Solution of Problem of Stable Extraction of Features with Using of Measure of Informativeness

Alexander E. Lepskiy[1]

National Research University - Higher School of Economics
Moscow, Russia
(e-mail: alex.lepskiy@gmail.com)

Abstract. In the paper we present a new notion of stochastic monotone measure and its application to image processing. By definition, a stochastic monotone measure is a random value with values in the set of monotone measures and it can describe a choice of random features in image processing. In this case, a monotone measure describes uncertainty in the problem of choosing the set of features with the highest value of informativeness and its stochastic behaviour is explained by a noise that can corrupt images.
Keywords: Image processing, Features extraction, Stochastic monotone measure.

1 Introduction

Let $\Omega = \{\omega_i\}_{i=1}^{n}$ be a set of features that correspond to an image X. To achieve the highest productivity and stable working of a pattern recognition system, it is necessary to choose a sufficiently small subset of features in Ω with the highest information value. There are some very well-known methods that can give us features with the highest information value based on the method of principal components, discriminant analysis and so on [2], [3]. But these methods based on linear algebra fail to take into account structural characteristics that are inherent to geometrical information. In this situation, information measures can be used. By definition, an information measure μ is a set function defined on the power set 2^{Ω} of Ω that for each $A \in 2^{\Omega}$ shows an information value of features in A.

As a rule the features used to describe the pattern are defined with varying degrees of imprecision. The nature of the imprecision may be different. The sets of features are an elements of some a probability space in the classical setting. For example, if the pattern is a discrete plane curve that extracted on the image and features are some characteristics of curve points (eg, feature is a estimation of curvature in given point of discrete curve [4]) then a random character of features (eg, curvature) will be due to noise of image. In this work we will consider monotonous measures of informativeness that are defined on the all subsets of the set of random features. Then monotonous measure will be a random variable $\mathrm{M}(A)$ for each fixed set of random features A. In this case the expectation $\mathbf{E}\left[\mathrm{M}(A)\right]$ be characterize the level of informativeness of representation A and the variance $\sigma^2 [\mathrm{M}(A)]$ be characterize the level of stability of representation to noise pattern. Then there is the problem of finding the most stable and informative representation A of the pattern X. In this article mentioned task will be discussed and resolved in the case when any random feature depends from some other features. This task is investigated for the most popular measures of informativeness for contour image – measure of informativeness by length.

2 Monotone Geometrical Measure of Informativeness

Measures of informativeness can be effectively used in image processing as shown in [1]. In image processing, the most important features are that do not depend on illumination of a scene and orthogonal transformations. Contours of images and their characteristics, for example, curvature of smooth curves can be such features. However, in reality, we have digitized curves that are given by some ordered sets of points. These curves can be corrupted by noise. This means that we can use only some statistical estimates of curvature [4]. A problem of choosing an optimal polygonal representation of a contour consist in finding such a representation that preserves geometrical characteristics of contour and will be statistically efficient. This choice can be produced by using measure of informativeness that are axiomatically defined as follows.

Let X be an initial closed contour given by an ordered finite set points, i.e. $X = \{x_1, ..., x_n\}$, where $x_i \in R^2$, $i = 1, ..., n$. We identify with any nonempty subset $B = \{x_{i_1}, ..., x_{i_m}\}$ a contour generated by connecting points with straight lines starting from points x_{i_1}, x_{i_2} and ending by points x_{i_m}, x_{i_1}.

A geometrical measure of informativeness $\mu : 2^X \to [0, 1]$ is a set function that has to obey the following properties:

1. $\mu(\emptyset) = 0$, $\mu(X) = 1$;
2. $A, B \in 2^X$ and $A \subseteq B$ implies $\mu(A) \le \mu(B)$;
3. let $B = \left\{x_{i_1}, ..., x_{i_{k-1}}, x_{i_k}, x_{i_{k+1}}, ..., x_{i_m}\right\} \in X$ and neighbouring points $x_{i_{k-1}}, x_{i_k}, x_{i_{k+1}}$ belong to a straight line in the plane, then $\mu(B) = \mu\left(B \backslash \{x_{i_n}\}\right)$;
4. μ is invariant w.r.t. affine transformations in the plane such as parallel transferring, rotation and scaling.

