, high signal to noise ratio and high calibration accuracy, the visible short wave infrared hyperspectral imager adopts TMA telescope for imaging, and separates the wave band of the visible near infrared and short wave infrared with field separator. It adopts the spectrograph basing on convex grating for fine spectrophotometry light splitting after going through the slit field, and the spectral image signals gather in the CCD plane-array detector and short wave infrared HgCdTe focal-plane detector respectively, realizing the width of 60km and spectral coverage of 400nm-2500nm, with hyperspectral imaging of 320 wave bands in total.

The key component for spectrophotometry is high efficiency convex grating. Basing on the convex grating, the improved Offner structure is adopted for the spectrograph. The CCD scale is a backlighting high frame frequency transfer component with 2200×320 pixels, and the quantum efficiency is as high as 80%.The short wave infrared sensor is a high frame frequency HgCdTe focal-plane device with the scale of 2000×512 pixels.

The detector drive circuit and drive detector operate with set integration time and operation frame frequency, and the image signals of detector are read out through information acquisition circuit for processing of clamping, amplifying and smoothing. Signals are converted to 12bits digital signals through AD module and are transmitted to the data transmission subsystem of the satellite with the arranged format.

The rate of hyperspectral data is as high as 2.02Gbps. The spectrum are real-time

programmable in-orbit , the spectrum channels can be selected to be downloaded arbitrarily in the 320 wave bands of near infrared and short wave infrared, and the data can be chosen to be transmitted as the original data, lossless compression and 1:4 or 1:8 lossy compression data.

Main technical indicators of the visible short wave infrared hyperspectral imager are shown in Table 1.

Table 1 Technical indicators of the visible short wave infrared hyperspectral imager

No.	Project	Technical indicator
1	Spectral range	0.4μm~2.5μm
2	Spatial resolution	30m
3	Swath	60km
4	Spectral resolution	VNIR：5nm；SWIR：10nm
5	Absolute radiometric calibration accuracy	Better than 5%
6	Relative radiometric calibration accuracy	Better than 3%
7	Spectrum calibration accuracy	VNIR：0.5nm；SWIR：1nm
8	Method of Spectrophotometry	Optical grading light splitting +OFFNER
9	Quantization level	12bits

4. Main features of hyperspectral small satellite

Satellite platform selection principles: On the premise of task requirements satisfaction, mature and verified satellite platform shall be given priority to, and main subsystems of the platform shall be mature products with COTs. Requirements for the satellite platform are shown in the following table2.

Table 2 Specification requirements for satellite platform

Project	Indicator requirement
The total mass of the platform	Platform mass: <300kg Load-bearing mass: >250kg
Power consumption	The peak power is no less than 700W
Attitude control mode	zero momentum and three-axis stabilization
Pointing accuracy	Attitude pointing accuracy: ≤0.1°(three-axis, 3σ);
Stability	Attitude stability:≤0.001 ° / s (three-axis, 3σ);
Attitude maneuver ability	Attitude maneuver ability (rolling, Pitching): ±35°

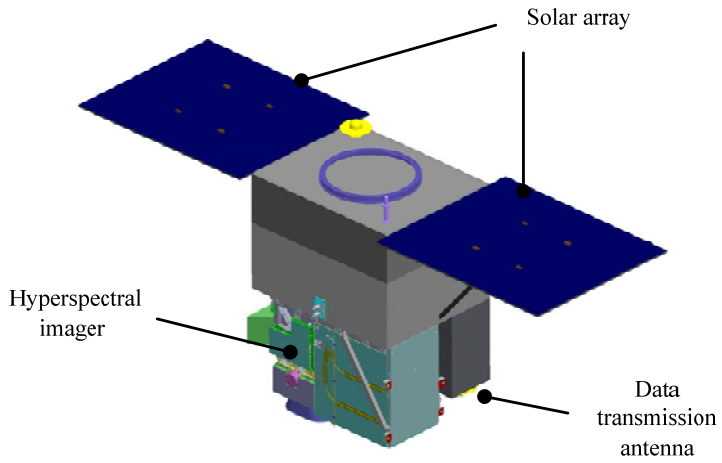


Figure 1 Configuration schematic diagram of hyperspectral small satellite in flight

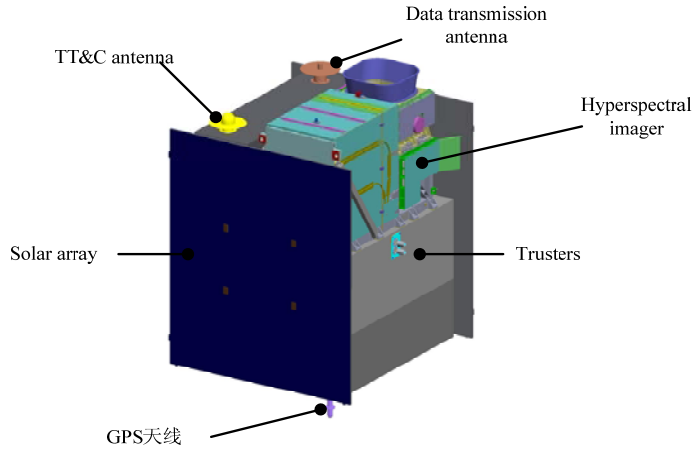


Figure 2 Configuration schematic diagram of the folding status of hyperspectral small satellite

5. Evaluation of hyperspectral small satellite

High resolution, multi-spectral and hyperspectral data is utilized in Asia-pacific small multi-mission satellite constellation. Remote sensing technology and various agriculture technologies are integrated and used for land resource survey, monitoring and analysis of agriculture and forestry, prediction of diseases and insect pests, and yield estimation of crops, etc. Besides, the monitoring and assessment of remote sensing technology on agriculture and forestry is shown in the following aspects, which can enhance the local economy in separate ways:

a) Agriculture and forestry resource utilization monitoring;
b) Vegetation Identification;
c) Crop area and ratio estimation;
d) Crop yield estimating;
e) Land survey data management and data post-processing;
f) Soil moisture and flood monitoring;
g) Drought condition of crops supervising;
h) Forestry planning;
i) Stand volume measurement;
j) Agricultural and forestry pest and disease damage monitoring.

High resolution,multi-spectrum and hyperspectral data of small multi-mission satellite constellation is also used for metallogenic belt comparative studies, mineral resource remote sensing surveys, mineral resource information and other application demonstrations, exploitation and production supervision, management and construction application systems. These information improve efficiency of mineral resource exploration for reducing expenditure and creating benefit.

This system is oriented for realizing the following functions in remote sensing mineral exploration:

a) Feature extraction of remote sensing geology;
b) Abnormal information extraction;
c) Extraction of ore-bearing mineralization geology information;
d) Remote sensing prospecting and modeling.

Conclusion

The application of space technologies especially the hyperspectral remote sensing have contributed significantly towards addressing issues relating to sustainable development. Specifically, disaster risk reduction and disaster management, as well as the environment and development in the region. This recognition covers the importance of comprehensive hazard and risk assessments, knowledge and information sharing, including reliable geo-referenced information and the importance of early warning systems. This final goal of such application is to reduce economic and social damage including the loss of human life.

Energy Management and Design of Vehicle Trajectory based on Optimized Section

Yang Ling-xiao, Min Chang-wan, Xiao Zhen, Zhang Ning-ning

(Science and Technology on Space Physics Laboratory, Beijing 100076, China)

Abstract: A new method of the optimal trajectory in target function is proposed that it is based on energy section for reentry gliding vehicle, and the trajectory with minimum range is solved by this design. The study for minimum range is meaningful for engineering, it can also improve the applicability and agility for vehicle. In addition, the trajectory with minimum range described by this method is more feasible for engineering. The angle-of-attack and the bank angle as control variables are used to regulate energy and range at quasi equilibrium-gliding. Firstly, the angle-of-attack section for minimum lift-to-drag ratio is ensured by the relation of range and lift-to-drag ratio, then the secure boundary for flight corridor is built by restrictions in flight. Finally, by the influencing rule of the D-e section and range, the D-e section is optimized for minimum range in corridor. Moreover, energy-managing is achieved by cooperating lateral motion, and the optimized D-e section is tracked to prove the practicability of programming trajectory with minimum range.

Key words: Optimization; Minimum range; Quasi-equilibrium gliding; Tracking

Introduction

Near-space vehicle is commonly between 20km to 100km from ground for flight. The vehicle traced back to the thirties of the twentieth century, and now the design of this vehicle is evolving and improving ceaselessly. Thereby, more type branch is getting to be expanded and fulfilled. HTV-2 which is manufactured by America is representative, this vehicle is taken as the study object.

The trajectory optimization is the key of study for vehicle, in respect that navigation, control and guide are based on the all-sided design of trajectory. At the same time, the ability of range and maneuverability, and trajectory trait is also provided the direction. The study for trajectory optimization mostly concentrated on the following aspects: the minimum total decalescence, the maximum longitudinal range and lateral range, the maximum full range, the shortest time for flight and so on. The work for the design of trajectory with minimum range is not very familiar, however, these study results are requisite for engineering. Hence, the design of trajectory with minimum range is used to validate this method in paper.

Combining minimum range with maximum range can plot out the maximum attainable area, that can offer arrival extent and attack ability for different vehicles to compare. The practicable scope and flexibility can be intensified by the study of minimum range, that will contribute on the effective mode of energy management for flight and the design of trajectory mode for close quarters.

1 Problem Formulation

In case of the initial height, velocity and flight-path angle of gliding is definite, the energy of flight is influenced by changing the trajectory mode, regulating the attack angle or using the bank angle for lateral motion. Moreover, the form and movement rule of optimal trajectory for the given

objective function should be found.

1.1 Dynamics

The trajectory optimization for near-space gliding vehicle is studied in this paper. The gliding phase has no power in flight, and sideslip angle is 0. Thus the three-degree-of-freedom (3DOF)

dynamics of an entry vehicle over a nonrotating spherical Earth are described by the following dimensionless equations of motion:

$$\begin{cases} \dot{r} = V\sin\gamma \\ \dot{\theta} = \dfrac{V\cos\gamma\sin\psi}{r\cos\phi} \\ \dot{\phi} = \dfrac{V\cos\gamma\cos\psi}{r} \\ \dot{V} = -D - g\sin\gamma \\ \dot{\gamma} = \dfrac{L\cos\sigma}{V} - (\dfrac{g}{V} - \dfrac{V}{r})\cos\gamma \\ \dot{\psi} = \dfrac{L\sin\sigma}{V\cos\gamma} + \dfrac{V}{r}\cos\gamma\sin\psi\tan\phi \end{cases} \tag{51}$$

Where $x = (r, \theta, \phi, V, \gamma, \psi)$ is the state value, r is the radial distance from the center of the Earth to the vehicle, θ is longitude, ϕ is latitude, V is velocity, γ is flight-path angle, ψ is course angle. The terms D and L are the aerodynamic drag and lift accelerations in g.

Because the power for vehicle is all from boosting phase, the flight mission with certain range is achieved by consuming energy. The transformation of kinetic energy and potential energy is described by the combination of height and velocity, using energy parameter e defined to be:

$$e = \frac{1}{r} - \frac{v^2}{2} \tag{52}$$

Where r is normalized by the radius of the Earth ($r = R / R_0$) , and v is also normalized by $v = V / \sqrt{R_0 g_0}$.

1.2 Flight constrains

The flight span for near-space vehicle is large, thus more constrains are considered in flight. The more constrains are required, the less safety area of trajectory is bounded for flight. The security and feasibility in flight are guaranteed by path constrains for flight in the aerosphere with long time; the design of trajectory is ensured by control constrains in area of flight ability; the accuracy of target arriving is achieved by terminal constrains.

Examples of the path constraints in the gliding process are heat rate, normal aerodynamic-load factor, and dynamic pressure, these are normalized as expressed in the following:

$$\begin{cases} \dot{Q}_s = \dfrac{C_1(\rho)^{0.5}}{\sqrt{R_d}} v^{3.15} \le \dot{Q}_{s\max} \\ n_y = L\cos\alpha + D\sin\alpha = \dfrac{\rho v^2 S_t (C_L\cos\alpha + C_D\sin\alpha)}{2} \le n_{y\max} \\ q = \dfrac{1}{2}\rho v^2 \le q_{\max} \\ \dot{\gamma} = (\dfrac{v^2}{r} - g)\cos\gamma + L = 0 \end{cases} \tag{53}$$

Where $\dot{Q}_{s\max}, n_{y\max}, q_{\max}$ are the peak value of the path constraints.

The longitudinal and lateral dynamics are considered in company for the design of trajectory with minimum range, the longitudinal state is controlled by angle-of-attack, and the lateral state is used in determining the bank angle, so the control variable is $u = (\alpha, \sigma)$. The boundary condition of control variable is the restriction for the angular rate and the corresponding α and σ as follow:

$$\begin{cases} \alpha \le \alpha_{\max} \\ |\sigma| \le \sigma_{\max} \\ |\dot{\alpha}| \le \dot{\alpha}_{\max} \\ |\dot{\sigma}| \le \dot{\sigma}_{\max} \end{cases} \tag{54}$$

By way of getting target with accuracy, the terminal conditions for the trajectory, typically in the form of the final value of height, velocity and course angle.

$$\begin{cases} |h_f - h_{ref}| \le \Delta h \\ |V_f - V_{ref}| \le \Delta V \\ \Delta\psi \le \Delta\psi_{ref} \end{cases} \tag{55}$$

2 The Effect Factor for Range

A majority of flight throughout entry flight phase is quasi equilibrium gliding. Quasi equilibrium gliding is that the required flight-path angular rate is almost 0 for lasting gliding at the altitude, which commands the balance for lift, centrifugal force and gravity. Then the key factors will be studied, which affect range at quasi equilibrium gliding, through finding the optimal states for these factor to minish range.

2.1 The relation for range to lift-drag ratio

When the vehicle is at quasi equilibrium gliding (γ=0) ,

$$\begin{cases} \dfrac{dV}{dt} = -D \\ \dfrac{V^2}{r} + L = g \end{cases} \tag{56}$$

From the equation, the range at quasi equilibrium gliding is represented by:

$$S = \frac{k}{2}\ln(\frac{gR - V_k^{\,2}}{gR - V_0^{\,2}})R_0 \tag{57}$$

Where k is lift-drag ratio, V_k is the final velocity, V_0 is the beginning velocity. It is known that k, V_0, V_k are to determine the gliding range, so that the gliding range is only related to k when the beginning and final velocity is definite. Minishing range is dependent on assuming the little lift-drag ratio trajectory. Based on the least lift-drag ratio trajectory, it can be taken the minimum range trajectory at quasi equilibrium gliding. In case of dependence only on longitudinal maneuver, angle-of-attack, height and velocity have influence on the lift-drag ratio.

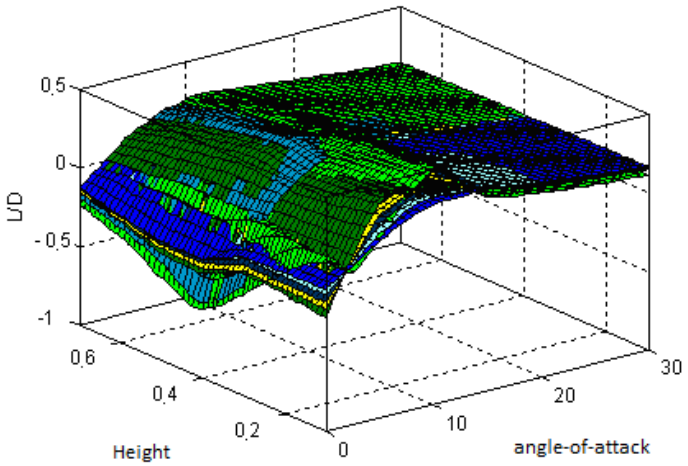


Fig. 2 The lift-drag ratio at different angle-of-attack, height and mach number(normalization)

The scope for angle-of-attack at safety flight ability is restricted more by bound conditions, then these constraints form the viable bound for angle-of-attack in flight that is $6^\circ - 20^\circ$. From the analysis of Figure 1, and the results for different height and mach number in valid bound are compared, it is observed that maximal angle-of-attack (20°) can get the minimum lift-drag ratio and minimum range in diversified height and mach number. Moreover, large angle-of-attack can not only minish range, but also heighten for reducing the dynamic pressure and heat rate at quasi equilibrium gliding, so that it is easier to suffice the constraints in flight.

2.2 The relation for range to energy

The flight state is transformed by lift and drag in gliding. Lift affects on height, and that works on potential energy; vehicle reduces velocity gradually by drag, moreover, it has effect on kinetic energy. So the design of trajectory with less range relies on energy managing in flight.

The expression for r and v is normalized by Eqs.(1) in the following:

$$\begin{cases} \dot{r} = v\sin\gamma \\ \dot{v} = -D - g\sin\gamma \end{cases} \tag{58}$$

By taking the derivative to e from Eq.(2), we have:

$$\frac{de}{dt} = v(-D - g\sin\gamma) + \frac{v\sin\gamma}{r^2} \tag{59}$$

When the vehicle is at quasi equilibrium gliding, $\gamma \approx 0, \sin\gamma \approx 0$ .According to (9), we have:

$$\dot{e} = Dv \tag{60}$$

By taking the derivative to range, we have:

$$\dot{s} = v\cos\gamma \tag{61}$$

Under the quasi equilibrium gliding, we can approximate γ at 0. So we have $\cos\gamma = 1$ and $\frac{ds}{de} = \frac{1}{D}$. The integral expression of range is as followed:

$$s \approx \int_{e_0}^{e_f} \frac{1}{D(e)} de \tag{62}$$

According to (12), we find that gliding phase is absolutely controlled by energy managing. On the assumption for the initial and terminal energy in gliding is definite, range is expressed by the D-e profile in gliding phase. According to the same energy, the higher the drag acceleration is, the faster the height and velocity of vehicle reduce, the shorter the range is. To some degree, Eq.(12) can estimate range by drag acceleration to energy.

3 Program to The Optimal trajectory

3.1 Design of Safe flight corridor

From (12), we find that range is determined by D-e profile. The Optimization of the trajectory with the specific mission is translated into the optimization of D-c profile under the corresponding objective function. So we need to mark out the feasible flight area for the D-e profile according to the constraint model.

According to the Eq.(2) for energy parameter and the expressions for drag acceleration, we have :

$$\begin{aligned} D &= \frac{\rho v^2 S_t C_D}{2m} \\ v &\approx \sqrt{2\left(1/\overline{r} - e\right)} \end{aligned} \tag{63}$$

Taking the Eq.(13) to the Eq.(3) for constraint models, we get the feasible safe D-c profile under the constraints:

$$\begin{cases} D_{\dot{Q}} = \dfrac{C_D S_t \dot{Q}_{s\max}{}^2 R_d}{2mk_Q^2 C_1} \cdot \dfrac{1}{(\dfrac{2}{\overline{r}} - 2e)^{2.15}} \\ D_{ny} = \dfrac{n_{y\max} C_D}{(C_L \cos\alpha + C_D \sin\alpha)} \\ D_q = q_{\max} \dfrac{S_t}{mg_0} C_D \\ D_{eg} = \dfrac{C_D}{C_L}\left(-\dfrac{1}{\overline{r}^2} + \dfrac{2e}{\overline{r}} \right) \end{cases} \tag{64}$$

When $D_{\dot{Q}_s}$ is the limit of constrained heat rate, D_{ny} is the limit of normal aerodynamic-load factor, D_q is the limit of dynamic pressure, D_{eg} is the boundary of equilibrium gliding. From Eq.(14) we can find that the boundary limit of the drag acceleration is related to not only the energy and the peak value of constraint conditions, also the C_D、C_L and α. Moreover, C_D、C_L is mainly affected by α. When the profile of attack angle is determined, we can calculate the flight boundary for vehicle gliding under the constraints from Eq.(14). This boundary contains the area which is the safety corridor of the complete gliding flight. The profile of trajectory which belongs to this area under the constraints can ensure the realization of flight .By (12),we can find that boundary determined by(14)can influence the range extremum of the feasible trajectory.

3.2 Optimization of the drag acceleration-energy flight profile

When the attack angle α and the flight corridor were confirmed, we can optimize the D-e profile for the arbitrary objective function in the flight corridor. The trajectory calculated by this profile is the optimal trajectory controlled by the objective function.

In order to optimize the D-e profile, we need to transform kinematics equations into the model defined by energy. Eq.(1) is built by the independent variable of time, we can get the kinematics model that is the function of energy by Eq.(10) for the relation of energy and time.

$$\begin{cases} \dfrac{dr}{de} = -\sin\gamma(\dfrac{1}{D}) \\ \dfrac{d\theta}{de} = -\dfrac{\cos\gamma\cos\psi}{r\cos\phi}(\dfrac{1}{D}) \\ \dfrac{d\phi}{de} = -\dfrac{\cos\gamma\sin\psi}{r}(\dfrac{1}{D}) \\ \dfrac{d\gamma}{de} = (g - \dfrac{V^2}{r})\dfrac{\cos\gamma}{V^2}(\dfrac{1}{D}) - \dfrac{1}{V^2}(\dfrac{L}{D}\cos\sigma) \\ \dfrac{d\psi}{de} = \left[\dfrac{\cos\psi\tan\phi}{r}(\dfrac{1}{D}) - \dfrac{L\sin\sigma}{V^2}(\dfrac{1}{D}) \right]\cos\gamma \end{cases} \tag{65}$$

If the initial state (v_0, r_0) and the terminal state (v_f, r_f) of the gliding flight are known, we can calculate the energy e_0 and e_f. According to Eq.(2), we can calculate the energy parameter

e_i of arbitrary state(v_i, r_i). We can get the maximum value $D_{\max}(e_i)$ and the minimum value $D_{\min}(e_i)$ of the drag acceleration corresponding to e_i by the boundary of flight corridor from Section 3.1 .The energy region $[e_0 \quad e_f]$ is divided to $n+1$ sections by n points $\{e_1 \quad e_2 \quad \dots \quad e_n\}$ .If $e_1 = e_0, e_n = e_f$, according to $x = \{D_1 \quad D_2 \quad \dots \quad D_n\}$ that is the optimization variable, the constraint condition is $D_{\min}(e_i) \le D_i \le D_{\max}(e_i)$, $i = 1,\dots,n$, then the designed trajectory can always be in the flight corridor by satisfying constraints. Meanwhile, the initial and terminal states and other constraints in flight are contented.

