

Применение MPEG-7 для классификации и поиска визуальных данных

Авторы:

Кирилл Вихровский

Авторы:

Алексей Игнатенко

Для решения этих задач используют два основных метода:

1. Контентно-зависимый поиск (CBR). Он опирается на использование низкоуровневой информации и визуальных данных в качестве запроса. Большинство контентно-зависимых поисков основываются на сравнении таких характеристик как гистограмма цветов, цветовое размещение, текстура или форма.
2. Поиск на основе метаданных (MBR). Данный поиск использует более высокоуровневую информацию, опираясь на ключевые слова или свободную текстовую аннотацию. В данной статье описываются средства стандарта MPEG-7 [1], которые могут быть использованы для решения задачи контентно-зависимого поиска.

Стандарт MPEG-7

Стандарт является членом семейства стандартов MPEG, но в отличие от MPEG 1,2 и 4 описывает не формат кодирования, а описание контента мультимедийных данных. В качестве языка текстуального представления MPEG-7 использует расширенную схему XML и состоит из трех основных элементов: дескриптора, схемы описания и языка описания определений. Дескриптор является представлением характеристики элемента

метаданных. Схема описания формируется комбинацией дескрипторов и определяет структуру и семантику взаимодействия между дескрипторами. Язык описания определений позволяет создавать новые схемы описания и, возможно, дескрипторы и обеспечивает расширение и модификацию существующих схем описания [2],[3].

Визуальные дескрипторы для изображений и 3D - моделей

Дескрипторы MPEG-7 могут объединяться как в более общие визуальные дескрипторы, так и выступать в качестве специфического описателя конкретной характеристики объекта. Последние могут включать в себя цвет, текстуру, форму и характеристики движения (для видео последовательностей). В данной статье сконцентрируемся на наиболее общих дескрипторах, которые могут использоваться в большинстве приложений.

Визуальные дескрипторы цвета

Цвет – одна из наиболее широко используемых визуальных характеристик в контентно-зависимом поиске. Цветные характеристики сравнительно устойчивы к изменению цветов фона и независимы от размера изображения и ориентации объекта в сцене. Тем не менее, не существует универсального цветового дескриптора, показывающего высокие результаты для любых приложений. В результате чего, комитетом MPEG набор дескрипторов был стандартизован. При этом каждый дескриптор имеет приемлемую функциональность в своем классе задач. Вкратце опишем каждый дескриптор. В качестве более подробной информации, рекомендуется [4],[5].

Цветовое пространство(CS): Понятие цветового пространства используется в других описаниях, базирующихся на цвете, и предоставляет возможность взаимодействия между различными дескрипторами цвета. Поддерживаются следующие пространства RGB, YCrCb, HSV, HMMD (hue-min-max-diff), матрица линейного преобразования

с учетом RGB, монохромное. HSV – известное цветовое пространство, широко использующееся в приложениях, работающих с изображениями. HMMD – новое цветовое пространство, введенное MPEG, используется только в дескрипторе структуры цвета (CSD), описанном ниже.

Масштабируемый цвет (SCD): Один из базовых дескрипторов характеристик цвета, описывающий распределение цвета в изображении. SCD является гистограммой цветов в цветовом пространстве HSV, которая кодируется с помощью преобразования Хаара. Ее двоичное представление является масштабируемым с точки зрения числа каналов и числа бит, характеризующих значение точности характеристики данных. Точность сравнения (которое производится по норме L1) возрастает с увеличением числа бит, используемых для описания. Данный дескриптор может быть использован для эффективного поиска в базе однородных данных.

Доминантный цвет (DCD): Этот дескриптор цвета является наиболее удобным для представления локальных характеристик объекта или изображения в целях высокоскоростного поиска, где для предоставления цветовой информации достаточно малого числа цветов. Могут использоваться и полные изображения, например, изображения флагов или цветных торговых марок. Квантование цвета используется для получения малого числа характерных цветов в каждой области/изображении. Дескриптор состоит из дискретизируемых цветов, их процентной распределенности в области, пространственной когерентности и цветового изменения.

Цветовое распределение (CLD): Этот дескриптор был разработан для описания пространственного распределения цветов в областях произвольной формы. Он может быть использован в задаче сравнения наброска с изображением, которое не поддерживается другими цветовыми дескрипторами. Этот дескриптор может использоваться для всего изображения или для любой его части.