Emphasize that axioms 1, 2 have been introduced by Sugeno for fuzzy measures (see [7]). Consider several ways for defining geometrical measures of informativeness [1].

a) Suppose that the length of an original contour is not equal to zero and a function $L(A)$ gives us the length of subcontour $A \in 2^X$. Then an measure of informativeness defined by contour length is $\mu_L(A) = \frac{L(A)}{L(X)}$.

b) Suppose that the domain limited by an original contour is convex, and a function $S(B)$ determines the area bounded by an subcontour $A \in 2^X$. Then an measure of informativeness defined by contour area is $\mu_S(A) = \frac{S(A)}{S(X)}$.

c) Let $w(x,A)$ be an estimate of information value of the part of a contour in a neighbourhood of point $x \in A$ in a subcontour $A \in 2^X$. Then an average measure of informativeness is defined by $\mu(A) = \frac{\sum_{x\in A} w(x,A)}{\sum_{x\in X} w(x,X)}$, where $w(x,A)$ has to be defined for any non-empty contour $A \in 2^X$ and $\mu(\emptyset) = 0$ by definition. It is easy to see that the introduced geometrical measure of informativeness μ_L and μ_S can be considered as average measure of informativeness. For example, for μ_L function $w(x,A)$ can be defined by $w(x,A) = |x-y|$, where y is the next neighbouring point in contour A; in case of μ_S function $w(x,A)$ can be defined by $w(x,A) = S(O,x,y)$, where O is the centroid of area, bounded by contour A, and $S(O,x,y)$ is the area of triangle with vertices in points O,x,y.

As it was mentioned above, points consisting of the original contour accumulate the influence of several random factors, and they can be therefore considered as random values. This means that the observed measure of informativeness can be viewed as a random monotone measure. This opens new directions of research.

3 Stochastic Average Monotone Measure of Informativeness

In real situations, values $w(x,A)$ can be considered as random values, because an original contour is corrupted by a noise. To stress this, we denote these values by capital letters as $W(x,A)$. In this case we have the measure of informativeness $\mathrm{M}(A) = \frac{\sum_{x\in A} W(x,A)}{\sum_{x\in X} W(x,A)}$. Then there is the problem of finding the most stable and informative representation $A \in 2^X$ of the pattern X for which the expectation $\mathbf{E}[\mathrm{M}(A)]$ will be most and the the variance $\sigma^2[\mathrm{M}(A)]$ will be least. If $W(x,A) = W(x)$ and random values $W(x)$, $x \in X$, are independent random variables then the measure of informativeness $\mathrm{M}(A)$ is additive and its values are random. For example, this probabilistic model can be used if values $W(x)$ can be conceived as estimates of absolute value of curvature and this estimation is produced by disjoint neighbourhoods of points $x \in X$ [4]. We emphasize that stochastic additive measures have been already investigated in the literature (see e.g. [6]). In this additive case the problem of finding the most stable and informative representation is investigated in [5].

Now we will be investigated the important case when the value $W(x,A)$ depends on two neighbouring points. For example, the geometrical measures of informativeness μ_L and μ_S are satisfied this condition.

Let $X = \{x_1,...,x_n\}$ be an original contour and let vertices be ordered by their indices. So if we consider any subcontour $A \in 2^X$, then the order defined on A is assumed to be generated by the order on X and given by indices in the representation $A = \{x_{i_1},...,x_{i_m}\}$, where $i_1 < ... < i_m$. So for any

$A \in 2^X$ we can identify its elements by their indices and write $x_k(A) = x_{i_k}$ if $k \in \{1, ..., m\}$ and $A = \{x_{i_1}, ..., x_{i_m}\}$, where $i_1 < ... < i_m$. We can also consider any integer index k assuming that $x_k(A) = x_l(A)$ if $l \equiv k(\text{mod} m)$. To work with such indices, we use a mapping π defined by $x_k(A) = x_{\pi_A(k)}$. We suppose that $W(x_k(A), A) = W(x_k(A), x_{k+1}(A))$, $k = 1, ..., |A|$, i.e. the value $W(x_k(A), A)$ depends on two neighbouring points $x_k(A), x_{k+1}(A)$. Further, for simplicity reasons, we denote $W(x_k(A), x_{k+1}(A)) = W_{k,k+1}(A)$. Then an average monotone measure and stochastic average monotone measure have a view

$$\mu(A) = \frac{\sum_{k=1}^{|A|} w_{k,k+1}(A)}{\sum_{j=1}^{|X|} w_{k,k+1}(X)}, \mathrm{M}(A) = \frac{\sum_{k=1}^{|A|} W_{k,k+1}(A)}{\sum_{j=1}^{|X|} W_{k,k+1}(X)} \tag{1}$$

correspondingly. We call M a stochastic monotone measure of informativeness if $W_{k,k+1}(A)$, $A \in 2^X$ are random variables. In this case M has random values. In this section we find estimates of numerical characteristics of M assuming that random variables $W_{k,k+1}(A)$, $W_{l,l+1}(A)$ are independent if $|l-k| > 1$. This situation appears if we suppose that x_k, $k = 1, ..., n$, are also independent random variables.

