

Разработка алгоритмов и программной подсистемы для подготовки видеолекций «Автослайд»

А. М. Гинцяк¹, Д. С. Баринов², Н. А. Абрамов³,

А. В. Никитина⁴, О. В. Ростова⁵

Санкт-Петербургский политехнический университет
Петра Великого (СПбПУ)

¹aleksei.gintciak@spbpu.com, dmitriy.barinov@spbpu.com,

³nikolay.abramov@spbpu.com, ⁴anna.nikitina@spbpu.com,

⁵o.2908@mail.ru

Я. М. Сомов

ООО «Лекториум»

somov@lektorium.tv

Аннотация. Описывается процесс разработки алгоритмов и сценариев обработки видеопотока с целью обнаружения, локализации, распознавания и классификации элементов в кадре, а так же программной среды для монтажа видеолекций с презентационными материалами «Автослайд». Данная система реализует названные алгоритмы и в автоматическом режиме выполняет временную синхронизацию видеопотока с презентационными материалами. Предполагается, что использование программного продукта Автослайд позволит на порядок уменьшить временные затраты монтажера при создании видеолекций.

Ключевые слова: видео; видеолекции; кадр; камера; классификация; обнаружение; образование; распознавание.

I. АКТУАЛЬНОСТЬ

Основными проблемами при подготовке образовательных программ являются требующие тщательной подготовки и планирования действия профессиональных операторов, режиссеров и монтажеров. В настоящее время для повышения успешности проведения монтажа видеопотока с нескольких камер, выполнения различных операций с видеоматериалом, применения спецэффектов, выполнения анализа видеоматериала в режиме реального времени, а также для включения в потоковую обработку данных текстовых и анимационных вставок, использования интерактивного досок необходима разработка автоматических программных средств монтажа, обеспечивающих снижение затрат времени на обработку видеоматериала.

Разрабатываемая программа предназначена для развития дистанционных форм образования в высшем и дополнительном профессиональном, школьном и дошкольном образовании. Основным конкурентом разрабатываемой автоматизированной системы съемки и монтажа является работа живых операторов и видеомонтажеров. В случае автоматизации себестоимость снижается, так как 80% расходов по видеосъемке – это оплата работы специалистов.

Разработка системы интеллектуального анализа видеолекций включает в себя следующие задачи:

- анализ этапов процесса подготовки дистанционного образовательного онлайн-курса с лекционными видеоматериалами;
- анализ алгоритмов синхронизации, предобработки и обработки видеопотоков;
- разработка комплекса алгоритмов и реализующего его программного обеспечения для автоматизированного монтажа видеолекций со слайдами презентаций.

II. ПОДГОТОВКА ЛЕКЦИОННЫХ МАТЕРИАЛОВ

Для создания видеолекции необходимо профессиональное техническое сопровождение, реализация которого возможна через структуры университета, осуществляющие профессиональную съемку, монтаж, создание анимации, графики, размещение на образовательном портале университета для возможности использования в обучении или с привлечением профессиональной организации, осуществляющей данную деятельность.

От получения заказа на создание видеолекции до внедрения в образовательный процесс через размещение на образовательном портале университета проходит несколько этапов: работа над сценариями лекций курса, составление графика записи лекций, разработки стиля курса, организация съемок, процедура съемок, черновой монтаж, экспертиза, корректура, внесения правок в черновой вариант видеолекций, чистовой монтаж, «упаковка» курса, размещение на образовательном портале университета, предоставление образовательных услуг [1]. Высокое качество образовательного контента обеспечивается только в случае наличия на каждом из этапов квалифицированного персонала.

В процессе научно-исследовательской работы был проведен анализ структуры и особенностей одного из наиболее

Работа выполнена при финансовой поддержке ФСИ, проект №213ГРНТИС5/35945 от 07.08.2017.

лее трудоемкого этапа подготовки видеолекций – процесса монтажа.

Монтаж видеолекций осуществляется в соответствии с принципами статической сборки [2], [3]. Статическая сборка – это система монтажа, основанная на статических изображениях слайдов с альфа-каналом (опционально) из подготовленных к конкретному ролику презентаций, накладывающихся поверх основного видео и анимационных переходов, которыми служат пресеты анимации в формате .mogrt.

