

СОДЕРЖАНИЕ

МЕНЕДЖМЕНТ И СИСТЕМЫ КАЧЕСТВА ОБРАЗОВАТЕЛЬНЫХ УЧРЕЖДЕНИЙ

Е.В. КАПИЛУС, Е.Б. ГАФФОРОВА
Некоторые аспекты эффективного функционирования
системы менеджмента качества вуза 2

А.В. ПАШКЕВИЧ
Основные аспекты создания системы менеджмента качества образования
в контексте компетентностного подхода в массовой школе 8

И.М. АРТАМОНОВ, Т.П. БАЛАЕВА
Особенности подготовки к сертификации системы менеджмента качества вуза
в сжатые сроки на примере МАИ 13

ВОПРОСЫ ПОДГОТОВКИ СПЕЦИАЛИСТОВ

Н.В. БОЙКО, В.В. МИХАЙЛОВ
Особенности предоставления образовательных услуг и существующие предпосылки
снижения качества подготовки специалистов 17

М.С. КАДАЦКАЯ, О.Е. КРАВЧЕНКО
Формирование компетентностной модели выпускника 20

А.А. СОБОЛЕВСКИЙ
Система оценки компетенций студента с использованием данных веб-сервисов 22

М.А. БЕЛОВ, О.Е. АНТИПОВ
Контрольно-измерительная система оценки качества обучения в виртуальной
компьютерной лаборатории 28

В.В. БРАТИЩЕНКО
Параметрическая модель экзаменационных оценок 32

ОБЩИЕ ПРОБЛЕМЫ ОБРАЗОВАНИЯ

А.А. КУЗЬМИНА
Публичные сети как система информационной и консультационной
взаимопомощи вузов в интеграционных процессах в сфере образования 35

ИННОВАЦИОННЫЙ МЕНЕДЖМЕНТ

С.Ю. КОВАЛЕНКО
Интеллектуальный капитал как основа становления инновационной экономики 39

И.Б. ТЕСЛЕНКО, О. Б. ДИГИЛИНА
Государственно-частное партнерство как инструмент формирования экономики
инновационного типа 42

КАЧЕСТВО И ИПИ(CALS)-ТЕХНОЛОГИИ

КАЧЕСТВО: РУКОВОДСТВО, УПРАВЛЕНИЕ, ОБЕСПЕЧЕНИЕ

А.В. АНДРОНОВ, Н.А. КОСИЛОВ, А.М. МЕРКУЛОВ, П.А. ПЕТРИКОВ
Проектирование беспроводной компьютерной сети с заданным качеством обслуживания ... 46

ПРИБОРЫ, МЕТОДЫ И ТЕХНОЛОГИИ

Д.А. КОРОЛЕВ
Методы и средства автоматизации многокамерной видеосъемки событий 49

Д.Ю. ПРОФЕРАНСОВ
Оценка пропускной способности узла коммутации 59

ЭКОНОМИКА И УПРАВЛЕНИЕ

Е.В. СЕМИНА
Концептуальная, формальная и структурно-функциональные модели
системы управления рисками факторинга 60

С.И. БАЦУНОВА
Швейцария в новых контурах финансовой карты мира 64

Н.Н. ЩУРОВ
Особенности оценки земельных участков при изъятии для государственных нужд 68

Д.А. Королев

МЕТОДЫ И СРЕДСТВА АВТОМАТИЗАЦИИ МНОГОКАМЕРНОЙ ВИДЕОСЪЕМКИ СОБЫТИЙ

В статье предлагается метод создания как автоматизированного, так и автоматического многокамерного съемочного комплекса, задействующего широкий спектр известных технологий. В рамках проекта предполагается алгоритмизация работы операторов и режиссера, создание нового графического интерфейса режиссера и предметной онтологии различных языков человеко-машинного взаимодействия, а также разработка виртуальной среды моделирования съемочных ситуаций для обучения съемочного комплекса.

Ключевые слова: многокамерная видеосъемка, видеосюжет, виртуальная среда моделирования

D.A. Korolev

METHODS AND TOOLS FOR AUTOMATION OF THE MULTI-EVENTS VIDEO

In this paper a method of automated and automatic multi-camera shooting is proposed with using a the wide range of modern information technologies. The project includes algorithmization of camera-work and the director's commands, a new director's graphical user interface and ontology of human-machine interaction languages, as well as the development of virtual simulation environment for teaching the developed system.

Keywords: multichamber video shooting, video topic, the virtual environment of modeling

Современное видеопроизводство, если оставить вне рассмотрения музыкальную индустрию и кино, в значительной части занимается записью и монтажом “говорящих голов”. Большинство ток-шоу, пресс-конференций и, в некоторой степени, мыльные оперы – это типовые по съемочному сценарию видеопрограммы. Добавим сюда многочасовые конференции, транслируемые в интернет, и растущий объем записываемых и транслируемых лекций и иных событий подобного визуального ряда.

Для съемки таких видеосюжетов (назовем их типичными), тем не менее, задействуется значительный человеческий ресурс – каждая камера управляется оператором, камер ставится обычно от трех до восьми, операторами и монтажом руководит режиссер. Это не считая инженеров и редакторов. И, зачастую, работа оператора заключается в том, чтобы держать заданный режиссером план одного человека или нескольких попеременно на протяжении всей передачи. В то же время, действия режиссера порой сводятся к переключению заранее наведенных камер с учетом известных правил монтажа в нужные моменты – по событию или для разбавления монотонной речи.

В сложившейся ситуации заметно значительное влияние традиционного уклада, сложившегося на заре телевидения, но очевидно, что это богатое поле для различного рода нововведений, о которых и пойдет речь в этой статье.

Поставим себе цель – создать съемочный комплекс нового поколения. Задачи останутся теми же – обеспечить запись или трансляцию типичных сцен разговорного жанра, художественные приемы сохра-

ним те же, что и при традиционном способе съемки.

Для примера возьмем конференцию: зрительный зал с одной стороны, президиум и трибуна – с другой. Докладчики выходят к трибуне из зала или президиума, обращаются к зрителям, на экране демонстрируются слайды. Выступления завершаются серией вопросов из президиума или зала, люди в президиуме иногда берут слово.