We see that $\mathrm{M}(A) = \frac{\xi}{\eta}$, where $\xi = \sum_{k=1}^{|A|} W_{k,k+1}(A)$ and $\eta = \sum_{j=1}^{|X|} W_{k,k+1}(X)$. The following lemma is used for estimating $\mathbf{E}[\mathrm{M}(A)]$ and $\sigma^2[\mathrm{M}(A)]$.

Lemma 1. *Let ξ and η be random variables that taking values in the intervals l_ξ, l_η respectively on positive semiaxis and $l_\eta \subseteq ((1-\delta)\mathbf{E}[\eta], (1+\delta)\mathbf{E}[\eta])$, $l_\xi \subseteq (\mathbf{E}[\xi] - \delta\mathbf{E}[\eta], \mathbf{E}[\xi] + \delta\mathbf{E}[\eta])$. Then it is valid the following formulas for mean and variance of distribution of $\frac{\xi}{\eta}$ respectively*

$$\mathbf{E}\left[\frac{\xi}{\eta}\right] = \frac{\mathbf{E}[\xi]}{\mathbf{E}[\eta]} + \frac{\mathbf{E}[\xi]}{\mathbf{E}^3[\eta]}\sigma^2[\eta] + \frac{1}{\mathbf{E}^2[\eta]}\mathbf{Cov}[\xi,\eta] + r_1, \tag{2}$$

$$\sigma^2\left[\frac{\xi}{\eta}\right] = \frac{1}{\mathbf{E}^2[\eta]}\sigma^2[\xi] + \frac{\mathbf{E}^2[\xi]}{\mathbf{E}^4[\eta]}\sigma^2[\eta] - \frac{2\mathbf{E}[\xi]}{\mathbf{E}^3[\eta]}\mathbf{Cov}[\xi,\eta] + r_2, \tag{3}$$

where $\mathbf{Cov}[\xi,\eta]$ is a covariation of random variables ξ and η, i.e. $\mathbf{Cov}[\xi,\eta] = \mathbf{E}[(\xi - \mathbf{E}[\xi])(\eta - \mathbf{E}[\eta])]$; r_1, r_2 are the residuals those depends on numerical characteristics of ξ and η. It being known that $|r_1| \le \frac{\delta}{1-\delta} \cdot \frac{\mathbf{E}[\xi]+\mathbf{E}[\eta]}{\mathbf{E}^3[\eta]}\sigma^2[\eta] \le \frac{\mathbf{E}[\xi]+\mathbf{E}[\eta]}{(1-\delta)\mathbf{E}[\eta]}\delta^3$, $|r_2| \le C\delta^3$.

This lemma is proved with help of expanding the function $\phi(x,y) = \frac{x}{y}$ into a Taylor series at the point $(\mathbf{E}[\xi], \mathbf{E}[\eta])$.

We will use formulas (2) and (3) without their residuals. Respective values $\tilde{\mathbf{E}}[\mathrm{M}(A)] = \mathbf{E}[\mathrm{M}(A)] - r_1$, $\tilde{\sigma}^2[\mathrm{M}(A)] = \sigma^2[\mathrm{M}(A)] - r_2$ we will call by estimations of numerical characteristics.

Introduce the following notation: $S(A) = \sum_{i=1}^{|A|} \mathbf{E}[W_{i,i+1}(A)]$, $K(A,X) = \sum_{i=1}^{|A|} k_i^X(A)$, where $k_i^X(A) = \sum_{j=1}^{|X|} \mathbf{Cov}[W_{i,i+1}(A), W_{j,j+1}(X)]$, $A \in 2^X$.