If the design of trajectory is requested to achieve optimization for certain performance, range constant $s = s_f$ will be took as one of the constraints and the objective function should reach the extremum under the constraints. When we want to get the trajectory that has minimum(maximum) terminal velocity, the target that need to be optimized is

$$\min J = \pm V(t_f) \tag{66}$$

The trajectory which make vehicle absorb the least heat during the reentry flight is as follow:

$$\min J = \int_{t_0}^{t_f} \dot{Q} dt \tag{67}$$

When programming the trajectory that range is the optimized target, range is took as the objective function. The objective function of the minimum(maximum) range is as follow:

$$\min J = \pm S \tag{68}$$

Using the pseudo-spectral method to optimize the D-e profile, we can solve the optimal problem by dispersing it and transforming into the nonlinear programming problem. This method has the same convergence rate as exponential function and a higher precision for results, moreover, it is not sensitive to the initial value.

3.3 Tracking of the programming profile

After getting the $D_{ref}(e)$, the bank angle σ is calculated by designing the tracking control rule. In actual trajectory integral, the actual drag acceleration $D(e)$ should be controlled to track $D_{ref}(e)$. The tracking result can prove if the programming optimum trajectory can be achieved.

Dynamics equation by normalization that belongs to aircraft vertical plane is as followed

$$\begin{cases} \dot{r} = v\sin\gamma \\ \dot{v} = -D - \dfrac{\sin\gamma}{r^2} \\ \dot{\gamma} = (v^2 - \dfrac{1}{r})\dfrac{\cos\gamma}{vr} + \dfrac{D}{v}u \\ \dot{s} = v\cos\gamma \end{cases} \tag{69}$$

Where $x = \{r \quad v \quad \gamma\}$ is the state variable of system, $y = D(x)$ is the output variable of system, $u = \frac{C_L}{C_D}\cos\sigma$ is the control variable and $D(x) = D_{ref}(x)$ is the controlled target. , $y = D(x)$, $u = \frac{C_L}{C_D}\cos\sigma$ and $D(x) = D_{ref}(x)$.

According to the theory of feedback linearization, we need take the derivative of the system output $D(x)$. Using the above data and the formula containing drag acceleration $D = \rho S_t C_D v^2 / 2m$ and the formula that estimates atmosphere density $\rho = 1.752e^{-R_0(r\text{-}1)/hs}$ (atmosphere constant is $hs = 6700\text{m}$), we can get the relationship which is the second order differential between input and output D by linearization processing:

$$\ddot{D} = a_D(x) + b_D(x)u = \upsilon \tag{70}$$

Above-mentioned Eq.(20) is the result of the dynamic equations with feedback linearization. So we can devise the linear control rule to actualize the tracking of the prospective output $D_{ref}(x)$.

$$e_D = D - D_{ref} \tag{71}$$

Where D_{ref} is the referenced drag acceleration by optimizing, in other words, that is the perfect output which needs to be tracked. The designed new input υ is described as

$$\upsilon = \ddot{D}_{ref} - 2\zeta\omega_n\dot{e}_D - \omega_n^2 e_D \tag{72}$$

i.e.

$$\ddot{e}_D + 2\zeta\omega_n\dot{e}_D + \omega_n^2 e_D = 0 \tag{73}$$

Where ζ is the damping coefficient of system, ω_n is the natural frequency. Through choosing suitable ω_n and ζ, we can achieve tracking of the actual output $D(x)$ to the expectant output $D_{ref}(x)$.

Finally, the control variable u is expressed as follow:

$$u = -\frac{1}{b_D(x)}[a_D(x) - \ddot{D}_{ref} + 2\zeta\omega_n\dot{e}_D + \omega_n^2 e_D] \tag{74}$$

Due to $\sigma = \arccos(\frac{u}{L/D})$, the absolute value of bank angle σ can be calculated.

4 Validation by Simulations

The research for the optimal trajectory with minimum range is unfamiliar now, but it is very significative for project. Besides, it has contribution for the design of trajectory which can manage

energy and range, and how to manage. Therefore, the trajectory with minimum range is simulated to validate the method in this paper.

It is analyzed that the 20° angle-of-attack is satisfied with constraints better, and that can get the minimum range among the viable angle-of-attack from Section 3.1 and Fig.1. So we have identified that $\alpha = 20^\circ$ with minimum lift-drag ratio, the $D - e$ profile for flight corridor is established from Eq.(15) with angle-of-attack section.

From the initial and terminal velocity and height in gliding phase, which were normalized, it is easily known that energy parameters are $e_0 = 0.6641, e_f = 0.9633$. Depending on the requirement of the vehicle parameter, it is ascertained that the maximum path constraints are $\dot{Q}_{s\max} = 900 k\mathrm{W/m}^2, n_{y\max} = 3, q_{\max} = 50000\mathrm{Pa}$. Combining with bound at quasi equilibrium gliding, the gliding corridor can be described by the bound conditions and Eq.(15), as Fig.3.

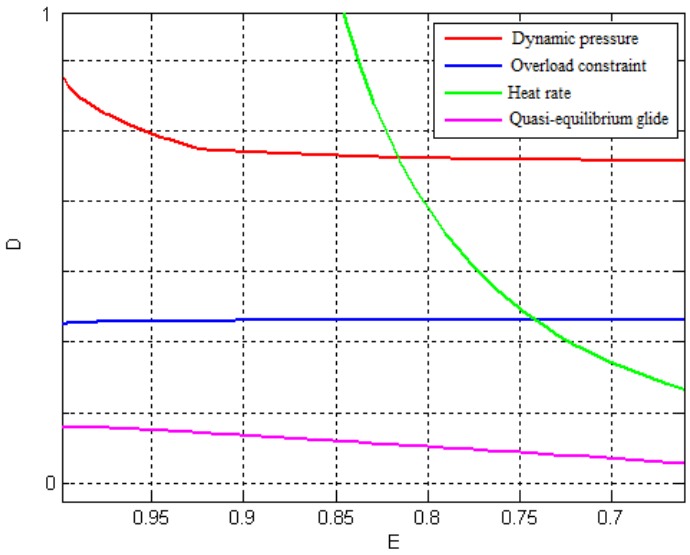


Fig.2 The D-e profile for flight corridor in 20° angle-of-attack (normalization)

Using the pseudo-spectral method to optimize the $x = \{D_1 \quad D_2 \quad \ldots \quad D_n\}$ inside the flight corridor boundary, the optimized D-e profile can achieve the trajectory with minimum range under the constraints. From Fig 4, we find that the optimized D-e profile is close to the upper boundary of flight corridor. The reason of that is when according to the same energy, the higher the drag acceleration is, the faster the height and velocity of vehicle reduce, and there will be the flight with shorter range.

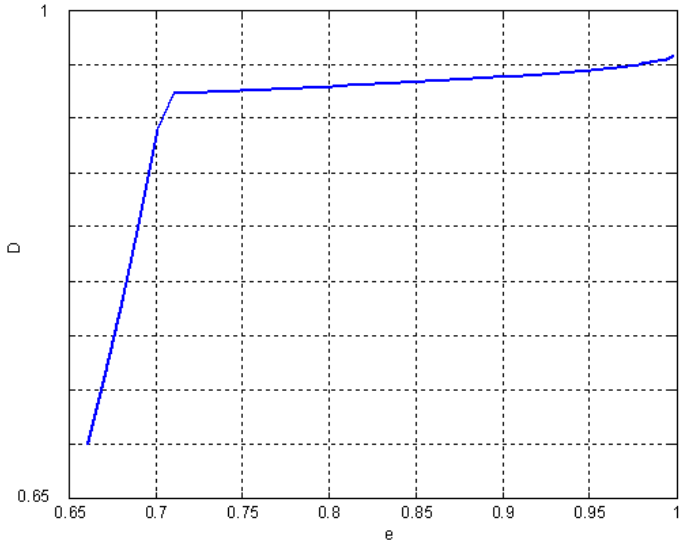


Fig.3 The optimal D-e profile for trajectory with minimum range (normalization)

In order to prove the trajectory with minimum range can be achieved, we track and take energy management of the optimized D-e profile with minimum range by lateral maneuver. The bank angle is made as the control quantity, the drag acceleration is took as the output quantity and $D(x) = D_{ref}(x)$ is deemed to be the controlled target. Fig 5 shows the tracking result for the trajectory with minimum range. The requirement that the constraints of heat rate, overload and dynamic pressure are guaranteed. According to the test results, we can get the minimum range during the gliding is 1900km.

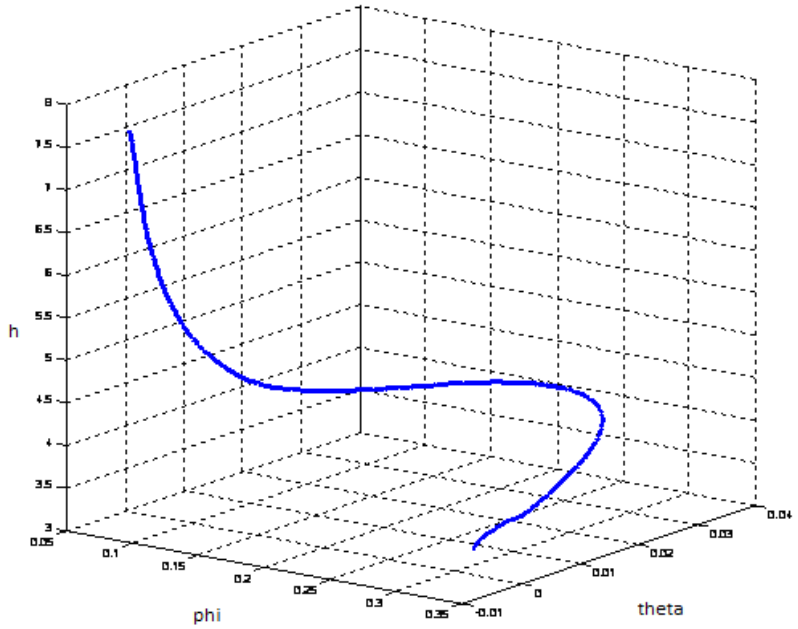


Fig.4 The 3D trajectory with minimum range (normalization)

The trajectory with minimum range can be tracked by lateral maneuver, that proves the trajectory designed in this paper can be better achieved. Through the comparation of the terminal optimal trajectory and the H-V profile of path constraints, the trait of the optimal trajectory with minimum range is that vehicle, which controlled by angle-of-attack and the bank angle, will reach

the lower boundary in H-V profile of constraints as soon as it can. Then the trajectory that is close to the lower boundary will achieve minimum range in the constraints.

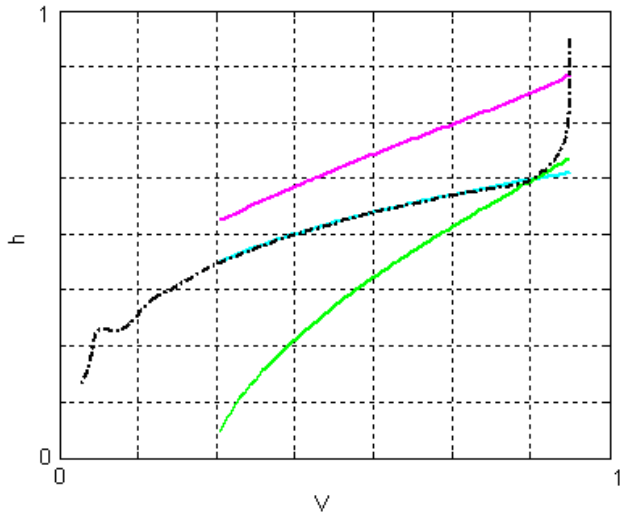


Fig.5 The H-V profile for the trajectory with minimum range and the bound conditions (normalization)

5 Conclusion

The goal of this work is to develop a method for the optimal trajectory at the specific target function. Then the feasibility of the method is validated by the simulation of trajectory for entry vehicle with minimum range. On the assumption that the initial and terminal energy is assured, the feasible verge for flight corridor is designed by bound conditions, and the $D-e$ profile for trajectory with minimum range is optimized in corridor, then the trajectory is tracked. Through the study, it has the following conclusions:

1) Range is determined by lift-drag ratio at quasi equilibrium gliding, and drag acceleration to energy profile which is designed can manage and regulate the range.

2) Minishing range can be actualized by the angle-of-attack at less lift-drag ratio, and also consuming energy through the lateral motion by controlling the bank angle.

3) The designed profile of trajectory with minimum range is influenced by the bound conditions.

4) The optimal trajectory with minimum range in constraints is close to the boundary for heat rate and overload.

5) The tracking results can prove that the method of trajectory optimizing in this paper has well workability and attainability.

The study for the optimal trajectory for entry gliding vehicle with minimum range also has contribution on the design of trajectory pattern with flight in close quarters, and the modes of managing energy.

Simulation Analysis and Experimental Study joint payload connection

CUI Shenshan ZHANG Tao CUI Zhandong WU Di XU Jian

(China Academy of Launch Vehicle Technology, Beijing 100076, China)

Abstract: In the space launch process, all kinds of payloads through connectors connected with the carrier, connectors under enormous concentrated load, is a key component determines the success or failure of the task. In this paper, project type application requirements, design of a payload connection joint structure, through the use of composite materials with variable thickness laminate design, will focus on the joint force evenly spread throughout the main load-bearing structure, at the same time it gives the main structural dimensions and layer plywood overlay information. Then MSC.Patran based on the connectors were meshing, define composite properties, complete finite element model creation. MSC.Nastran uses linear solver compression connectors under load static strength analysis and calculation, the carrying capacity of the connection to obtain displacement connector and strain distribution, confirmed the failure mode of the structure, estimated the connection joints. Finally, the connectors were tested to verify the structure, completed the destruction test under compressive loads to determine the carrying capacity of the structure. Through simulation analysis and experimental verification combination of connectors structures were static strength assessment, verify the weak link structure for the subsequent design and improvement provide a reference.

Keywords: payload, connectors, simulation analysis, experimental research

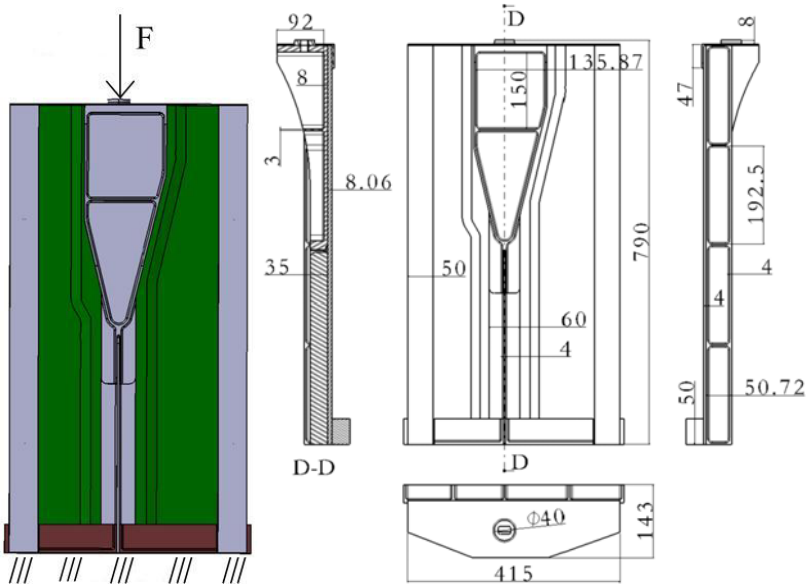
Introduction

Application of advanced composite materials, lightweight aircraft structures, Integration and high performance plays a vital role[1], which has many excellent characteristics specific strength and specific stiffness, performance can be design, can significantly lose weight and significantly improve the structure of anti-fatigue properties. To further enhance the material properties, the need for composite laminate thickness of the partition optimize the design, but the thickness of the mutation (stiffness mutation) causes stress concentration, prone to delamination, causing the structure ahead of destruction[2], so the design variable thickness of the laminated board should be in thickness changes ply drop zones or increasing transition designed to avoid stiffness mutation.

Structure should be able to withstand extreme loads and stresses comprehensive role in the anticipated operating environment without causing failure or breakage[3]. This paper presents a payload connection joint structure, its use of composite materials with variable thickness laminate will focus payload spread. The use of common business software MSC.Patran / Nastran finite element model is completed and analyzed to create calculated to obtain the connection displacement connector and strain distribution patterns to predict the structural damage. Then the structure of the connectors static strength test verification, complete destruction test under compressive loads to determine the carrying capacity of the structure for the subsequent design and improvement provide a reference.

1 Connectors for Test Piece Design

Connectors for test piece design is shown in this article 1 (a), the bottom of the test piece support clamped, concentrated load is applied at the top boss. Wall of composite structures, remaining structural components made of aluminum alloy 2A14 manufacturing. Composite wall thickness of the reinforcing region 8.06mm, the thickness of the general area 4.18mm; longitudinal stringers stud thickness 4mm, the edge strip width 60mm, high 35mm; loading surface of the boss having a diameter of 40mm circular, as shown in Figure 1 (b). Composite panels mainly high-performance unidirectional tape, to improve damage tolerance and maintaining a continuous smooth outer surface layer, the surface of the test piece laying a ± 45 ° woven, test piece ply sequence shown in Table 1.



（a） （b）

Fig.1 Connector connection schematic

Tab.1 Details ply composite parts information

Location	Overlay direction	Number of plies	Layer thickness/ mm	Total thickness/ mm
the general area	[(+45)/(0)/+45/0/-45/90/45/0/-45/+45/0/-45/0]s	26	0.22/0.15	4.18
the reinforcing region	[(+45)/(0)/(+45)/(0)/+45/-45/0/+45/-45/0/90/0/+45/-45/0/+45/-45/0/+45/-45/0/+45/-45/90/0]s	50	0.22/0.15	8.06

2 Material

The composite of the test piece used for the carbon fiber unidirectional tape and woven, metallic material 2A14, the mechanical properties of the material parameters shown in Table 2 and Table 3 below.

Tab.2 Composites Mechanical Parameters

item	Unidirectional tape	Woven
0°Tensile modulus /GPa	164	73.6
0°Tensile Strength /MPa	2200	983
Main Poisson's ratio	0.36	0.06
90°Tensile modulus /Gpa	9.05	71.6
90°Tensile Strength /MPa	66.3	930
Poisson ratio	0.03	0.06
0°Compressive strength /MPa	1160	590
0°Compression modulus /GPa	169	79.8
90°Compressive strength /MPa	218	543
90°Compression modulus /GPa	10.9	81
0°Bending modulus /GPa	130	61
0°Bending strength /MPa	1520	993
90°Bending strength /MPa	105	922
90°Bending modulus /GPa	8.28	55.8
shear modulus /GPa	5.69	4.64
Shear strength/MPa	101	123
0°Interlaminar shear strength /MPa	98.6	70.4

Tab.3 Mechanical properties of metallic materials parameters

Material Name	E (GPa)	μ	Density （kg/m3）	Yield Strength （MPa）	Ultimate strength （MPa）
2A14	70	0.3	2800	375	440

3 Meshing

As used herein, MSC.Patran connectors with MSC.Nastran for numerical simulation. Because of the various components of the test piece are shell structure, so as to Quad4 based mesh, mesh side length of about 5mm, including a total of 30,930 nodes, 29,756 units. After you create a grid in need of direction check, select Verify in Meshing window within the Action item, select Element Method item inside, select Normals within Test items, click Apply to confirm the unit method meets the design requirements, finite element mesh as shown Fig 2.

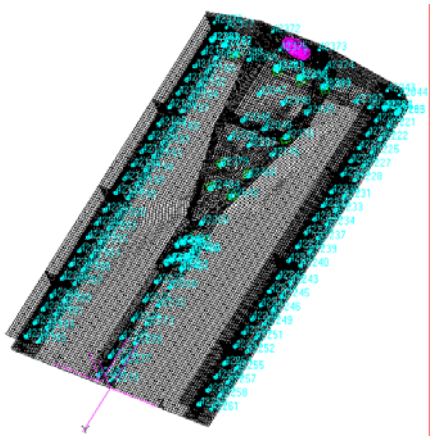


Fig.2 Finite element mesh

4 Modeling Composite Laminates

Plywood complete composite layer is created, we first need to define the single-layer composite material properties. Choose 2d Orthotropic Material on the user interface, fill material information were created and named unidirectional tape and woven property, complete the single layer of material creation. Second, the definition of laminate plies, named laminates in Composite-Laminate interface. In the laminate definition window, select the laminate type, and then click the name of the shop fill material layer, layer thickness and ply orientation, complete laminate created. Finally, given shell element properties, the Properties window in the cell, such as choosing 2D-Shell, name the attribute, select Thin, Laminate, Standard Formulation in Options. In the Properties window, select the definition of laminate have been created, specify the reference coordinate system, complete attribute definition. Variable thickness of the laminated composite panel wall thickness distribution shown in Figure 3.