Рассмотрим задачи, которые встают перед нами на пути к полной автоматизации съемки.

1. Управление PTZ-видеокамерой

Роботизированные камеры имеют дистанционное управление, команды передаются по локальной сети, возможно управление специальным контроллером (джойстиком) или при помощи программных средств.

На практике оказывается, что используемые в доступном рыночном сегменте PTZ-видеокамеры не могут обеспечивать “художественную” съемку, если использовать их со стандартным программным интерфейсом. Все дело в том, что вращение головки по горизонтали (панорама, pan), вертикали (наклон, tilt) и трансфокация объектива (zoom) осуществляются при помощи электроприводов, которые потребляют достаточно много энергии. Поскольку эти камеры проектировались изначально для нужд видеонаблюдения, в задачи которых входит не столько производить художественные произведения, сколько их охранять, то стали возможны допущения, позволившие снизить пиковое электропотребление устройства: при обращении через стандартный API

камера не использует все три привода одновременно. Она может быстро сменить направление съемки и изменить фокусное расстояние объектива, но сделает это в два или три шага, а не одним плавным движением, как это сделал бы видеооператор.

Итак, основной задачей низкоуровневого управления PTZ-камерой станет **создание драйвера художественной панорамы** (в понятие панорамы включим также трансфокацию).

2. Распознавание образов, планов и композиция кадра

Чтобы построить композицию кадра или установить нужный план, оператор должен определить объект съемки. Для человека это дело привычное, но компьютер этому навыку еще следует научить. Под распознаванием здесь мы понимаем следующее:

1. Положение лица и рук человека в кадре;
2. Размер лица относительно кадра;
3. Направление головы, в детализированном случае – взгляда;
4. “Узнавание” человека.

Рассмотрим каждый из этих навыков.

Зная **положение человека в кадре**, можно построить композицию кадра. Наиболее известное правило композиции – “**правило третей**” (англ. источник) – выделяет на экране 4 точки (рис. 1), наша задача состоит в том, чтобы сопоставить лицо (глаза) человека и эти точки на экране для построения правильной композиции. На этом месте читатель, имеющий опыт фото- или видеосъемки, найдет массу возражений и исключений, но не будем торопиться, конечный результат ещё далек и судить пока рано.

Существует и техническая задача, выполняемая при помощи распознавания объектов в кадре. Часто случается, что в кадре оказывается, например, два объекта, находящихся на разном расстоянии от камеры. При этом в фокусе оказывается обычно объект, который или а) ближе к центру кадра, или б) ближе к камере. Так обычно настроен автофокус видеокамеры. В нашем случае, если работает не автономная камера, а съемочный комплекс, мы можем предположить ситуацию, когда задумка режиссера будет противоречить алгоритму автофокуса, и наша задача – создать для режиссера (или программы, выполняющей его роль) инструмент для указания ведомого объекта в кадре. Например, если камере дано указание следить за данным объектом, а во время съемки в кадр попал кто-нибудь еще, то камера не должна перестраивать фокус с ведомого объекта до получения соответствующего указания.

Становится возможным прием, легко доступный операторам при помощи профессиональных камер:

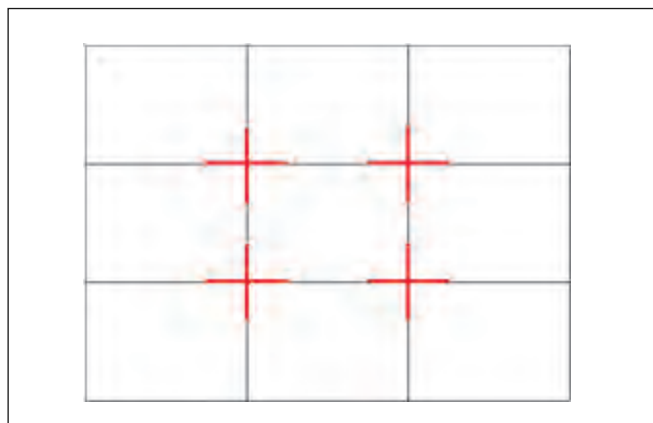


Рис. 1. Правило третей

внутрикадровый монтаж перефокусировкой объектива.

Зная **размер лица**, мы можем судить о **крупности плана**. Классификация крупности планов известна с начала XX века, кинематограф и телевидение используют удобные им классификации (отечественный, американский). Например, в кино известны планы Льва Кулешова: 1 – деталь, 2 – крупный, 3 – первый средний, 4 – второй средний, 5 – общий, 6 – дальний (рис. 2).

Поскольку крупность планов считается от размеров человека в кадре, а пропорции у большинства людей относительно близки, то крупность плана можно вычислять из размера головы человека в кадре (обычно голова присутствует во всех кадрах в рассматриваемом нами жанре).

Для определения плана мы не можем полагаться на **полный рост** человека, так как он виден камере только на общем или дальнем планах, однако важно отслеживать положение рук и, по возможности, ног: как видно на рисунке, в некоторых случаях средние планы удобно определять, опираясь не только на размер головы, но и на **размер рук** и их попадание в кадр. Так, сидящий за столом человек на первом среднем плане должен входить в кадр с руками, если их видно на столе.

В результате измерения относительной величины головы на 2-6 планах и относительной величины человека на 5-6 планах, где эти измерения возможны, различные (и немногочисленные) источники дают большой разброс. Так, среднеквадратичное отклонение величины головы на первом среднем плане превышает 200 единиц, при среднем значении по пяти планам – 77. Для полного роста характерны меньшие различия, здесь отклонения не превышают 0,2.

