

## СРАВНИТЕЛЬНЫЙ АНАЛИЗ СИСТЕМ КОНТЕКСТНОГО ПОИСКА ИЗОБРАЖЕНИЙ

Шозда Н.С.

Кафедра ПМИИ, ДонГТУ  
[shozda@r5.dgtu.donetsk.ua](mailto:shozda@r5.dgtu.donetsk.ua)

### Abstract

*Natalia S. Shozda. Comparison of content-based image retrieval system. With the growth of distributed image databases the ability to search for images based on their content is becoming increasingly important. The use of comprehensive textual annotation provide one method for image retrieval. The searching of images by their visual content complement the text-based approaches. Visual features of the images describe their content better, than textual annotation. This page allows to search for images based on the content from databases located in IBM, Virage and Columbia University servers.*

### Введение

Необходимость в разработке систем контекстного поиска изображений обусловлена большим объемом и неструктурированностью баз данных (БД) изображений, в связи с чем выполнять индексирование по самому изображению достаточно сложно. Традиционным подходом к работе с такими БД является снабжение изображения текстовой аннотацией, по которой и выполняется индексирование и поиск. Однако текстовое описание изображений, особенно включающих большое количество объектов, субъективно и зависит от человека, составляющего это описание. Более новым подходом к решению указанной проблемы является контекстный поиск, который выполняется по содержимому изображения в соответствии с запросом пользователя, где указываются визуальные характеристики искомого изображения, либо изображение-образец поиска (в этом случае должны быть найдены изображения, «сходные» с образцом). После выполнения запроса пользователю предлагаются несколько изображений, в наибольшей степени удовлетворяющих запросу. При этом содержимое изображения представляется цветом, текстурой, взаимным расположением закрашенных областей и другими характеристиками (иногда их называют метаданными), которые определяются в момент занесения изображения в БД и хранятся вместе с ним. Последующий поиск выполняется именно по этим характеристикам; доступ к самому изображению осуществляется только при его отображении.

Среди систем контекстного поиска изображений, разработанных в настоящее время, можно назвать Virage, QBIC, VisualSeek. Данная работа посвящена анализу этих систем с точки зрения их архитектуры, методов индексирования и поиска, а также подходов к определению подобия изображений.

### 1. Архитектура систем

Архитектуры систем контекстного поиска изображений во многом сходны (рис.1). Как правило, в этих системах при занесении изображения в БД и при вводе образца поиска выполняется обработка изображения (определение его характеристик) и индексирование. При последующем поиске анализируется созданная индексная

структура данных. В работе системы контекстного поиска изображений можно выделить следующие шаги:

- заполнение БД, когда выполняются определение характеристик изображения, индексирование по этим характеристикам, запоминание метаданных и изображений;
- обработка запроса, заключающаяся в определении характеристик изображения-образца поиска и собственно поиске результатов.

Очевидно, что система, функционирующая по этой схеме, должна включать модули заполнения БД и поиска в ней данных, а также специальные методы индексирования. При этом сами данные могут быть представлены в виде традиционной объектно- связной модели.

В работе системы QVIS можно выделить следующие основные функции:

- Заполнение БД изображениями. При необходимости выделяются фрагменты изображения и изображение снабжается текстовым описанием и ключевыми словами.
- Вычисление характеристик изображения и занесение их в БД.
- Обработка запроса. К этому этапу относятся формулирование и выполнение запроса, а также просмотр результатов. При необходимости запрос уточняется и выполняется повторно.

Для реализации этих функций в состав системы включены средства хранения изображений и видео, вычисления их характеристик и индексирования по вычисленным характеристикам. В интерфейсе приложения реализованы задание параметров запроса и отображение результатов поиска в порядке убывания их подобия запросу. Модуль обработки запроса взаимодействует с остальными частями системы.