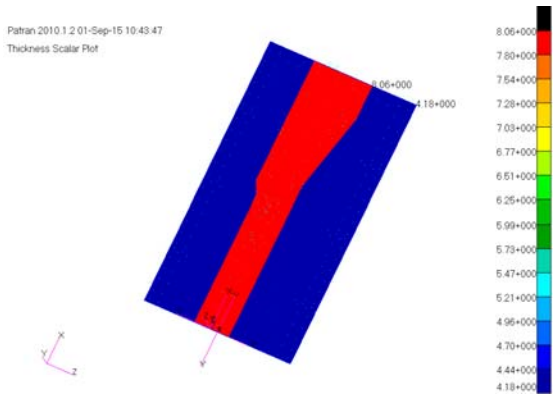


Fig.3 Thickness distribution

5 Failure Criteria

Isotropic material different from that for the anisotropic composite, the strength in different directions is completely different. As used herein, Tsai-Wu failure criterion for composites forecast, which was expressed as follows [4]:

$$F\,I = F_1\sigma_x + F_2\sigma_y + F_{11}\sigma_x^2 + F_{22}\sigma_y^2 + 2F_{12}\sigma_x\sigma_y + F_{66}\tau_{xy}^2 \qquad (1)$$

among them:

$$F_1 = \frac{1}{TX} - \frac{1}{CX}; F_2 = \frac{1}{TY} - \frac{1}{CX}; F_{11} = \frac{1}{TX \times CX}; F_{22} = \frac{1}{TY \times CY};$$

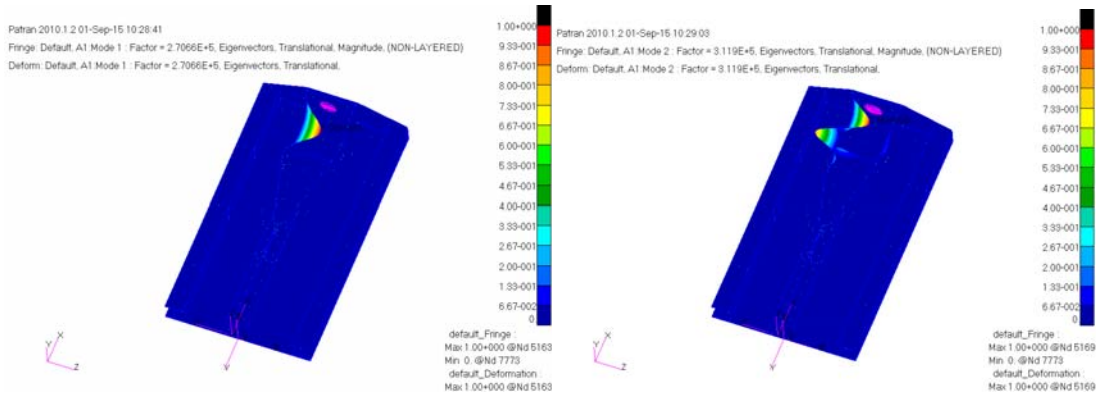
$$F_{12} = \frac{IXY}{\sqrt{TX \times CX \times TY \times CY}}; F_{66} = \frac{1}{SXY^2}$$

When $F\,I \geq 1$, composites failure occurs.

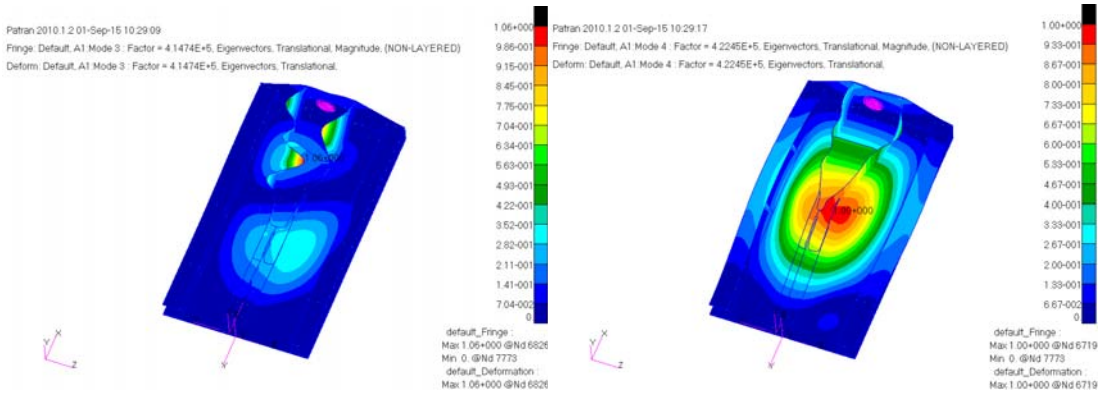
6 Simulation Results

Using MSC.Nastran finite element analysis software Buckling solver (SOL105) linear buckling analysis, when applied to the total load opening panel structure 1N of the first order

buckling factor 277660N, buckling load F = 1 × 277660N = 277.66kN. Buckling located at stud connectors, shown in Figure 4.



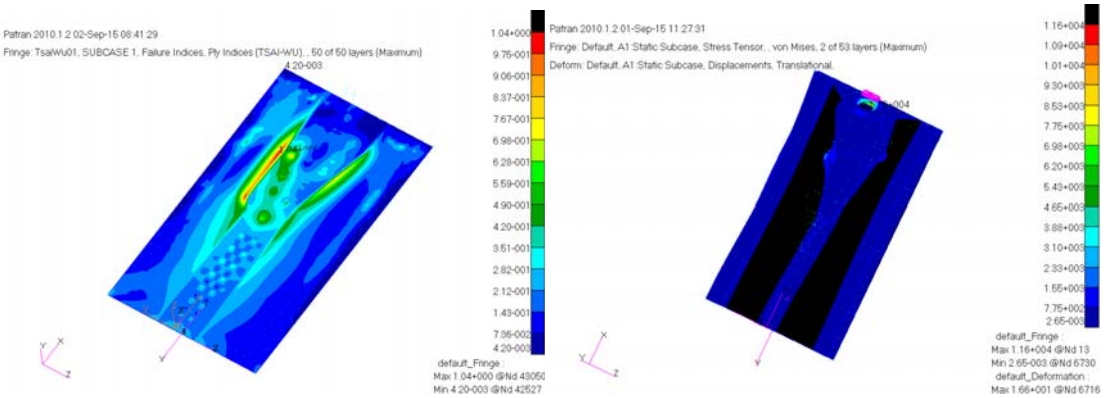
（a）The first order buckling　　（b）Second Order buckling



（c）Third-order buckling　　（d）Fourth Order buckling

Fig.4 Buckling schematic

In the first order buckling load of the structure load analysis, Tsai-Wu criterion maximum coefficient of 1.04, as shown in Figure 5 (a), the coefficient is very close to 1, the composite member substantially in a critical state of destruction. At this aluminum structure of the maximum stress 310MPa, FIG. 5 (b), the aluminum alloy has not yet reached the yield stress. Past experience shows, Tsai-Wu criterion is generally more conservative, and the actual buckling failure load often slightly lower than the calculated value, and the test piece wall thickness variation at the use of the transition layer decreases design, will help ease the stress concentration. Comprehensive analysis, test pieces will be the first buckling damage, as shown in this case maximum displacement 16.6mm, as shown in Figure 6.



（a）　　（b）

Fig.5 Structural response under buckling loads

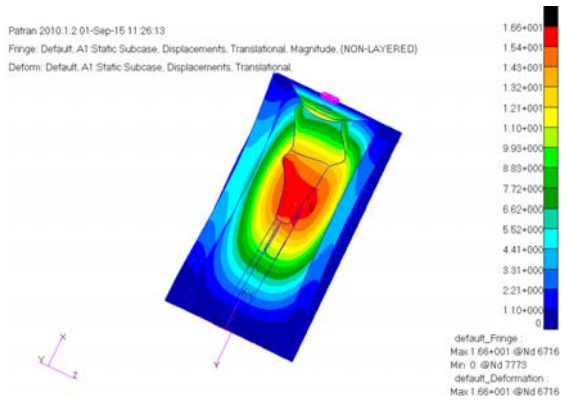


Fig.6 Structure buckling under the load displacement contours

7 Experimental Verification

First, pre-test to check the entire test system is in good condition, pre-test load to 50kN. After uninstalling official test conducted formal tests until the test piece is loaded progressively destroyed, the test pieces as shown in Figure 7.

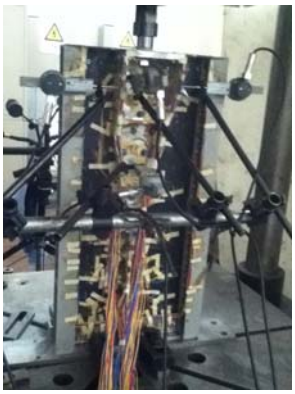


Fig.7 Test pieces

Test pieces of strain measurement point arrangement shown in Figure 8, the test pieces in the front, the back stick, studs, etc. Total Paste 30 strain gauges. Test pieces of displacement measurement points are arranged as shown in Figure 9.

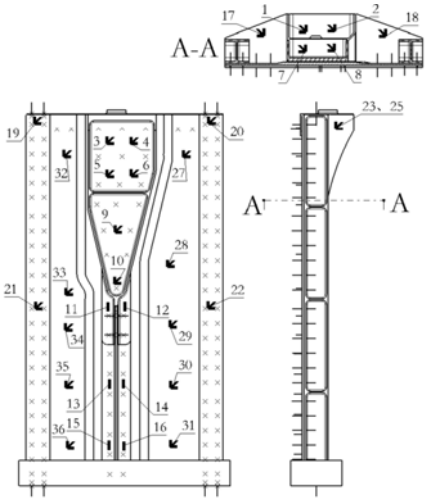


Fig.8 Strain measurement point arrangement

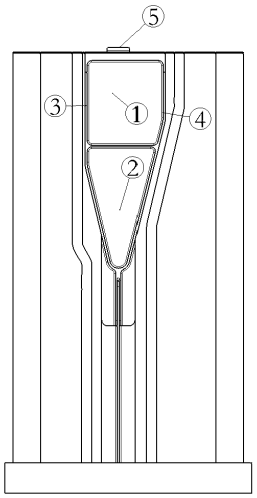


Fig.9 Displacement measurement point arrangement

8 Test results

189.25kN loaded when a huge sound, connectors stud buckling curve shown in Figure 10, the final destruction of the test piece occurs, a sharp decline in the carrying capacity, the end of the test.

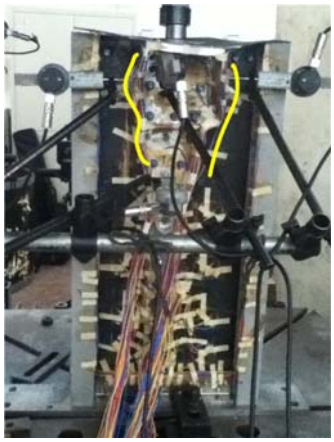


Fig.10 Test pieces undermine state

Connectors test piece Load-displacement curves shown in Figure 11, where the abscissa is loaded load value, a load of said negative compressive load, the ordinate for the displacement measurement point, said positive and negative displacement measuring direction.

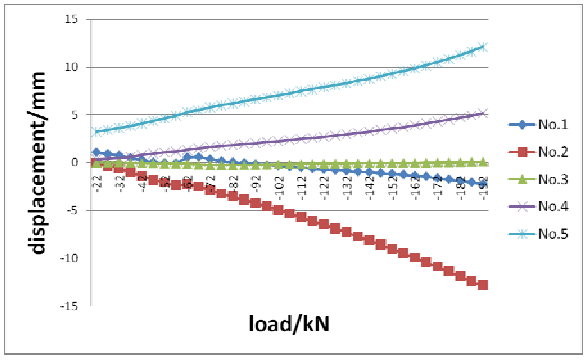


Fig.11 Connectors test pieces load-displacement curve

After the composite panel maximum measured strain 3665με, lower than its failure strain, and experimental observation Composite Panels no damage, indicating that the composite material with variable thickness wall thickness variations reasonable transition area, meet the design requirements.

9 In Conclusion

(1) This connection joint structure for the payload, using MSC.Patran / Nastran to create a static strength finite element analysis model, to connectors analysis provides a feasible engineering simulation analysis method;

(2) The application of the finite element model analysis, the connecting joint under compression load displacement and strain distribution, and accurately predict the main failure mode of buckling failure;

(3) In this paper, ply drop approach to change the thickness of the composite laminate stiffness mutation problem, test results show that the method is reasonably practicable, be able to meet the design requirements.

reference

[1] li Zhengneng. Learn aircraft structure [M],Bei Jing：Beijing University of Aeronautics and Astronautics Press，2010，P382.

[2] Yang Naibin,Zhang Yining. Composite aircraft design [M], Bei Jing：Aviation Industry Press，2002，P106.

[3] Peng Xiaobo. Design and Implementation of US Space Shuttle [M], Bei Jing：China Aerospace Press，2015，P130.

[4] Yang Jian,Zhang Pu,Chen Huohong. New MD Nastran Finite Element tutorial examples [M], Bei Jing：Machinery Industry Press，2008，P224.

Mass Optimization Design of Separation Devices for Pint-size Load Based on Multi-Island Genetic Algorithm

Zhao Shanshan, Zhang Wen, Zhong Jiehua, Li Jing

(China Academy of Launch Vehicle Technology, Beijing, 100076)

Abstract: The pint-size loads, such as small satellite, have property of high function density and fast new technology application. However, when applied in pint-size loads, traditional separation technology has such problems like huge separation impact, low reliability, large weight, etc. Traditional design and analysis methods of separation devices are also not fully suitable for pint-size loads. In this paper, model of separation devices mass optimization is established based on multi-island genetic algorithm. The operation results indicate that the mass of separation mechanism is decreased by 30.7%, which shows that the multi-island fully optimizes the performance of separation devices. The proposed optimization approach obtains a better application.

Key Words： Pint-size Load; Separation Device; Optimization Design; Multi-Island Genetic Algorithm

Tether Deployment Control Scheme Analysis and Design for Electrodynamic Tethered Upper-Stage Deorbit System

Zhang Feng, Wang Shu-ting, Wang Xiao-ding, Wu Sheng-bao, Shen Lin

(R&D Center, China Academy of Launch Vehicle Technology, Beijing, 100076)

Electrodynamic tethered deorbit technology is a novel and effective way for the removal of abandoned upper-stages or satellites. Tether deployment stage is of great importance for this technology. This paper handles the tether deployment control problem of the electrodynamic tethered system. To do so, based on "bead" philosophy, a rigid-flexible coupling dynamics is firstly formulated for the tether deployment stage. Further, to investigate the impact of the upper-stage attitude on tether deployment stability, requirement analysis for the initial attitude of the upper-stage is undertaken by numerical simulations, which indicates that, besides the tether libration control, the attitude control is necessary for the upper-stage as well. As a consequence, a coupled control scheme is then proposed: 1) for initial eject phase, attitude motion of the upper-stage is controlled regardless of tether libration; 2) and for sequential deployment phase, not only is attitude motion of upper-stage stabilized, but also is tether libration controlled by using Kissel algorithm. The following numerical simulation shows the effectiveness of the proposed control scheme.

Keywords: Electrodynamic tethered deorbit system; Tether deployment; Control scheme design

1 Introduction

Due to the long-term manned space activities, the increasing space debris greatly threatens not only the on-orbit spacecraft but also the limit space resource[1]-[6]. To remove space debris like abandoned satellites, upper-stages, recent year has witness many ways including propulsion based deorbit method and space manipulator based capturing[1], but these ways require high cost and fuel consumption.

The electrodynamic tethered deorbit technology is a novel way to remove abandoned spacecraft like upper-stages or unusable satellites. To explain this method, once a spacecraft fails, or has completed its mission and no longer required, a conductive tether with a small weight will be deployed. At both ends of the tether, a means of providing electrical contact with the ambient plasma will be provided to enable current to be transmitted to and from the ionospheric plasma[2]-[6]. The electrodynamic interaction of the conductive tether moving at orbital speeds across the Earth's magnetic field will induce current flow along the tether and meanwhile generate a drag force (the Lorentz force) on the spacecraft to deorbit far more rapidly than it would due to atmospheric drag alone. Compared with other deorbit ways, the electrodynamic tethered deorbit technology has the advantage of low mass, compact size, fuel efficiency (little or no propellant required) and ease of use[5].

Tether deployment is a key stage for the whole deorbit process and contains two sub-stages: 1) eject phase, which is to separate the payload from the upper-stage; and 2) deployment phase, which is to deploy the whole tether after the eject phase.

This paper handles the tether deployment control problem of the electrodynamic tethered system. To do so, based on "bead" philosophy, a rigid-flexible coupling dynamics is formulated for

the tether deployment stage. Further, the requirement analysis of the initial attitude of the upper-stage is undertaken, which indicates that the attitude control of the upper-stage is necessary is necessary besides the tether libration control. Thus, a coupled control scheme is proposed: 1) for initial eject phase, the attitude of the upper-stage is controlled regardless of the tether libration; and 2) for sequential deployment, not only is attitude motion of upper-stage stabilized, but also is tether libration controlled by using Kissel algorithm. The following numerical simulation shows the effectiveness of the proposed control scheme.

2 Dynamic modeling for tether deployment

As mentioned in the above section, a payload M is separated and ejected from the upper-stage R to deploy the whole tether. As shown in Fig. 1, two coordinate frames are considered. The inertial frame O_1-XYZ is with the origin O at the Earth's center. The X axis of the frame directs to the point of vernal equinox, the Z axis aligns with the Earth's rotational axis, and the Y axis completes a right-hand coordinate system. The orbital frame O_2-xyz is with the origin O_2 at the center mass of the system, the x axis directs from the Earth's center to the spacecraft, the y axis aligns with the normal of the orbital plane, and the z axis completes a right-hand coordinate system.

The upper-stage is considered as a cylinder with length l and the moment of inertia J. The payload is regarded as a mass point and linked to the bottom center of the upper-stage via a viscoelastic tether with stiffness EA, density μ_L, and nominal length L.

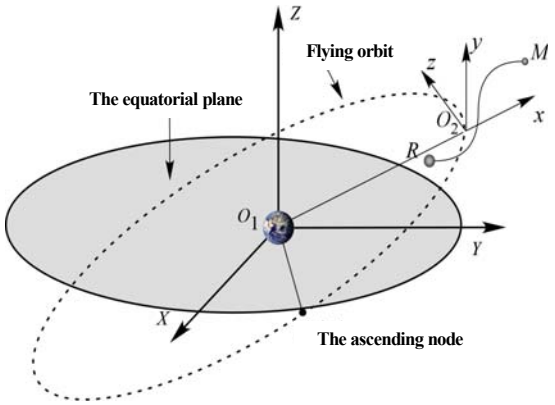


Fig.1 The sketch of the electrodynamic tethered deorbit system

The deorbit dynamics is composed of two parts: the tether deployment dynamics and the attitude dynamics of the upper-stage.

(1) The tether deployment dynamics

To ensure high simulation accuracy, the "bead"-philosophy is utilized to describe the tether dynamics, where the tether is considered as a sequence of lumped masses connected by viscoelastic dampers. For simplicity, denote the upper-stage, and the payload as the node 0 and the node n+1, respectively. Meanwhile, the sequence of lumped masses is represented as the node 1, the node 2,…, the node n.

Fig.2 Discretized tether

Based on Newton's law, the translational dynamics of the upper-stage and the payload can be described by

$$m_{R(M)}\ddot{\boldsymbol{r}}_{R(M)} = \boldsymbol{F}_{R(M)} + \boldsymbol{T}_{R(M)} + \boldsymbol{P}_{R(M)} \tag{1}$$

where $r_{R(M)}$ is the translational vector of the upper-stage or the payload, $F_{R(M)}$ and $T_{R(M)}$ are, respectively, the gravitational force and the tension force affecting the upper-stage or the payload, $P_{R(M)}$ is the perturbation force.

The dynamics of the deployed nodes satisfies

$$m_i\ddot{\boldsymbol{r}}_i = \boldsymbol{F}_i + \boldsymbol{T}_i + \boldsymbol{P}_i, \quad i = n_R + 1, \cdots, n - n_M \tag{2}$$

where $\boldsymbol{T}_i = \boldsymbol{T}_{i,i-1} + \boldsymbol{T}_{i,i+1}$, $\boldsymbol{T}_{i,i-1}$ and $\boldsymbol{T}_{i,i+1}$ are the tension force acting on the node I from the forward node i-1 and the backward node i+1, respectively; meanwhile, $T_{i,i-1} = EA(\eta_{i,i-1} - 1 + \alpha_d\dot{\eta}_{i,i-1})$, $\eta_{i,i-1}$ denotes the elongation of the segment $L_{i,i-1}$ between the node i and the node i-1, P_i is the perturbation force acting on the node i.

The perturbation force usually results from the aerodynamic drag, gravitational perturbation and electrodynamic force, which in fact hardly leads to a great impact since the tether deployment duration is significantly shorter than the whole deorbit process. Thus, these perturbation forces are neglected during the tether deployment stage.

(2) Attitude dynamics of the upper-stage

The kinematics of the upper-stage is given by

$$\dot{\boldsymbol{q}} = \frac{1}{2}\boldsymbol{W}\bar{\boldsymbol{\omega}} \tag{3}$$

where quaternion is utilized to describe the attitude of the upper-stage and thus avoid the singularity problem, $\boldsymbol{q} = (\lambda_1, \lambda_2, \lambda_3)^{\mathrm{T}}$ is the vector part of the attitude quaternion, $\bar{\boldsymbol{\omega}} = \left(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3\right)^{\mathrm{T}}$ denotes the angular velocity of the upper-stage respect to the orbital frame, and the matrix W satisfies

$$\boldsymbol{W} = \begin{bmatrix} \lambda_0 & -\lambda_3 & \lambda_2 \\ \lambda_3 & \lambda_0 & -\lambda_1 \\ -\lambda_2 & \lambda_1 & \lambda_0 \end{bmatrix}$$

Furthermore, the dynamics of the upper-stage is given by

$$\boldsymbol{J}\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times (\boldsymbol{J}\boldsymbol{\omega}) = \boldsymbol{M} \tag{4}$$

where $\boldsymbol{M} = \left(M_x, M_y, M_z\right)^{\mathrm{T}}$ is the external torque vector, $\boldsymbol{\omega} = \left(\omega_1, \omega_2, \omega_3\right)^{\mathrm{T}}$ denotes the angular velocity vector of the upper-stage respect to the inertial frame, satisfying

$$\boldsymbol{\omega} = \bar{\boldsymbol{\omega}} + \boldsymbol{\omega}_r = \bar{\boldsymbol{\omega}} + \boldsymbol{A}_{\lambda}{}^{\mathrm{T}}(0, 0, \dot{v})^{\mathrm{T}} \tag{5}$$

where $\boldsymbol{\omega}_r$ and $\dot{\nu}$ are the orbital angular velocity and the time derivation of the true anomaly, A_λ is the transformation matrix from the body frame to the orbit frame.

