

УДК 681.3

Кластеризация кадров для сегментирования видеоданных

Вовк О.Л.

Донецкий национальный технический университет
vovk.olga@gmail.com

Abstract

Vovk O.L. The Shot Clusterization for Video Data Segmentation. This article is devoted to video data segmentation. In this work is proposed to use clusterization for find the video information fragments. Experimental estimation of video data segmentation by the offered method is given.

Введение

Задача распознавания видеоданных наиболее часто формулируется в двух следующих запросах [1]:

- найти видеофильмы, имеющие ключевой стоп-кадр, похожий на образец запроса;
- упорядочить видеоматериалы коллекции по их подобию заданному видеообразцу.

Эти запросы возникают в системах видеонаблюдения, анализа видеоданных, тематической классификации видеоинформации. Для корректного удовлетворения таких запросов видеоданные предварительно разбиваются на сегменты (отдельные фрагменты) видеоинформации.

Цель работы – модификация методов сегментации видеоданных путем использования методов статистической кластеризации для устранения избыточности результатов.

В соответствии с поставленной целью можно выделить основные задачи данной работы:

- постановка задачи анализа видеоданных;
- выделение недостатков существующих методов сегментирования видеоинформации;
- разработка метода сегментирования на основе иерархической технологии кластеризации;
- оценка качества сегментирования видеоданных предложенным методом.

Постановка задачи анализа видеоданных

На вход системе распознавания видеоданных подается видеофайл, который в результате сегментирования разбивается на последовательность видеофрагментов. Для каждого полученного видеофрагмента выделяется ключевой кадр (выделение значимого кадра видеофрагмента будем называть абстрагированием), для которого считаются различные характеристики: визуальные примитивы (характеристики цвета, формы, текстуры объектов изображения ключевого кадра), аудио примитивы, временные примитивы

и примитивы движения. В случае, если на вход системе подается отдельный видеокадр, то для него считаются только визуальные примитивы, которые затем сравниваются с визуальными примитивами ключевых кадров видеофрагментов анализируемой коллекции видеоматериалов.

В зависимости от формулировки запроса на выходе системы распознавания видео может быть изображение ключевого кадра или ранжированный набор видеоданных.

Общая схема поиска видеофильма с ключевым кадром, подобным кадру запроса, приведена на рис.1. Общая схема ранжирования видеофильмов по их подобию видеообразцу показана на рис.2.

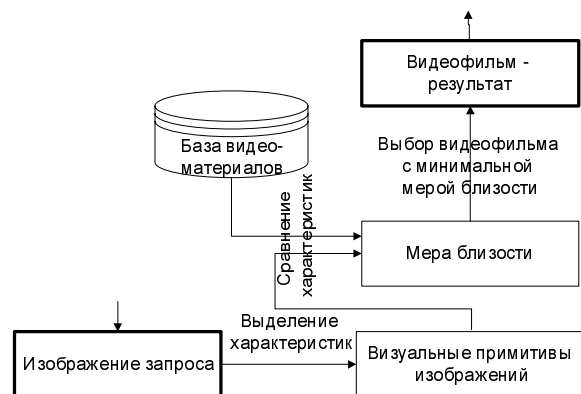


Рисунок 1 – Общая схема поиска видеофильма по кадру запроса

В случае решения задачи, общая схема которой приведена на рис.1, сегментирование производится только на стадии индексации видеофильмов базы данных. При решении задачи, основные этапы которой показаны на рис.2, сегментирование производится и для всех видеофильмов базы данных и видеофильма – запроса поиска.

При решении любой из рассмотренных задач анализа видеоданных особое внимание необходимо обратить на качество выделения

сегментов (процент пропущенных границ сегментов), так как от него напрямую зависит эффективность решения всей задачи.