На рис. 1 представлены изображения и их гистограммы. Основываясь на цветовом распределении, два левых изображения являются наиболее схожими.

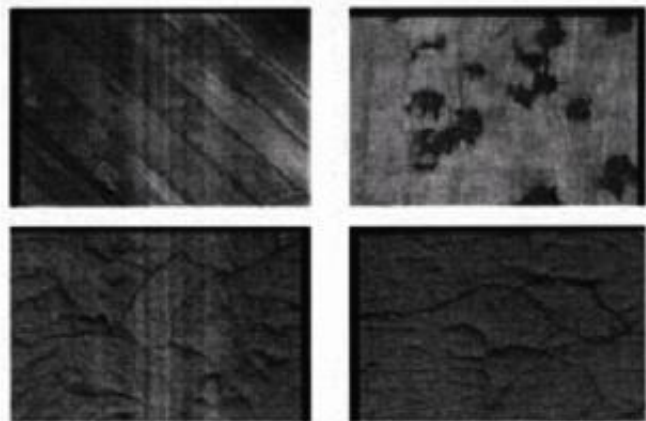
Структура цвета(CSD): Основное предназначение данного дескриптора – выделить локальные цветовые характеристики в изображениях. Для этой цели структурированным блоком 8×8 изображение сканируется методом скользящего окна. При каждом сдвиге окна считается количество конкретных цветов попавших в него. Таким образом, создаем цветовую гистограмму. Дескриптор структуры цвета обеспечивает дополнительную функциональность и улучшенный поиск, базирующийся на подобию естественных изображений.

Последние два дескриптора в качестве составного дескриптора дают неплохие результаты для задачи классификации изображений в базе данных (выделение всех сходных изображений в более чем 95% случаев). Для улучшенного поиска к данным дескрипторам можно добавить дескриптор краевой гистограммы, описанный ниже. В этом случае мы будем иметь нулевую вероятность получения ошибочного “сходного” изображения [8].

Визуальные дескрипторы текстуры

Текстура относится к визуальным характеристикам, которые имеют свойство однородности или неоднородности, в зависимости от наличия многочисленных цветов или интенсивности в изображении. Она содержит важную структурную информацию о поверхностях и окружающей их среде. Описывая текстуры изображений и 3D-объектов с помощью

соответствующих дескрипторов текстур, мы получаем мощное средство для сравнения и поиска. Для поиска основанного на сходстве дескрипторы текстуры могут быть использованы как независимо, так и в комбинации с другими MPEG-7 дескрипторами.



На рис.2 нижние изображения похожи с точки зрения дескрипторов текстуры MPEG-7, верхние значительно различаются

Однородная текстура(HTD): Однородная текстура представляет собой важный визуальный примитив для поиска и просмотра в большой коллекции сходно выглядящих образов. Изображение может рассматриваться как мозаика однородных текстур. Таким образом, текстурные характеристики, соответствующие определенным областям, могут быть использованы для индексации визуальных данных. Вычисление этого дескриптора базируется на фильтрации посредством набора фильтров Габора, настроенных на определенные ориентации и масштаб.

Неоднородная текстура - Краевая гистограмма(EHD): Для описания изображений с неоднородной текстурой, в MPEG-7 введен дескриптор краевой гистограммы. Этот дескриптор описывает пространственное распределение краев, отчасти как дескриптор цветового распределения(CLD). Вычисление этого описателя основывается на делении изображения на 16 неперекрывающихся блоков равного размера. Затем для каждого блока вычисляется информация о пяти краях: вертикальном, горизонтальном, 45°, 135° и неориентированном. Так как

края играют важную роль для восприятия изображения, данный дескриптор может использоваться для поиска изображений по образцам или наброскам, в особенности для изображений с нерегулярными краями. Наилучшие характеристики системы поиска изображения могут быть достигнуты путем использования полу-глобальных и глобальных гистограмм, получаемых непосредственно из дескриптора. К тому же свойства системы поиска изображения могут быть существенно улучшены при комбинировании данного дескриптора с другими (например, с дескриптором гистограммы цвета).

