Использование визуальных особенностей восприятия компонент цветовой модели HSI при поиске изображений по содержанию

©Волосных Д.Ф.

МГТУ им. Н. Э. Баумана dmitry.volosnykh@gmail.com

Аннотация

В данной работе представлены результаты участия системы PIRS в дорожке поиска изображений по визуальному подобию в рамках семинара РОМИП'2010. Представлены базовые соображения, основанные на визуальных свойствах компонент цветовой модели HSI. Предлагается способ устранения проблемы с точкой разрыва компоненты цветового тона.

1 Введение

Развитие аппаратной базы вычислительных машин уже сегодня позволяет хранить большие коллекции изображений разного характера. Поиск необходимых изображений через просмотр пусть даже их пиктограмм стал утомительным занятием. Необходимы какие-то средства автоматического извлечения искомых изображений. Одним из таких средств является поиск изображений

по содержанию (по визуальному подобию), когда пользователем предоставляется изображение-образец.

2 Описание системы

2.1 Выбор цветовой модели

HSI — цилиндрическое представление цветовой модели RGB, состоящее из трёх компонент: цветовой тон (hue), насыщенность (saturation), интенсивность (intensity). Выбор цветовой модели HSI обусловлен тем, что её компоненты ближе к понятиям, которыми фигурирует человек при описании характеристик того или иного цвета. К тому же, каждая из компонент независима от остальных двух [1]. По этим причинам цветовая модель HSI часто используется в алгоритмах компьютерного зрения.

2.2 Наблюдаемые свойства цветовой модели HSI

При работе с тестовой коллекцией из фотографий без единой темы и разного качества (ROMIP_FLICKR'08) были замечены следующие особенности.

Для всего изображения компонента цветового тона даёт хоть и грубое — представление о наличии и положении областей, соответствующих некоторой поверхности, например, имеющей один цвет, но неравномерное освещение из-за собственной конструкции. Или же поверхность может иметь точки близких (хотя на самом деле разных) цветов вследствие фактурного шума (артефакты кожи человека, царапины и другие мелкие дефекты на чём-либо и т.п.) или шума, вызванного технологическими недостатками датчика, с помощью которого формировалось изображение. Т.о. различия цветовых тонов для разных точек таких поверхностей существенно меньше различий в их цвете. Компонента интенсивности, в свою очередь, сохраняет детали изображения. Её иногда используют при применении методов обработки полутоновых изображений. И, наконец, компонента насыщенности выделяет контуры и границы, а также, интуитивно, придаёт изображению ощущение объёма.

В качестве хорошего примера перечисленных явлений можно привести изображение 00000010/255875412.jpg (см. Рис. 1).

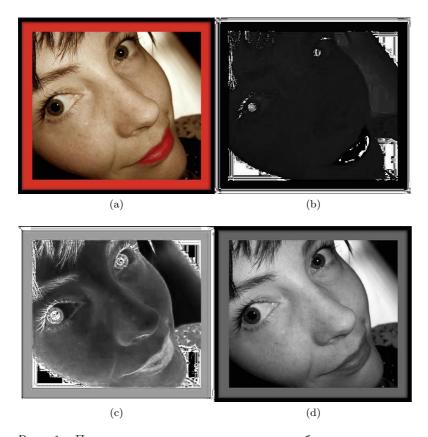


Рис. 1: Пример удачного представления изображения в цветовой модели HSI. (а) Исходное цветное RGB-изображение 00000010/255875412.jpg. HSI-компоненты изображения: (b) цветовой тон, (c) насыщенность, (d) интенсивность.

Сто́ит обратить внимание на область губ дамы весьма недурной наружности на этом изображении: на Рис. 1b хорошо заметны участки с существенно различными значениями компоненты цветового тона, хотя на исходном изображении (Рис. 1a) очевидна однородность цветового тона губ. В разделе 2.3.2 рассказывается, как можно бороться с этой проблемой.

Понятно, что при работе с полутоновыми изображениями существенна лишь компонента интенсивности.

2.3 Индексируемые признаки изображений

Наблюдения, изложенные в разделе 2.2, приводят к следующим соображениям.

