

# Поиск изображений по визуальному подобию на основе цветовой модели HSV

© Крешихин Д.Н.

Рязанский Государственный  
Радиотехнический Университет  
pelou@yandex.ru

## Аннотация

В статье описана поисковая система hsvise, участвовавшая в дорожке поиска изображений по визуальному подобию семинара РОМИП'2009. Рассмотрены принципы и подходы к построению системы, приведено краткое математическое описание. Разобраны и проанализированы результаты оценки системы.

## 1. Введение

Целью исследований было создание поисковой системы для базы изображений объёмом порядка  $10^4 - 10^5$ . При этом ключевым ограничением являлась вычислительная сложность алгоритмов, которая бы позволяла использовать такую систему на не самых современных персональных компьютерах. К сожалению, успехи в построении систем поиска изображений по визуальному подобию значительно уступают некоторым другим областям информационного поиска[1], поэтому при такой постановке задачи требования к качеству поиска были довольно мягкими.

## 2. Описание системы

В качестве цветовой модели изображения была выбрана схема HSV, которая представляет каждый пиксель тремя компонентами: тон(hue), насыщенность(saturation) и значение(value). Это позволяет

описать каждое изображение тремя вещественными функциями от координат  $h(\bar{x})$ ,  $s(\bar{x})$  и  $v(\bar{x})$ , где  $\bar{x}$  - двумерный вектор. При этом каждая из компонент до некоторой степени независима, т.к. схожесть двух изображений по тону необязательно отвечает их полной схожести.

Предпочтение было отдано в пользу HSV вместо других моделей, таких как RGB и CMYK, по двум причинам. Во-первых, потому что устройство HSV ближе к человеческому восприятию, поскольку RGB и CMYK определяют цвет как комбинацию базовых цветов [2], в то время как компоненты HSV отображают информацию о цвете в более привычной для человека форме: какой именно цвет, насыщенный он или тусклый, светлый или тёмный. Во-вторых, HSV позволяет использовать при определении схожести изображений меры, допускающие быстрые алгоритмы обработки.

Итак, для сравнения двух изображений можно использовать степень схожести их компонент. Пусть два изображения имеют компоненты  $v(\bar{x})$  и  $v'(\bar{x})$ , тогда высшей степенью схожести будет следующая ситуация:

$$v(\bar{x}) - v'(\bar{x}) = 0 \text{ для } \bar{x} \in D, \quad (1)$$

где  $D$  - область изображения. В общем случае требовать полное сходство бессмысленно, поэтому достаточно ограничиться малым значением некоторой меры различия. Можно предложить большое количество мер различия компонент, например:

$$d(v, v') = \max_{\bar{x} \in D} |v(\bar{x}) - v'(\bar{x})|, \quad (2)$$

$$d(v, v') = \int_D (v(\bar{x}) - v'(\bar{x}))^2 d\bar{x}, \quad (3)$$

$$d(v, v') = \int_D (v(\bar{x}) - v'(\bar{x}))^4 d\bar{x} \text{ и т.п.} \quad (4)$$

Допустим, что изображение  $v'(\bar{x})$  является немного преобразованной копией изображения  $v(\bar{x})$ . Пусть преобразование можно описать некоторым оператором  $A\{\}$ , тогда  $v'_{\bar{\alpha}} = A\{\bar{\alpha}, v\}$ , где  $\bar{\alpha}$  - параметры преобразования, тогда задача сравнения компонент сводится к исследованию следующей зависимости:

$$d(\bar{\alpha}) = d(v, v'_{\bar{\alpha}}) = d(v, A\{\bar{\alpha}, v\}). \quad (5)$$

Мера (3) обладает определённым преимуществом, т.к. может быть сведена к взаимной корреляции:

$$d(\bar{\alpha}) = \int_D (v(\bar{x}) - v'_{\bar{\alpha}}(\bar{x}))^2 d\bar{x} = \int_D v(\bar{x})^2 - 2v(\bar{x})v'_{\bar{\alpha}}(\bar{x}) + v'_{\bar{\alpha}}(\bar{x})^2 d\bar{x} =$$

$$= \int_D v(\bar{x})^2 d\bar{x} - \int_D 2v(\bar{x})v'_{\bar{\alpha}}(\bar{x}) + \int_D v'(\bar{x})^2 = \int_D v(\bar{x})^2 d\bar{x} + \int_D v'_{\bar{\alpha}}(\bar{x})^2 dx - 2R(\bar{\alpha}) \quad (6)$$