Then the formulas for $\tilde{\mathbf{E}}\,[\mathrm{M}(A)]$ and $\tilde{\sigma}^2\,[\mathrm{M}(A)]$ based on (2) and (3) can be written in the form

$$\tilde{\mathbf{E}}\,[\mathrm{M}(A)] = \tfrac{S(A)}{S(X)} + \tfrac{S(A)}{S^3(X)} K(X,X) - \tfrac{1}{S^2(X)} K(A,X), \tag{4}$$

$$\tilde{\sigma}^2\,[\mathrm{M}(A)] = \tfrac{1}{S^2(X)} K(A,A) + \tfrac{S^2(A)}{S^4(X)} K(X,X) - \tfrac{2S(A)}{S^3(X)} K(A,X). \tag{5}$$

4 Stochastic Measure of Informativeness by Contour Length

Assume that an original contour is corrupted by noise. In this case, $X = \{x_k + \mathbf{n}_k\}_{k=1}^m$, $x_k \in R^2$ and $\mathbf{n}_k = (\xi_k, \eta_k)$ are random variables. Suppose also that ξ_k, η_k, $k = 1, ..., m$, are independent, normally distributed and such that $\mathbf{E}\,[\xi_k] = \mathbf{E}\,[\eta_k] = 0$, $\sigma^2\,[\xi_k] = \sigma^2\,[\eta_k] = \sigma^2$, $k = 1, ..., m$. In this section we consider a monotone stochastic measure M defined by contour length. This measure has the view of (1), where $W_{k,k+1}(A) = |x_{k+1}(A) + \mathbf{n}_{k+1}(A) - x_k(A) - \mathbf{n}_k(A)|$. We investigate such characteristics of $\mathrm{M}(A)$ as $b\,[\mathrm{M}(A)] = \tilde{\mathbf{E}}\,[\mathrm{M}(A)] - \mu(A)$ and $\tilde{\sigma}^2\,[\mathrm{M}(A)]$. In general, the random variable $\sum_{k=1}^{|A|} W_{k,k+1}(A)$ is not satisfied to conditions of Lemma 1. However the probability of large deviations of random length of noisy polygonal line from non-noisy length will be small if the variance of noise is small. Therefore we assume that the random length satisfied approximately to conditions of Lemma 1. Suppose that $W_{k,k+1}(X)$, $k = 1, ..., m$, are independent random variables. This requirement can be satisfied by the choice of some subcontour (basic contour) from the initial contour.

4.1 Numerical Characteristics of Random Variable $W_{k,k+1}(A)$

We give asymptotic formulas for $\mathbf{E}\,[W_{k,k+1}(A)]$ and $\sigma^2\,[W_{k,k+1}(A)]$. These formulas proved with help expansion of the random variable $W_{k,k+1}(A)$ by Taylor formula.

Proposition 1. *The following asymptotic equalities are valid*

$$\mathbf{E}\,[W_{k,k+1}(A)] = l_k \left(1 + \tfrac{\sigma^2}{l_k^2} + \tfrac{\sigma^4}{2l_k^4} + O\left(\tfrac{\sigma^6}{l_k^6}\right)\right),$$

$$\sigma^2\,[W_{k,k+1}(A)] = 2\sigma^2 \left(1 - \tfrac{\sigma^2}{l_k^2} + O\left(\tfrac{\sigma^4}{l_k^4}\right)\right),$$

where $l_k = |x_{k+1}(A) - x_k(A)|$.

Corollary 1. *It is true the equality*

$$S(A) = \sum_{k=1}^{|A|} \mathbf{E}\,[W_{k,k+1}(A)] = L(A) + \sigma^2 \sum_{k=1}^{|A|} |x_{k+1}(A) - x_k(A)|^{-1} + \sigma O\left(\tfrac{\sigma^3}{l^3}\right),$$

where $L(A)$ *is the length of contour* A *without an influence of noise, i.e.* $L(A) = \sum_{k=1}^{|A|} |x_{k+1}(A) - x_k(A)|$, *and* $l = \min_k |x_{k+1}(A) - x_k(A)|$.

Next we compute the covariance $\mathbf{Cov}\left[W_{k-1,k}(A), W_{k,k+1}(A)\right]$ between random variables $W_{k-1,k}(A)$.

Let $\mathbf{l}_i = \mathbf{l}_i(A) = x_{i+1}(A) - x_i(A)$ be a i-s segment-vector of polygon A and $\alpha_i = \alpha(x_i) = \left(\mathbf{l}_{i-1} \overset{\wedge}{,} \mathbf{l}_i\right)$.