В соответствии с основными методическими рекомендациями по разработке образовательного контента (в том числе и для национальной платформы открытого образования [4]), в процессе создания дистанционных образовательных онлайн-курсов выделяются следующие этапы:

- подготовка;
- разработка фирменного стиля;
- монтаж;
- монтаж нетипичных роликов;
- хранение и архивирование.

III. АНАЛИЗ АЛГОРИТМОВ

Проведенный анализ алгоритмов синхронизации видеоданных с двух камер сверхвысокого разрешения и алгоритмов временной синхронизации видеопотоков с презентационными материалами выявил перспективность использования в этих целях алгоритма, идентичного используемому в системах дистанционной проверки экзаменов. Предполагается, что выбор этого алгоритма обеспечит максимальную точность синхронизации среди всех рассмотренных [5], [6].

Проведенный анализ существующих подходов и алгоритмов предобработки видеоданных с целью повышения качества, корректировки цветовых моделей и моделей освещенности позволяет сделать вывод об отсутствии алгоритма, полностью соответствующего требованиям по комплексной предобработке видеоданных. Было принято решение разработать интеграцию нескольких алгоритмов и предложить их программную реализацию на базе метода гомоморфной фильтрации [7] и компиляции алгоритмов с адаптивными масками обработки [8]–[10].

Проведенный анализ используемых алгоритмов потоковой обработки данных с целью обнаружения, локализации, распознавания и классификации элементов в кадре в различных предметных областях показал, что данная функциональность является зависимой от области применения алгоритмов, требований по скорости обработки (средняя скорость обработки существующих решений – 15 кадров в сек), что требует дополнительной проработки возможности комплексного подхода, обеспечивающего высокую скорость обработки (до 30 кадров в секунду) и точность обнаружения, локализации, распознавания и классификации – как людей (для выделения в кадре лектора), так и прямоугольных элементов (области презентации для автозамены при плохом качестве). В качестве ба-

зовых подходов были выбраны метод MIL [11], [12], каскад Хаара [13].

В ходе исследования алгоритмов было определено, что на данный момент не существует единого комплекса алгоритмов, способного лечь в основу системы интеллектуального анализа видеолекций.

IV. РАЗРАБОТКА АЛГОРИТМОВ

В разработанной системе интеллектуального анализа видеопотока для реализации функциональности по обработке видеопотока можно выделить два важных структурных элемента:

- обработка кадров со слайдами;
- выделение лектора.

Обработка кадров со слайдами предназначена для обработки видеофайлов и слайдов презентации для определения координат углов слайдов в кадре и покадровой замены слайдов.

Входными данными являются:

- слайды презентаций;
- видеофайл лекции.

Выходными данными является массив координат углов слайда для каждого кадра.

Для реализации обработки кадров со слайдами были использованы следующие алгоритмы:

- поиск по ключевым точкам на слайдах (BRIEF) [14];
- поиск по ключевым точкам на видео (ORB) [15];
- фильтрация (фильтр Калмана) [16].

На первом этапе обработки кадров со слайдами осуществляется поиск элементов на слайдах с помощью дескриптора ключевых точек BRIEF. Дескриптор обеспечивает распознавание одинаковых участков изображения, которые были получены с разных углов обзора. В окрестности точки выбирается некоторым образом множество пар пикселей, и на данном множестве строится набор бинарных тестов, представляющих собой сравнение интенсивности в соответствующих пикселях. Результаты тестов объединяются в битовую строку, которая и является итоговым дескриптором ключевой точки.

На втором этапе обработки осуществляется поиск ключевых точек слайдов на кадрах и определение углов слайдов с помощью дескриптора ORB, который состоит из улучшенных алгоритмов FAST [17] и BRIEF.

В завершение обработки с помощью фильтра Калмана убираются несоответствия в масштабе слайдов, т.е. проводится синхронизация слайдов по размеру. Фильтр Калмана использует динамическую модель системы, известные управляющие воздействия и множество последовательных измерений для формирования оптимальной оценки состояния [18]. Алгоритм состоит из двух повторяющихся фаз: предсказание и корректировка. На первом рассчитывается предсказание состояния в следующий момент времени. На

втором, новая информация корректирует предсказанное значение. Схема алгоритма фильтрации представлена на рис. 1.



