

ТЕХНИЧЕСКИЕ НАУКИ*Никитин Илья Константинович*

аспирант

Московский Авиационный Институт

(Национальный Исследовательский Университет)

г. Москва

ЭЛЕМЕНТЫ АССОЦИАТИВНОГО ПОИСКА ПО ВИДЕО

Аннотация: за последние несколько лет объем носителей мультимедийной информации вырос в несколько раз. В сложившейся ситуации необходим простой и гибкий поиск неструктурированным мультимедийных данных. В статье предлагается обзор существующих методов ассоциативного поиска по видео.

Ключевые слова: анализ видео; аннотирование видео; видео-поиск; кадры; классификация видео; нечеткие дубликаты видео; ранжирование видео; сцены; съёмки.

Введение

Существует смысловой разрыв между низкоуровневыми данными в видео и потребностями пользователя. Тем не менее, большинство методов поиска по видео следуют парадигме прямого отображения низкоуровневых характеристик на смысловые понятия. Этот подход требует предварительной обработки данных. Результаты такого отображения не будут устойчивы. Без учета конкретной предметной области задача кажется неразрешимой. Последнее время стало появляться много клипов с очень схожим содержанием (нечеткие дубликаты видео). Задача эффективной идентификации нечетких дубликатов играет ключевую роль в задачах поиска, защите авторских прав, и многих других. Для анализа большого объема видео-данных применяют ассоциативный поиск. В англоязычной литературе ассоциативный видео-поиск называют «content based video retrieval» (CBVR) – поиск по содержанию.

Ассоциативный поиск используется для автоматического реферирования видео, анализа новостных событий, видеонаблюдения, и в образовательных целях [1]. Видео содержит в себе несколько видов данных. Авторы [2] и [3] выделяют четыре вида.

1. Метаданные – заголовок, автор и описание;
2. Звуковая дорожка.
3. Тексты полученные при помощи технологии оптического распознавания символов (OCR);
4. Визуальная информация кадров видео.

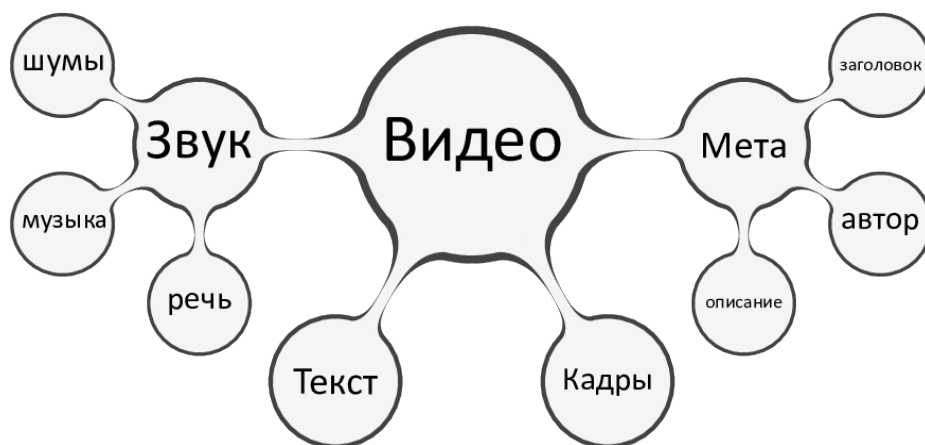


Рис 1. Структура видео

Видео обладает комплексностью [4]. Комплексность видео состоит в возможности автора выразить мысли, используя по крайней мере два информационных канала. Каналы могут быть визуальными, звуковыми или текстовыми.

Таблица 1

Некоторые научные работы по ассоциативному поиску видео

<i>Год</i>	<i>Работа</i>	<i>Тема</i>
2008	[5]	Сегментация видео
2011	[6]	
2010	[7]	Автоматическое реферирование Видео
2012	[8]	
2012	[9], [10], [11]	Индексация видео
2013	[12]	
2014	[13]	
2012	[14]	Комплексный ассоциативный поиск
2013	[15]	
2011	[16], [17]	Представление видео
2012	[18], [19], [20]	Смысловой ассоциативный поиск
2012	[21], [19]	Аннотирование видео
2012	[22]	Видео-поиск по движению
2011	[23]	Ранжирование видео
2012	[21]	
2010	[24]	Классификация видео
2011	[25]	
2012	[26], [27]	

В работе [21] дают хороший обзор аннотации видео. В работе [23] описываются свежие исследования методов ранжирования видео. Ассоциативный поиск видео состоит из следующих шагов.