Очевидно, что полагаться на имеющиеся данные по крупности планов невозможно, но можно взять их за основу. Существуют и косвенные признаки каждо-

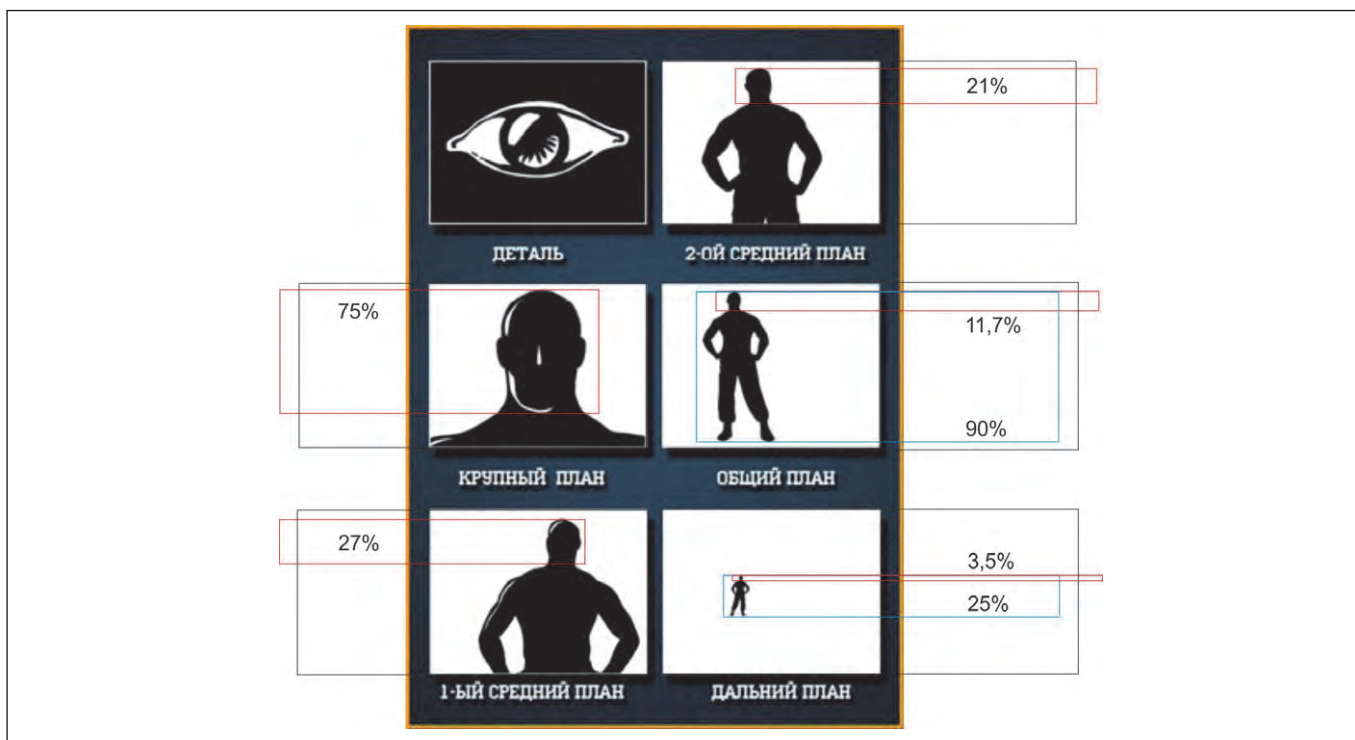


Рис. 2. Деление планов по крупности

го плана, именно они помогут скорректировать крупность для достижения искомого визуального впечатления. Так, на втором среднем плане, например, должны полностью помещаться руки – эта задача может быть решена распознаванием, но не головы, а всего тела. Используя такие признаки, можно компенсировать погрешности распознавания плана по размеру головы в кадре (для 2-6 планов).

Также важно отметить, что ориентироваться на полный рост ненадежно – в отличие от головы, которая не меняет свою форму, человек часто наклоняется, садится или попадает на такой план, где не виден полный рост.

Для построения правильной композиции кадра также нужно знать **направление поворота головы**, а лучше – взгляда. Например, если у человека голова повернута вправо, то его обычно следует размещать в левой стороне кадра.

И, наконец, **узнавать человека** нужно для того, чтобы обучить систему. До сих пор мы не рассматривали автоматизированный съемочный комплекс как интеллектуальную систему, но в наших планах – замена режиссера, а для этого потребуется несколько больше, чем просто брать нужный план и переключать на другой, совместимый с предыдущим. Как мы увидим дальше, навык узнавания и запоминания человека понадобится съемочному комплексу для тонкой настройки на выступающего и предугадывания поведения других “подобных” выступающих.

Таким образом, задача распознавания, применительно к данному проекту, выражается в следующем:

1. Поиск объекта по указанным координатам (координаты получаются с датчика движения).
2. Выделение лица, его размера в кадре и направления. Это позволит сделать привязку к крупности планов.

3. Работа с датчиком движения

Игровая индустрия подарила миру устройства, весьма полезные для широкого круга задач. Среди них контроллер движения Kinect. Он распознает до семи человек перед собой и следит за положением их рук, ног и головы. Этих данных достаточно, чтобы иметь представление о происходящем на сцене и передавать обработчику координаты людей, находящихся на сцене.

Полученные координаты пересчитываются для каждой камеры с учетом смещения в пространстве. Благодаря наличию общей картины сцены и способности камер распознавать эти объекты после наведения на них, режиссер может оперировать не координатными командами, а объектными указаниями, например, “следить за докладчиком”.

Система распознавания (и узнавания), работающая в связке с видеокамерой, дополняет данные, получаемые от датчика движения (в нашем случае используется Microsoft Kinect).

Основные задачи в рамках этой части проекта:

1. Определение и трекинг координат людей, их голов и рук в видимой области (сцене);
2. Передача этих координат с пересчетом для каждой установленной в системе видеокамеры, с учетом её смещения в пространстве относительно датчика;
3. Логирирование всех движений, в привязке к идентификаторам людей, для дальнейшего анализа и персонализации моделей поведения.

4. Алгоритмизация режиссуры

Режиссерская работа априори считается искусством и не подлежит автоматизации. Так ли это? Режиссура кино – вероятнее всего да, но стоит ли отказываться от возможности автоматизировать рутинные процессы. Как уже говорилось выше, жесткие алгоритмы поведения съемочного комплекса не рассматриваются в качестве единственной основы для принятия решений, предусмотрено обучение нейросети на примерах и при помощи компьютерного моделирования ситуаций в виртуальной среде. Но основа для принятия операторских и режиссерских решений также существует.

Рассмотрим основные правила, которые следует заложить в алгоритмы съемки и монтажа, чтобы съемочный комплекс смог начать вести пригодную для демонстрации съемку.