Рис. 1. Схема алгоритма фильтрации Калмана

Фильтр Калмана позволяет сглаживать движение между кадрами и масштаб слайдов, что приводит к минимизации дребезга в видеопотоке.

При реализации выделения лектора в видеопотоке осуществляется обнаружение, выделение и дальнейшее преобразование лектора к виду маски, которая будет накладываться на видео.

Входные данные:

- видеофайл лекции;
- массив координат углов слайдов для каждого кадра.

Выходные данные - бинарная маска, характеризующая положение лектора.

Система включает в себя программную реализацию алгоритмов вычисления набора метрик для сравнения областей:

- расстояние Хэмминга;
- среднеквадратичная дистанция;
- CCOR;
- CCOEFF.

Бинарная маска лектора передается в градациях серого, что позволяет выполнять цветокоррекцию. После получения координат маски лектора в каждом кадре, программа заменяет часть слайдов, которую лектор не перекрывает. Для финальной цветокоррекции генерируется файл в формате Final Cut Pro XML. Входные параметры - входное видео, набор изображений, для каждого из которых должны быть указаны 4 координаты для локализации изображения, и номера фрейма (или фреймов).

Следующим этапом является проектирование и разработка программного обеспечения, реализующего описанный комплекс алгоритмов.

V. ПРОЕКТИРОВАНИЕ И РАЗРАБОТКА СИСТЕМЫ

В качестве основного инструмента разработки был использован язык C++, который включает в себя библиотеку компьютерного зрения OpenCV.

Структурно подсистема «Автослайд» представляет собой набор различных модулей, не связанных между собой жёсткой иерархией:

- веб-фронтэнд – обеспечивает функционирование веб-интерфейса;
- планировщик задач – служит для извлечения информации о задачах из базы данных, передачи задачи обработчикам видео, подготовки задачи для передачи обработчикам и запуска обработки видео;
- обработчики видео – выполняют видеообработку;
- СУБД – обеспечивает хранение данных о задачах, пользователей, системной информации (mariadb);
- СХД – служит для хранения файлов и предоставления общего доступа к ним (NFS).

Подобная организация продиктована отсутствием необходимости построения сложных иерархических зависимостей и абстракций.

VI. РЕЗУЛЬТАТЫ РАБОТЫ СИСТЕМЫ

Разработанная система интеллектуального анализа видеолекций используется для автоматизированного монтажа видеолекций с презентационными материалами. Результат работы системы продемонстрирован на рис. 2 (исходный видеокادر) и рис. 3 (видеокادر после обработки).



Рис. 2. Исходный видеокادر



Рис. 3. Видеокادر после обработки

Презентационные материалы на исходном видеокадре не являются контрастными, даже крупный текст с трудом различим. Плохая цветопередача также является следствием как низкого качества проектора, так и освещения. Данные проблемы могут быть решены с использованием разработанной системы интеллектуального анализа видеолекций.

Обработанный видеокадр, в отличие от исходного, содержит хорошо различимые презентационные материалы. Контрастность и цветопередача в этом случае не зависят от качества проектора, качества съёмки и освещённости в аудитории.

VII. ЗАКЛЮЧЕНИЕ

В статье описаны комплекс алгоритмов и процесс создания системы интеллектуальной обработки видеолекций, включающий в себя следующие шаги:

- анализ этапов процесса подготовки дистанционного образовательного онлайн-курса с лекционными видеоматериалами;
- анализ алгоритмов синхронизации, предобработки и обработки видеопотоков;
- разработка комплекса алгоритмов обработки видеокадров со слайдами и выделения лектора;
- проектирование и разработка системы для автоматизированного монтажа видеолекций со слайдами презентаций.

Результаты апробации разработанной системы и анализа результатов автоматической обработки видеокадров позволяют сделать вывод о повышении качества видеолекций за счёт лучшей различимости текста и графических объектов на презентационных материалах.

В настоящий момент подана заявка на государственную регистрацию программы для ЭВМ, номер заявки АААА-Г18-618040400050-0.

СПИСОК ЛИТЕРАТУРЫ

[1] Чубаркова Е.В. Видеолекции – перспективы для совершенствования и развития новых образовательных технологий /Е.В. Чубаркова, А.О. Прокубовская // Инновации в профессиональном и профессионально-педагогическом образовании: материалы 20-й Всероссийской научно-практической конференции, 22-23 апреля 2015 г., г. Екатеринбург / Рос. гос. проф.-пед. ун-т. Екатеринбург, 2015. Т. 1. С. 347-350.