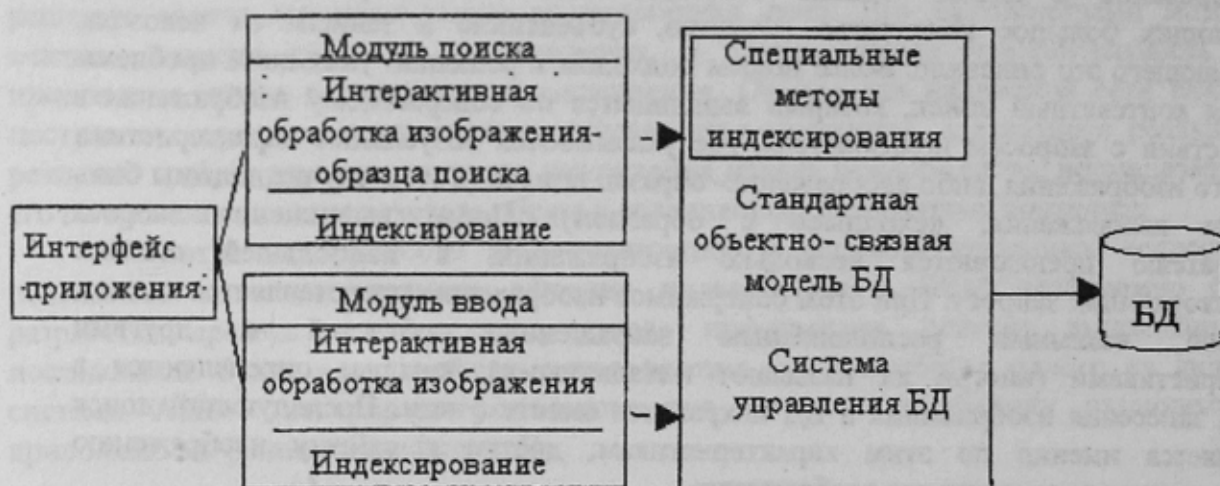


Рисунок 1 Обобщенная архитектура системы контекстного поиска изображений

Система VisualSeek, выполняющая контекстный поиск изображений и видео в WWW, позволяет выполнять запрос по цвету, текстуре и взаимному расположению областей изображения. В дальнейшем предполагается реализация поиска по композиции изображения и встроенному тексту. В описываемой системе реализованы автоматическое выделение областей изображения и признаков, эффективное представление признаков, извлечение сжатых данных, индексирование. Пользователю предоставлена возможность выполнять поиск изображения как по локальным признакам областей, так и по глобальным признакам изображения. Для запроса первого типа пользователь указывает расположение закрасенных областей путем позиционирования области на сетке запроса. Возвращаемые изображения могут быть



использованы для глобального запроса. В составе системы можно выделить следующие компоненты:

- Графический интерфейс с пользователем;
- Приложение- сервер для обработки запросов, выполняющее анализ строки запроса и запись ее в файл, регистрацию запроса, преобразование цветовых параметров, заданных в запросе, в пространство HSV. После этого выполняется открытие БД и индексов и собственно выполнение запроса.
- Сервер поиска изображений
- Архив изображений
- Индексные файлы, определяющие упорядочение изображений по их визуальным характеристикам.

В модели данных для представления визуальной информации в технологии *Virage* используется трехуровневое представление содержимого изображения: сырое изображение, обработанное изображение (уровень объектов изображения), уровень примитивов (характеристик) изображения. Технология *Virage* построена вокруг ядра, называемого *Virage Engine* и работает на уровне объектов изображений. В состав ядра входят три функциональные части для анализа изображений, сравнения изображений и управления.

## 2. Индексирование изображений

Индексирование изображений является средством ускорения поиска данных в БД и является одной из самых важных частей любой системы поиска изображений. Суть индексирования в том, что создается специальная структура данных (как правило, древовидная), соответствующая упорядоченности изображений, хранящихся в БД, по критерию поиска и содержащая информацию об их фактическом расположении. Таким образом, доступ к самому изображению осуществляется только для его отображения после завершения поиска. К индексированию по цветовым характеристикам изображения можно выделить два подхода, использующиеся для запросов различных типов: индексирование по глобальному распределению цветов и локальное цветовое индексирование. Различие между этими методами в том, что индексирование по глобальному распределению позволяет сравнивать изображения целиком, тогда как локальное индексирование дает возможность сравнивать ограниченные области в пределах изображения.

Разработчики системы *QVIC* большое внимание уделяют в процессе индексирования идентификации объектов изображения. Разработчиками предложены полностью автоматические методы идентификации объектов для ограниченного класса изображений и полуавтоматические- для остальных. После выделения объектов для каждого из них и для изображения в целом вычисляются цветовые, текстурные характеристики изображения и характеристики его композиции. В качестве цветовых характеристик *QVIC* вычисляет цветовую гистограмму из  $k$  элементов (как правило,  $k=64$ ), а также средний цвет (R, G, B) и его координаты в пространстве Munsell [8, 9] для каждого объекта и изображения в целом с целью уменьшения количества цветов (для исходного изображения это количество равно 16М).