1. Композиция кадра. Статика

а. **Правило третей.** Кадр делится на три части по вертикали и горизонтали, ключевые точки находятся в пересечениях линий раздела. Для съемки человека обычно используются верхние.

б. **Продолжение взгляда.** Следует оставлять больше места в направлении, куда повернута голова или направлен взгляд. Применительно к правилу третей, следует разместить лицо (глаза) в верхней левой точке пересечения линий, если он смотрит вправо, и в верхней правой – если смотрит влево. Если смотрит в камеру, то следует руководствоваться наиболее выигрышным фоном.

с. **Целостность кадра.** Стараться не резать кадр по лицу соседа, на средних планах не отрезать руки, если они согнуты или находятся поверх стола.

2. Динамическая композиция

а. **Панорамирование** следует производить из статического положения и заканчивать статическим, плавно набирая скорость и плавно останавливаясь. Движение камеры во время панорамы должно быть равномерным.

б. Следует избегать **одновременного панорамирования и изменения фокусного расстояния**, поскольку при постоянной угловой скорости враще-

ния камеры, визуально скорость изменится.

с. **Направление панорамы** не должно меняться в пределах одного кадра, кроме задачи слежения за объектом.

д. При **монтаже двух панорам** направления должны быть одинаковыми.

3. Монтажные переходы

а. **Монтаж по крупности планов.** При склейке двух планов разной крупности следует избегать соседних и слишком разных. Правильный переход – через один шаг.

б. **Монтаж по направлению съемки.** При переключении камер следует сохранять направление движения или взгляда. Человека на сцене не следует снимать с противоположных углов зала, лучше выбрать ракурсы от центра и в одну из сторон.

с. **Монтаж по фону.** При съемке одного объекта несколькими камерами желательно, чтобы фон, видимый с одной камеры, частично был виден с других ракурсов. Предпочтительное перекрытие – 30%.

4. Динамика сцены

а. В зависимости от типа события, **средняя длина кадра** может варьироваться. Для лекций и докладов кадр будет длиннее, для концертов и диалогов – короче. Пределы длин кадров устанавливаются для различных типов событий индивидуально, но в целом лежат в пределах от 3 до 30 секунд.

б. **Смена кадров** производится или в такт музыке (если идет музыкальная передача), или по смыслу изложения, или по смене выступающего. При сопровождении презентацией – по смыслу или по обращению выступающим внимания зрителей на презентацию.

5. Сопутствующая информация

а. При трансляции в интернет возможно **сопровождение видеоряда контекстной информацией**, взятой из сети. Эта тема не входит в рассмотрение в данной статье, но в режиссерском плане она также может присутствовать, будем иметь её в виду.

б. **Титры.** Обычно титры появляются в начале и конце выступления. Во время длинных выступлений желательно держать титры с именем и должностью (организацией) выступающего на экране или показывать их с некоторой периодичностью.

5. Создание персонализированных моделей поведения актеров

Для создания модели актера съемочный комплекс должен уметь узнавать людей в кадре, анализировать их поведение и систематизировать результаты анализа для дальнейшего применения с конкретным актером или с другими, проявляющими подобные характеристики.

Допустим, на конференции сегодня два доклада: первым выступает импульсивный молодой человек, временами бегающий по сцене к экрану и размахивающий руками. Следом за ним выступает престарелый профессор, сохраняя одну позу на протяжении всей своей речи. Режиссер и операторы интуитивно понимают по первым движениям выступающих, что можно ожидать от них дальше — крупный план молодого подвижного докладчика будет опасным, велик риск промахнуться, когда он резко отойдет в сторону, зато есть множество событий, по которым можно и нужно будет переключить камеру. В случае со статичным профессором все наоборот — можно выбирать любую крупность плана, но придется искать повод для переключения камеры или переключать без повода, просто чтобы зритель не заскучал.

Но самое интересное начнется на следующий день: если операторы и режиссер запомнили особенности поведения каждого из докладчиков, то, когда те снова поднимутся на сцену, съемка будет вестись с учетом полученного опыта.

Другой пример — запись цикла лекций, который читает один профессор. В таком случае обучение съемочного комплекса на тестовых записях позволит избежать брака при чистой съемке.

Персонализация моделей поведения имеет свое развитие, выходящее за пределы потребностей съемочного комплекса, но, вероятно, востребованное в более широком контексте его применения. По аудиовизуальным данным о выступающем, можно провести классификацию его психотипа. Однако полезность и точность именно такого подхода вызывает сомнения, так как решаемые прикладные задачи являются нетипичными для такой классификации, и более перспективный метод — использование кластеризации с определением заранее неизвестного числа кластеров. Такое исследование возможно только при накоплении достаточной статистической выборки и в настоящее время не проводится.

Если на той же конференции, приведенной ранее в пример, на сцене появляется еще один профессор, по своему поведению в первые минуты “похожий” на предыдущего, которого система уже “научилась” снимать наилучшим образом, то в начале выступления система сопоставляет его поведение с имеющейся “библиотекой” и выбирает наилучшую стратегию съемки, чтобы соблюсти баланс выразительности и “безопасности” кадров.

На практике более значимое применение моделей актера — это изменение типа события. Если утром на сцене шла официальная часть со статичными выступлениями, а вечером там же состоялся концерт,

то профиль съемки и динамика монтажа изменятся. Имея заготовленные модели и зная их аудиовизуальные признаки, комплекс должен автоматически выбрать подходящий профиль съемки. Разумеется, нельзя исключать и ручной режим переключения профилей.

Обучение съемочного комплекса по различным кейсам можно проводить на реальных примерах, но это дорогостоящий способ и он не позволяет охватить все виды событий. Для первоначального обучения может применяться заранее подготовленный и тегированный набор записей различных событий (лучшие практики), а также компьютерное моделирование различных ситуаций (кейсы). Рассмотрим этот подход подробнее.

Подытожим задачи в этом разделе:

1. Разработать набор критериев оценки поведения людей на сцене на базе аудиовизуальной регистрации их действий.
2. Опираясь на “узнавание” людей при помощи видеокамер и датчика движения, а также на лог действий каждого из зарегистрированных ими людей, составить обобщенную модель поведения для каждого человека.
3. Провести кластеризацию полученных данных, с выявлением психотипов выступающих.