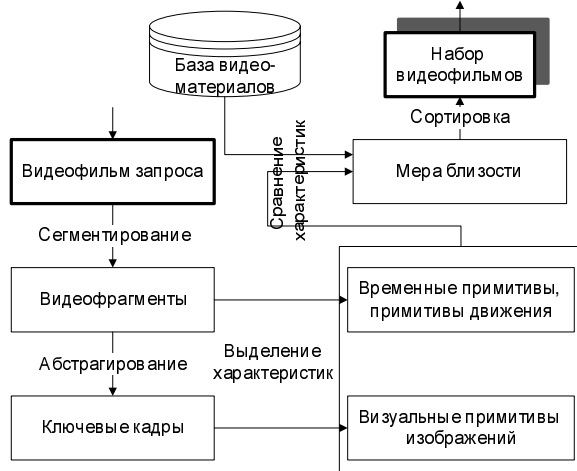


Рисунок 2 – Общая схема ранжирования видеофильмов

Методы сегментирования видеоданных

В связи с большим объемом видео-файлов для организации эффективного поиска данных с удовлетворительными показателями полноты и точности, а также для обеспечения быстрого предоставления пользователю релевантной информации имеет смысл индексировать каждый фильм не как единое целое, а как последовательность логически самостоятельных частей – видеофрагментов [2]. Задача сводится к определению границ видеофрагментов, они могут быть связаны с точками монтажа, изменением положения снимающей камеры и т.п. Формально задачу можно поставить так [3]: на вход подается упорядоченный набор кадров, необходимо выделить из них последовательность номеров, каждый из которых соответствует началу нового фрагмента.

Временное сегментирование может выполняться путем автоматического анализа изображения. Рассмотрим существующие методы сегментирования [4-7].

Попарное сравнение.

Для нахождения отличия I и $I+1$ кадров (изображений) видеопоследовательности производится попарное сравнение характеристик каждого пикселя кадра I и соответствующего кадра пикселя $I+1$. Мету близости двух кадров в общем виде можно записать как:

$$d(I, I+1) = \sum_{j=1}^N d(I_j, (I+1)_j),$$

где I_j – характеристика j -ого пикселя кадра I , N – общее число пикселей кадра. В качестве метрики d наиболее часто используют Евклидову метрику [4].

В случае, если

$$d(I, I+1) > \varepsilon, \quad (1)$$

то говорят, найдено начало нового видеофрагмента. Параметр ε подбирается исходя из априорной информации о видеофильме.

Сравнение регионов.

Сравниваемые кадры I и $I+1$ предварительно разбиваются на регионы, для которых производятся вычисления математического ожидания и дисперсии характеристик (цветовых или интенсивности). Далее производится вычисление расстояний между отдельными регионами на уровне полученных значений математического ожидания и дисперсии:

$$d(I, I+1) = \sum_{k=1}^M d(I_k, (I+1)_k),$$

где I_k – характеристика (математическое ожидание или дисперсия) k -ого региона кадра I , M – общее число выделенных для кадра регионов.

Считается, что найдено начало нового фрагмента в случае, если условие (1) удовлетворяется для математического ожидания и дисперсии анализируемых кадров.

Глобальные гистограммы.

Построение глобальных гистограмм производится путем разделения исходного пространства цветов кадра на некоторое дискретное количество. Высота гистограммы – количество пикселей кадра определенного цвета. В данном случае расстояние между кадрами вычисляется путем суммирования различий между гистограммами отдельных цветов.

Поиск начала нового фрагмента аналогичен методу попарного сравнения – в его основе удовлетворение условия (1).

Локальные гистограммы.

В основе данного подхода – комбинация методов сравнения регионов и метода глобальных гистограмм. Т.е. исходные сравниваемые кадры разбиваются на регионы, а для отдельных регионов вычисляются гистограммы. Сравнение кадров производится на уровне сравнения гистограмм соответствующих регионов.

Рассмотренные выше методы сегментирования видеoinформации показывают, что основной их недостаток – большое количество попарных сравнений кадров. Причем каждое сравнение кадров сводится к сравнению большого количества пикселей, регионов, гистограмм.

Также необходимо отметить, что при использовании рассмотренных методов невозможно учесть степень влияния отдельных характеристик на результат кластеризации: суммирование расстояний по всем характеристикам между отдельными парами объектов приводит к невозможности учета «большого» отличия между отдельными

характеристиками в силу суммирования с «маленькими» отличиями по другим характеристикам.

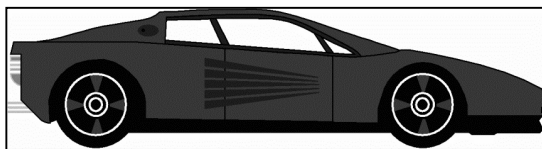
Кроме того, особое внимание при разработке метода сегментирования видеофрагментов необходимо уделить факту, что ложное обнаружение границы лишь увеличит требуемые ресурсы, а пропуск действительной границы сделает невозможным корректное вычисление оптического потока, который является одной из основных характеристик при поиске видеоинформации [3].