As a consequence, Eqs. (1)-(4) constitute the rigid-flexible coupling dynamics of the deorbit system during the tether deployment stage. It should be noted that, as the tether deploys, the nodes will be gradually removed from the upper-stage, and thus the system performs time-varying.

3 Requirement analysis of the initial attitude of the upper-stage

To ensure tether deployment stability, both tether libration and attitude motion of upper-stage should be within an admissible range. To this end, on one hand, tether libration is able to be regulated by tension force via an appropriate algorithm; on the other hand, in order to minimize the hardware change for the control system of the upper-stage, a straightforward strategy is not to control the attitude motion of the upper-stage. However, this requires a proper and strict initial attitude condition of the upper-stage for the tether deployment stage. This requirement analysis will be undertaken in this section.

(1) Tether control algorithm

To proceed with the requirement analysis, an appropriate tension control algorithm should be developed to control the tether libration. In view of the applicability and computational complexity, Kissel algorithm is adopted and the tension force is determined by

$$T(s(t)) = 0.02(s(t) - s_c(t)) + 2\dot{s}(t) + 3ms(t)\dot{\nu}^2 \tag{6}$$

where m denotes the mass of the payload, s(t) is the deployed tether length and the $s_c(t)$ represents the nominal deployment trajectory, satisfying

$$s_c(t) = s_0 + s_f \cdot \left\{ 1 - \exp\left[-t \cdot \frac{\ln(s_f / s_0)}{t_f} \right] \right\} \tag{7}$$

where s_0 and s_f are the initial and final deployed tether length, respectively, t_f is the pre-determined deployment time.

It can be observed from Eq.(6) that the Kissel controller is indeed a PD-type controller and $3ms(t)\dot{\nu}^2$ is a feed-forward item.

(2) Simulation and analysis

To proceed, suppose that the system operates in a circular orbit with inclination 0°, the full tether length is 5km, the stiffness coefficient is EA=105N, the damping coefficient is α=0.05, the mass of the payload is 40kg, the initial eject velocity is 1m/s and the eject direction angle is π/8 rad, Set the number of the nodes as 5 and the deployment time is determined as t_f=30000s. The terminal of the eject phase corresponds to the time that the payload flying to the local vertical.

The parameters of the upper-stage are set as M=4000kg, l=11m, D=3.35m, and the moment of inertia are given as

$$J_x = 1.3\times10^4\,\mathrm{kg\cdot m^2}, \quad J_y = J_z = 6.5\times10^4\,\mathrm{kg\cdot m^2}$$

Based on the above control strategy and parameter settings, for various initial conditions of tether deployment (different ω and θ, ω-the initial angular velocity; θ-the deviation angle between the vertical axis of the upper-stage and the local vertical), numerical simulations are given to

investigate the influence of the initial attitude of the upper-stage. It should be noted that, during the tether deployment, once the vertical axis of the upper-stage deviate the local vertical over 90°, the tether is prone to be twined, which leads to the system instability. Thus, this can be viewed as the stability criterion of the tether deployment.

Fig. 3 shows the variation of the maximum derivation angle θ_{max} during tether deployment under various initial attitude conditions of the upper-stage, where the x axis in Fig. 3 denotes the initial derivation angle $\theta(0)$, satisfying

$$\lambda_0 = \left[\cos\frac{\theta(0)}{2} \quad 0 \quad 0 \quad \sin\frac{\theta(0)}{2} \right]^{\mathrm{T}}$$

It can be concluded from the simulation results that, the initial attitude condition has a great impact on the tether deployment stability. If the attitude motion of the upper-stage is not controlled, the stability margin is small and thus the requirement of the initial attitude of the upper-stage performs rigorous. To explain this, from Fig. 3, once the initial angular velocity is over 0.2°/s, the tether deployment is unstable whatever the initial deviation angle is.

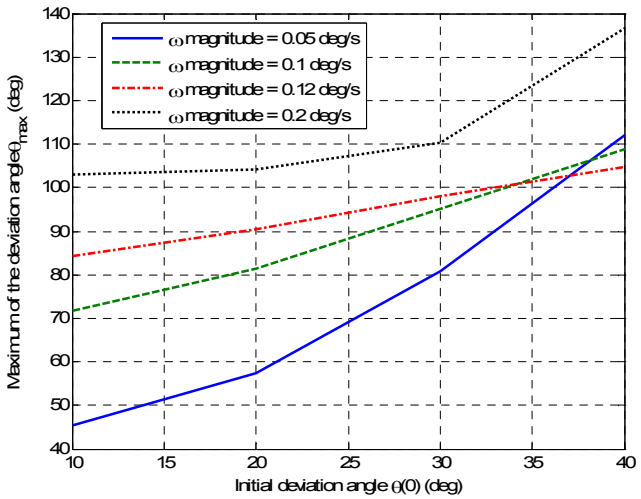


Fig.3 The requirement analysis of the initial attitude condition of the upper-stage

4 Control scheme design and analysis

(1) Control scheme design

Based on the analysis in the above section, to avoid the impact of the initial attitude of the upper-stage on the tether deployment stability, the attitude control is necessary for the upper-stage during the tether deployment stage. To handle this, a coupling control scheme is proposed as follows:

● Eject phase: Tether libration is not controlled, but the attitude motion of the upper-stage is controlled by RCS system;

● Deployment phase: Tether libration and the attitude motion of the upper-stage are both controlled.

The tether libration control still adopts Kissel algorithm as shown above, and the attitude control scheme are given as follows.

The proposed RCS system is composed of four 300N-nozzles and four 150N-nozzles, with the configuration shown in Fig. 4, and all the nozzles are fixed on the bottom plane of the upper-stage.

Nozzle 1#~4# are 300N-nozzles and used for pitch- and yaw-channel control, while Nozzle 5#~8# are 150N-nozzles and used for roll-channel control.

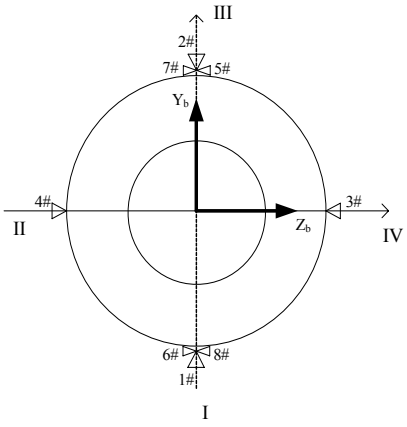


Fig.4 Nozzle configuration (End view)

The nozzle control commands are designed by using phase-plane algorithm, i.e.,

$$k_q = \begin{cases} 1, & e \ge \theta_{Dq} \\ 1, & e \ge (1-h)\theta_{Dq}, \dot{e} < 0 \\ 0, & -\theta_{Dq} < e \le (1-h)\theta_{Dq}, \dot{e} < 0 \\ 0, & -(1-h)\theta_{Dq} \le e < \theta_{Dq}, \dot{e} > 0 \\ -1, & e < -(1-h)\theta_{Dq}, \dot{e} > 0 \\ -1, & e \le -\theta_{Dq} \end{cases} \quad q = \gamma, \psi, \varphi \tag{8}$$

where θ_D and h are parameters to be determined, $k_q (q = \gamma, \psi, \varphi)$ is the command input of the nozzle, and e satisfies

$$e = q + k_{qd}\dot{q}, \quad q = \gamma, \psi, \varphi, \tag{9}$$

where $k_{qd} (q = \gamma, \psi, \varphi)$ is the design parameter for each channel.

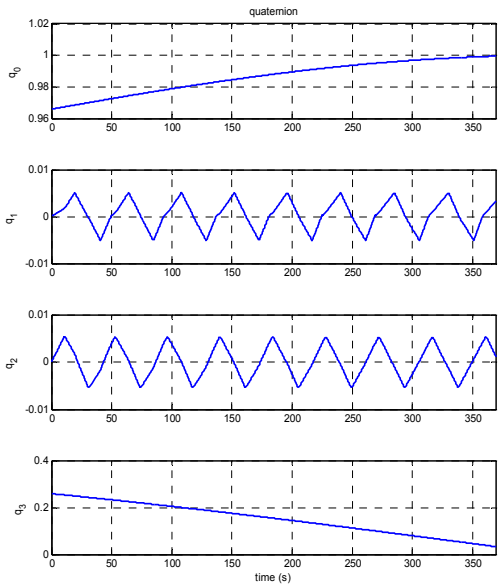
(2) Simulation and analysis

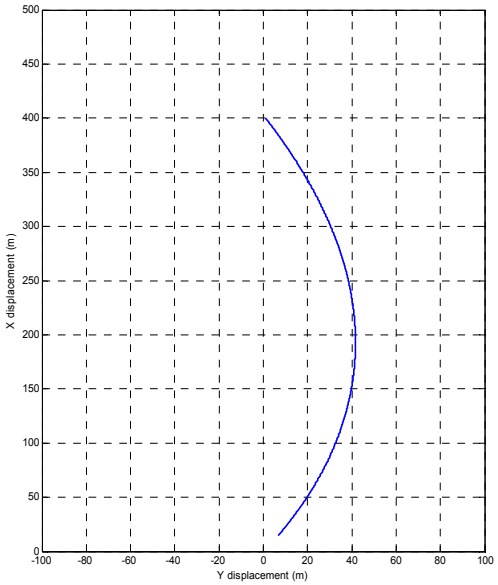
A. Eject phase

System parameters are the same as those in the above section. Initial attitude of the upper-stage are

$$\boldsymbol{\lambda}_0 = \left[\cos\frac{\pi}{12} \quad 0 \quad 0 \quad \sin\frac{\pi}{12} \right]^{\mathrm{T}}, \quad \bar{\boldsymbol{\omega}}_0 = \left[0.001 \quad 0.001 \quad 0.0011 \right]^{\mathrm{T}} - \boldsymbol{\omega}_r$$

Fig. 5 shows the attitude variation of the upper-stage in quaternion form and the trajectory of the payload in the xy-plane.





(a) Attitude quaternion of the upper-stage (b) Payload trajectory in the orbital plane

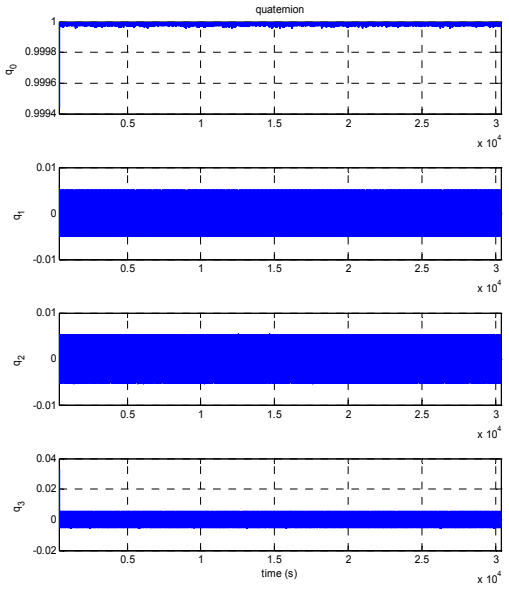
Fig.5 Simulation results in the eject phase

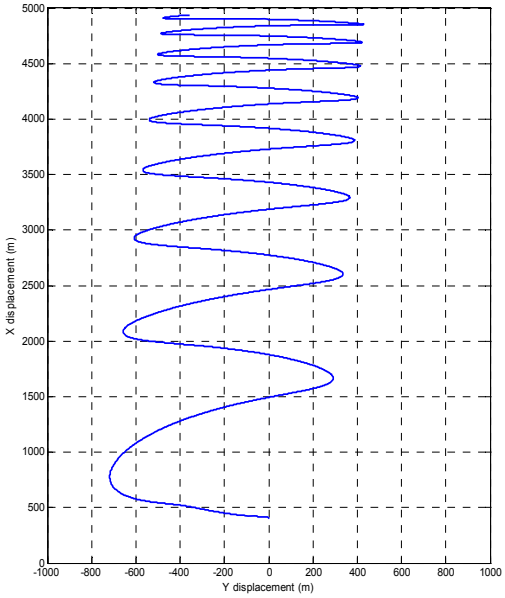
It can be seen from the figure that the payload flies back to the local vertical in about 370s, during which the attitude motion of the upper-stage performs stable under the attitude control.

B. Deployment phase

The maximum tension force is set as 2N, and other parameters are set as above.

Fig. 6 shows the attitude variation of the upper-stage in quaternion form and the trajectory of the payload in the xy-plane. Fig. 7 illustrates the variation of the derivation angle between the vertical axis of the upper-stage and the local vertical.





(a) Attitude quaternion of the upper-stage (b) Payload trajectory in the orbital plane

Fig.6 Simulation results in the eject phase

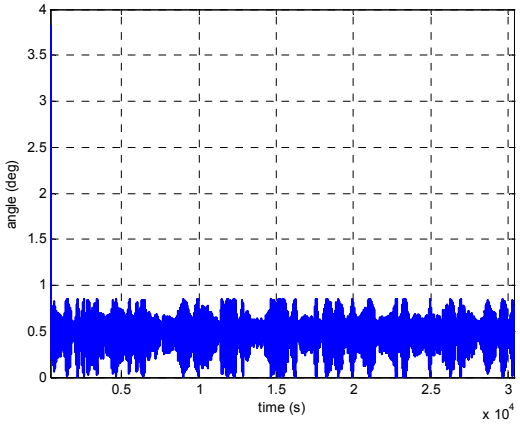


Fig.7 The variation of the deviation angle of the upper-stage

It can be seen from the figure that the tether deployment is stable. The tether is gradually deployed to the full length under Kissel control law, and the libration is controlled within 0.1rad. Meanwhile, the attitude motion of the upper-stage performs stable and converges to the desired values.

5 Conclusion

Electrodynamic tethered deorbit technology is a novel and effective way for removal of abandoned upper-stages. This paper handles the tether deployment control problem of electrodynamic tethered systems. Based on “bead” philosophy, a rigid-flexible coupling dynamics is formulated for the tether deployment stage. Further, the requirement analysis of the initial attitude of the upper-stage indicates that the attitude control of the upper-stage is necessary besides the tether libration control. Thus, a coupled control scheme is then proposed: 1) for initial eject phase, the attitude of the upper-stage is controlled regardless of the tether libration; and 2) for sequential deployment, not only is attitude motion of upper-stage stabilized, but also is tether libration controlled by using Kissel algorithm. The following numerical simulation shows the effectiveness of the proposed control scheme.

References

[1] G. Rouleau, S. Verma, I. Sharf, and E. Martin, Vision-based tracking and trajectory generation for robotic capture of objects in space, AIAA Guidance, Navigation and Control Conference, California, USA, Aug, 2005. AIAA 2005-6446

[2] R. Forward, R. Hoyt, Application of the terminator tether electrodynamic drag technology to the deorbit of constellation spacecraft, AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Cleveland, USA, July, 1998. AIAA98-35302.

[3] V. Aslanov, V. Yudintsnev, Dynamics of large space debris removal using tethered space tug, Acta Astronautica, 2013, 91(10): 149-156.

[4] R. Stevens, W. Baker, Optimal control of a librating electrodynamic tether performing a multirevolution orbit change, Journal of guidance, control and dynamics, 2009, 32(5): 1497-1507.

[5] H. Wen, D. Jin, H. Hu, Advances in dynamics and control of tethered satellite systems, Acta Mechanica Sinica, 2008, 24(3): 229-241.

[6] J. Sanmartin, E. Lorenzini, M. Martinez-Sanchez. Electrodynamictether applications and constraints, Journal of Spacecraft and Rockets, 2010, 47(3): 442-456.

Development and Discussion of the Air Launching Technology

Du Linfei, Wu Jihua

(Ninth System-Designing Institute,CASIC,Wuhan 430040,China)

Abstract: Recently, rapid satellite networking and fast space response become present popular trends based on the diversity and flexibility of the requirement of space tasks. With the design conception of reducing weight and size of satellites, more and more people do research on the launching technology which has lower cost and faster mobility in the field. The air launching technology gains wide attention because of its excellent mobility and low cost based on the recycling of the vehicle platform. In this paper, the development situation domestic and abroad is introduced and some key issues of air launching technology is discussed. In addition, the advantages of air launching in the satellite networking is clarified and several pieces of assumption are displayed concerning the development of the air launching technology.

Introduction

Rapid satellite networking, fast information support and rapid satellite supplement have become extremely important technologies recently with the military warfare intelligentized and the space resource highly informationized[1]. Usual methods with fixed launch tower can not satisfy the requirement because of the limitation of launch field, launch window, etc. Consequently, mobility launch is drawing great attention which can be divided into three sorts according to the launch platform: land based launch, sea based launch and air based launch. Air based launch means throwing the rocket at a given height before the ignition via platforms such as airplanes and airships. Air based launch has better mobility and lower cost which makes itself possess significant meaning.

Domestic and International Development

International Development

America put forward the concept of air based launch in the 1950s. In the year 1990, the rocket called Pegasus was successfully launched which brought a brilliant start in the field of air based launch[2]. Thereafter, QuickReach , ALASA ,Bigbird appeared which represented landmarks of airbased launch. American researchers verified the platforms of the aero-transport, battle-plane and specially-made plane which accomplished the technology of lower suspension, inner fill and so on. The transportation has grown from initial kilograms to afterward tons ans supports space fast response strongly.

Japan also raised air based launch into the agenda in recent years, such as NanoLauncher and ALSET[3], ie air launch system enabling technology in 2010 and 2011 respectively, which may choose C-130 as the launch platform and send the effective load ranging 100kg to 200kg into the LEO less than 24 hours.

Similarly, France ,German and Spain put up the Aidebaran scheme in 2008[4]. The battle plane called ‘wing’ is intended to be selected as the launch platform, which is promising to send the 300kg effective load into the LEO and is to accomplish the demo at the end of 2015.

In 1999, Russia put up the Polyot program planning to send the 4500kg effective load into the LEO with the platform of AN-124. Nevertheless, the progress is extremely slow for a time because of the limitation of the domestic economic situation. Fortunately, Russia restarted the program in 2012 and established cooperation in the research with Ukranie, Indonesia and German.

Typical models

Development by OSC(Orbital Sciences Corporation) and Hercules, Pegasus is the first air based rocket whose length is 16.9 meters and wingspan is 6.7 meters and it can send the effective loads weighing 450kg into LEO. Pegasus was first launched in 1990 using the reconstructed B52 bombing plane as the launch platform and adopted the low suspension method. The flight flow is as follows. First, the plane carried the rocket to a height of 13km. After the free drop of 5 seconds, the ignition was touched off and sent the satellite weighing 191kg into the LEO eventually. The first successful performance brought a brilliant opening in the air based launch field. Thereafter, Pegasus executed approximately 40 tasks with the L-1011 platform and the success rate is up to 92.5% which makes itself a typical model.

'QuickReach' was developed by American Air Launch Company whose length is 20.12 meters, diameter is 2.21 meters and the effective load in LEO is up to 450kg[5]. The C-17A transport is selected with inner fill method and the rocket slides along the slides via gravity. After that, stabilization parachute is opened to drag the rocket to keep the attitude stable. Ignition begins after the signal is received at the time of 6.3s and the cable is burned down sending the effective load into objective orbits. Up to now, three mode releases has been accomplished. Unfortunately, the program was abandoned in 2007 finally because America Flight Operational Quality Assurance consider the liquid rocket is of low reliability.

In 2011, American Defense Advanced Research Projects Agency released the 'ALASA' project mainly used in the nano-satellite fast launch field[7]. The F15 battle plane is chosen as the platform which can suspend 45kg effective loads underneath without modification. When the plane rises up to 7.5km, it releases the rocket and then the ignition starts to keep the rocket continue rising. ALASA is a project that the total whole time is less than 24 hours from the plan being made to reaching the goal and the single cost is less than 1 million dollars. The first test flight is intended to be carried on at the end of 2015.

BigBird is a mode that America is devoting to whose wingspan is tend to be 117m, length is 72m and the fastest speed is 850kn/h[8]. The rocket is suspended between planes and is launched at the height of 9.1km, which can send a satellite weighing 6.1t to LEO. The first flight test is intended to be carried on in 2016.

Domestic Development

Up till now, the research of air based launch rest on the theoretical analysis and no public reports is found yet. CASC, CASIC, Beihang University and Harbin Engineering University start some trials in this field with the method of theoretical analysis and numerical simulation[9].

Key Technology

The mature application of the air based launch cannot work without key issues being solved. The first main problem is the installation and separation of the rocket and the second is the stability of the attitude. Thus, these issues are worth exploring[10].

Installation and Separation

The detailed methods contain underneath suspension, inner fill, back carry and drag behind. The same thing is how to promise the rocket released stability and safety, accomplish control of the attitude before ignition and keep the plane safe no matter what the method is.

Underneath suspension means the rocket is carried under the wing or airframe. This scheme is easy to be fulfilled but the size of the rocket is limited strictly. Meanwhile, the aerodynamic shape is affected because of the suspension.

Inner fill is a way that puts the rocket into the cabin. When needs separation, the hatch is open and the rocket slips out under gravity and the drag of the stability parachute. Although there is no influence on the aerodynamic characters and the carrying capability is fully used, the reliability is low and the separation has a tremendous effect on the attitude of the rocket which makes the adjustment more difficult.

What back carry means is that the rocket is fixed on the back of the plane and the separation is accomplished by relying on the life force via adjusting the wings of the rocket. This method also has an effect on the aerodynamics but has little limitation on the size which allows the rocket to carry heavier loads.

To accomplish drag behind, the rocket is dragged by the cabins behind the plane and the separation is finished with burning down the cables. In addition, this method sounds to be easy and has little limitation on the size of the rocket but still has an effect on the aerodynamics. Moreover, the take off and land is very difficult and the route of the technology is extremely rough.