Proposition 2. *We have*

$$\mathbf{Cov}\left[W_{k-1,k}(A), W_{k,k+1}(A)\right] =$$
$$= -\sigma^2 \cos\alpha_k \left(1 - \left(\tfrac{1}{l_{k-1}^2} + \tfrac{\cos\alpha_k}{2l_{k-1}l_k} + \tfrac{1}{l_k^2}\right)\sigma^2 + \ o\left(\tfrac{\sigma^2}{l^2}\right)\right),$$

where $l_k = l(x_k) = |x_{k+1}(A) - x_k(A)|$, $l = \min\{l_{k-1}, l_k\}$.

Calculate the covariance $K(A,X) = \sum_i k_i^X(A)$ between the all segments of polygon A and all segments of basic polygon X with help of last corollary. Let $\alpha(x)$ ($\beta(x)$) be an inner angle of polygon A (polygon X) in vertex x, $\gamma(x)$ be an angle between the vectors $x_{+1}(A) - x$, $x_{+1}(X) - x$, where $x_{+1}(A)$ ($x_{+1}(X)$) is the next point w.r.t. x in the contour A (contour X).

Corollary 2. *We have*

$$K(A,X) = 4\sigma^2 \sum_{x\in A} \cos\tfrac{\alpha(x)}{2} \cos\tfrac{\beta(x)}{2} \cos\left(\gamma(x) + \tfrac{\alpha(x)-\beta(x)}{2}\right) + \sigma^2 o\left(\tfrac{\sigma}{\Delta(A)}\right)$$

for $A \in 2^X$, *where* $\Delta(A) = \min_i\{\Delta_i(A)\}$.

4.2 The Numerical Characteristics of Stochastic Measure of Informativeness by Length

We will find numerical characteristics of stochastic measure of informativeness by length using the results of the previous item. We will consider the application of calculated characteristics to solution of task of finding polygonal representation of curve that is stable to noise.

At first we formulate the theorem about expectation of stochastic measure of informativeness by length $\mathbf{E}\left[\mathrm{M}(\cdot)\right]$ on 2^X that follows from equality (4), Corollaries 1, 2.

Theorem 1. *The asymptotic equality*

$$\tilde{\mathbf{E}}\left[\mathrm{M}(A)\right] = \tfrac{L(A)}{L(X)} + \tfrac{C_1(A)}{L^2(X)}\sigma^2 + o\left(\tfrac{\sigma^2}{\Delta^2(A)}\right), A \in 2^X \qquad (6)$$

is true, where

$$C_1(A) = -L(A)\sum_{x\in X}|\mathbf{l}_x|^{-1} + L(X)\sum_{x\in A}|\mathbf{l}_x|^{-1} + 4\tfrac{L(A)}{L(X)}\sum_{x\in X}\cos^2\tfrac{\beta(x)}{2} -$$
$$-4\sum_{x\in A}\cos\tfrac{\alpha(x)}{2}\cos\tfrac{\beta(x)}{2}\cos\left(\gamma(x) + \tfrac{1}{2}\alpha(x) - \tfrac{1}{2}\beta(x)\right).$$

Notice that the asymptotic formula for bias $b\,[\mathrm{M}(A)] = \tilde{\mathbf{E}}\,[\mathrm{M}(A)] - \mu(A)$ of stochastic measure of informativeness by length with noise follows from (6):

$$b\,[\mathrm{M}(A)] = \tfrac{C_1(A)}{L^2(X)}\sigma^2 + o\left(\tfrac{\sigma^2}{\Delta^2(A)}\right), \quad A \in 2^X,$$

and also $C_1(X) = 0$.

Similarly, we will find the asymptotic formula for variance of stochastic informational measure by length with help of formula (5), Corollaries 1, 2.

Theorem 2. *The asymptotic equality*

$$\tilde{\sigma}^2\,[\mathrm{M}(A)] = 4\tfrac{C_2(A)}{L^2(X)}\sigma^2 + o\left(\tfrac{\sigma^2}{\Delta^2(A)}\right), \quad A \in 2^X$$

is true, where

$$C_2(A) = \sum_{x\in A}\cos^2\tfrac{\alpha(x)}{2} + \tfrac{L^2(A)}{L^2(X)}\sum_{x\in X}\cos^2\tfrac{\beta(x)}{2} -$$

$$-2\tfrac{L(A)}{L(X)}\sum_{x\in A}\cos\tfrac{\alpha(x)}{2}\cos\tfrac{\beta(x)}{2}\cos\left(\gamma(x)+\tfrac{1}{2}\alpha(x)-\tfrac{1}{2}\beta(x)\right).$$

The value of random error (the variance of stochastic informational measure) characterizes the degree of stability of informational measure of curve with respect to level of curve noise. We can put the task about finding of polygonal representation of fixed cardinality $A \in 2^X$, $|A| = k$, which minimized the value of variance of stochastic informational measure by length. As can be seen from Theorem 2 the polygonal representation

$$A = \arg\min_{A\in 2^X,\,|A|=k} C_2(A)$$

is a solution of indicated task for not great level of curve noise σ.