2.3.1 Признак на основе компоненты цветового тона

Для компоненты цветового тона целесообразно попробовать использовать признак, аналогичный предлагаемому в статье [2] признаку MPEG-7 CLD (Color Layout Descriptor). В рамках данного эксперимента был разработан и применён признак HLD (Hue Layout Descriptor), основная идея которого полностью заимствована у MPEG-7 CLD (за исключением применения дискретного преобразования косинусов и нелинейного квантования его коэффициентов):

- 1. исходное изображение делится регулярной сеткой на 64 блока (8 в длину и 8 в ширину);
- 2. для каждого блока находится среднее арифметическое цветового тона.

Получившиеся 64 числа и составляют признак HLD. Для определения расстояния между двумя такими признаками использовалась формула, так же подсмотренная в [2], но без извлечения квадратного корня:

$$\operatorname{distHLD}(h', h'') = \sum_{i} w_{i} \cdot \operatorname{distHue}^{2}(h'_{i}, h''_{i}), \tag{1}$$

где h',h'' — признаки HLD, w_i - весовые коэффициенты каждого из блоков.



Рис. 2: Пример проблемы компоненты цветового тона модели HSI. (a) Исходное цветное RGB-изображение 00000010/255886715.jpg. (b) компонента цветового тона того же изображения в цветовой модели HSI (обратите внимание на левый верхний угол).

2.3.2 Построение функции distHue

Необходимость вводить функцию расстояния distHue вместо модуля разницы обусловлена тем, что значение цветового тона принадлежит отрезку $[0,H_{max}]$ (например, $H_{max}\in\{1,2\pi,360^\circ\}$). Причём, отсчёт начинается от красного тона. Соответственно, близкие тона красного могут иметь существенно различные числовые значения. На данный эффект уже обращалось внимание в разделе 2.2 (см. Рис. 1). Особенно хорошо эту проблему видно на Рис. 2.

Поскольку цветовая модель HSI цилиндрическая, то и 0, и H_{max} отвечают красному тону, т.е. отрезок $[0,H_{max}]$ можно свернуть в окружность, соединив его концы. Для определённости, на Рис. 3 считаем, что цветовые тона h_1 и h_2 таковы: $h_1 \leq h_2$.

На Рис. За область значений цветого тона представлена в виде отрезка. В этом случае функция расстояния между цветовыми тонами имеет вид:

$$\Delta h_{12} = \Delta (h_1, h_2) = |h_1 - h_2|, \tag{2}$$

что соответствует отрезку, изображённому пунктирной линией. При таком подходе близкие тона красного могуть иметь сильно различающиеся числовые значения.

На Рис. 3b та же область представлена в виде окружности. В таком случае за расстояние между двумя точками естественнее

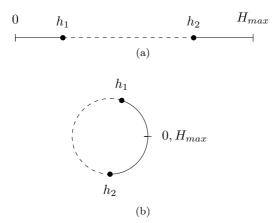


Рис. 3: Построение функции distHue. (a) Представление области значений цветого тона в виде отрезка. (b) Представление той же области в виде окружности.

брать длину кратчайшей из двух дуг, на которые они разбивают окружность. На рисунке это, очевидно, сплошная дуга. Тогда:

distHue
$$(h_1, h_2) = \min \{ \Delta h_{12}, H_{max} - \Delta h_{12} \}.$$
 (3)

К сожалению, с точки зрения наглядности (при просмотре полутонового изображения, соответствующего компоненте цветового тона) устранение этого недостатка не представляется возможным. Однако, это не мешает нам использовать выработанную функцию 3 для более корректной оценки близости цветовых тонов.

2.3.3 Признак на основе компоненты интенсвности

Для компоненты интенсивности признак IHM (Intensity Histogram Moments) формировался существенно проще: строилась гистограмма интенсивностей, для которой вычислялись первые три момента. Для оценки расстояния в таком пространстве признаков использовалась простая формула:

distMoments
$$(\mu', \mu'') = \sum_{i=1}^{3} v_i \cdot |\mu'_i - \mu''_i|,$$
 (4)

где μ', μ'' — признаки ІНМ, v_i - весовые коэффициенты каждого из моментов.