Attitude Control after Release

After the rocket is released, it lacks motivation so that the position, velocity and attitude have high uncertainly. In case of emergency, the plane needs to reach the safe area before the ignition of the rocket. Meanwhile, the initial condition has a tremendous effect on the orbit injection accuracy. Thus, measures should be taken to ensure the stability of the attitude.

One method is to add the stability parachute which means opening the parachute after release to keep it stable. Although it is easy to accomplish, this method is passive so that the control precision is hard to ensure and the disturbance is occurred to the plane. Therefore, starting from the active control, air injection mechanism can be designed to control the attitude so that the attitude get adjusted timely and the precision can be guaranteed. However, the pessimistic mass increases which exerts an influence on the carry ability of the rocket.

Real-time Trajectory Planning

The trajectory is sensitive to initial condition in air based launch and the rocket is easily affected by the lateral wind and release deviation. Thus, methods shall be taken to measure the position, velocity and attitude precisely in time and then the trajectory planning can be accomplished by comparing the measurement with the standard. Traditional methods such as variable parameter design and muti-objective optimization are still hard to satisfy the needs of the launch and more methods which have better convergence should be explored.

Advantages in the Fast Response

Rapid satellite networking, fast information support and rapid satellite supplement all require the launch with good quality of mobility. On the condition of the preparation of 24 hours or 48 hours, air based launch can shorten the time tremendously because of the following advantages[11].

This method is rarely limited to the road condition and geographic position and the space is wide.

It has a strong existence ability and can maintain itself freely no matter attack or defense and the concealment is excellent.

The plane can provide a high speed to rocket because of itself so that the fuel can be saved and the carry ability can be raised greatly.

Without the launching site on the ground, the maintenance cost can be saved that the ensure system can be carried on the platform.

Forecast

Based on the various advantages, there is a significant meaning in exploring air based launch technology domestically. On the condition of our scientific research ability and industrial manufacturing level, mature technology can be used to design China's air based launch system with the following three steps.

First, the middle-launch platform. When comes to the launch platform, H6 can be selected with some modification mainly under the consideration of carry ability, technological statement and economic indicator and the aviation industry. Referring to the rocket, many mature solid rockets domestically can be used with some modification. The installation method can choose underneath suspension to deal with 100kg effective loads. As far as the control of the attitude, air injection mechanism and changeable wings can be used together so that the excellent aerodynamics can be guaranteed.

Second, the small-launch platform. Based on the middle-launch platform, technology with lower cost, faster response and higher load rate can be developed. ALASA is a typical representation in the field of small effective loads which can send a 45kg satellite into LEO in less than 24 hours and the cost is less than 1 million per launch. Such innovation can be drawn. Small battle planes with good mobility quality can be used as the platform so that the air based launch can quickly win its status in the field of the nano satellite and commercial space,especially in the cost control and mobility time.

Third, the large-launch platform. Based on the small-launch platform, innovation can be made further to elevate the transport ability. Large transports being modified, the inner fill method can be developed to lessen the limitation of the size of the rocket. Meanwhile, the LEO mass can be advanced to more than 500kg promisingly.

Reference

[1] Li Li, Shen Huairong. Discussion about the technology of air emergency flexible launch in spacecraft wartime[J]. Journal of the Academy of Equipment Command and Technology .2003,14(4):39-42.

[2] Orbital Science Corporation. Pegasus user's guide [EB / OL] .http ://www . orbital . com/ Pegasus / Pegasus-guide.pdf.2014-03-17.

[3] Yuchin, Takashia, Seijim, et al. ALSET-Japanese air launch system concept and test plan[C].//AIAA.27th annual AIAA/USU Conference on Small Satellites.Tokyo, Japan:SSC,2013-V-7.

[4] Talbotc, Pilarg, Ludgerf. 'Aldebaran', a launch vehicle system demonstrator[C]// AIAA .7th Responsive Space Conference. Los Angeles, California, USA:AIAA,2009-1005.

[5] Karpova, Ivanovr, Kovalevskym. Air launch aerospace international project[C] // ICAS. 27th International Congress of the Aeronautical Sciences. Moscow, Russia: ICAS, 2010-D1.3.3.

[6] Garyc. Quickreach responsive launch system[C]// AIAA.4 th Response Space Conference. Los Angeles, California, USA-AIAA, 2006-2003.
[7] Defense Advanced Research Projects Agency. Report of the horizontal launch study[R]. Washington,D.C, USA: Department of the Air Force.2011: 24-41.
[8] America is combining the largest plane in the world which can launch the rocket in the air[EB/OL]. http://m. bendibao.com/show673099.html.
[9] Li Li, Shen Huairong. The analysis on a kind of plane-rocket combination and separation scheme of launching small satellite in the air[J]. Journal of the Academy of Equipment Command and Technology, 2002.13(6):43-46.
[10] Yang Huabao, Ma Xin. Air launching technology and its key technologies[J]. Fire Control and Command Control, 2009,33(9): 15-17.
[11] Pang Yong. Research on combined environmental adaptability for ship-based and air-based missile[J]. Tactical Missile Technology, 2004,(6):39-43.

Experimental Research of Regression Rate of N_2O and HTPB Hybrid Rocket Motor

GUO Fang, LUO Zhan-fei, Zhao Hui-jie, Yang Fan, Zhang Ming-liang

(The 41st Institute of No.6 Academy China Aerospace Science&Industry Corp., Hohhot 010010, China)

Abstract: Regression rate behavior of hybrid rocket motor was investigated through motor experiments using the propellant of N_2O and HTPB. Results of regression rate were analyzed by three kinds of regression rate models: the regression rate relies on the oxidizer flux- $\dot{r} = aG^{0.8}$; the regression rate relies on the chamber pressure- $\dot{r} = bp^n$; the regression rate relies both on the oxidizer flux and the chamber pressure- $\dot{r} = aG^{0.8}bp^n/(aG^{0.8} + bp^n)$. Through comparative analysis, a fuel regression rate formula was fitted - $\dot{r} = 1.2929\,p^{0.21734}$. That is to say the regression rate is nearly independent of the oxidizer flux but varied markedly with the chamber pressure. Then the mechanism was discussed using the heterogeneous reactions theory. Instantaneous regression rate and the interior ballistic performance calculation were produced based on the fitted fuel regression rate formula, and the theoretic curves fit well with the experiment results, which validate the method and regression research in this paper.

Key word: Hybrid rocket motor; Fuel regression rate; Motor experiment

Nomenclature

a testing constant

$\overline{A}_p$ average value of the passage area during the combustion time, m^2

b testing constant

B' parameter

C_O concentration of the oxidizer absorbed

C_F^S concentration of the base at the fuel surface

C_m monolayer capacity

C_F thrust coefficient

C^* characteristic exhaust velocity, mm/s

d_{po} grain diameter before combustion, mm

d_{pf} grain diameter after combustion, mm

E energy of desorption

F_0 initial motor thrust, N

F_i motor thrust in the moment i, N

$\overline{G}$ average value of the oxidizer mass flux, kg/s·m^2

K_1 parameter

k chemical reaction rate constant

k' constant

k'' constant

$\overline{\dot{m}}$ average value of the oxidizer mass flow rate, kg/s;

$\dot{m}_0$ initial oxidizer mass flow rate from the flowmeter, kg/s

$\dot{m}_i$ test result of the oxidizer mass flow rate from the flowmeter in the moment i, kg/s

L length of the gain, mm

$\dot{m}_{p0}$ initial solid-fuel mass flow rate, kg/s

$\dot{m}_{pi}$ solid-fuel mass flow rate in the moment i, kg/s

n testing constant

P pressure

$\overline{p}$ average value of the chamber pressure, MPa

p_i test result of chamber pressure from the pressure sensor, MPa

R gas constant

r rate of the reaction

$\dot{r}$ regression rate or burning rate

$\overline{\dot{r}}$ average value of the solid-fuel regression rate during the combustion time, mm/s

$\dot{r}_i$ regression rate in the moment i

r_{p0} initial radius of the grain, mm

r_{pi} grain radius in the moment i, mm

T absolute temperature

t_b combustion time, s

t_0 beginning time, s

t_f end time, s

Z_m number of molecules adsorbed per unit area in a complete monolayer

θ fraction of the surface covered

v oscillation frequency of absorbed molecules

ρ_f solid fuel density, kg/m^3

Δt time step length, s

1. Introduction

The key parameter influencing the design of hybrid fuel grains, and, therefore, the overall hybrid motor design, is the solid-fuel regression rate. Probably the most important influential theory so far explaining regression rate behavior was developed by Maxman and Gibert[1], Maxman et al.[2], and Maxman[3][4]. They based their approach on turbulent combustion over a flat-fuel surface. According to their study, the solid-fuel regression rate is governed primarily by convective heat transfer from flame to the fuel surface. The local mass flux G is the primary operating parameter governing regression rate in this theory. At about the same time Maxman and his coworkers performed their work, Smoot and Price[5][6][7] conducted a large number of experiment which indicated that, at the relatively low pressure and high mass fluxes, the solid-fuel regression rate in hybrids employing oxidizers display a strong dependency on chamber pressure, but only a weak dependence on mass flux.

Recently in China, most of the researchers focus in the relationship between the regression rate and the mass flux[8] instead of the relationship between the regression rate and the chamber pressure. In this paper, regression rate behavior of hybrid rocket motor was investigated through motor experiments using the propellant of N_2O and HTPB. Results of regression rate were analyzed by three kinds of regression rate models. Through comparative analysis, a fuel regression rate formula was fitted. Then the mechanism was discussed using the heterogeneous reactions theory. Instantaneous regression rate and the interior ballistic performance calculation were produced based on the fitted fuel regression rate formula.

2. Test Motors & Test Result

The test motor consists of the injector, the ignition, the aft-chamber, the nozzle and the other affiliated parts (Shown in Fig. 21). The motor combustion case is made of the alloy steel and the insulation is made of the erosion resistant composites. The joint structure between the cap and the aft-chamber, the aft-chamber and the post-chamber, the post-chamber and the nozzle is flange. O-ring is used for sealing. The solid-fuel is based on HTPB with the tube grain and the oxidizer is N_2O. Five motors are tested and one of them is abandoned because of the abnormal data. The others are labeled as Motor Ⅰ, Ⅱ, Ⅲ, Ⅳ. Table 23 is the test result of hybrid motors.

Fig. 21 Test Motor

Table 23 Test Result

Parameter		I	II	III	IV
d_{po}	mm	30	30	30	30
d_{pf}	mm	32.18	30.76	33.14	32.19
t_b	s	9.99	9.84	11.13	11.04
$\overline{p}$	MPa	3.64	2.76	2.58	2.39
$\overline{\dot{m}}$	kg/s	0.50	0.56	0.48	0.52
$\overline{A}_p$	m^2 (10^{-3})	1.98	1.8	2.08	1.98
$\overline{G}$	kg/s·m^2	298	352	291	325
$\overline{\dot{r}}$	mm/s	1.72	1.57	1.63	1.56

The combustion time t_b :

$$t_b = t_f - t_0$$

The average value of chamber pressure $\overline{p}$:

$$\overline{p} = \frac{\int_{t0}^{tf} p_i dt}{t_b}$$

The average value of oxidizer mass flow rate $\overline{\dot{m}}$:

$$\overline{\dot{m}} = \frac{\int_{t0}^{tf} \dot{m}_i dt}{t_b}$$

The average of the passage area during the combustion time:

$$\overline{A}_p = \frac{\pi(d_{po}^2 + d_{pf}^2)}{8}$$

The average value of the oxidizer mass flux during the combustion time $\overline{G}$:

$$\overline{G} = \frac{\overline{\dot{m}}}{\overline{A}_p} = \frac{8\int_{t0}^{tf} \dot{m}_i dt}{t_b \pi (d_{po}^2 + d_{pf}^2)}$$

The average value of the solid-fuel regression rate during the combustion time $\overline{\dot{r}}$:

$$\overline{\dot{r}} = \frac{d_{pf} - d_{p0}}{2t_b}$$

3. Fitting of Regression Rate Data

According to the influence factor, the combustion can be divided into three areas:

For the turbulent flow range, the regression rate was predicted to be a function of the 0.8 power of total specific flow rate and independent of pressure:

$$\dot{r} = aG^{0.8}$$

For intermediate flow-rate region, variation of HTPB regression rate with pressure change was irregular:

$$\dot{r} = \frac{aG^{0.8}bp^n}{aG^{0.8} + bp^n}$$

At higher flow rates, pressure influence increases:

$$\dot{r} = bp^n$$

3.1 Model A -- $\dot{r} = aG^{0.8}$

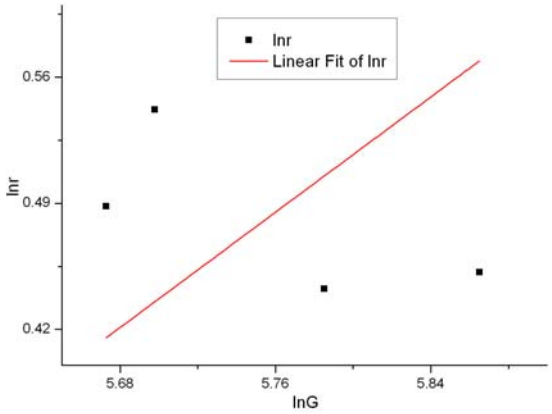


Fig. 22 Fitted Curve Plot

Given that the model is not convenient for fitting, the value is built by evaluating the logarithm of current value (Shown in Table 24):

$$\ln\dot{r} = \ln a + 0.8\ln G$$

Table 24 Parameters

Data Source		I	II	III	IV
Input X	$\ln G$	5.67	5.86	5.85	5.67
Input Y	$\ln\dot{r}$	0.54	0.45	0.49	0.44

The least square method (LSM) is used to solve the parameters. The result is

$$\ln a = -4.12291$$

That is,

$$\dot{r} = 0.0162G^{0.8}$$

According to Fig. 22, the residual is relatively large as the fitting result of this model.

3.2 Model B -- $\dot{r} = bp^n$

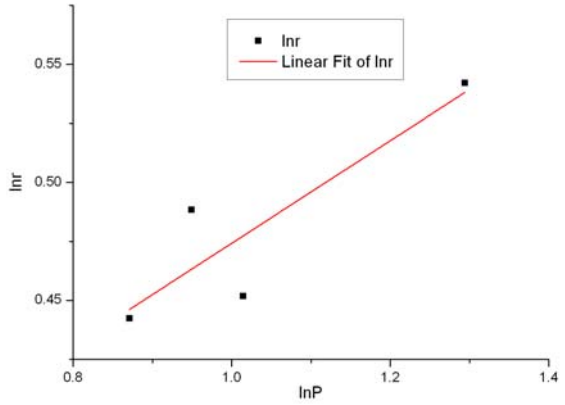


Fig. 23 Fitted Curve Plot

The value is built by evaluating the logarithm of current value (Shown in Table 25):

$$\ln\dot{r} = \ln a + n\ln p$$

Table 25 Parameters

Data Source		I	II	III	IV
Input X	$\ln p$	1.29	1.01	0.95	0.87
Input Y	$\ln\dot{r}$	0.54	0.45	0.49	0.44

LSM is used to solve the parameters. The result is

$$\ln a = 0.25691$$
$$n = 0.21734$$

That is,

$$\dot{r} = 1.2929\,p^{0.21734}$$

According to Fig. 23, the result shows that the model data agrees well with this model.

3.3 Model C -- $\dot{r} = aG^{0.8}bp^n/(aG^{0.8} + bp^n)$

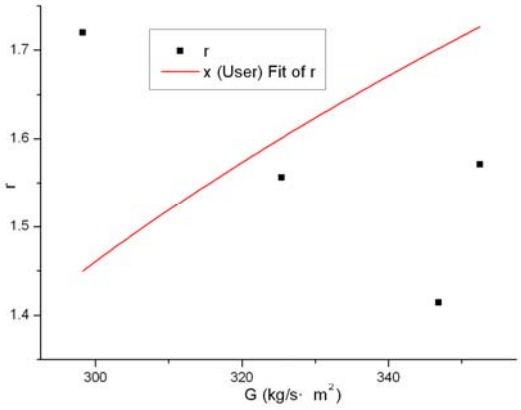


Fig. 24 $\dot{r}$ vs. G Fitted Curve Plot

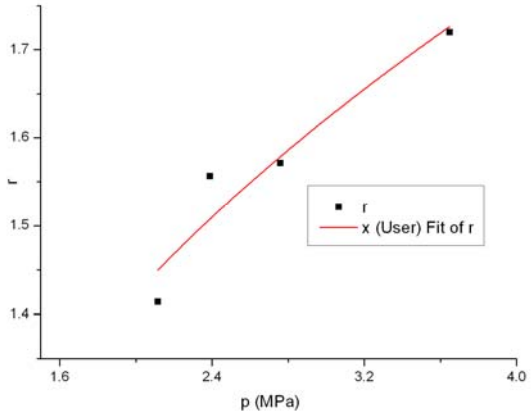


Fig. 25 $\dot{r}$ vs. p Fitted Curve Plot

Nonlinear curve fit is employed in this model. In this regression rate model, we used the oxidizer mass flux G and the chamber pressure p as the independent variables and the solid-fuel regression rate $\dot{r}$ as the dependent variable. In order to reduce the number of iteration, we used the results of last two models as initial values: $a_0 = 0.0162$, $b_0 = 1.2929$, $n_0 = 0.21734$. The result is

$$a = 1.64687e21,\ b = 1.14026,\ n = 0.32045$$

That is,

$$\dot{r} = \frac{1.64687e21G^{0.8} \times 1.14026\,p^{0.32045}}{1.64687e21G^{0.8} + 1.14026\,p^{0.32045}}$$

According to the result, a is much larger than b. That is to say bp^n in the denominator can be ignored. $G^{0.8}$ crosses out from the top, so we're left with an answer in p^n. According to the Fitted Curve Plot, the pressure is fitted much better than the mass flux, which draws a same conclusion with the analysis.

3.4 Comparison of Models

The results of three models:

$$\dot{r} = 0.0162G^{0.8}$$

$$\dot{r} = 1.2929p^{0.21734}$$

$$\dot{r} = \frac{1.64687e21G^{0.8} \cdot 1.14026p^{0.32045}}{1.64687e21G^{0.8} + 1.14026p^{0.32045}}$$

The fitted regression rate with three models and their errors are shown in Table 26. We can see that the error of Model B-- $\dot{r} = bp^n$ is minimal. Therefore, we choose $\dot{r} = 1.2929p^{0.21734}$ as the model to fit the solid-fuel regression rate in this test condition. That is to say the regression rate is nearly independent of the oxidizer flux but varied markedly with the chamber pressure.

Table 26 Comparison of Models

Model	$\bar{\dot{r}}$ (mm/s)				
	Ⅰ	Ⅱ	Ⅲ	Ⅳ	Ave.
measured $\bar{\dot{r}}$	1.72	1.57	1.63	1.56	1.62
A	1.55	1.77	1.51	1.66	1.62
Error(%)	-10	12	-7	6	9
B	1.71	1.61	1.59	1.56	1.62
Error(%)	0	3	-3	0	1
C	1.73	1.58	1.55	1.51	1.59
Error(%)	0	-9	-15	-12	9
Ave.	1.71	1.63	1.64	1.61	1.65
Error(%)	-1	4	1	4	2

4. Pressure Dependence of Hybrid Fuel Burning Rate[9]

Smoot and Price observed the pressure dependence of regression rate at high flow rate. At lower flow rates, the regression rate is independence of pressure and is governed entirely by the turbulent heat transfer in the boundary layer. At higher flow rates, the Reynolds numbers are higher. The diffusion rate becomes high, and the process is governed by the kinotics of heterogeneous chemical reactions at the interface.

In the present case, we initially have a heterogeneous reaction followed by constructive reactions. We make the following assumptions, which normally are valid:

1) Adsorption-desorption equilibrium is established quickly;

2) Observed rate is determined by the amount of surface covered by reacting molecules and by the specific velocity of the surface reactions;

3) The lowest step is the rate-determining step.

The adoption process would be the rate determining slowest step. Hence, the rate of the reaction would be given by

$$r = kC_O C_F^S$$

C_O can be estimated easily on the basis of Langmuir adsorption isotherm, according to which θ, the fraction of the surface covered, is given by

$$\theta = \frac{C_O}{C_m} = \frac{K_1 p}{K_1 p + Z_m \nu \exp(-E/RT)}$$

The value of average velocity as obtained from Maxwell's law of distribution of velocity has been used further

$$K_1 = N/(2\pi MRT)^{1/2}$$

Thus, we have

$$C_o = \frac{C_m B' p}{1 + B' p}$$

Where $B' = K_1 \exp(E/RT)/Z_m \nu$

It should be noted that B' would be constant at constant temperature.

The number of molecules striking per unit area per unit time is related to pressure P when the Maxwell-Boltsmann distribution law is used. In the actual rocket conditions, where the situation is far removed from the equilibrium, this would be related to P^n, where n is any index ($0 < n > 1$), so that for such a case

$$C_o = C_m B' p^n/(1 + B' p^n)$$

Since C_F^S is a constant and is given by

$$r = k' B' p^n/(1 + B' p^n)$$

Where k' is a constant, and is given by

$$k' = C_F^S C_m$$

Hence,

$$\dot{r} = r/\rho_f = \frac{k'}{\rho_f}\left[\frac{B' p^n}{1 + B' p^n}\right]$$

The equation predicts that, at very high pressure, when a complete monolayer has been formed,

$\dot{r} = k''$, since $B' P^n >> 1$

When $B' P^n < 1$, we would have

$$r = k'' B' p^n$$

Further, we make $b = k'' B'$

$$r = b p^n$$

This is the formula when the flow rate is relatively high and is also used as the model we chose to fit our solid-fuel regression rate.