Визуальные дескрипторы формы

Форма объектов изображения для большинства приложений предоставляет мощное средство для поиска и сравнения в базе изображений. Типичные примеры таких приложений используют бинарные изображения символов, объектных контуров и 2D/3D границ. В системах поиска и классификации 3D объектов и изображений часто критичным является требование инвариантности дескриптора формы относительно масштабирования, поворота и переноса описываемого объекта. Дескриптор может 2D или 3D природы в зависимости от конкретного приложения. В общем случае, 2D описатели формы могут быть разделены на две группы: базирующиеся на контуре и базирующиеся на областях. Первая группа дескрипторов использует информацию только о границах объектов для описания сходных контурных характеристик. Дескрипторы, базирующиеся на областях, используют информацию о форме всей области в целом для извлечения полезных характеристик в случае анализа схожих пространственных распределений пикселей объектов. Применимость дескрипторов, базирующихся на областях или на контурах, зависит от конкретного приложения и характеристик объекта. Опишем визуальные дескрипторы MPEG-7[6].

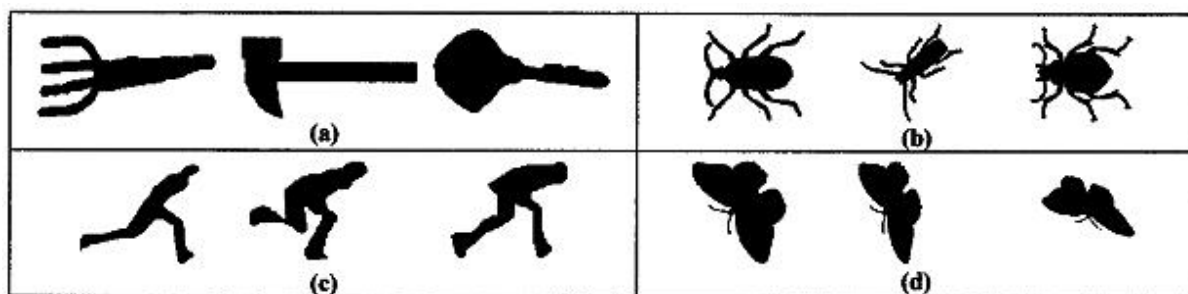
Форма, базирующаяся на областях(ART - Angular Radial Transformation): Данный дескриптор используется для объектов, которые могут быть наилучшим образом описаны, основываясь на выделении областей, нежели их контуров. И принадлежит классу методов инвариантов момента. Идея инвариантов момента заключается в использовании

моментов, базирующихся на областях, которые являются инвариантными относительно преобразований как свойство формы. Данный дескриптор является очень компактным (140битов/область), а также очень устойчивым к сегментационному шуму.



На рис.3 изображения являются схожими в своей группе относительно MPEG-7 ART.

Форма, базирующаяся на контуре(CBSD): Для описания объектов, форма которых наилучшим образом может быть определена информацией о контуре, можно использовать MPEG-7 дескриптор формы, базирующийся на контуре. Данный дескриптор описывает представление контура, основанное на кривизне масштабируемого пространства (CCS), включая эксцентриситет и кривизну оригинальных и фильтруемых контуров. Средний размер дескриптора – 122 бита/контур.

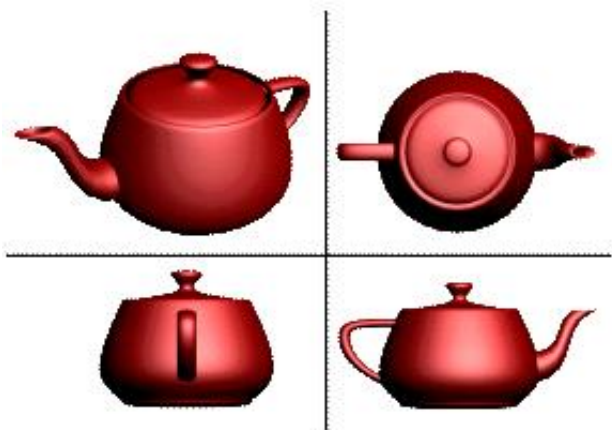


На рис.4(b)-(d) изображены сходные объекты, при сравнении дескриптором MPEG-7 CBSD. Изображения (a) являются сходными относительно областей, но различными с точки зрения сравнения контуров.

Дескриптор 3D-формы: Дескриптор полезен для сравнения физических и

виртуальных 3D объектов и основывается на понятии спектра формы. Грубо говоря, спектр формы определяется как гистограмма индекса формы, вычисленной по всей 3D поверхности объекта. Сам индекс представляет собой меру локальной выпуклости каждой локальной 3D области.