1. Анализ временной структуры видео – деление видео на фрагменты, которое включает обнаружение границ съёмок.
2. Определение характеристик фрагментов.
3. Извлечение информации из характеристик.
4. Аннотация видео, построение семантического индекса.
5. Обработка пользовательского запроса и выдача результата.
6. Обратная связь и переранжирование результатов для улучшения поиска характеристик



Рис 2. Схема поиска по видео

Деление видео

Деление видео включает в себя обнаружение границ съёмок, извлечение ключевых кадров, сегментацию сцен и аудио.

Обнаружение границ съёмок

Видео делят на фрагменты по времени. В качестве таких фрагментов могут выступать съёмки. Съёмка (кинематографический кадр, монтажный план) – отрезок киноплёнки, на котором запечатлено непрерывное действие между пуском и остановкой камеры, или между двумя монтажными склейками.

Для обнаружения границ съёмок, как правило, сначала извлекают визуальные характеристики каждого кадра. Затем, на основе выделенных признаков, оценивают сходство между кадрами. Границы съёмок определяют по смене неоднородных кадров. В работе [4] описаны параметры смены кадров и ошибки выделения на основе глобальных и локальных характеристик для обнаружения съёмок и классификации.

Извлечение ключевых кадров

Среди кадров одной съёмки есть избыточность. Для ее уменьшения выделяют кадры, которые наиболее полно отражают содержание съёмки.

При извлечении ключевых кадров используют различные характеристики: цветовые гистограммы; края; очертания; оптические потоки. Способы извлечения подразделяются на шесть категорий: последовательное сравнение; глобальное сравнение; на основе ссылочных кадров; на основе кластеризации; на основе упрощения кривых; и на основе объектов или событий [28].

При последовательном сравнении ключевой кадр сравнивают с другими кадрами до тех пор, пока не будет найден «сильно отличный». Для сопоставления кадров используется цветовые гистограммы [29].

В работе [30] описано создание средней гистограммы без канала прозрачности. С помощью такой гистограммы описывается цветовое распределение кадров в съёмке.

Сегментация сцен

Сегментация сцен также известна как деление сюжета на блоки. Сцена представляет собой группу смежных съёмок. Эти съёмки связаны между собой конкретной темой или предметом. Сцены обладают семантикой более высокого уровня чем съёмки.

Существует три способа сегментации сцен: деление по ключевому кадру; деление на основе объединения визуальной и звуковой информации; деление по фону.

При делении сцен по ключевому кадру каждая съёмка представляется набором ключевых кадров. Для кадров выявляют их характеристики. Близкие по

времени кадры с близкими характеристиками группируют в сцены [31]. Далее, используя сравнение блоков ключевых кадров, вычисляют сходство между съёмками. Ограничение деления по ключевому кадру заключается в том, что кадры не могут эффективно представить динамическое содержание съёмки.

Съёмки в пределах сцены, как правило, связаны динамическим развитием сюжета в пределах сцены, а не сходством ключевых кадров.

При одновременном анализе звуковой и визуальной информации сменой сцен считают границу съёмки, где содержимое обоих каналов изменяется одновременно. Для определения соответствия между этими двумя наборами сцен используют алгоритм поиска ближайшего соседа с ограничением по времени [32]. К минусам подхода следует отнести сложность определения связи между аудио сегментами и визуальными съёмками.

Деление сцен по фону основано на гипотезе, что съёмки, принадлежащие к одной сцене часто имеют один и тот же фон. Для восстановления фона каждого кадра используют объединение близких по цвету пикселей в одноцветные прямоугольные области. Сходство съёмок определяют с помощью оценки распределения цвета и текстуры всех фоновых изображений в кадре. Для управления процессом группировки съёмок применяют кинематографические правила [33].