Использование кластеризации для сегментирования видеоданных

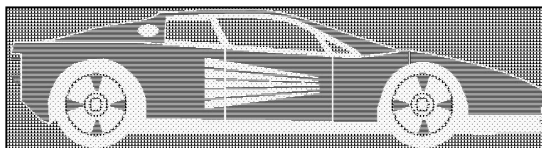
В основе предлагаемого метода сегментирования видеоинформации – выделение кластеров отдельных кадров.

Для кластеризации изображений кадров в данной работе используется статистический иерархический агломеративный метод кластеризации, который подробно описан в [8,9].

Основное преимущество кластеров перед регионами (согласно методу сравнения регионов, описанному выше) – отсутствие требования территориальной «неразрывности» пикселей. Т.е. кластер – это группа пикселей с подобными характеристиками, расположение пикселей не имеет значения. За счет этого количество кластеров будет значительно меньше количества регионов и, таким образом, количество сравниваемых объектов при анализе двух кадров также уменьшится. Пример кластеризации изображения согласно выбранного метода приведен на рис. 3.



Исходное изображение



Изображение с выделением трех кластеров

Рисунок 3 – Пример выделения кластеров кадра

Процедуру нахождения начала нового видеофрагмента в видеофильме на основе кластеризации кадров можно представить последовательность следующих шагов.

– Для кадров I и $I+1$ производится выделение отдельных кластеров на основе цветовых характеристик пикселей. Множество выделенных кластеров кадра I обозначим

$C_I = \{c_{I,1}, c_{I,2}, \dots, c_{I,Mi}\}$ (где Mi – количество кластеров кадра I), кадра $I+1$ – $C_{I+1} = \{c_{I+1,1}, c_{I+1,2}, \dots, c_{I+1,Mi1}\}$ (где $Mi1$ – количество кластеров кадра $I+1$).

– Нахождение соответствующих кластеров сравниваемых кадров. Кластер $c_{I,t}$ кадра I соответствует кластеру $c_{I+1,p}$ кадра $I+1$, если удовлетворится условие:

$$d(c_{I,t}, c_{I+1,p}) = \min_{j=1, Mi; k=1, Mi1} d(c_{I,j}, c_{I+1,k}).$$

При расчете расстояния d учитываются все необходимые характеристики кластеров: цветовые характеристики, характеристики формы, характеристики текстуры, характеристики местоположения. Формулы для расчета приведенных характеристик даны в [10].

– Вычисление нормализованного расстояния между рассматриваемыми кадрами. Нормализация производится путем введения коэффициентов «значимости» отдельных кластеров. Под коэффициентом значимости будем понимать долю занимаемой l -тым кластером площади изображения s_l . Т.е.

$$s_l = \frac{K_{C_{l,l}}}{P},$$

где $K_{C_{l,l}}$ – количество пикселей l -того кластера кадра I , P – общее количество пикселей кадра. Таким образом, выполняется условие: $\sum_{l=1, Mi} s_l = 1$.

Тогда формулу для вычисления нормализованного расстояния можно записать в виде:

$$d(I, I+1) = \sum_{t=1, Mi} s_t \cdot d(c_{I,t}, c_{I+1,p}).$$

Как отмечалось ранее, в качестве метрики d наиболее часто используется Евклидово расстояние, которое в нашем случае можно записать в виде:

$$d(c_{I,t}, c_{I+1,p}) = \sqrt{\sum_{j=1, J} (R_j c_{I,t} - R_j c_{I+1,p})^2},$$

где R_j – j -ая характеристика кластера, J – общее число анализируемых характеристик кластера. Также необходимо провести стандартизацию отдельных характеристик кластеров, так как различные характеристики имеют различные единицы измерения. Методы стандартизации характеристик объектов подробно описаны в [11].

– Проверка выполнения условия начала нового фрагмента видеоданных (1).

Оценка качества выделения сегментов видеоинформации

Для экспериментальной оценки предложенного метода и описанного ранее метода сравнения регионов, анализируемые видеофильмы были разбиты на две группы: мультипликационные фильмы и видеофильмы. В каждую группу были включены по 5 фильмов

различного качества (разрешение и скорость видеопотока) и содержания.

Границы видеосегментов определялись экспертом.

В анализируемых методах в качестве характеристик кластеров (регионов) учитывались характеристики цвета (в пространстве цветов RGB) и характеристики формы кластеров (первый и второй моменты инвариации).

Параметр ε был подобран экспериментально, исходя из условия минимизации количества пропущенных границ видеофрагментов: $\varepsilon = 0.1$.