5. Application

5.1 Instantaneous Regression Rate

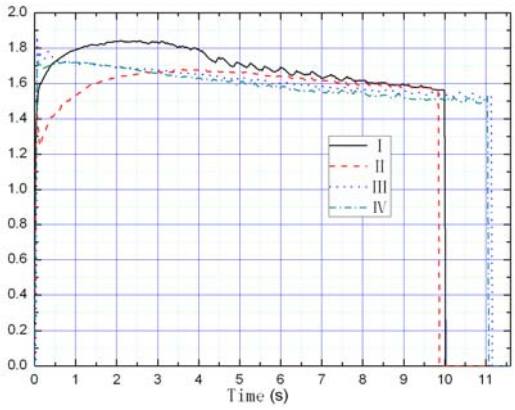


Fig. 26 Instantaneous Regression Rate

The fitted instantaneous regression rate (use Model $\dot{r} = 1.2929\, p^{0.21734}$) is shown in Fig. 26. We can see that the instantaneous regression rate is getting lower while the time's passing.

5.2 Hybrid Motor Thrust

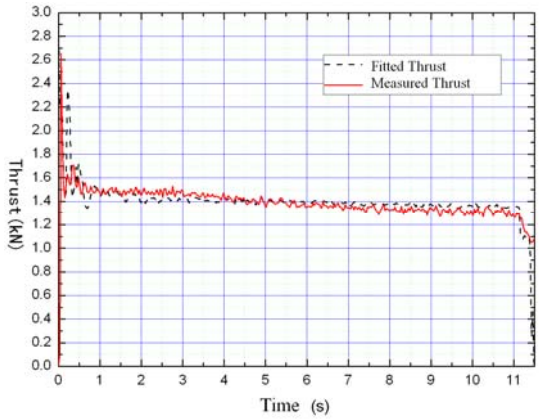


Fig. 27 Measured and Fitted Thrust

Use the iterative method, the initial condition is

$$\begin{cases} r_{p0} = 15mm \\ \dot{m}_{p0} = 0 \\ \dot{m}_0 = \dot{m}_{O0} + \dot{m}_{p0} \\ F_0 = C_F C^* \dot{m}_0 \end{cases}$$

The iterative process is

$$\begin{cases} r_{pi} = r_{p(i-1)} + \dot{r}_{(i-1)} \bullet \Delta t \\ \dot{m}_{pi} = \pi(r_{pi}^2 - r_{p(i-1)}^2) L \rho_f / \Delta t \\ \dot{m}_i = \dot{m}_{Oi} + \dot{m}_{pi} \\ F_i = C_F C^* \dot{m}_i \end{cases}$$

Where $\dot{r}_i$ is fitted by $\dot{r} = 1.2929\, p^{0.21734}$. Fig. 27 is the comparison of the measured and fitted motor thrust. We can see that the theoretic curves fit well with the experiment results, which validate the method and regression research in this paper.

6. Conclusion

1) Results of regression rate were analyzed by three kinds of regression rate models. Through comparative analysis, a fuel regression rate formula was fitted - $\dot{r} = 1.2929 p^{0.21734}$. The regression rate is nearly independent of the oxidizer flux but varied markedly with the chamber pressure.

2) The mechanism was discussed using the heterogeneous reactions theory.

3) The theoretic curves fit well with the experiment results, which validate the method and regression research in this paper.

Reference

[1] Marxman, G. A., and Gilbert, M., "Turbulent Boundary Layer Combustion in the Hybrid Rocket," Ninth International Symposium on Combustion, Academic Press, New York, 1963, pp.371-383.

[2] Marxman, G. A., Wooldridge, C. E., and Muzzy, R. J., "Fundermentals of Hybrid Boundary Layer Combustion," Heterogeneous Combustion, edited by H. G. Wolfhard, I. Glassman, and L. Green, Jr., Vol. 15, AIAA Progress in Astronautics and Aeronautics, Academic Press, New York, 1964, pp. 485-521.

[3] Marxman, G. A., "Combustion in the Turbulent Boundary Layer on a Vaporizing Surface," Tenth Symposium (International) on Combustion, Combustion Inst., Pittsburgh, PA, 1965, pp. 1337–1349.

[4] Marxman, G. A., "Boundary-Layer Combustion in Propulsion," Eleventh Symposium (International) on Combustion, Combustion Inst., Pittsburgh, PA, 1966, pp. 269–289.

[5] Smoot, L. D., and Price, C. F., "Regression Rates of Nonmetalized Hybrid Fuel Systems," AIAA Journal, Vol. 3, No. 8, 1965, pp. 1408–1413.

[6] Smoot, L. D., and Price, C. F., "Regression Rates of Metalized Hybrid Fuel Systems," AIAA Journal, Vol. 4, No. 5, 1965, pp. 910–915.

[7] Smoot, L. D., and Price, C. F., "Pressure Dependence of Hybrid Fuel Regression Rates," AIAA Journal, Vol. 5, No. 1, 1966.

[8] Tian Hui，Wu Jun-feng，Yu Nan-jia，LI Jun-hai，Cai Guo-biao, "Experimental Reaearch of Regression Rate of N2O and Metalized HTPB Hybrid Rocket Motor", Journal of Propulsion Technology, Vol. 35, No. 3, 2014.

[9] Rastogi, R. P. and Deepak, D., "Pressure Dependence of Hybrid Fuel Burning Rate," AIAA Journal, Vol. 14, No. 7, 1976, pp. 988–990.

STU-2: A CubeSat Mission for Earth Observation and Marine/Air Traffic Monitoring

Shufan Wu, Wen Chen, Caixia Chao, Chuanxin Zhang, Zhongcheng Mu, Tao He, Zhiqiang Jia, Ziyi Wu, Ya Dai

Shanghai Engineering Centre for Microsatellite, Chinese Academy of Science
Haike Road 99, Pudong District, 201203 Shanghai, P.R.China
shufan.wu@mail.sim.ac.cn

Wenhe Liao, Xiang Zhang

Nanjing University of Science and Technology
200 Xiaolingwei Street, 210094 Nanjing, P.R.China

Abstract: The STU-2 mission is with 3 CubeSats aiming for polar region Earth observation and marine/air traffic monitoring, as well as new technology demonstration. It carries three main payloads, an optical camera for observing icing situation in polar region, an AIS receiver for monitoring marine traffic, and an DS-B receiver for collecting aircraft traffic information, allocated into three independent CubeSats, with the idea of distributed payloads in spacecraft networks. The 3 CubeSats are interconnected via inter-satellite-links. This paper presents the mission targets and scenario, the satellite system design, the innovative technologies and instruments or devices being used, and the new technologies to be demonstrated on-board. On 25th Sept. 2015, the three CubeSats have been successfully launched into the expected orbit. All of them have gone through the in-orbit commissioning phase, reached into a good health status, with some in-orbit data well received.

KeyWords: Nano Satellite, CubeSat, Satellite networking and constellation, Northern sea route, AIS, ADS-B.

I. Introduction

In 1999, the CubeSat concept was developed by California Polytechnic State University and Stanford University, which sets up a set of standard or specification to help universities worldwide to perform space science and exploration[1]. A CubeSat is a type of miniaturized satellite for space research that usually has a volume of exactly one liter (10cm cube), has a mass of no more than 1.33 kg, and typically uses commercial off-the-shelf components (COTS) for its electronics. Since then, the CubeSat standard has been worldwide accepted and inspired many universities to develop their own CubeSat project and technologies, which further produced a lot of spin-off high-tech companies dedicated to the CubeSat community. The standard 10cm×10cm×10cm basic CubeSat is called a '1U' CubeSat meaning one unit. CubeSats are scalable along any one axis, by 1U increments, based on that, CubeSats such as '2U' CubeSat, '3U' CubeSat, and 'nU' CubeSat have been built and launched, or under developments. The CubeSat standard has brought the NanoSat world into a 'CubeSat Era'[2,4,5].

In 2014, the world has seen 158 nano/microsatellites (below 50kg, and most of them are CubeSats) being launched, representing an increase of nearly 72% compared to 2013[3]. Among them, 107 were built for commercial purposes and missions, showing a big increasing in commercial application sector for CubeSats. It proves that after a decade's development, CubeSat has evolved from purely educational tools towards a useful platform for technology demonstration and scientific instrumentation as well as commercial application, and has walked out of university labs into commercial sector targeting for both technologies and applications.

CubeSat has been found applications mainly in three major sectors[4,6], education and training, technology demonstration, and commercial/operational applications, where the applications sector is becoming more and more weighted in on-going Nano satellite missions. In 2014 CubeSat has been selected as one of the 10 top scientific achievements or breakthroughs by the Science magazine. Looking at the application sector of CubeSats, two lines of movements can be observed, one is the ever increasing performance and capacity with individual CubeSat, the other is the usage of network and/or constellation of multiple CubeSat carrying on either same payloads to enhance re-visit frequency and coverage or different payloads to perform more complicated tasks. A good example is the Planet Labs Flock family, which has been already launched 97 3U CubeSats[3], and plans to launch 131 CubeSats by mid 2015[7]. With a 3U CubeSat, it can achieve Earth observation images at meters level ground resolution. Meanwhile, thanks to its low cost, a network of a few hundred CubeSats become affordable and commercially feasible. Even large constellations of satellites are also proposed by multiple companies, as example, SpaceX and OneWeb announced their plans to pursue sizeable constellation of 4025 and 648 satellites respectively (they are not necessarily CubeSats but believe to be within the catalogue of nano/micro satellites for cost consideration).

For CubeSat network and/or constellation, two directions are further diverged, one is to increase the number of satellites to enhance coverage or re-visit frequency, where all the satellites carry the same payloads with similar function, the other is to use a network and/or constellation with multiple CubeSats carrying different payloads to perform an integrated task or service. The former direction has been demonstrated already by the mentioned examples with Planet Labs, SpaceX, and OneWeb, while the later direction is more oriented towards distributed payloads or multiple functional satellites forming a network to provide integrated services, which is foreseen an important direction for many application programs with low-cost CubeSat. This paper presents a small CubeSat constellation or network, named the STU-2 mission (also named the TW-1 mission during the development phase), to explore the integrated services for Earth observation and marine/air traffic information collection and monitoring, based on the low-cost and off-the-shelf CubeSat technologies. It consists of 3 CubeSats, by one 3U CubeSat and two 2U CubeSats, carrying different payloads and instruments, being launched into an LEO SSO orbit, forming an along-trace satellite network and/or constellation. The main tasks and objectives of this mission are listed in the following:

- ✧ Monitoring the sea ice situation, especially for the North and South polar regions, with an on-board optical camera,
- ✧ Monitoring global marine traffic information with an on-board AIS receiver,
- ✧ Monitoring global air traffic information with an on-board ADS-B receiver,
- ✧ Demonstration of CubeSats inter-satellite link and networking technologies,

✧ Demonstration of new technologies, components and subsystems for CubeSat, including a SoC dual-band GPS/Beidou receiver, a MEMS cold-gas micro-propulsion system, and a multi-channel inertial sensor.

This mission has been launched into an SSO orbit at 481km, by piggy-backing the Long March 11 with its maiden flight on Sept 25th 2015, in Jiuquan, China. This paper will discuss the mission and the satellite design, the innovative instruments or devices being used, and the new technologies to be demonstrated on-board.

II. Mission Requirements Analyses

The performance and capacity of an individual CubeSat is always constrained by its limited volume, mass, and especially available power onboard. While many space missions need multiple functions linked each other to be provided for different customers to meet multiple needs. In the mission to be discussed in this paper, the following three functions and payloads are required to perform the mission task

- ✧ Monitoring the icing situation for the North and South polar regions
- ✧ Monitoring marine traffic information with an on-board AIS receiver,
- ✧ Monitoring air traffic information with an on-board ADS-B receiver.

In addition, the following new technologies and products are planned to be demonstrated through this mission

- ✧ Demonstration of CubeSat inter-satellite link based on CubeSat Space Protocol(CSP)[11];
- ✧ Demonstration of CubeSats adaptive networking technologies based on Gamalink[12]
- ✧ Demonstration of new technologies for CubeSat, including a SoC dual-band GPS/Beidou receiver, a MEMS cold-gas micro-propulsion system, and a multi-channel inertial sensor.

II.I Polar region icing situation awareness

The amount of ice in polar arctic regions has been receding progressively every summer as a result of global climate changes in the past decades. For the northern artic, these changes have conveniently allowed commercial vessels to pass this area in summer, emerging a marine route called “Northern Sea Route”, which attracts high interests from shipping companies. For example, the Northern sea route allows a vessel sailing from Dalian in China to Rotterdam in Europe to complete the journey in about two-thirds the distance of a typical route via the Strait of Malacca and the Suez Canal, as illustrated in Fig. 1, and in only half the distance of a route via the Cape of Good Hope.

Fig 1. Northern Sea Route

Using these Northern sea routes means faster shipments, reduced fuel expenses and a smaller impact on the environment. However, it is essential to monitor the ice situation in northern arctic region to provide navigation for the Northern sea route, for vessels to avoid dangerous icebergs floating in the sea even in mid of summer. The microsatellite WINSat-1 was dedicated to this purpose to monitor the icing situation in arctic region, helping provide navigation information for vessels in the Northern sea route[8].

Apart from commercial vessel voyage, both arctic and Antarctic regions are of great values and interests of scientific investigation. The XueRong icebreaking research vessel has made more than a dozen voyage to both arctic and Antarctic regions. In 2012, it has made a voyage through the arctic ocean to Europe amid the record ice melt, passing through the Bering Strait before joining the westbound on the Northern Sea Route. It provides annual resupply to Antarctic Zhongshan Station at Prydz Bay. In particular, it participated the rescue operation of Akademik Shokalskiy during Dec 2013 and Jan 2014, where several icing breaking science vessels were involved. For all these marine journey and operation in arctic and Antarctic regions, information on icing situation is very useful to help navigate the sailing route.

So the first mission requirement is trying to meet all these practical requirements for both the commercial cargo vessels to navigate through Northern Sea Route, and for the science research vessels to explore arctic and Antarctic regions. An appropriate camera is needed as one prime payload to meet these requirements.

II.II Marine Traffic Monitoring from Space

Marine Automatic Identification System (AIS) Class A transceiver, with 12W output power, is used for ships over 300 ton and all passenger ships, to broadcast its identification, position, course and speed to other ships and land station to avoid collisions. Smaller ships often carry low-cost Class B transceivers with 2W output power. AIS is helpful to ensure the safety of ships, for example in case of hijacking, and for search and rescue operations in case of emergencies. Ships fitted with AIS receivers and transponders can be tracked by AIS base stations located along coast lines or, when out of range of terrestrial networks, through a growing number of satellites that are fitted with special AIS receivers. AIS transponders have horizontal range that is highly variable, but typically only up to about 100km. They reach much further vertically – up to about 1000km, which makes it possible to use satellite-based AIS receiver to collect their signals, though AIS is not initially meant for satellite-based application. A global information system is, therefore, only possible via space-based receivers.

There is apparently a growing interest in satellite-based AIS, which was first proposed in 2003 during the IAA 4th Earth Observation of Small Satellite Conference. In 2004, FFI of Norway further investigated the feasibility of AIS messages receiving onboard small satellites, and produced prototypes. In 2005 and 2006, Norway launched two 1U CubeSat, NCube-1 and NCube-2, with space-based AIS receiver for test. In 2010, Norway launched AISSAT-1, which is a Nanosatellite and weighs 6 kg, sized by 20x20x20 cm. AISSAT-1 works on polar orbit with a height of 600 km, and successfully collects and transmits downwards all the ship information about Norway's surrounding sea in a great amount. Besides, the U.S launched TACSAT-2 in 2006, which carried AIS receiver to monitor ships within 2000 sea miles for United States' Coral Guard. After that, countries including Canada, Japan, and multiple European countries followed to study space-based AIS receiver, and have built multiple regional AIS system as a result. ORBCOMM became the first commercial service provider of satellite AIS since early 2009. Its next 18 ORBCOMM Generation-2 (OG-2) satellites plan to carry AIS receivers.

This second mission requirement is trying to further enhance the marine traffic monitoring in arctic and antarctic region, and meanwhile, to explore technical approach and commercial potentials of satellite-based AIS monitoring network.

II.III Air Traffic Monitoring from Space

The Automatic Dependent Surveillance – Broadcast (ADS-B) system is today a standard equipment on civil aircraft, transmitting periodically data packages containing information on key data such as aircraft ID, position, altitude and intent. It is designed for terrestrial based ground station to monitor air traffic flow in certain regions. Space based ADS-B is the idea to place sensitive receivers on board satellites in orbit, which can receive ADS-B packages and relay them to the relevant ground stations. This idea becomes very attractive and interesting in recent years for micro/nano satellite application.

In April 2013, ESA launched PROBA-V, weighing 140 kg at an altitude of 820 km in the orbit of SSO. It demonstrated for the first time a space based ADS-B receiver. With antenna FOV wider than 500 km, PROBA-V's ADS-B receiver coverage is a lot bigger than that of ground ADS-B system. It is capable to receive around 12,000 ADS-B messages in two hours. Furthermore, ADS-B receiver has been considered as payloads of Iridium Next, who plans to launch from 2015 to 2017 66 LEO communication satellites, distributed in 6 orbital planes (11 satellite per orbital plane). Carrying ADS-B receiver for each satellite, Iridum Next is supposed to enable any single satellite to monitor planes within sights as long as 9 minutes. Produced by Harris, ADS-B receiver for Iridium Next is capable to simultaneously handle 10,000 targets, and even meets the requirements of ADS-B message collection in heavy-transported space region. This project is supposed to achieve global ADS-B message collection and centralized data processing. It is expected that related business and service supports could follow the project to be further developed better.

In November 2013, GOMSpace launched GOMX-1, the first CubeSat that carries an ADS-B receiver[9], produced with COTS components. GOMX-1 proves the feasibility of CubeSats to receive ADS-B messages in space, and provides valuable space data.

The third mission requirement is meant to explore the related technologies and potential application of the space-based ADS-B idea, to help enhance air traffic management from space.

II.IV New Technology/Product Demonstration

CubeSat community is fast developing with new technologies, components, and products. In orbit demonstration (IOD) of new technologies and products is also one of the main application

fields of CubeSat. A few new technologies and/or products are carried on-board the STU-2 satellites for IOD purposes.

Orbit determination is always needed for any mission. For satellite constellation or network, it is more essential to have exact orbit information of each satellite. A chip-sized dual band GPS/Beidou receiver is carried on-board STU-2A and 2C satellites, which is featured with very small size and low power consumption. On-board STU-2B, two more different dual-band GPS/Beidou receivers are also carried on-board for IOD purposes.

Micro-propulsion system is essential for any orbit correction or maintenance. CubeSat can play more functions and tasks when being placed as a constellation or formation network, where orbital maintenance is a basic building block. STU-2A has carried a MEMS based cold-gas micro-propulsion module, developed by the NanoSpace Ltd from Sweden[10], as its first in-orbit flight demonstration.

A multi-channel inertial sensor with the type ADIS16448， is carried on-board STU-2C, to explore its performance in-orbit and its potential usage for CubeSat missions. It is a commercial product includes a triaxial gyroscope, a triaxial accelerometer, a triaxial magnetometer, and pressure sensors.

Another important technology, which enables satellite networking in space, is the inter-satellite communication and the adaptive networking capabilities of satellites in space. The STU-2 mission will explore this new technology by using this minimum small constellation of 3 satellites in space. The adaptive networking technology is explored by the newly developed Gamalink[12], as an experimental payload for IOD. Meanwhile, CubeSat Space Protocal (CSP)[11] will be used to explore inter-satellite communication between 3 CubeSats in this mission

Further, aiming to enhance the space application ability of the CubeSat, the MEMS cold-gas micro-propulsion module on-board STU-2A will be demonstrated in orbit for the rendezvous and formation flying technologies between two CubeSats in a later stage of the mission.

III. Mission Design with distributed payloads on multiple CubeSat

As each of the main payloads in this mission has been demonstrated to be feasible with a CubeSat platform, thus the mission design of STU-2 follows the idea of using multiple CubeSats to carry distributed payloads, forming a small constellation or network in space, to jointly perform all the required tasks. Therefore, three CubeSats are designed for this mission, i.e. one 3U CubeSat STU-2A, one 2U CubeSat STU-2B, and one 2U CubeSat STU-2C, each carries different mission payloads and testing payloads, as listed in Table 1.

Table 1 Mission design and satellites configuration

CubeSat	STU-2A	STU-2B	STU-2C
SatelliteBus	3U CubeSat	2U CubeSat	2U CubeSat
Main Payloads	Camera	AIS Receiver	ADS-B Receiver
New Technology	BD/GPS, Micro-Prop. Gamalink	GPS/BD, Gamalink	BD/GPS Inertial Sensor

The mission scenario of STU-2 is illustrated in Fig.2. three CubeSats forms a small network of satellite to take pictures of the Earth surface, to monitor marine traffic via AIS signal collection, and

to collect aircraft flight information via ADS-B receiver, meanwhile, three satellites can communicate each other through inter-satellite links and/or network capability provided by Gamalink and CSP protocol.

In this mission, the three main payloads are distributed at three separate CubeSats, instead of one satellite. The advantages of this distribute approach can be listed below:

- Multiple functions to be performed by multiple satellites
- Very low cost thanks to the available commercial products for CubeSats
- Extendable for more functions with additional satellites being added later on
- Flexible for future extension of the network to enhance re-visit frequency

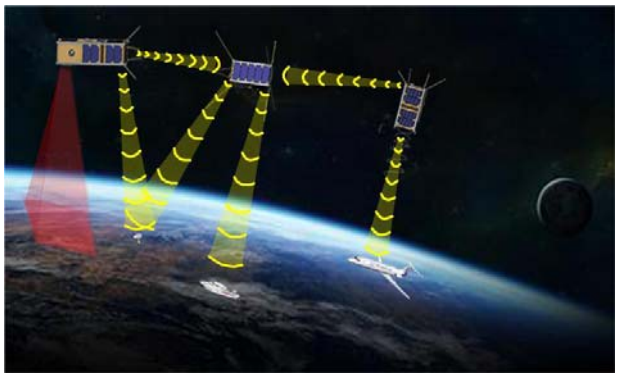


Fig. 2 STU-2 Mission Scenario

IV. Satellite Design and Mission Implementation

IV.I STU-2A satellite design

STU-2A is a 3U CubeSat with body mounting solar panel, passive thermal control except for battery where a heater is applied to ensure certain temperature. 3-axis attitude stabilization and control were designed based on fine Sun sensor, Star tracker, reaction wheels, and micro-propulsion. Considering the payload data requirement, STU-2A carries an S-band transmitter with 125 kbps data rate. The whole satellite picture before integrating into the POD is shown in Fig.3.