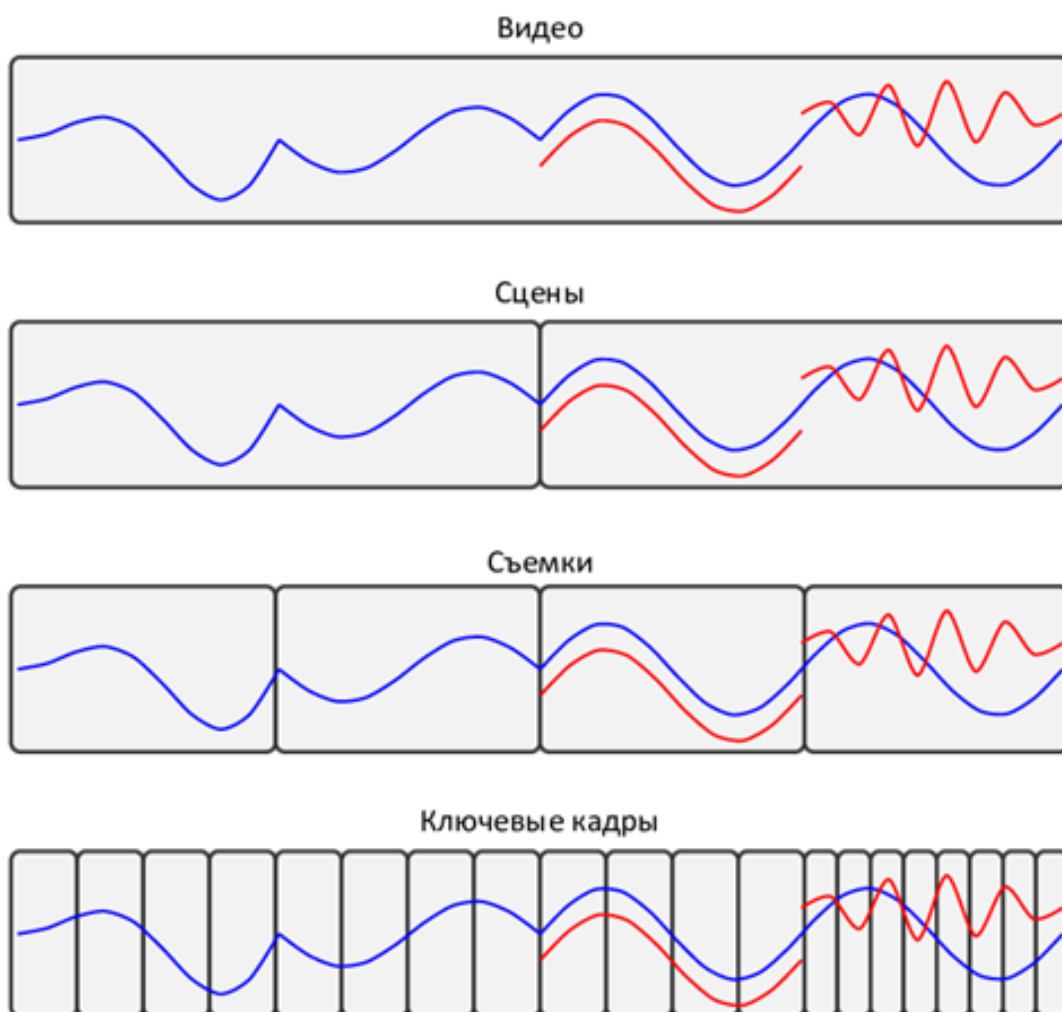


Рис 3. Схема визуальной сегментации видео на примере движения точек.

Временная ось отложена по горизонтали

Сегментация звука

Звуковая дорожка – богатый источник информации о содержании для всех жанров видео.

Как показано в лингвистической литературе, границы «высказываний» выделяются интонационно. На существенные изменения темы обычно указывают: длинные паузы, изменения тона и более изменением амплитуды колебаний.

На основе дерева принятия решений выбирают типичную интонационную функцию. После чего, лексическая информация извлекается с помощью Скрытых Моделей Маркова (НММ) и статистических моделей языка.

Выделение признаков

Из полученных частей видео выделяют признаки. К признакам относят: характеристики ключевых кадров; объекты; движение в кадре; характеристики аудио и текста.



Рис 4. Классификация признаков для выделения при ассоциативном видео-поиске

Характеристики ключевых кадров

Выделяют цветовые, текстурные, формовые, краевые характеристики.

Цветовые характеристики включают в себя цветовые гистограммы, цветовые моменты, цветовые коррелограммы, смесь Гауссовых моделей. При выделении локальной цветовой информации изображения разбивают на блоки 5 на 5 [35].

Текстурными характеристиками называют визуальные особенности

поверхности некоторого объекта. Они не зависят от тона или насыщенности цвета объекта. Текстурные характеристики отражают однородные явления в изображениях. Для выделения текстурной информации из видео применяют фильтры Габора [36].