6. Обучение съемочного комплекса

Для обучения съемочного комплекса могут применяться:

- жесткие алгоритмы;
- специально адаптированные примеры реальных съемок;
- компьютерное моделирование ситуаций.

Базовые алгоритмы могут определять лишь основные правила поведения комплекса. Адаптироваться к различным ситуациям поможет обучение на примерах и моделях. Примеры дают более близкую к жизни картину, но требуют ручной подготовки материалов и тегирования, моделирование позволяет имитировать любую ситуацию, но сохраняется условность моделирования.

После обучения комплекс проходит проверку и оценку: его решения оцениваются экспертами, дающими обратную связь для закрепления или переобучения.

Рассмотрим подробнее моделирование ситуаций и процесс обучения комплекса.

Общая схема выглядит следующим образом (рис. 3).

В качестве виртуальной среды используется игровой движок, в котором создается помещение, аналогичное по своим значимым параметрам месту предстоящей съемки (если обучение проводится целена-

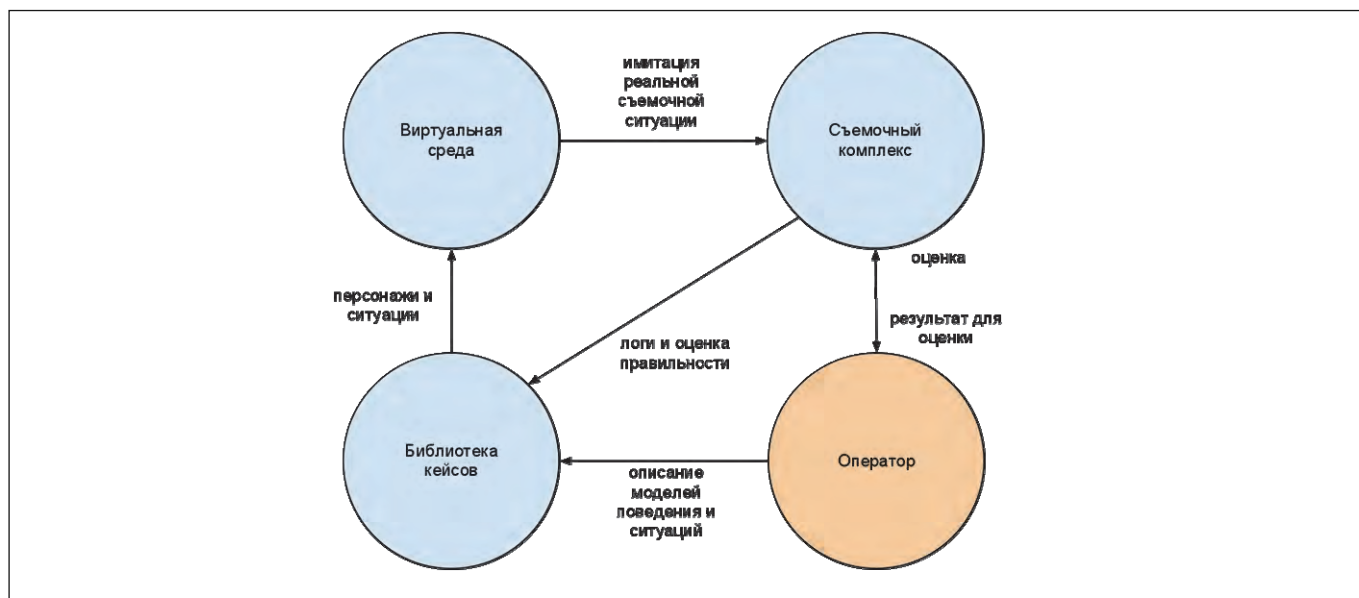


Рис. 3. Обобщенная модель тренировки съемочного комплекса на виртуальном тренажере