Дескриптор 2D/3D формы: В общем случае, форма 3D объекта может быть описана конечным числом 2D форм, которые представляют собой вид объекта с разных углов обзора. Таким образом, 2D виды 3D объекта могут быть описаны с помощью дескрипторов 2D форм. Сравнение 3D объектов можно производить путем сопоставления пар соответствующих 2D проекций рассматриваемых объектов. Вообще говоря, наибольшая производительность для 3D форм достигается при использовании дескриптора формы, базирующейся на контуре. Дескриптор также известен как дескриптор световых полей. Предоставляет достаточно устойчивый поиск в базе объектов с различными разрешениями [10].



Заключение

В данной статье были описаны основные визуальные дескрипторы в стандарте MPEG-7, которые могут быть эффективно использованы для контентно-зависимого поиска и фильтрации: поиска в базе данных по изображению, поиска по 3D/2D проекциям трехмерного объекта, поиска изображения и 3D объекта по наброску, автоматической классификации данных в базе. Дескрипторы извлекаются с помощью соответствующих

методов и могут быть сохранены, а также передаваться отдельно от мультимедийного контента, который они описывают. Дескрипторы предоставляют широкие возможности для сравнения изображений и 3D моделей, основываясь на характеристиках цвета, текстуры и формы объекта. Использование дескрипторов MPEG-7 в приложении гарантирует совместимость с другими приложениями, базирующихся на их использовании, одновременно предоставляя гибкие возможности для разрабатываемого приложения. В стандарте определяются только алгоритмы извлечения характеристик из объектов, но никак не фиксируются меры, в терминах которых должно проводиться сравнение дескрипторов.

Помимо контентно-зависимого поиска MPEG-7 предлагает достаточно большие возможности для поиска на основе метаданных, которые не рассматривались в данной статье. Для подробного изучения читателю предлагается [1],[9].

Литература

1. MPEG-7 Overview. <http://www.mpeg-7.com> ^[1] (Industry Focus Group)
2. B. S. Manjunath, P. Salembier, and T. Sikora. Introduction to MPEG-7: Multimedia Content Description Standard. Wiley, April 2002
3. S.-F. Chang, T. Sikora, and A. Puri. Overview of MPEG-7 Standard. EURASIP Journal on Applied Signal Processing, vol. 2002, issue 4 (April 2002), pp.343 - 353
4. A.Yamada, M. Pickering, S.Jeannin, L. Cieplinski, J.-R. Ohm, and M. Editors, Eds. MPEG-7 Visual Part of Experimentation Model Version 8.0. ISO/IEC JTC1/SC29/WG11 #N3673, La Baule, October 2000
5. B. S. Manjunath, J.-R. Ohm, V.V.Vasudevan, and A.Yamada. MPEG-7 color and texture descriptors. IEEE Tr. CSVT, p. 703--715, June 2001
6. M. Bober. MPEG-7 Visual Shape Descriptors. IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, Issue 6, pp.716-719, 2001
7. Thomas Sikora. The MPEG-7 Visual Standard for Content Description. IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, issue 6, pp.696-702, 2001
8. Horst Eidenberger. A Video Browsing Application Based on Visual MPEG-7

Descriptors and Self-Organising Maps. International Journal of Fuzzy Systems, Vol. 6, No. 3, pp. 124-137, 2004

9. Pei-Jeng Kuo, Terumasa Aoki, Hiroshi Yasuda. Building Personal Digital Photograph Libraries: An Approach with Ontology-Based MPEG-7 Dozen Dimensional Digital Content Architecture. Proceedings Computer Graphics International, pp.482- 489, 2004
10. Alberto del Bimbo and Pietro Pala. Content-Based Retrieval of 3D Models. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP), Volume 2 , Issue 1 (February 2006), pp.20 - 43

Дополнительная информация

Ссылка:

Кирилл Вихровский, Алексей Игнатенко. Применение MPEG-7 для классификации и поиска визуальных данных. *Компьютерная графика и мультимедиа*. Выпуск №5(1)/2007.

<http://cgm.computergraphics.ru/content/view/161>

Выпуск:

Выпуск №5(1)/2007 [2]

1. <http://www.mpeg-7.com/>
2. <http://cgm.computergraphics.ru/issues/issue15>