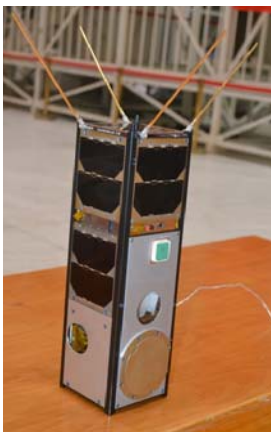


Fig.3 STU-2A

STU-2A main payload is an optical camera for polar region observation for icing situation, with a resolution of 100m and a swath of 200km from an orbit of 481km. Because of limited data downlink rate, the camera is expected to operate in inconsecutive mode rather than consecutive video mode. Meanwhile, it is expected that the two sequential images should overlap a bit to have a continuous observation of the polar region. By preliminary estimation, two sequential images might

be captured with an interval of 10-20s. And based on the preliminary mission analysis, the time accessing to Arctic is 7-11 minutes per orbit. It is expected that the camera could operate for 5-10 orbits during a day. Based on onboard data storage volume, downlink access windows, transmitter downlink rate, and expected satellite operations, the final specifications for flight model are given in the following table. The flight model is shown in Fig. 4.

Table 2 Camera Specification

Items		Specifications
Structure	Mass	466g
	Dimension	90×90×85mm3
Electrics	Power	<8.2 W (ave) <9.5W (peak,<10ms)
Observation	Revolution	93m
	Swatch	217.8km
Optics	Band range	470~650nm
	Lens	40.5mm
	FOV	2w≥30°
	MTF	0.23
Imaging	Frame frequency	1/5，1/10，1/15fps
	Quantization Digit	8bits
	Operation temp	-25°~45°

Fig.4 Flight model for Camera

The other payloads for technology demonstration include a dual band BD/GPS receiver, micro-propulsion module, and an integrated communication module (the Gamalink module). The dual band BD/GPS receiver is an in-house developed system on a chip (SOC) receiver, with a size of 22.4 mm 17 mm 2.2 mm, mass of 4g, power consumption of 0.5W, and operation temperature range of [-40 85] deg. It can achieve a position accuracy of 5m in horizontal and 10m in altitude, a velocity precision of below 1 m/s, with a sensitivity of below -130dB. Its cold-start time is within 1 minute and warm start time within 10s. Its physical picture is illustrated in Fig 5.

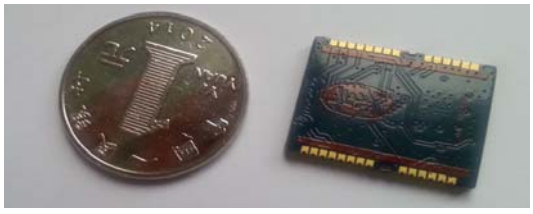


Fig. 5 Dual band BD/GPS receiver

The micro-propulsion module in Fig.6 is a newly developed cold-gas propulsion system based on MEMS technology. It is contributed by NanoSpace from Sweden[10], based on its long-history space engineering heritage, the main specifications are given in the following.

- Four 1mN thrusters with closed loop thrust control
- Thrust resolution: <10μN
- Propellant: Butane
- Total impulse: 40Ns
- Isp: 90s
- Size: 10x10x5cm^3
- Mass including the propellant: 330g
- Power consumption: 3.2 W

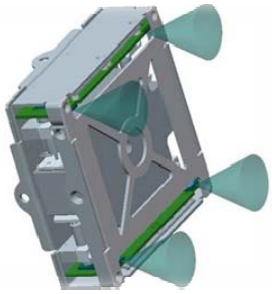


Fig.6 Micro-propulsion module

The GAMALINK is also a new and advanced component developed at the Tekever Ltd in Putogeese, which is mainly based on two novel technologies: Software-defined Radio (SDR) and Mobile wireless Ad hoc NETworks (MANETs)[12]. The GAMALINK device for the space segment includes the ad hoc network capability for Inter-Satellite Link (ISL), a GPS receiver and a radio-based attitude determination algorithm.

Based on the mission objectives and tasks assigned to STU-2A, the satellite system configuration is specified as given in Table 3. A brief introduction for every subsystem comes along in the following.

Table 3 STU-2A satellite specification

Subsystem	Item	Specification
Structure	Dimension [mm]	340.5x100x100
ADCS	Attitude Knowledge	1°(3σ)
	Pointing Accuracy	2°(3σ)
	Pointing Stability	0.1°/s
Thermal	Internal temperature	-10℃ ~ +35℃
EPS	Bus voltage	13.2 V ~ 16.8V
	Battery properties	2.6 Ah，1 Year
TT&C	Frequency	UHF(435-438 MHz）
	Modulation	2-FSK
	Uplink	4.8 kbps

Subsystem	Item	Specification
	Downlink	4.8 kbps
S-band transmitter	Date rate	125kbps
	Frequency	2.425GHz
	Modulation	QPSK
	BER	<10-6
OBC	Process capacity	20 MIPS
	Process storage	RAM >2 M, Flash>256 K

For ADCS subsystem, three-axis stabilization and control are implemented by applying reaction wheels combined with the magnetic control to enhance the TW-1A ability. The ADCS consists mainly of magnetometer, magnetorquer being integrated in the solar panels, coarse fine Sun sensor, fine Sun sensor, micro Star tracker, and 3 reaction wheels, as well as a micro-propulsion module for technology demonstration.

For data transmission, an S-band transmitter is used, working at a frequency around 2.42GHz, providing a data rate of 125kbps.

For TM/TC, an UHF band transmitter is used at a frequency around 435 MHz. The data rate is chosen to at 4.8kbps to leave more communication budget margin, though the selected transmitter has a much higher capacity.

The OBC is using a mature product, NanoMind A712D from GOMSpace, which is an efficient system for space applications with limited resources, e.g. for CubeSat or nano-satellite missions. It provides on-board a 3-axis magnetometer to sense the Earth's magnetic field and coil drivers that can be used to implement attitude control based on magnetic sensing and actuation.

For EPS, a space-proven COTS product for CubeSat is chosen. Its photovoltaic power can converse up to 30W. Maximum power point tracking (MPPT) of panels is adopted for optimal power generation. It has 3 input channels which can set the power-point with 3 different options independently. The measures for battery under-voltage and over-voltage protection are taken. The number of the regulated power buses has two: 3.3V@5A and 5V@4A. Three main electrical and mechanical interfaces are included, which are separation-switch interface, Remove-Before-Flight-pin interface and I2C interface.

IV.II STU-2B satellite design

STU-2B is a 2U CubeSat, with its main payload being an AIS receiver, to receive marine traffic information of the ground. The AIS receiver is provided by the Satlab Ltd from Denmark， with a heritage from the AAUSAT3. The CubeSat-level AIS receiver is a fully self-contained SDR, which is suitable for LEO satellite missions. Weighing less than 50 g and using 0.7 W during full operation load, this versatile SDR offers the best performance possible given the typical constraints of a CubeSat.

The satellite bus is mainly developed by a team from the Nanjing University of Science and Technology. To enhance on-board power supply, body-mouting solar cells and two deployable small solar panels are used, as shown in Fig.7. The momentum-biased three-axis stabilization and control is implemented via two momentum wheels mounted on one axis (with one-failure redundancy) combined with the magnetic control strategy. For OBC and EPS and TT&C subsystem, the same products with STU-2A were chosen. The main design specifications of STU-2B are given in Table 4.

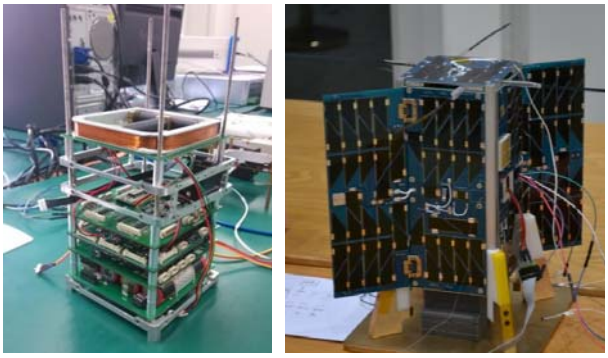


Fig.7 STU-2B

Besides the main payload of an AIS receiver, it carries also two new models of dual band GPS/BD receiver, which are different version from that on-board STU-2A, mainly for technology demonstration and new product development/verification.

Table 4 STU-2B satellite specification

Subsystem	Item	Specification
Structure	Dimension envelope	239 x 100 x 100 mm^3
ADCS	Attitude Knowledge	5°(1σ)
	Pointing Accuracy	10°(1σ)
	Pointing Stability	0.5°/s
Thermal	Internal temperature	-10℃ ~ +35℃
EPS	Bus voltage	6.4V ~ 8.4 V
	Battery properties	5.2 Ah，1 Year
TT&C	Frequency	UHF（435-438 MHz）
	Modulation	2-FSK
	Uplink	4.8 kbps
	Downlink	4.8 kbps
OBC	Process capacity	20 MIPS
	Process storage	RAM >2 M, Flash>256 K

IV.III STU-2C satellite design

STU-2C in Fig.8 is a standard 2U CubeSat, with its main payload being an ADS-B receiver, to monitor civil airplane flying within the nadir space region of the satellite. The payload is provided by GOMSpace, upgraded from the heritage and experiences of the GOMX-1 mission[11].

For ADCS, the pure magnetic control scheme is implemented with the magnetic coils being integrated in the solar panel. The design specification for STU-2C is given in Table 5.

The other payloads for demonstration include a dual band BD/GPS receiver and a multi-channel inertial sensor. The BD/GPS receiver is using the same chip sized receiver as that in STU-2A. The inertial sensor is using the commercial component of ADIS16448 with including a triaxial gyroscope, a triaxial accelerometer, a triaxial magnetometer, and pressure sensors. Each sensor has its own dynamic compensation formulas that provide accurate sensor measurements. The

ADIS16448 is packaged in a module that is approximately 24.1 mm × 37.7 mm × 10.8 mm and has a standard connector interface.

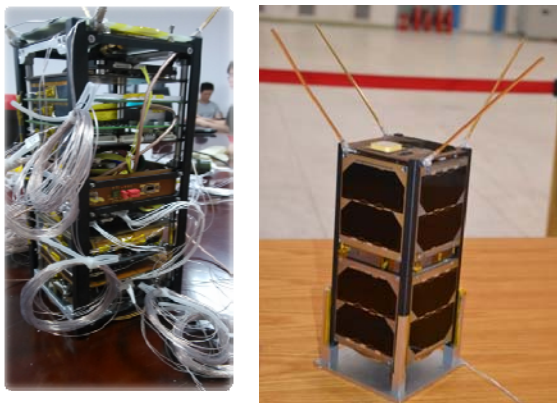


Fig.8 STU-2C

Table 5 STU-2C satellite specification

Subsystem	Item	Specification
Structure	Dimension envelope	239 x 100 x 100 mm^3
ADCS	Attitude Knowledge	5°(1σ)
	Pointing Accuracy	10°(1σ)
	Pointing Stability	0.5°/s
Thermal	Internal temperature	-10℃ ~ +35℃
EPS	Bus voltage	12.0V ~ 16.8V
	Battery properties	2.6 Ah，1 Year
TT&C	Frequency	UHF（435-438 MHz）
	Modulation	2-FSK
	Uplink	4.8 kbps
	Downlink	4.8 kbps
OBC	Process capacity	20 MIPS
	Process storage	RAM >2 M,Flash>256 K

IV.IV Design Summary

The overall design and final implementation results of the three satellites are summarized in Table 6.

Table 6 Main specifications for flight model

ITEM	STU-2A (3U)	STU-2B (2U)	STU-2C (2U)
Size [mm]	343.3×114×114	239×114×114	239×114×114
Temperature[°C]	[-5 +25]	[0 +40]	[-8 +36]
Bus Voltage Battery	12.0~16.8V 2.6Ah Lithium	6.4~8.4V, 5.2Ah Lithium	12.0V~16.8V 2.6Ah Lithium
ADCS	3 Wheel control AME: 0.2°(3σ) APE:1.8deg RPE: 0.03deg/s	Momentum biased AME:4deg APE: 8deg RPE: 0.04 deg/s	Magnetic Control AME: 4deg APE: 13 deg RPE:0.04 deg/s
Propulsion	45g Butane Thrust: 4mN Total Imp. 40Ns	None	None
Data	S-band,	None	None

ITEM	STU-2A (3U)	STU-2B (2U)	STU-2C (2U)
Transmitter	125Kbps		
Mass [kg]	2.9kg	2.2	1.7
Power Cons. [W]	2.9	2.9	2.0
On-board computer	Processor: ARM7, Speed: 20MIPS， Operating System: FreeRTOS Data Bus: I2C, Connector: PC104		
TMTC	UHF-band，Uplink: 4.8kbps，Downlink: 4.8kbps		

V. Launch Campaign and in-orbit results

The STU-2 mission with three CubeSats has been launched successfully into space on Sept 25th, 2015, on-board a newly developed small launch vehicle in China, the LoangMarch-11 launcher, towards the target LEO orbit of SSO, with an altitude of 481km, inclination of 97 deg, and LTDN at 8:00am. At the moment of writing the paper, all three CubeSats are at their check-out operation and in-orbit commissioning mode. Fig 9 illustrated the final configurations of the three CubeSats ready for launch.

On the second day after the launch, Sept 26th 2015, when the STU-2C passes over Shanghai ground station at 18:15pm, the ADS-B receiver on board the CubeSat was switched on via tele-command for first trial. Immediately the STU-2C captured ADS-B signals transmitted from 16 aircraft flying under the satellite. After one orbit, when the STU-2C passes over Shanghai at 19:50pm, the STU-2C has received 51665 ADS-B signals, transmitted from 405 different aircraft flying along the region below that satellite orbit. During the in-orbit commission phase till Oct 19 2015, ADS-B signals from more than 10000 aircraft have been successfully collected and transmitted downwards by the STU-2C

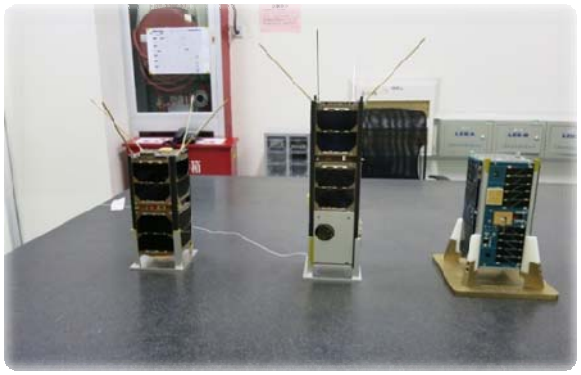


Fig.9 STU-2 CubeSats ready for launch

Figure 10 illustrated the distribution of these ca 10000 aircraft distributed worldwide, which show a very busy air traffic in south-east Asia and North Atlantic region and west Europe, including a lot of flights over the Pacific and Atlantic oceans, and flights over Arctic circle, while much less flights are observed in Africa, South America, and Australia as well. Also very few flights are seen in some sensitive or impractical regions, like North Korea, Libya, Antarctic region and Tibet plateau, etc. These observation fit well with the practical situation

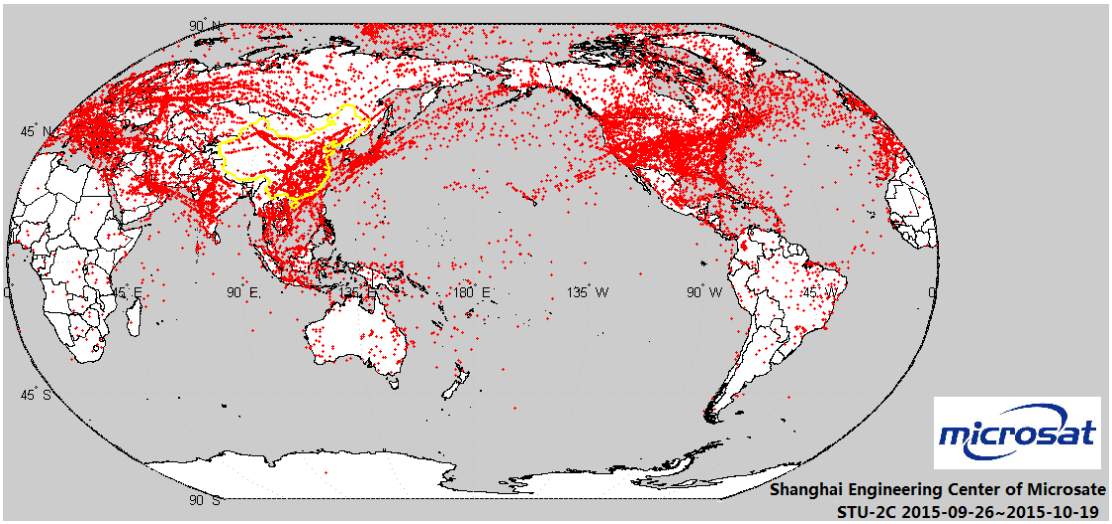


Fig.10 STU-2C captured aircraft distribution within one orbit on Sept 26th 2015

VI. Summary

This paper presents a small Nano satellite constellation, the STU-2 mission, to explore the integrated services for Earth observation and marine/air traffic information collection and monitoring, with distributed payloads among three CubeSats, based on the low-cost and off-the-shelf CubeSat technologies. It consists of 3 CubeSats, with one 3U CubeSat and two 2U CubeSats, carrying different payloads and instruments, which have been launched into an SSO LEO orbit at 481km on Sept 25 2015, forming an along-trace satellite network and/or constellation. The main tasks and objectives of this mission are listed in the following:

- Monitoring the sea ice situation, especially for the Northern and Southern polar regions, with an on-board optical camera,
- Monitoring global marine traffic information with an on-board AIS receiver,
- Monitoring global air traffic information with an on-board ADS-B receiver,
- Demonstration of CubeSat inter-satellite link based on CubeSat Space Protocol (CSP);
- Demonstration of CubeSat adaptive networking technologies based on Gamalink
- Demonstration of new technologies for CubeSat, including a SoC dual-band GPS/Beidou receiver, a MEMS cold-gas micro-propulsion system, and a multi-channel inertial sensor.

VII. ACKNOWLEDGMENT

This mission is designed and being implemented by a consortium led by the Shanghai Engineering Centre for Microsatellite in China, together with partners including the Nanjing University of Science and Technology (contributed for STU-2B platform), the GomSpace from Denmark (ADS-B and AIS receiver, as well as CubeSat components), the NanoSpace from Sweden (contributed the MEMS micro-propulsion module), and the Tekever Space from Portugal (contributed the Gamalink). The authors would like to express sincere appreciation and many thanks to these partners, together with them this mission has been made feasible and successfully implemented.

References

[1] ‘CubeSat Design Specification’, California State Polytechnic University, http://www.CubeSat.org/images/ developers/cds_rev12.pdf

[2] D. DePasquale, A.C. Charania, ‘ Nano/Microsatellite Launch Demand Assessment 2011’, http://www.sei.aero/ eng/papers/uploads/archive/SpaceWorks_NanoMicrosat_Launch_22Nov2011_revA.pdf

[3] Buchen, E., “2015 Small Satellite Market Observations” http://159.226.251.229/videoplayer/ SpaceWorks_Small_ Satellite_Market_Observations_2015.pdf, Retrieved in May 2015

[4] Su, R., Zhang, K., Song, H. “Summarization of Very Small Satellite Development”, Spacecraft Engineering (a Chinese Journal), Vol 22, No.6, pp104-111, Dec 2013

[5] Wu S. and Chen H., “Micro/Nano Satellite Technologies and Applications in China”, IAC.13.B4.1.5, 64th International Astronautical Congress, Sept. 2013, Beijing, China

[6] Bouwmeester, J., Guo, J., “Survey of worldwide pico- and nansatellite missions, distributions, and subsystem technology”, Acta Astronautica, 67(2010) 854-862

[7] Anon, https://en.wikipedia.org/wiki/Planet_Labs, retrieved in Aug 2015

[8] Anon. WINSat-1, http://en.axelspace.com/solution/wnisat1/, retrieved in Aug 2015

[9] Alminde, L. Kaas, K., Bisgaard, M., Christiansen, J., Gerhardt, D., “GOMX-1 Flight Experience and Air Traffic Monitoring Results”, SSC14-XII-7, 28th Annual AIAA/USU Conference on Small satellites, Aug 2014

[10] Anon, CubeSat MEMES Propulsion Module, http://www.sscspace.com/about-the-ssc-group/ ssc-companies/nanospace/publications-2, Retrieved in Aug 2015

[11] Anon. Cubesat Space Protocol, https://en.wikipedia.org/wiki/Cubesat_Space_Protocol, Retrieved in Aug 2015

[12] Rodrigues, P., Oliveira, A., Mendes, R. “SDR-Based Ad Hoc Space Networks (SASNETs)”, IAC-11-B2.1.3, 62nd IAC, Cape Town, South Africa, 2011

Epsilon Launch Vehicle for Small Satellites

Yasuhiro MORITA

Japan Aerospace Exploration Agency (JAXA), Sagamihara/Tsukuba, Japan

The Epsilon launch vehicle, the newest version of Japan's solid propulsion rocket, successfully had its first flight in September 14, 2013, carrying "HISAKI", the world first extreme ultra-violet planetary telescope satellite, onboard. It should be emphasized that the JAXA appreciates the advantages of combined benefits of the standardized small satellites and the Epsilon's highly efficient launch system to increase the level of space activities. The development of the Epsilon is featured by its universal design, which means a cost effective, user friendly and ultimately efficient launch system. The enhanced user-friendly character reached at the world-leading level both in the orbit injection accuracy and in the riding comfort while the mobile launch control made a revolution in the space launch system. This paper reveals the novel concept of the Epsilon rocket that was realized by the first flight results.