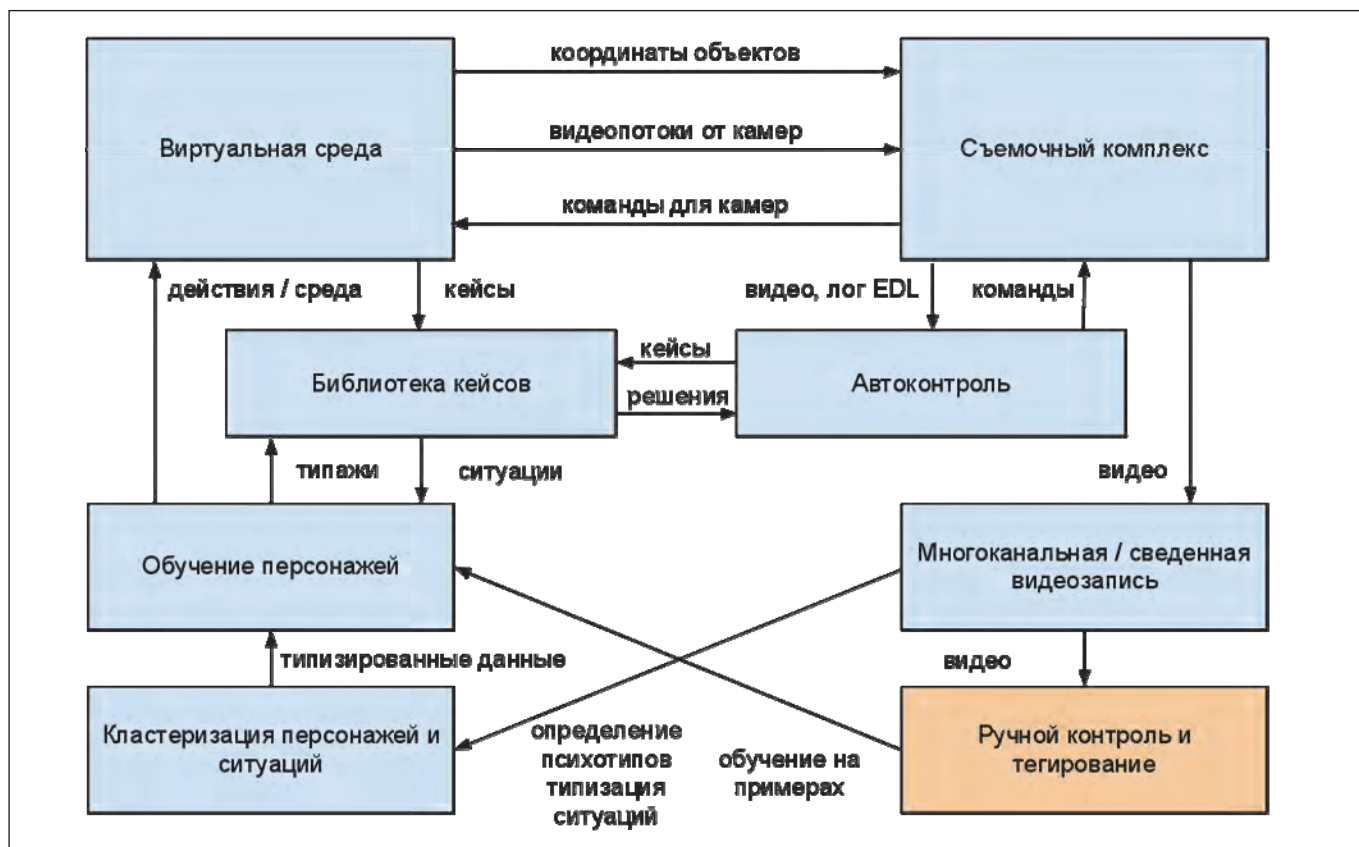


Рис. 4. Детализированная схема обучения съемочного комплекса

правленно перед съемкой), или берутся типовые помещения, если производится первоначальное обучение системы. В созданное виртуальное пространство помещаются персонажи, имеющие поведение, характерное для заданной ситуации (конференция, ток-шоу и т.д.). В помещении размещаются вир-

туальные камеры и виртуальный датчик, при этом "датчик" выдает такой же набор данных об "актерах" в виртуальном помещении, какой выдавал бы физический датчик в обычном зале, а камеры выдают видеопоток в том же формате, в каком его выдают настоящие камеры, и выполняют инструкции, как

это делают настоящие PTZ-камеры. То есть, для модуля обработки съемочного комплекса виртуальный и реальный набор датчиков и камер являются идентичными.

Далее отрабатываются как штатные ситуации (шестичасовая конференция с докладами по 10 минут и вопросами из зала), так и различные выходы за формат события (внезапная активность на сцене, бурная дискуссия с перекрестными выступлениями множества участников, потеря связи с одной из камер, уронили датчик, и т.д.).

Предложенный метод обучения интересен тем, что позволяет задать не алгоритм действия на все случаи жизни, а моделировать сами эти случаи, а потом скорректировать по необходимости действия системы при помощи обратной связи. Также появляется возможность, которой нет в реальной жизни – комплекс может “пристреляться” к помещению еще до попадания туда. Причем, благодаря моделированию, могут быть отработаны все возможные варианты – например, выход на сцену артиста из всех возможных точек. Все это производится без участия человека, комплекс обучается самостоятельно, задача человека состоит лишь в том, чтобы оценить решения управляющей системы в критических местах. Для этого должен быть предусмотрен инструмент оценивания при просмотре записи, в том числе в режиме ускоренного воспроизведения.

В итоге мы получаем систему, в которой компьютер учит компьютер. На входе закладываются не алгоритмы поведения, а сами ситуации, смоделированные визуально в виртуальном помещении.

В развитие такого подхода можно предложить съемку постановочных сцен, когда постановщик моделирует в движке снимаемую сцену, съемочный комплекс делает несколько версий съемки (возможно использование различных профилей, учитывающих вариацию динамики монтажа и съемки, реакцию на слова или речевые всплески), после чего выбранный вариант принимается для производственной съемки и дальше реальные актеры отыгрывают ту же сцену, а съемочный комплекс записывает ее с учетом заранее выбранного съемочного сценария. Это значительно упростит съемку мыльных опер и рекламных роликов, где запись проводится по дублям и заранее написанному тексту, и раскадровкам. При этом время актеров не тратится на творческие поиски, а съемка ведется с минимумом технических сбоев.

Подытоживая, отметим следующие задачи:

1. Алгоритмизировать существующие правила, выделив пределы допустимых погрешностей и областей применения. Показать достаточность

введенного базового набора правил для начала работы съемочного комплекса без обучения.

2. Построить нейросеть для обучения съемочного комплекса на примерах.
3. Создать на базе игрового движка модель с интерфейсами “управляемая видеочамера” и “датчик движения” для обучения съемочного комплекса путем воспроизведения смоделированного в виртуальном пространстве типового (нетипичного) поведения людей на мероприятиях определенного типа.

7. Языки управления

Для режиссерского управления съемочным процессом исторически существовали голосовой, визуальный и кнопочный языки. Для голосового обычно предусмотрена служебная связь, визуальный – руками можно показать оператору, что ему делать, а основной инструмент режиссера многокамерной записи – это видеомикшер, где сводятся все входящие сигналы и получается итоговая передача, микшер управляется кнопками и движками.

Мы создаем принципиально новую схему работы съемочного комплекса, но сохраняем суть выполняемой работы и предусматриваем возможное наличие режиссера (в таком случае автоматическое переключение каналов не задействуется).

Чтобы понять задачу создания языков управления, нужно сформулировать роль режиссера в работе съемочного комплекса.

- Видеокамеры умеют брать план и вести объект. Они знают, что такое композиция кадра и могут делать правильные панорамы.
- Автоматика не позволит переключиться на расфокусированную камеру, на несовместимый план или не готовую камеру.
- Съемочный комплекс распознает людей и может обеспечить многоракурсную разноплановую съемку выступающего, быстрое наведение на других людей благодаря датчику движений, отслеживающему состояние всей сцены.
- Даже фокус, который в автоматическом режиме иногда уходит с объекта съемки, контролируется системой трекинга – камера не просто направлена в нужную сторону, она именно знает, что она снимает, и фокус корректируется с учетом этого знания, а не по формальным признакам.
- Экспозиция камер оценивается и корректируется централизованно.