[2] Бегенева Е.И. Техника построения мегатекста языкового учебного курса: параллельный и ассоциативный монтаж. // Научный вестник Воронежского государственного архитектурно-строительного

университета. Серия: современные лингвистические и методико-дидактические исследования. 2015. №2 (26). С. 93–105.

[3] Никишина В.Б., Запесоцкая И.В., Кузнецова А.А. Технология создания видеолекций: мифы и реальность // Современные проблемы науки и образования. 2017. №4. С. 140.

[4] Требования и рекомендации по разработке онлайн-курсов, публикуемых на национальной платформе открытого образования. Подготовлен рабочей группой Совета Министерства образования и науки Российской Федерации по открытому образованию. Утверждён заседанием Правления Ассоциации «НПОО», протокол №2 от 23.07.2015.

[5] Объединение видеофрагментов с нескольких камер и синхронизация их по времени. [Электронный ресурс: <https://habrahabr.ru/post/277179/>] (дата обращения 17.01.2018).

[6] Синхронизация кадров видеоизображения с частотой обновления экрана. [Электронный ресурс: [https:// software.intel.com/ru-ru/articles/video-frame-display-synchronization](https://software.intel.com/ru-ru/articles/video-frame-display-synchronization)] (дата обращения 17.01.2018).

[7] Oppenheim A.V., Shafer R.W., Stockham Jr.T.G. Nonlinear ltering of multiplied and convolved signals, Proc. IEEE, 1968, Vol. 56, No. 8, pp. 1264–1291.

[8] Mantiuk R., Daly S., Kerofsky L. Display adaptive tone mapping // ACM Trans. on Graph, vol. 27(3), 2008. P.1-10.

[9] Kober V., Mozerov M., Alvarez-Borrego J., Ovseyevich I. A. Rank image processing using spatially. Adaptive neighborhoods. Pattern Recognition and Image Analysis, 2001, Vol. 40, No. 6, pp. 971–983.

[10] Kober V., Mozerov M., Alvarez-Borrego J. Nonlinear filters with spatially-connected neighborhoods. Optical Engineering, 2001, Vol. 40, No. 6, pp. 971–983.

[11] Babenko B., Yang M. H., Belongie S. Robust object tracking with online multiple instance learning //IEEE transactions on pattern analysis and machine intelligence. 2011. T. 33. №. 8. С. 1619-1632.

[12] Y. Freund and R.E. Schapire, “A Decision-Theoretic Generalization of Online Learning and an Application to Boosting” J. Computer and System Sciences, vol. 55, pp. 119-139, 1997.

[13] Bradski G., Kaehler A. Learning OpenCV: Computer vision with the OpenCV library. "O'Reilly Media, Inc.", 2008.

[14] Calonder, Vincent Lepetit, Christoph Strecha, Pascal Fua, “BRIEF: Binary Robust Independent Elementary Features”, 11th European Conference on Computer Vision (ECCV), pp. 778–792, 2010.

[15] Ethan Rublee, Vincent Rabaud, Kurt Konolige, Gary Bradski: "ORB: an efficient alternative to SIFT or SURF", Computer Vision (ICCV), IEEE International Conference on. IEEE, pp. 2564 – 2571, 2011.

[16] Сирота А.А., Иванков А.Ю. Блочные алгоритмы обработки изображений на основе фильтра Калмана в задаче построения сверхразрешения. // Компьютерная оптика. 2014. №1 (38).

[17] Rosten, Edward, Tom Drummond: "Machine learning for high-speed corner detection", 9th European pp. 430–443, 2006.

[18] Mager V.E., Belousov V.E., Desyatirikova E.N., Polukazakov A.V., Ivanov S.A., Pochneva I.V. “Information processing algorithm at creation of optimum structure of the self-adjusted technical system in quality parameters”, in Proc. of 2017 IEEE 2nd International Conference on Control in Technical Systems, CTS 2017; St. Petersburg; Russian Federation; 25-27 October 2017, Publisher: IEEE, pp.118-121. DOI: 10.1109/CTSUS.2017.8109503