Полученные результаты приведены на рис. 4.

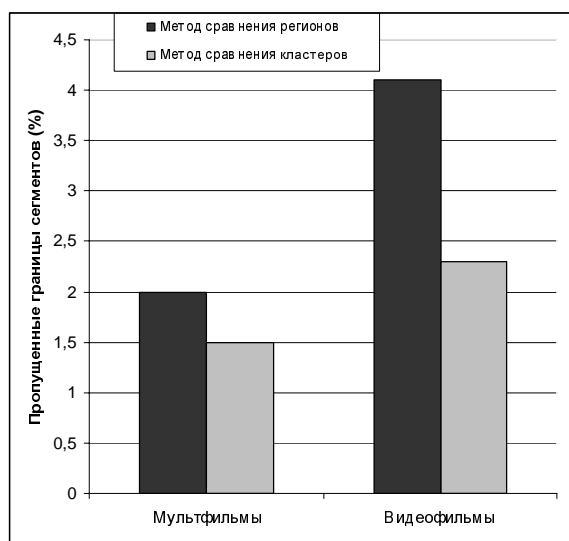


Рисунок 4 – Результаты сравнения метода сравнения кластера и метода сравнения регионов.

Приведенные результаты обоснованы учетом различных характеристик выделенных кластеров (анализируются характеристики цвета и формы), нормализацией расстояний между кадрами, «неразрывностью» пикселей кластеров.

Выводы

В работе предложен новый подход к сегментированию видеоданных, в основе которого кластеризация кадров видеoinформации. Кроме того, предлагается нормализовать кластеры для учета приоритетов кластеров большей площади и стандартизировать характеристики кластеров для приведения единиц измерения характеристик к одному масштабу.

Результаты экспериментов показали преимущество рассматриваемого метода: при сегментировании видеоданных предложенным методом в среднем на 1,25 % меньше пропущенных границ сегментов по сравнению с анализируемым методом сравнения регионов.

В дальнейшем планируется

усовершенствование предложенной модели путем применения аппарата мультимножеств для учета разнородных характеристик кластеров, а также расширение тестовых данных для получения более обоснованных результатов.

Литература

1. Н.С. Байгарова, Ю.А. Бухштаб. Проект «Кинолетопись России»: представление и поиск видеoinформации // I Всероссийская конференция “Электронные библиотеки”, Санкт-Петербург, 1999. – с.209-215.
2. Gupta A., Jain R. Visual Information Retrieval // Communications of the ACM. – 1997. – vol. 40, №5. – P. 70-79.
3. Байгарова Н.С., Бухштаб Ю.А., Евтеева Н.Н. Современная технология содержательного поиска в электронных коллекциях изображений. – Институт прикладной математики им. М. В. Келдыша РАН, <http://www.artinfo.ru/eva/EVA2000M/eva-papers/200008/Baigarova-R.htm>
4. Bradley L. Pyburn, B.S. Analysis of the Applicability of Video Segmentation to Unmanned Aerial Vehicle Surveillance Video. – USA, 1999. – 109 p.
5. T. Lin, H.J. Zhang. Media segmentation system and related methods. – USA: US_Patent4, 2004. – 18 p.
6. A. C. Loui, D. Gatica-Perez. Video structuring by probabilistic merging of video segments. – USA:US_Patent7, 2007. – 17 p.
7. M. Cooper, T. Liu, E. Rieffel. Video Segmentation via Temporal Pattern Classification // IEEE Transactions on Multimedia. – 2007. – vol. 9. – p. 610-618.
8. Башков Е.А., Вовк О.Л. Оценка эффективности нового статистического иерархического агломеративного алгоритма кластеризации для распознавания регионов изображений // Системні дослідження та інформаційні технології. – Інститут прикладного системного аналізу НАН України, Київ. – 2005. – №2. – С. 117-130.
9. Вовк О.Л. Новый поход к выделению визуально подобных цветов изображений // Проблемы управления и информатики. – Институт кибернетики В.М. Глушкова НАН Украины, институт космический исследований, Киев. – 2006. – №6. – с.
10. Wang J. Z., Li J. Wiederhold G. SIMPLicity: Semantics-Sensitive Integrated Matching for Picture Libraries // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2001. – vol. 23, №9. – P. 947-963.
11. Лепский А.Е., Бронеvич А.Г. Математические методы распознавания образов: Курс лекций. – Таганрог: ТТИ ЮФУ, 2009. – 152 с.

Поступила в редколлегию 10.03.2009