Key Words: Rocket Science, Solid Propulsion, Mobile Launch Control

1. Introduction

The first launch of the Epsilon rocket, Japan's newest version of solid rocket launcher, was conducted in September 2014 carrying a ultra-violet planetary telescope satellite, HISAKI, into its elliptic orbit (Fig. 1). It should be emphasized that the JAXA strongly appreciates the advantages of combined power of the standardized small satellites and the Epsilon's highly efficient launch to increase the level of space activities. The launch site of the Epsilon rocket is the Uchinoura Space Center (USC), the home of Japan's solid rockets, which was modified to be more efficient although it had been a highly compact launch complex. The efficient Epsilon launch vehicle and the compact USC will serve as the most powerful tool that contributes to small missions (tentatively, maximum 1.2 ton into LEO and 450kg into SSO, as of the first flight, that will be increased more than 30% by the evolution program now underway for the second flight).

The development of the Epsilon is featured by its universal design, which means a cost effective, user friendly and ultimately efficient launch system. Quite recently, the space transportation is getting into real business. In this situation, the most important thing, of course, must be the cost and performance. But is it enough? The answer is No. So what else? It is the universal design concept of the Epsilon launch vehicle including additional values that should satisfy both customers and stake holder at a higher level. From the point of view of payload, such additional value can be associated with the user-friendly character, which means emotionally the hospitality, or OMOTENASHI spirit to respond any special customer requests as much as possible while physically the ridding comfort and operability. The first flight results demonstrated that the enhanced user-friendly character reached at the world-leading level both in the orbit injection accuracy and in the riding comfort. In addition, payload can be accessible mechanically until 3hours before the launch. The additional value associated with the space transportation technology should be technical innovation to make us to open up the future. The world first mobile launch control, established through the first flight, made a revolution in the space launch system[1-6]. This converted science fiction to the science fact.

Such novel ideas will not only make the Epsilon rocket far ahead of the world but will also bring transportation technologies to a higher level. They will become the world standard. In addition, the simplification of such launch control should be the technology that is indispensable to future reusable rocket systems. In this way, the concept of Epsilon is beyond the scope of mere solid-propellant rockets. Rather the Epsilon rocket aims at achieving the innovative transportation technologies that can be equally applicable to liquid fuel rockets as well as future space transportation systems.

Now that the first launch was finished, the most important is what the next step should be beyond the Epsilon. JAXA has already started the Enhanced Epsilon program and the target of the next innovation is higher performance and lower cost with the goal at a challenging level. The development is already underway to evolve the Epsilon rocket and enhance the space transportation technologies in the larger scale. The Enhanced Epsilon will have the launch capacity increased by more than 30% and it will be applied to the second flight of Epsilon that is scheduled to be carried out next year.

Fig. 28: The Epsilon launch vehicle had its first lift-off.

2. Universal Design of Epsilon and First Flight results

2.1 Vehicle Configuration of Epsilon

As a next generation launch vehicle, the Epsilon rocket has unique emphasis in development on:

- ✓ Optimization of the system for better cost and performance,
- ✓ Enhancement of user friendly character for more customer satisfaction, and
- ✓ Revolution of the launch system for more efficient launch.

The configuration of the Epsilon rocket is a three-staged solid propellant vehicle, having a 450 kg payload capacity into a SSO at 500 km height (Table 1). As already mentioned, it will be increased more than 30% by the evolution program now underway to be applied to the second flight. Each of the first and the second stages has 3-axis attitude control capability while the third stage is just spin-stabilized for more simplicity. In addition to this standard configuration, the newly designed upper stage rocket, Post Boost Stage (PBS), can be onboard to increase the orbit injection accuracy while the special vibration attenuating capability can be added to the Payload Attach fitting (PAF) to get better riding comfort. On the other hand, the Mobile Launch Control has been established to eliminate the inefficiency of the old-fashioned launch operations. Those aspects of Epsilon will be detailed in the following sections.

Table 1: Representative specifications of Epsilon. Note that EE denotes the Enhanced Epsilon.

Items	Specifications
Configuration	3-stage solid propellant launcher with optional PBS (Post-Boost Stage)
Length/Diameter	24 m (26 m for EE)/ 2.5 m
Lift-off mass	92 ton (95 ton for EE)
Launch Capacity	SSO (500 km): 450 kg (590 kg for EE)
Next Generation Technologies	Autonomous Checkout System Mobile Launch Control
First Flight	in 2013 (EE in 2016）
Launch Site	Uchinoura Space Center（USC)

2.2 Compact and High-performance Rocket Design

2.2.1 Simpler manufacture

Our strategy here is to get lower cost, better performance, and higher reliability and one of the keys to success is a revolution of manufacture, which is aimed at converting manufacturing processes to more efficient ones and reducing the number of components and parts at a significant level. Note that one of the key factors pushing up the price of the rockets is the inefficient large scale manufacturing process. A good example for simpler process can be directed to the motor case: the manufacture of the filament wounded motor case used to utilize a complicated autoclave curing that needs to apply high pressure. Is has been converted to mere oven that cures in an easier way under normal pressure. A typical example having reduced number of parts is the payload fairing: it has been made one integrated structure of the half shell formation consisting of a half cone and a half cylinder although they used to be a build-up structure that required a large scale labor and time duration resulting in higher cost. Now they are designed to be solidly constructed together into a single component (Fig. 2). This seems like a small step but may lead to a big leap to reform the current manufacturing process to a much more efficient one. The next step will lead to a conversion of the current metallic rocket structures to integrated CFRP.

Fig. 2: Integrated structure of the payload fairing.

2.2.2 General-purpose product

Another key to success to get lower cost, better performance, and higher reliability is to avoid using special parts and materials that are only dedicated to space applications and to make the most use of industrial general-purpose product. A representative example is the payload fairing again (Fig. 3). For the tip portion of the fairing that is exposed to the most severe thermal condition, the high-performance 3-D C/C material for space use was utilized. Now it has been replaced by general-purpose silicon form for industrial use that is designed to survive over the limited flight time. Such strategy is also taken to innovate onboard avionics systems. The idea is to utilize commercial

parts: those latest components having smaller, lighter and inexpensive characteristics. Note, the same trend can be observed in some satellite development. Of course, special consideration should be necessary for rocket applications with respect to parts reliability because of the limited redundancy of rocket guidance and control subsystem. As the first step of this strategy, the current mechanical relays of PSDB (Power Sequence Distribution Box) on-board the upper stage will be replaced by semiconductor relays for the Enhanced Epsilon to be launched next year.

2.2.3 Collaboration with different cultures and industries

Another key to success to get lower cost, better performance, and higher reliability is the concept of "the space open to other industries" , which means flexible thinking to collaborate with ideas of different cultures and advanced and/or common technologies of other industries that have not been involved in space development at significant level. The space related technologies are sometimes considered highly advanced. It is definitely true for some key components such as liquid fuel engines and solid rocket propellants. But not always. General-purpose components such as avionics systems utilize absolutely outdated technologies for higher reliability. Changing this situations will turn our future much better. The mobile launch control can be considered a good role model of this strategy. It is realized in collaboration with the latest AI&IT technologies that are becoming commonplace in the medical field and the automotive industry.

Fig. 3: Nose cone of the payload faring: general-purpose Silicon form.

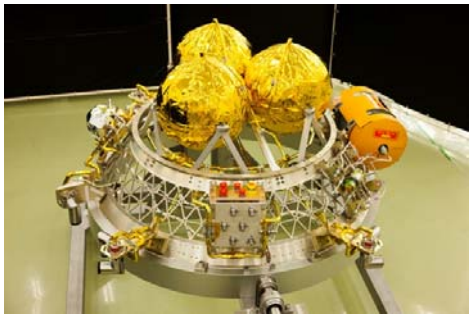


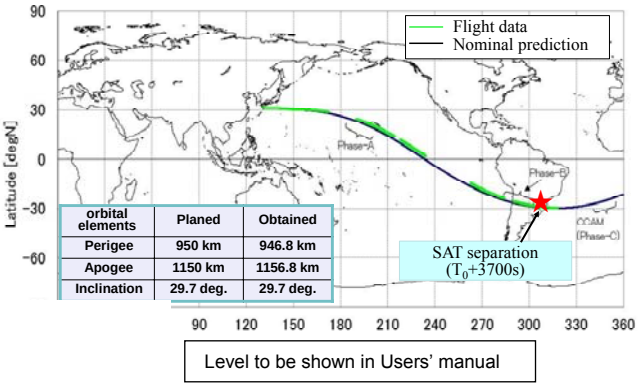
Fig. 4: Post Boost Stage (PBS)

2.3 Enhanced User Friendly Character

The topic is associated with the user friendly character that is of the world leading level as well. More versatile orbital maneuverability and accurate orbit injection can be achieved by an optional tiny post-boost stage (PBS) that can be installed as the fourth stage rocket (Fig. 4). It utilizes tiny 50 N hydrazine engines, similar to the H2A attitude control engines. By using this option, a wide variety of orbits, including solar synchronous orbit, that small satellites require, can easily be reached. In addition, the accuracy of trajectory can be increased to as high as that of the liquid propellant rockets (Table 2). [7-8] The effectiveness of the optional PBS was successfully demonstrated by the flight data of the Epsilon's first launch. Despite the third stage rocket of Epsilon is simply spin-stabilized, possibly resulting in relatively large orbital dispersion, it was reduced by the PBS engines firing twice: above the Hawaii islands and over Argentina. As a result,

the final trajectory obtained was virtually aligned with the planned one at the separation of the payload. The injection error remained less than 7 km as measured at the apogee (1156.8 km as compared to nominal 1150 km), which can be considered at the world-leading level (Fig. 5).

To further enhance the user-friendly characteristics, a special payload attachment fitting (PAF) has been available to lower the level of high frequency vibration (around 50 Hz) that is caused by the oscillatory burning of the first stage solid rocket booster (SRB-A). The mechanism consists of a multi-layered structure of rubbers and thin metals, having lower axial rigidity, to isolate the high frequency vibration. The flight data validated the effectiveness of the vibration attenuator. The lateral vibration level to be shown in the users' manual will be as small as 0.3 G, which is better than or equivalent to those of rival launchers (Fig. 6). Thus, the vibrational environment of Epsilon can be considered significantly comfortable and competitive worldwide.



orbital elements	Planed	Obtained
Perigee	950 km	946.8 km
Apogee	1150 km	1156.8 km
Inclination	29.7 deg.	29.7 deg.

Fig. 5: Trajectory of the first flight of the Epsilon.

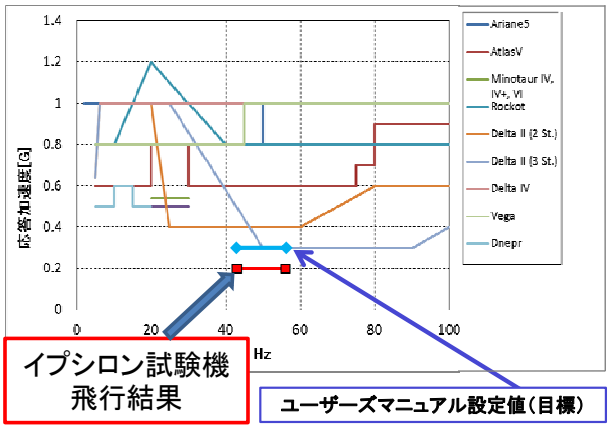


Fig. 6: Sinusoidal vibration level of the Epsilon.

Another aspect that small satellites will most welcome is a reduction in the acoustic environment. Note that the acoustic vibration caused by solid propellant rockets is relatively severe as compared to that by liquid propellant vehicles due to its higher thrust at ignition. For the Epsilon rocket, the acoustic environment is lowered by modifying the configuration of ground facilities in a special way: a combination of two independent approaches: elevating the vertical location of the launch pad by as much as 12 m to increase the distance from the ground; and making a slope underneath the launch pad to deflect the direction of the exhaust flow by 90º. The designs was validated using a refined method of CFD analysis[9] as well as a series of static firing tests using a scale model of the launch facilities. The effectiveness of the modified ground facilities was also far

better than our original expectation and the level to be put in the users' manual can be at the world best (Fig. 7). In this way, the Epsilon rocket has become the most comfortable launch vehicle for small satellites all over the world.

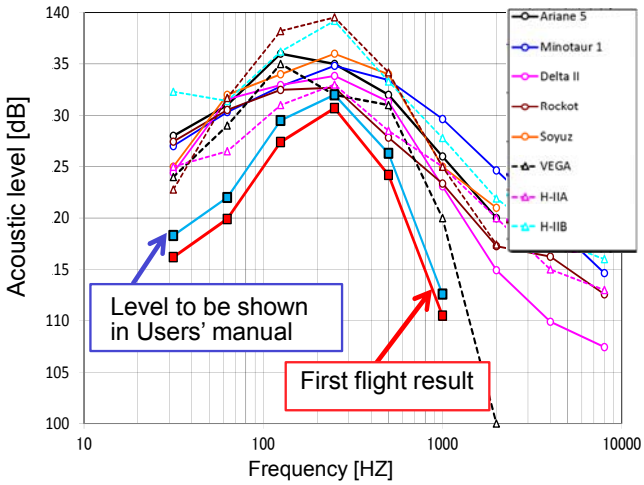


Fig. 7: Acoustic environment for payload Epsilon.

2.4 Evolution of Launch System and Mobile Launch Control

More important aspect of the Epsilon development is an evolution of the launch system, which is intended to dramatically increase the availability of the rocket and to reform the current launch system into a responsive launch by improving the operation performance to the highest standard of the next-generation. The key to success is an innovation of the onboard avionics architecture[10-11]. A newly developed onboard component, the Responsive Operation Support Equipment (ROSE), is designed to act as a centralized data processing system that gathers data from all the other onboard avionics and transmits them to the ground based launch control system through a high-speed network. In addition, the Miniature Ordonnance Checker (MOC) has been developed to conduct check-out of the ignition circuits onboard.

Then, what happens when we apply this novel avionics architecture to the launch control? Until now, the launch control room contains tons of ground support systems and lots of workers involved, like a carnival. From now on, the combination of ROSE/MOC makes it possible to conduct the launch control anytime, anywhere in the world simply by using a couple of PCs that are designed to be highly intelligent so as to automatically checkout the data sent from onboard. Owing to this endeavor, the launch control of Epsilon can be performed remotely, safely, and instantaneously. This is called the Mobile Launch Control, absolutely the next generation technology or you can say, a realization of the science fiction to the science fact. (Fig. 8).

Fig. 8: Epsilon's launch control room.

Table 2: Specifications of orbital injection accuracy.

Accuracy in SSO@500km	**Configuration**	
	Standard	PBS onboard
Perigee height (km)	±25	±20
Apogee height (km)	±100	±20
Inclination (degrees)	±0.6	±0.2

Now, Epsilon has become the first launcher in Japan that is controlled from outside the restricted area, resulting in safer and more comfortable launch operations. Due to these endeavors, the time needed for the ground operation is dramatically reduced[12] and it will take only 9 days to launch since the first stage stand-on.

The intelligent checkout system consists of two different functions: automatic judgement of static data by comparing with threshold; and autonomous diagnosis of dynamic data by pattern matching. The automatic function makes it possible to reduce the number of operators in the block house and the time for the launch control. On the other hand, the autonomous character makes it possible to reduce the number of engineers involved and the time required for the dynamic evaluation that is usually conducted at early stage of launch operations. As of the first launch, this function remains kind of prototype and is planned to be established step-by-step by accumulating enough data to the database through the up-coming launches.

3 Next Step and Beyond

3.1 Enhanced Epsilon

Now that the first flight was conducted almost perfectly, the most important is what the next step will be. As already mentioned, an upgrade program, the Enhanced Epsilon, is underway for higher performance and lower cost mainly by renewing the entire second stage and part of the third stage. The second stage motor case will be made larger and lighter. The combined effect of more propellant and less mass will result in higher performance. In addition, the expanded motor case will be put out of the payload fairing to make space for payload.

On the other hand, the manufacture process of the second stage motor case will be made simpler for lower cost by converting the tree-layer insulation to combined single-layer formation, which can be considered one of the endeavors to reduce the number of the parts and components. As for the third stage, some of its components will be made lighter for higher performance. For example, the current mechanical relays of PSDB (Power Sequence Distribution Box) of the stage will be replaced by semiconductor relays, as already stated as the strategy to utilize latest parts with smaller and lighter character. Due to these efforts, the Enhanced Epsilon will be able to carry about 590 kg into SSO at 500km altitude (30% increase as compared to the current Epsilon) and have the payload capacity expanded by more than 15%. The Enhanced Epsilon is tentatively scheduled to be launched in JFY 2016 (Fig. 9).

Fig. 9: Possible second mission of the Epsilon.

3.2 Mobile Tracking Control

Beyond the mobile launch control of the first generation Epsilon, the intelligence of rockets can be further enhanced for the next generation Epsilon. The final goal can be autonomous flight safety control, which means the rockets are designed to detect their own trajectories and conduct safety operations by themselves if necessary. This will eliminate the expensive tracking radars and their associated facilities, thus making the entire launch complex as simple as only a mobile telemetry station: an extreme mobility of the launch system. Such a high level of autonomous character will lead to a diversity of launch systems because it doesn't need a gigantic launch complex any more. To realize this revolutionary concept of the mobile tracking control, a step-by-step approach is planned to be taken (Table 3). That is, the telemetry data of onboard navigation system will substitute those of the tracking radar as well. Now the integrated navigation package, named RINA, is under development and will be onboard for technology demonstration at the second launch of Epsilon.

Table 3: An idea of a step-by-step evolution of the autonomous flight safety control.

	Configuration	
	Semi-autonomous	Full autonomous
Year for Demonstration	2016	TBD
Ground facilities		
Tracking radar	—	—
Command system	○	—
Range safety computer	○	—
Telemetry system	○	○
Onboard systems		
Navigation sensor	○	○
Telemetry system	○	○
Command decoder	○	—
Onboard range safety computer	—	○
Note: ○: Required —: Not required		

4. Conclusion

This paper deals with the novel universal design concept of the Epsilon launch vehicle including additional values that should satisfy both customers and stake holder at a higher level. From the point of view of payload, such additional value can be associated with the user-friendly character, which means emotionally the hospitality, or OMOTENASHI spirit to respond any special

customer requests as much as possible while physically the ridding comfort and operability. The first flight results demonstrated that the enhanced user-friendly character reached at the world-leading level both in the orbit injection accuracy and in the riding comfort. In addition, payload can be accessible mechanically until 3hours before the launch. The additional value associated with the space transportation technology should be technical innovation to make us to open up the future. The world first mobile launch control, established through the first flight, made a revolution in the space launch system. This converted science fiction to the science fact. Such novel ideas will not only make the Epsilon rocket far ahead of the world but will also bring transportation technologies to a higher level. Beyond these achievements, the Enhanced Epsilon is now underway to make the Epsilon more versatile.

References

1) Morita, Y., Imoto, T. and Ohtsuka, H.: "Research on an Advanced Solid Rocket Launcher in Japan," 26th ISTS, paper No. 2008-g-02 (2008) / also published by the Japanese Rocket Society, Journal of Space Technology and Science, Vo. 24, No. 1, 2008, pp. 1-6.

2) Y. Morita, T. Imoto, H. Habu and H. Ohtsuka, "Japan's Next Generation Solid Rocket Launcher," 60th Congress of the International Astronautical Federation (IAC), Daejeon, Republic of Korea, 2009, Paper No. IAC-09.D2.1.9.

3) Y. Morita, K. Hori, T. Imoto, H. Ohtsuka, A. Fukuchi, and R. Akiba, "Advanced Solid Rocket Launcher and beyond," American Astronautical Society, Advances in the Astronautical Sciences, Vol. 138, pp. 589-596, 2010.

4) Y. Morita, T. Imoto, and H. Ohtsuka, "Development of Japan's Next Generation Solid Rocket Launcher- the Epsilon Rocket," 61st Congress of the International Astronautical Federation (IAC), Prague, Czech Republic, 2010, Paper No. IAC-10.D2.1.7.

5) Y. Morita, T. Imoto, S. Tokudome and H. Ohtsuka, "Epsilon Rocket Launcher and Future Solid Rocket Technologies," 28th International Symposium on Space Technology and Science (ISTS), Okinawa, Japan, 2011, paper No. ISTS 2011-g-01.

6) Y. Morita, T. Imoto, S. Tokudome, and H. Ohtsuka, "Development Status of Japan's Epsilon Solid Rocket Launcher and its Evolution," 62nd Congress of the International Astronautical Federation (IAC), Cape Town, South Africa, 2011, Paper No. IAC-11.D2.1.6.

7) Tanaka, K., Otsuka, H., Saito, K., Tamura, T., Ishii, N., Maeda, Y., and Morita, Y.: "PBS (Post Boost Stage) Design of the Advanced Solid Rocket," 27th ISTS, paper No. 2009-a-19.

8) Y. Morita, "On the Guidance and Control System of Epsilon Solid Rocket Launcher," Proceedings of the 21st Workshop on Astrodynamics and Flight Mechanics, ISAS, 2011.

9) Tsutsumi, S., Fukuda, K., Takaki, R., Shima, E., Fujii, F. and Ui, K.: "Numerical Analysis of Acoustic Environment for Designing Launch-Pad of Advanced Solid Rocket," 26th ISTS, paper No. 2008-g-05.

10) Y. Morita, T. Imoto and H. Ohtsuka, "A Year to Launch: Japan's Epsilon Launcher and its Evolution," 63rd Congress of the International Astronautical Federation (IAC), Naples, Italy, 2012, Paper No. IAC-12.D2.1.8.

11) Kawase, M., Tamura, M., Sunami, K., Izumi, T. and Morita, Y.: "The Avionics System Design Concept for the Advanced Solid Rocket," Journal of Space Technology and Science, Vo. 24, No. 1, 2008, pp. 13-16.

12) Kato, K., Imoto, T., Miho, K. and Arakawa, S.: “The Autonomous Check-out System Concept Design for the Advanced Solid Rocket,” 26th ISTS, paper No. 2008-a-31.