Т.к.  $\int_D v(\bar{x})^2 d\bar{x} + \int_D v'_{\bar{\alpha}}(\bar{x})^2 dx > 0$ , то минимальному различию, исходя из (6), будет соответствовать максимальное значение  $R(\bar{\alpha})$ . Для таких преобразований, как сдвиг и вращение, корреляция  $R(\bar{\alpha})$  может быть вычислена с помощью быстрого преобразования Фурье, что благоприятно сказывается на быстродействии системы поиска.

Ниже на рисунке 1 приведена функциональная схема hsvise.

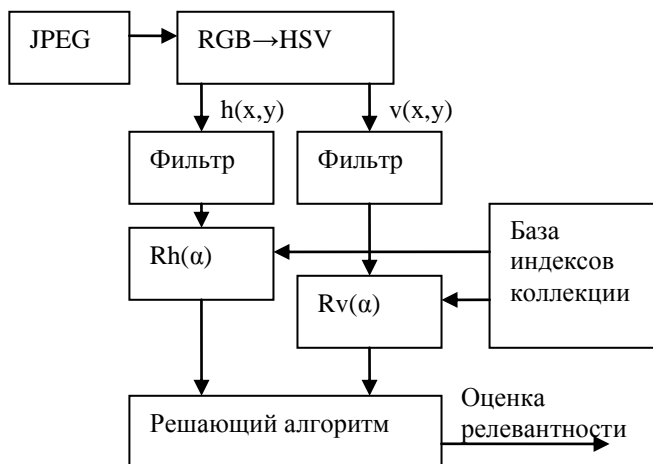


Рисунок 1 - функциональная схема hsvise

Алгоритм работы системы следующий: из входного jpeg-файла формируется массив в формате RGB, затем преобразуется в формат HSV. Для поиска использовались только две компоненты -  $h$  и  $v$  (т.к. процесс сравнения насыщенности двух изображений трудно свести к столь однозначной трактовке, в отличие от тона и яркости). Затем выделенные компоненты проходят фильтрацию и прореживания для сокращения вычислительных затрат и устранения незначительных деталей изображений.

Обработанные компоненты сравниваются с компонентами из базы индексов изображений коллекции. Далее на основе результатов сравнения решающий алгоритм вычисляет оценку релевантности. Наличие существенных выбросов функции  $R(\bar{\alpha})$  свидетельствует о

схожести композиций изображений и трактуется как признак локального сходства. Множественные выбросы  $R(\bar{\alpha})$  средней величины указывают на различие композиций, но наличие значительного количества схожих элементов, что трактуется как признак глобального сходства.

### 3. Результаты

Результаты участников дорожки сведены в диаграммы, представленные на рисунках 2,3,4,5. На всех диаграммах заметно, что в первую очередь выделяются результаты hsvise (на всех рисунках - третья колонка) и участника №4. При слиянии по правилу AND система hsvise немного опережает по полноте ответов участника №4, но иногда уступая по точности.

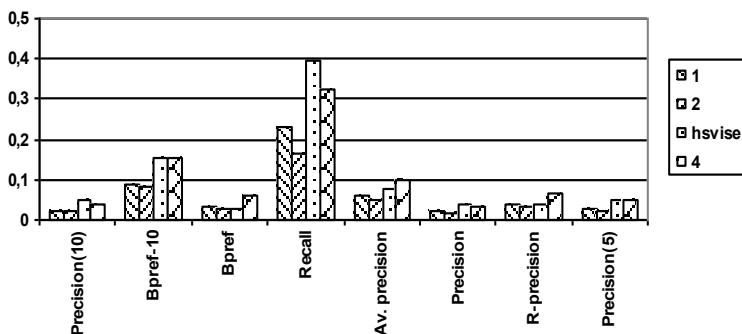


Рисунок 2 - результаты участников дорожки поиска по визуальному подобию, оценка AND-strong