То есть, режиссеру остается только переключать камеры и давать команды, что этим камерам снимать. Если с переключением камер все относительно привычно, то командовать камерами придется иначе, операторы-люди в комплексе не предусмотрены.

Учитывая наличие у видеокамер навыков трекинга и определения планов, а также наличие датчика, “ведущего” все объекты на сцене, становится возможным поднять уровень общения человека-режиссера с компьютерным съемочным комплексом до уровня общения с человеком-оператором – команды “взять крупный план ведущего” или “панорама от трибуны до Иванова” уже не выглядят научной фантастикой. Но здесь мы затрагиваем только голосовой интерфейс.

При общении с компьютером человек привык использовать клавиатуру и мышь, в последние годы все больше привыкает к сенсорным экранам. Емкостные сенсоры сделали подобное общение с компьютером удобным и функциональным, пригодным для широкого спектра применений. Специфика замены кнопочного ввода сенсорным дисплеем состоит в контекстной привязке элементов, отображаемых на экране и, следовательно, элементов контроля, к текущей ситуации. Если обычно пользователю предлагалась или стандартная, или специализированная клавиатура, реже – адаптированный для конкретных задач манипулятор, то с использованием сенсорного ввода мы можем выводить на экран любую информацию и получать команды в виде multi-touch команд. И в этом случае совсем не обязательно, чтобы управляющие элементы повторяли знакомые аппаратные кнопки.

Все, кто пользовался современными смартфонами с экранной клавиатурой, знают, что аппаратная клавиатура удобнее. В то же время, популярность этих устройств говорит сама за себя – несравнимо большая функциональность и принципиально новые возможности смартфонов делают этот недостаток не столь существенным, и помогает им в этом продуманный графический интерфейс. Следовательно, наша задача – создать пользовательский интерфейс, решающий задачи режиссера многокамерных съемок и трансляций, а также определить набор команд (это уже второй язык после голосового управления), реализуемый этим интерфейсом.

В заключение темы приведем самый древний язык – жесты. Казалось бы, зачем режиссеру делать пассы руками или вертеть головой, если можно просто использовать ранее описанные интерфейсы? Но всегда ли режиссер – единственный, кто управляет ходом эфира? Возможно, ведущему требуется дать знак, чтобы с него перевели камеру, или лектор просит показать следующий слайд – эти случаи встречаются постоянно, и всегда в зале или в аппаратной должен присутствовать внимательный человек, следящий за этими знаками. В нашем случае роль внимательного наблюдателя выделяется специально

созданному для этого устройству – датчику движений Kinect. Он с легкостью определяет жесты, и эти жесты могут быть интерпретированы в команды для съемочного комплекса или для презентационного компьютера. Наша задача – определить эти команды, и это будет третий язык, который нам понадобится в съемочном комплексе для управления ходом эфира или записи.

Важно отметить, что множества команд и их атрибутов в этих языках могут, но не должны совпадать. Для каждого интерфейса взаимодействия характерна своя область применения и свои сильные выразительные средства, свои ограничения. Языки следует составлять с учетом применимости в самых разных ситуациях.

К ограничениям нужно также добавить влияние внешней среды: при ведении трансляции с концерта, где играет громкая музыка, голосовое управление затруднено. Существуют технические средства снизить акустические помехи, но падение разборчивости команд повышает риски ошибок. Аналогичные факторы существуют и для остальных интерфейсов: яркое солнце, осадки или низкая температура мешают работать с сенсорным ЖК-экраном, а концертное освещение может создавать помехи работе датчика движения.

8. Развертывание комплекса

При установке съемочного комплекса на месте съемки существуют задачи как художественного, так и технического плана. Удачное размещение камер позволит давать хорошие планы и ракурсы, избегать попадания случайных людей или нефотогеничных объектов в кадр. В то же время, взаимное расположение камер и датчика движения должно быть известно обработчику съемочного комплекса с высокой точностью, для пересчета координат ведомых объектов, полученных датчиком для каждой из камер.

Встает задача автоматического определения расположения камер. В идеальном случае, находясь в прямой видимости, камеры и датчик должны автоматически определить свое расположение в помещении, без дополнительных маркеров и ручного ввода координат, основываясь на визуальной оценке размеров (а по ним – расстояния) видимых камер и датчика.

9. Третье измерение

Важным следствием использования автоматического управления камерами является возможность относительно легкого перехода на 3D-вещание. Поскольку вся необходимая для позиционирования информация уже доступна в системе, то возможно обеспечить бинокулярное сведение пар камер. Таким образом,

съемочный комплекс на 4 PTZ-видеокамеры сможет использоваться и как четырехкамерный и как двухкамерный в 3D-формате. Существенные отличия в обработке начинаются на этапе кодирования видео.

Существующие PTZ-видеокамеры имеют ширину несколько большую, чем стандартная стереобаза (65-70 мм). Для достижения корректного стереоэффекта их следует устанавливать на удалении от объектов съемки и вести съемку на увеличенном фокусном расстоянии, что в целом соответствует типовой съемочной ситуации, когда камеру ставят вдоль стен или в конце зала. Это также позволяет избегать слишком больших углов поворота камеры – так как поворачивается не вся стереопара, а только камерные головки; в стереопаре PTZ-камеры имеют ограничение по горизонтальному углу поворота, так как в кадр попадает соседняя камера. При съемке вблизи будет также сказываться разница в расстоянии до объекта.

Для съемок удаленных объектов стереобазу можно искусственно увеличивать как вручную, устанавливая камеры на расстоянии друг от друга, так и в автоматическом режиме, снабдив их приводом.

Заключение

В заключение подытожим темы для разработки, очерченные в ходе описания съемочного комплекса.

1. “Авто-оператор”. Трекинг указанного объекта (человека) с соблюдением заданной крупности плана и композиции.
2. Узнавание человека и запоминание его характерных действий.
3. Трекинг нескольких человек при помощи датчика Kinect, взаимодействие с камерами.
4. “Авто-режиссер”. Определение совместимых ракурсов и планов, выявление моментов переключения и панорамирования.
5. “Авто-установка”. Позиционирование камер в зале и определение расстояний и направления между камерами и датчиком.
6. 3D-Моделирование сцены по данным датчика и метаданным от режиссера.
7. Графический пользовательский интерфейс режиссера для управления съемочным комплексом.
8. 3D-Моделирование ситуаций и обучение по ним съемочного комплекса.
9. Накопление и кластеризация профилей выступающих для оптимизации операторских и режиссерских решений.
10. Предварительное моделирование постановочных сцен и автоматизация их съемки с реальными актерами.