Example 1. Let $X = \{x_1, ..., x_6\}$ be an ordered set of vertexes of regular 6-gon with length of segment is equal 1. Calculate the value $C_2(A)$ for various polygonal representations A of cardinality $|A| = 3$: $A_1 = \{x_1, x_3, x_5\}$, $A_2 = \{x_1, x_2, x_4\}$, $A_3 = \{x_1, x_2, x_3\}$ (Figure 1). Since $\beta(x) = \frac{2\pi}{3}$, $x \in X$, $L(X) = 6$, $\sum_{x\in X}\cos^2\frac{\beta(x)}{2} = 1.5$, then

$$C_2(A) = \sum_{x\in A}\cos^2\tfrac{\alpha(x)}{2} + \tfrac{L^2(A)}{24} - \tfrac{L(A)}{6}\sum_{x\in A}\cos\tfrac{\alpha(x)}{2}\cos\left(\gamma(x) + \tfrac{\alpha(x)}{2} - \tfrac{\pi}{3}\right).$$

Therefore $C_2(A_1) = 1.125$, $C_2(A_2) = 1.25$, $C_2(A_3) = \frac{56+22\sqrt{3}-5\sqrt{2}-3\sqrt{6}}{48} \approx 1.66$. Thus the contour A_1 is a most stable contour to noise pollution with respect to informational measure by length among of contours of cardinality is equal 3.

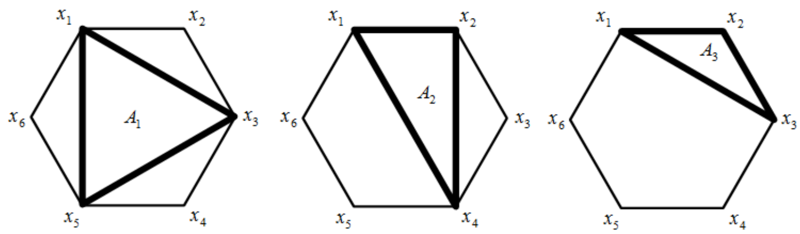

Fig. 1. Various polygonal representations A of X of cardinality $|A| = 3$.

5 Summary and Conclusions

The problem of finding of optimal stable pattern representation was discussed in this paper. We got geometrical conditions for finding a most stable contour to noise pollution with respect to measure of informativeness by length among of contours of equal cardinality. That geometrical conditions may be used to formulation and solution of others problems of finding of optimal stable pattern representation.

Acknowledgements

I would like to thank Andrew Bronevich for his helpful and stimulating comments on the manuscript of my paper. The study was implemented in the framework of The Basic Research Program of the Higher School of Economics in 2012. This work was supported by the grants 11-07-00591 and 10-07-00135 of RFBR (Russian Foundation for Basic Research).

References

1. Bronevich, A., and Lepskiy, A., "Geometrical fuzzy measures in image processing and pattern recognition", *Proc. of the 10th IFSA World Congress.* Istanbul, Turkey, 151–154 (2003).
2. Duda, R.O., Hart, P.E., and Stork, D.G., *Pattern Classification and Scene Analysis: Part I Pattern Classification*, John Wiley & Sons (1998).
3. Jolliffe, I.T., *Principal Component Analysis*, Series: Springer Series in Statistics. 2nd ed., Springer, NY (2002).
4. Lepskii, A.E., "On Stability of the Center of Masses of the Vector Representation in One Probabilistic Model of Noiseness of an Image Contour", *Automation and Remote Control*, 68, 75–84 (2007).
5. Lepskiy, A.E., "Stable Feature Extraction with the Help of Stochastic Information Measure", *Lecture Notes in Computer Science(LNCS)*, vol. 6744, Berlin / Heidelberg: Springer Verlag, 54–59 (2011).
6. Shiryaev, A.N., *Probability.* Graduate Texts in Mathematics, Springer Pub. (1995).
7. Wang, Z., and Klir G.J., *Generalized Measure Theory*, IFSR International Series on Systems Science and Engineering. 25, Springer (2009).