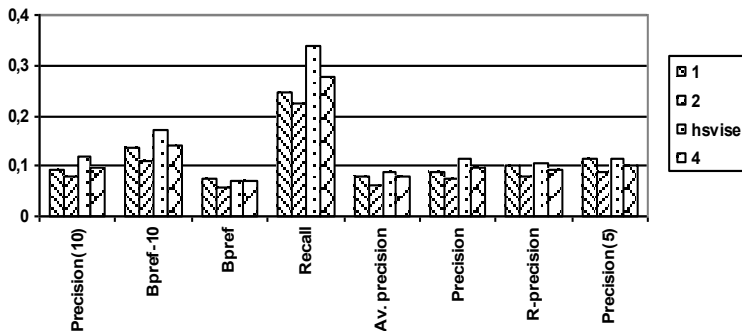


Рисунок 3 - результаты участников дорожки поиска по визуальному подобию, оценка AND-weak

По оценке AND-strong BPREF система hsvise получила самый низкий результат, возможно, что причиной является малое количество известных релевантных изображений при расчёте метрики (что указывается в [3]), т.к. во всех остальных случаях ни на BPREF, ни на BPREF-10 она уже не уступает другим участникам.

При слиянии результатов по правилу OR различия между участниками становятся не такими существенными, но при этом hsvise и участник №4 показывают немного лучшие результаты.

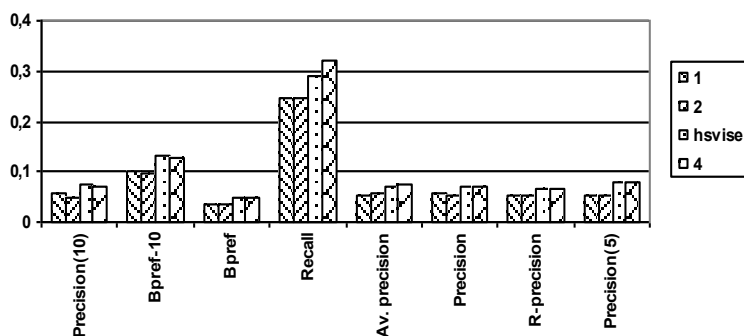


Рисунок 4 - результаты участников дорожки поиска по визуальному подобию, оценка OR-strong

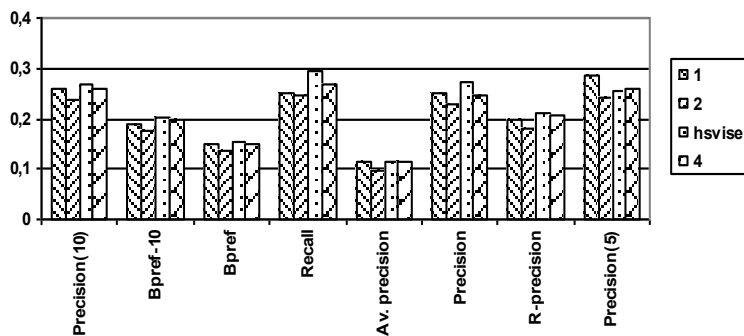


Рисунок 5 - результаты участников дорожки поиска по визуальному подобию, оценка OR-weak

#### 4. Заключение

Результаты участия hsvise на дорожке поиска по визуальному подобию РОМИП'2009 показали, что описанный в статье подход

вполне адекватно решает поставленную задачу, не уступая остальным участникам. Но, к сожалению, большинство метрик имеют малые абсолютные значения, что указывает на сложность решаемой задачи поиска изображений и на большой потенциал в исследовании этой проблемы.

В заключении автору хочется выразить огромную признательность организаторам РОМИМ'2009 и в особенности тем людям, с кем доводилось общаться и контактировать за время участия.

## **Литература**

- [1] Datta, Ritendra; Dhiraj Joshi, Jia Li, James Z. Wang . "Image Retrieval: Ideas, Influences, and Trends of the New Age". ACM Computing Surveys 40: pp 1-60, 2008.
- [2] Gonzalez, Rafael C.; Richard Eugene Woods. Digital Image Processing. Prentice Hall, 2007.
- [3] Приложение А. Официальные метрики РОМИП'2007-2008 // Труды РОМИП 2007-2008. Санкт-Петербург: НУ ЦСИ, 2008.

## **Content-based Image Retrieval Founded on HSV Color Model**

Kreshikhin D.N.

The paper describes retrieval principles that were used by hsvise-system for participation in content-based image retrieval track at ROMIP'2009.