2.3.4 Слияние данных

После того, как расстояния между всеми признаками рассматриваемой пары изображений I', I'' вычислены, они используются наиболее распространённым образом:

$$\operatorname{dist}(I', I'') = u_1 \cdot \operatorname{distHue}(h', h'') + + u_2 \cdot \operatorname{distMoments}(\mu', \mu''),$$
(5)

что является слиянием с весовыми коэффициентами u_i для каждого пространства признаков.

Итак, получена конечная функция для определения расстояния между парой изображений.

3 Результаты эксперимента

По результатам эксперимента (см. Рис. 4, 5) видно, что лидирующие позиции занимают участники хххх-1 и хххх-6 в то время как система PIRS показывает, как правило, средние результаты. Это говорит об адекватности выбранного подхода к задаче поиска изображений по содержанию, хотя, очевидно, предстоит большая работа по улучшению качества поиска.

4 Дальнейшие планы

Пользуясь наблюдениями раздела 2.2 и выводами в статье [3], в будующем планируется попробовать применить к компоненте насыщенности, как к полутоновому изображению, один из алгоритмов, учитывающих профиль яркости уровня изображения.

5 Заключение

Как и в труде [4] стоит обратить внимание на то, что подавляющее большинство систем-участников демонстрируют низкие абсолютные показатели. Поэтому проблема поиска изображений по содержанию остаётся открытой.

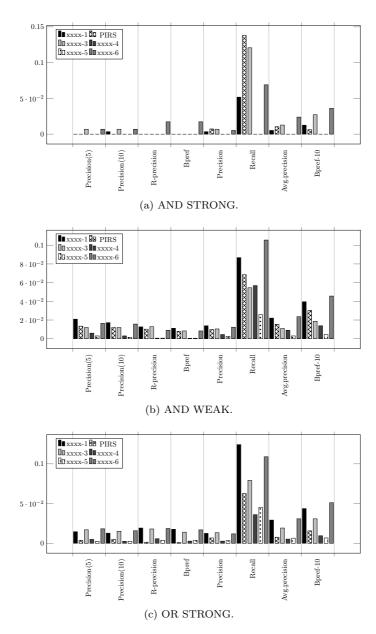


Рис. 4: Результаты систем-участников (начало).

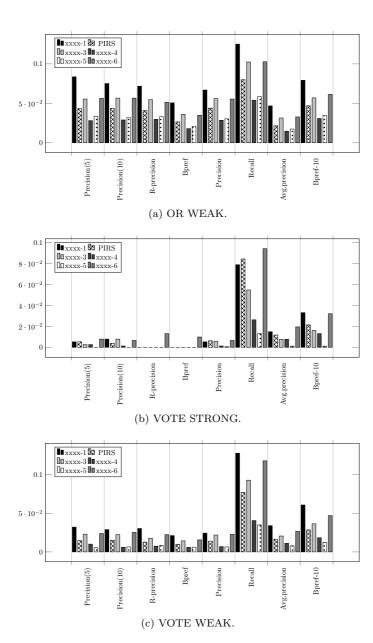


Рис. 5: Результаты систем-участников (продолжение).

Список литературы

- [1] Р. Гонсалес, Р. Вудс, *Цифровая обработка изображений*, Москва: Техносфера, 2000.
- [2] E. Kasutani, A. Yamada, The MPEG-7 Color Layout Descriptor: A Compact Image Feature Description For High-Speed Image-Video Segment Retrieval, IEEE conference, pp.674-677, 2001.
- [3] А. Мельниченко, А. Гончаров, ЛММИИ на РОМИП-2009: Методы поиска изображений по визуальному подобию и детекции нечетких дубликатов изображений, РОМИП'2009, стр.108–121, 2009.
- [4] Д. Н. Крешихин, Поиск изображений по визуальному подобию на основе цветовой модели HSV, POMИП'2009, стр.83— 88, 2009.

Exploiting visual perception properties of HSI color model components in content-based image retrieval.

©D. F. Volosnykh

Abstract

This paper presents results of participation in content-based image retrieval track at ROMIP'2010. Basic conclusions based on visual perception properties of HSI color model components are discussed. Solution of jump discontinuity issue with hue component is offered.