Контурные или формовые характеристики, описывают формы объектов в изображениях. Они могут быть извлечены из контуров или областей объектов.

На конференции TRECVID-2005 для получения пространственного распределения краев в задаче поиска по видео был предложен дескриптор гистограммы границ (EHD) [37].

Характеристики объектов

Такие характеристики включают параметры областей изображения, которые соответствуют объектам: основной цвет; текстуру; размер и т.д.

В работе [38] предложена система поиска лиц. По видео-запросу с конкретным человеком система способна выдать ранжированный список съёмок с этим человеком. Текстовая индексация и поиск приводят к расширению семантики запроса и делают возможным использование Glimpse-метода (agrep), для поиска нечеткого соответствия [39].

Характеристики движения

Характеристики движения ближе к смысловым понятиям, чем характеристики ключевых статических кадров и объектов. Движение в видео может быть вызвано движением камеры и движением предметов в кадре.

Движения камеры такие как «приближение или удаление», «панорамирование влево или вправо» и «смещение вверх или вниз» используются для индексации видео. Движения объектов на данный момент являются предметом исследований.

Звуковые характеристики

Преимущество аудио-подходов состоит в том, что они обычно требуют меньше вычислительных ресурсов, чем визуальные методы. Кроме того, аудио-записи могут быть очень короткими.

Многие звуковые характеристики выбраны на основе человеческого

восприятие звука. Характеристики аудио можно разделить на три уровня [33]:

- низкоуровневая акустика, такая как средняя частота для кадра,
- средний уровень, как признак объекта, например, звук скачущего мяча,
- высокоуровневые, такие как речь и фоновая музыка, играющая в

определенных типах видео.

В работе [40] используют блочные характеристики аудио. Аудио-поток при этом разделяется на отрезки в 2048 отсчетов. Для выделения таких характеристик применяют функцию Ханна и логарифмическую шкалу.

Анализ видео

Интеллектуальный анализ данных в больших базах видео стал доступен недавно. Задачи анализа видеoinформации можно сформулировать как выявление: структурных закономерностей видео, закономерностей поведения движущихся объектов, характеристик сцены, шаблонов событий и их связей, и других смысловых атрибутов в видео.

В работах применяют извлечение объектов – группировку различных экземпляров того же объекта, который появляется в различных частях видео.

Для классификации пространственных характеристик кадров применяют метод поиска ближайших соседей [42].

Обнаружение специальных шаблонов применяется к действиям и событиям, для которых есть априорные модели, такие как действия человека, спортивные мероприятия, дорожные ситуации или образцы преступлений [43].

Поиск моделей – автоматическое извлечение неизвестных закономерностей в видео. Для поиска моделей используют экспертные системы с безнадзорными или полуконтролируемым обучением.

Поиск неизвестных моделей полезен для изучения новых данных в наборе видео. Неизвестные образцы обычно находят благодаря кластеризации различных векторов характеристик.

Для выявления закономерностей поведения движущихся объектов используют n-граммы и суффиксные деревья. При этом анализируют последовательности событий по многократным временным масштабам.

Классификация видео

Задача классификации состоит в том, чтобы отнести видео к предопределенной категории. Для этого используют характеристики видео или результаты интеллектуального анализа данных.

Классификация видео – хороший способ увеличить эффективность видео–поиска. Семантический разрыв между низкоуровневыми данными и интерпретацией наблюдателя, делает ассоциативную классификацию очень трудной задачей.

Смысловая классификация видео может быть выполнена на трех уровнях [15]: жанры, события видео и объекты в видео.

Жанры

Жанровая классификация разделяет видео на подмножество соответствующее жанру и несоответствующее [11].

В работе [44] предложена классификация большого числа видео только по заголовку видео. Для этого использован поэтапный метод опорных векторов.

Видео классифицируют также на основе статистических моделей различных жанров. Для этого анализируют структурные свойства: статистику цвета, съёмки, движение камеры и объектов. Свойства используются, чтобы получить более абстрактные атрибуты стиля. К абстрактным атрибутам стиля можно отнести: панорамирование камеры и изменение масштаба, речь и музыку. Строят отображение этих атрибутов на жанры видео.