11. Автопозиционирование стереопары при 3D-съемке.

Каждая из перечисленных тем является самостоятельным направлением исследования и разработки со своими практическими результатами, применимыми безотносительно использования в рамках рассматриваемого съемочного комплекса.

Статья опирается на опыт создания систем видео-взаимодействия и успешные эксперименты по созданию автоматизированной камеры. Ряд зарубежных ученых проводил исследования и разработки. Наиболее полные материалы могут быть найдены на сайте Microsoft Research. Ниже перечислены материалы по основным разделам данной статьи.

ЛИТЕРАТУРА

К п. 1

- Yong Rui, Anoop Gupta and JJ Cadiz, [Viewing Meetings Captured by an Omni-Directional Camera](#). Collaboration and Multimedia Systems Group, Microsoft Research
- Bin Yu, Cha Zhang, Yong Rui, Klara Nahrstedt, [A Three-Layer Virtual Director Model For Supporting Automated Multi-Site Distributed Education](#)

К п. 2

- Соколов А.Г. Монтаж. Телевидение, кино, видео. Ч. 3. “Терминология и графический язык”. – М.: Изд. А. Дворников, 2000. 242с.
- Филиппов С.А. Киноязык и история. Краткая история кинематографа и киноискусства. – М.: Клуб “Альма Анима”, 2006. 208с.
- Yong Rui, Anoop Gupta, Jonathan Grudin, Liwei He, [Automating lecture capture and broadcast: technology and videography](#). Multimedia Systems. Digital Object Identifier 10.1007/s00530-004-0132-9 Springer Verlag 2004
- Yong Rui, Yunqiang Chen, [Better Proposal Distributions: Object Tracking Using Unscented Practice Filter](#). Collaboration and Multimedia Systems Group, Microsoft Research.
- Yunqiang Chen, Yong Rui, Thomas S. Huang, [JPDAF Based HMM for Real-Time Contour Tracking](#). Microsoft Research.
- Yunqiang Chen, Yong Rui, Thomas S. Huang, [Mode-based Multi-Hypothesis Head Tracking Using Parametric Contours](#). Microsoft Research.
- Yong Rui, Liwei He, Anoop Gupta and Qiong Liu, [Building an Intelligent Camera Management System](#). Collaboration and Multimedia Systems Group, Microsoft Research

- Wei Bian, Dacheng Tao and Yong Rui, Cross-Domain Human Action Recognition IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS PART B: CYBERNETICS, to appear. [pdf]

К п. 3

- Shahram Izadi, David Kim, Otmar Hilliges, David Molyneaux, Richard Newcombe, Pushmeet Kohli, Jamie Shotton, Steve Hodges, Dustin Freeman, Andrew Davison, and Andrew Fitzgibbon, KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera, ACM Symposium on User Interface Software and Technology, October 2011
- Richard A. Newcombe, Shahram Izadi, Otmar Hilliges, David Molyneaux, David Kim, Andrew J. Davison, Pushmeet Kohli, Jamie Shotton, Steve Hodges, and Andrew Fitzgibbon, KinectFusion: Real-Time Dense Surface Mapping and Tracking, in IEEE ISMAR, IEEE, October 2011
- Ross Girshick, Jamie Shotton, Pushmeet Kohli, Antonio Criminisi, and Andrew Fitzgibbon, Efficient Regression of General-Activity Human Poses from Depth Images, in ICCV, IEEE, October 2011

К п. 4

- M.N. Wallick, Y. Rui, L. He, A Portable Solution for Automatic Lecture Room Camera Management. Microsoft Research
- Yong Rui, Anoop Gupta, and Alex Acero. Automatically Extracting Highlights for TV Baseball Programs. Microsoft Research
- Yong Rui, Anoop Gupta and Jonathan Grudin. Videography for Telepresentations April 12, 2002

Technical Report MSR-TR-2001-92 Microsoft Research.

- Yong Rui, Anoop Gupta, Jonathan Grudin and Liwei He. Automating Lecture Capture and Broadcast: Technology and Videography. Microsoft Research.
- Yong Rui, Anoop Gupta and Jonathan Grudin. Videography for Telepresentations. October 2, 2001 Technical Report MSR-TR-2001-92 Microsoft Research.

К п. 6

- Oliver Willams, Andrew Blake, Roberto Cipolla, A Sparse Probabilistic Learning Algorithm for Real-Time Tracking. Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2Volume Set 0-7695-1950-4/03 IEEE 2003.
- Xuezheng Liu, Lei Zhang, Mingjing Li, Hongjiang Zhang, Dingxing Wang, Boosting Image Classification with LDA-based Feature Combination for Digital Photograph Management. Department of Computer Science and Technology, Tsingua University and Microsoft Research Asia, Beijing, China.

К п. 9

- Link: i2i – 3D Visual Communication / Microsoft Research, Cambridge.

*Королев Денис Александрович,
канд. техн. наук, доцент МГИЭМ
e-mail: denis.korolev@auditory.ru*

УВАЖАЕМЫЕ ЧИТАТЕЛИ !

Напоминаем вам, что продолжается подписка на журнал

КАЧЕСТВО. ИННОВАЦИИ. ОБРАЗОВАНИЕ

Подписку вы можете оформить:

- **через отделения связи**

каталог Агентства «Роспечать» – индекс 80620, 80621,

каталог «Пресса России» – индекс 14490,

- **через редакцию**

Дополнительную информацию можно получить

по телефону: + 7 (495) 916-28-07

УВАЖАЕМЫЕ АВТОРЫ !

**ОБРАТИТЕ ВНИМАНИЕ, ЧТО ЯВЛЯЯСЬ ПОДПИСЧИКОМ НАШЕГО ЖУРНАЛА,
ВЫ УСКОРИТЕ ПУБЛИКАЦИЮ СВОЕГО МАТЕРИАЛА**