В работе [27] для классификации жанров используется комбинация из четырех дескрипторов:

– *блоковый аудио дескриптор:*

1) захватывает локальную временную информацию;

– *дескриптор визуальной временной структуры:*

1) использует информацию о смене съёмки,

2) оценивает количество съёмки за определенный интервал времени («ритм» видео),

3) описывает об «активные» и «не активные» смены съемок;

– *дескриптор цвета:*

1) использует статистику распределения цвета, элементарных оттенков, цветовых свойств, и отношений между цветами;

– *статистика фигур контуров.*

Были проведены эксперименты на видеоматериалах общей продолжительностью 91 часе видео. Классификация проводилась на семи жанрах видео: мультфильмы, реклама, документальные фильмы, художественные фильмы, музыкальные клипы, спортивные соревнования и новости. Комплексный дескриптор позволил авторам достичь точности 87% – 100% и полноты 77% – 100%.

События

Событие может быть определено как любое явление в видео, которое может быть воспринято зрителем и одновременно играет роль для представления содержимого.

Каждое видео может состоять из многих событий, и каждое событие может состоять из многих подсобытий. Таким образом складывается иерархическая модель [45].

Объекты

Объектная классификация является самым низкоуровневым типом классификации. Съёмки классифицируют тоже на основе объектов. Объекты в съёмках представлены с помощью параметров цвета, текстуры и траектории. В работе [46] для кластеризации связанных съёмок используется нейронная сеть. Каждый кластер отображен на одну из 12 категорий. Объекты разделяются по положению в кадре и характеру движения.

Аннотирование видео

Процесс присваивания переопределённых смысловых понятий фрагментам видео называют аннотированием. Аннотирование видео подобно классификации, за исключением двух различий.

1. Для классификаций важны жанры, а для аннотирования понятия. Жанры

и понятия имеют различную природу, несмотря на то, что некоторые методы могут быть использованы в обоих задачах.

2. Классификация видео применяется к полным видео, а аннотируют обычно фрагменты [47].

Аннотирование, основанное на обучении, необходимо для анализа и понимания видео. Было предложено много различных способов автоматизации процесса.

Например, в работе [21] было разработано «быстрое полуконтролируемое графовое обучение на нескольких экземплярах» (Fast Graphbased Semi-Supervised Multiple Instance Learning – FGSSMIL). Алгоритм работает в рамках общей платформы для разных типов видео одновременно (спортивные передачи, новости, художественные фильмы). Для обучения модели используется небольшое число видео, размеченных вручную, и значительный объем не размеченного материала.

Заключение

Многие вопросы остаются открытыми и требуют дальнейшего исследования, особенно в следующих областях.

Большинство современных подходов индексации видео сильно зависят от предварительных знаний о предметной области. Это ограничивает их расширяемость для новых областей. Устранение зависимости от предварительных знаний – важная задача будущих исследований.

Современные подходы к смысловому поиску видео, как правило, используют набор текстов для описания визуального содержания видео. В этой области пока осталось много неразрешенных вопросов. Например, отдельной темой для исследования может быть эмоциональная семантика видео [15].

Эффективное использование информации о движении имеет большое значение для поиска видео. Важными задачами направления являются:

- способность различать движения фона и переднего плана;
- обнаружение движущиеся объектов и определять события в кадре;
- объединение статических характеристик и характеристик движения;

– построение индекса движения.

Интересными вопросами для исследования остаются:

– быстрый видео-поиск с помощью иерархических индексов;

– адаптивное обновление иерархической индексной модели;

– обработка временных характеристик видео во время создания и обновления индекса;

– динамические меры сходства видео на основе выбора статистических функций.

Список литературы

1. Dimitrova N. et al. Applications of Video-Content Analysis and Retrieval // IEEE MultiMedia. Los Alamitos, CA, USA: IEEE Computer Society Press, 2002. vol. 9, № 3. pp. 42-55.

2. Chung Y.Y. et al. Performance Analysis of Using Wavelet Transform in Content Based Video Retrieval System // Proceedings of the 2007 Annual Conference on International Conference on Computer Engineering and Applications. Stevens Point, Wisconsin, USA: World Scientific; Engineering Academy; Society (WSEAS), 2007. pp. 277-282.

3. Smeaton A.F. Techniques used and open challenges to the analysis, indexing and retrieval of digital video // Information Systems. 2006. vol. 32, № 4. pp. 545-559.

4. Nigay L., Coutaz J. A Design Space for Multimodal Systems: Concurrent Processing and Data Fusion // Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems. New York, NY, USA: ACM, 1993. pp. 172-178.

5. Haase K.B., Davis M.E., Davis M.E. Media Streams: Representing Video for Retrieval and Repurposing. 1995.

6. Kamabathula V.K., Iyer S. Automated Tagging to Enable Fine-Grained Browsing of Lecture Videos // 2012 IEEE Fourth International Conference on Technology for Education. Los Alamitos, CA, USA: IEEE Computer Society, 2011. vol. 0. pp. 96-102.

7. Fu Y. et al. Multi-View Video Summarization // Multimedia, IEEE

Transactions on. 2010. vol. 12, № 7. pp. 717-729.

8. Wang M. et al. Event Driven Web Video Summarization by Tag Localization and Key-Shot Identification // *Multimedia, IEEE Transactions on.* 2012. vol. 14, № 4. pp. 975-985.

9. Chen X., Hero A., Savarese S. Multimodal Video Indexing and Retrieval Using Directed Information // *Multimedia, IEEE Transactions on.* 2012. vol. 14, № 1. pp. 3-16.

10. Zha Z.-J. et al. Interactive Video Indexing With Statistical Active Learning // *Multimedia, IEEE Transactions on.* 2012. vol. 14, № 1. pp. 17-27.

11. Wu J., Worring M. Efficient Genre-Specific Semantic Video Indexing // *IEEE Transactions on Multimedia.* 2012. vol. 14, № 2. pp. 291-302.

12. Paul A. et al. Video Search and Indexing with Reinforcement Agent for Interactive Multimedia Services // *ACM Trans. Embed. Comput. Syst.* New York, NY, USA: ACM, 2013. vol. 12, № 2. pp. 25:1-25:16.

13. Asghar M.N., Hussain F., Manton R. Video Indexing: A Survey // *International Journal of Computer and Information Technology.* 2014. vol. 03, № 01. pp. 148-169.

14. Huurnink B. et al. Content-Based Analysis Improves Audiovisual Archive Retrieval // *Multimedia, IEEE Transactions on.* 2012. vol. 14, № 4. pp. 1166-1178.

15. Tamizharasan C., S C. A Survey On Multimodal Content Based Video Retrieval // *International Journal of Emerging Technology and Advanced Engineering.* Chennai, INDIA: Sri Sai Ram Engineering College. An ISO 9001:2008 Certified; NBA Accredited Engineering Institute, 2013. vol. 3.

16. Karpenko A., Aarabi P. Tiny Videos: A Large Data Set for Nonparametric Video Retrieval and Frame Classification // *Pattern Analysis and Machine Intelligence, IEEE Transactions on.* 2011. vol. 33, № 3. pp. 618-630.

17. Cheng X., Chia L.-T. Stratification-Based Keyframe Cliques for Effective and Efficient Video Representation // *IEEE Transactions on Multimedia.* 2011. vol. 13, № 6. pp. 1333-1342.

18. Jiang Y.-G. et al. Fast Semantic Diffusion for Large-Scale Context-Based

Image and Video Annotation // Image Processing, IEEE Transactions on. 2012. vol. 21, № 6. pp. 3080-3091.

19. Yu H.Q. et al. Using Linked Data to Annotate and Search Educational Video Resources for Supporting Distance Learning // Learning Technologies, IEEE Transactions on. 2012. vol. 5, № 2. pp. 130-142.

20. Andre B. et al. Learning Semantic and Visual Similarity for Endomicroscopy Video Retrieval // Medical Imaging, IEEE Transactions. 2012. vol. 31, № 6. pp. 1276-1288.

21. Zhang T. et al. A Generic Framework for Video Annotation via Semi-Supervised Learning // IEEE Transactions on Multimedia. 2012. pp. 1206-1219.

22. Chu W.-T., Tsai S.-Y. Rhythm of Motion Extraction and Rhythm-Based Cross-Media Alignment for Dance Videos // Multimedia, IEEE Transactions on. 2012. vol. 14, № 1. pp. 129-141.

23. Tian X. et al. Bayesian Visual Reranking // Trans. Multi. Piscataway, NJ, USA: IEEE Press, 2011. vol. 13, № 4. pp. 639-652.

24. Tahayna. B. et al. Optimizing support vector machine based classification and retrieval of semantic video events with genetic algorithms // Image Processing (ICIP), 2010 17th IEEE International Conference on. 2010. pp. 1485-1488.

25. Sargin M.E., Aradhye H. Boosting video classification using cross-video signals // Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference. 2011. pp. 1805-1808.

26. Lim J. et al. Adult movie classification system based on multimodal approach with visual and auditory features // Information Science and Digital Content Technology (ICIDT), 2012 8th International Conference on. 2012. vol. 3. pp. 745-748.

27. Ionescu B. et al. Video Genre Categorization and Representation using Audio-Visual Information // SPIE - Journal of Electronic Imaging. 2012. vol. 21(2).

28. Truong B.T., Venkatesh S. Video Abstraction: A Systematic Review and Classification // ACM Trans. Multimedia Comput. Commun. Appl. New York, NY, USA: ACM, 2007. vol. 3, № 1.

29. Zhang X.-D. et al. Dynamic Selection and Effective Compression of Key

Frames for Video Abstraction // Pattern Recogn. Lett. New York, NY, USA: Elsevier Science Inc., 2003. vol. 24, № 9-10. pp. 1523-1532.

30. Matsumoto K. et al. SVM-Based Shot Boundary Detection with a Novel Feature // 2012 IEEE International Conference on Multimedia and Expo. Los Alamitos, CA, USA: IEEE Computer Society, 2006. vol. 0. pp. 1837-1840.

31. Truong B.T., Venkatesh S., Dorai C. Scene Extraction in Motion Pictures // IEEE Trans. Cir. and Sys. for Video Technol. Piscataway, NJ, USA: IEEE Press, 2003. vol. 13, № 1. pp. 5-15.

32. Sundaram H., Chang S.-F. Video scene segmentation using video and audio features // Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on. 2000. vol. 2. pp. 1145-1148 vol.2.

33. Chen L.-H., Lai Y.-C., Mark Liao H.-Y. Movie Scene Segmentation Using Background Information // Pattern Recogn. New York, NY, USA: Elsevier Science Inc., 2008. vol. 41, № 3. pp. 1056-1065.

34. Repp S., Grob A., Meinel C. Browsing within Lecture Videos Based on the Chain Index of Speech Transcription // IEEE Transactions on Learning Technologies. Los Alamitos, CA, USA: IEEE Computer Society, 2008. vol. 1, № 3. pp. 145-156.

35. Yan R., Hauptmann A.G. A review of text and image retrieval approaches for broadcast news video // Inf. Retr. 2007. vol. 10, № 4-5. pp. 445-484.

36. Adcock J. et al. Fxpal experiments for trecvid 2004 // Proceedings of the TREC Video Retrieval Evaluation (TRECVID). 2004. pp. 70-81.

37. Hauptmann A.G. et al. Informedia at TRECVID 2003: Analyzing and Searching Broadcast News Video // Proceedings of the TRECVID Workshop. 2003.

38. Sivic J., Everingham M., Zisserman A. Person Spotting: Video Shot Retrieval for Face Sets // ACM International Conference on Image and Video Retrieval. 2005.

39. Li H., Doermann D. Video indexing and retrieval based on recognized text // Multimedia Signal Processing, 2002 IEEE Workshop on. 2002. pp. 245-248.

40. Seyerlehner K. et al. Using Block-Level Features For Genre Classification, Tag Classification And Music Similarity Estimation. 2010.

41. Su C.-W. et al. Motion Flow-Based Video Retrieval // Multimedia, IEEE

Transactions on. 2007. vol. 9, № 6. pp. 1193-1201.

42. Anjulan A., Canagarajah C. A unified framework for object retrieval and mining // IEEE Transactions on Circuits and Systems for Video Technology. 2009. vol. 19 (1). pp. 63-76.

43. Quack T., Ferrari V., Gool L. Video mining with frequent item set configurations // Int. Conf. Image Video Retrieval. 2006. pp. 360-369.

44. Jiang Y.-G., Ngo C.-W., Yang J. Towards Optimal Bag-of-features for Object Categorization and Semantic Video Retrieval // Proceedings of the 6th ACM International Conference on Image and Video Retrieval. New York, NY, USA: ACM, 2007. pp. 494-501.

45. Chang P., Han M., Gong Y. Extract Highlights From Baseball Game Video With Hidden Markov Models. 2002. pp. 609-612.

46. Hong G., Fong B., Fong A. An intelligent video categorization engine // Kybernetes. 2005. vol. 34, № 6. pp. 784-802.

47. Yang L. et al. Multi-modality Web Video Categorization // Proceedings of the International Workshop on Workshop on Multimedia Information Retrieval. New.