В системе *VisualSEEK* изображения индексируются по цветовым и текстурным характеристикам. Индексирование по глобальному распределению цветов основано на использовании цветовых гистограмм. Индексирование изображений по цветовым характеристикам локальных областей состоит из двух шагов: выделения областей и определения их цветовых характеристик. Получение характеристик отдельных

областей в системе VisualSeek основано на представлении цветов области в виде бинарного цветового вектора, показывающего, какие из возможных цветов фактически присутствуют в рассматриваемом фрагменте изображения. Цвет включается в цветовой вектор, если он присущ хотя бы 20% точек изображения. Объединяя двоичные вектора цвета и текстуры, получают новую объединенную характеристику, для которой создается отдельный индекс, в итоге повышающий эффективность поиска.

В системе Virage перед индексированием выполняется сглаживание изображения и повышение его контрастности. При определении каждой характеристики изображения (разработчики называют эти характеристики примитивами) используется предварительно обработанное изображение и в зависимости от его содержимого вычисляется специфический набор данных для этого примитива. Вектор вычисленных данных о примитивах сохраняется в специфической структуре данных. Подробную информацию об алгоритмах индексирования фирма Virage не предоставляет, известно только, что для определения каждой характеристики используется несколько методов.

### **3 Особенности поиска и методы определения подобия изображений.**

Как было сказано ранее, рассматриваемые исследования в области контекстного поиска сосредоточена на визуальных характеристиках- цвете, текстуре, композиции изображений.

Разработчики системы QVIS разделяют понятия «сцена» (или «изображение») и объект, являющийся частью изображения. Для объекта в системе вычисляются такие характеристики, как средний цвет, гистограмма цветов, текстура, форма, расположение, для изображения, кроме трех первых характеристик, определяется расположение областей внутри изображения. Для представления вычисленных характеристик используются вектора, включающие цвет, текстуру и композицию объектов изображения, выделяемых вручную. Текстурные и цветовые признаки также используются для описания содержимого всего изображения. Параметрами запроса являются процентный состав цветов, их расположение, текстура и композиция изображения и составляющих его объектов. Программа позволяет пользователю выполнить поиск только по ключевым словам, поиск по содержимому изображения, поиск, использующий как ключевые слова, так и содержимое изображения, поиск с параметрами (задаются процентное соотношение цветов, их расположение и ключевые слова). Каждый из методов поиска, используемый в QVIS, позволяет пользователю искать изображения по различным критериям. Используются следующие критерии определения подобия изображений:

- Поиск по ключевым словам- метод, принятый по умолчанию. Задание ключевых слов позволяет получить изображения, упорядоченные по возрастанию следующего параметра:  $1 - N_m / N_q$ , где  $N_m$ - количество ключевых слов, совпадающих в запросе и в изображении,  $N_q$ - количество ключевых слов запроса.
- Процентное соотношение цветов. Этот критерий позволяет найти изображения, содержащие те же цвета, что и образец поиска, и в том же соотношении.
- Распределение цветов позволяет найти изображения, в которых цвета, сходные с цветами образца, расположены в тех же областях изображения.

В VisualSEEK, разработке фирмы Advent, пользователь может строить визуальный образец запроса с выбором цветов, текстур и заданием формы, размера и пространственного расположения областей. При сравнении изображений может определяться точное совпадение, совпадение с наименьшими расхождениями



(наилучшее из возможных) и несовпадение изображений. В случае цветового сравнения точное совпадение означает, что вектора цветов должны быть абсолютно идентичны для образца поиска и найденного изображения. Совпадение с минимальными расхождениями означает, что система проверяет вектора цветов, близкие к искомому и определяет разницу между ними (расстояние), причем результаты поиска возвращаются в порядке возрастания этого расстояния. В случае пространственного сравнения поиск совпадения с минимальными расхождениями означает, что расстояние между изображениями определяется в зависимости от сдвига закрашенных областей относительно их расположения в образце поиска.

При поиске изображений, подобных одному из хранящихся в БД, в качестве критерия сходства изображений может использоваться одна из ряда специально вычисляемых величин, например, конъюнкция цветовых гистограмм, моменты цветовых гистограмм, конъюнкция закрашенных областей, евклидово расстояние между цветовыми гистограммами и другие. В зависимости от выбранного критерия сходства изображений поиск выполняется по-разному.

При работе с **Virage** у пользователя имеется возможность искать изображения с учетом таких примитивов, как цвет – обобщенное представление цветового содержимого изображения, композиция – пространственное размещение закрашенных областей изображения, текстуры в изображении (например, дерево, гранит, мрамор, облака имеют специфические текстурные характеристики) и структура – общие характеристики формы объектов в изображении. Каждый из этих примитивов имеет собственную линейную метрику (то есть каждому примитиву изображения соответствует числовая мера), при этом поиск подобия выполняется на основе «расстояний» между примитивами запроса и примитивами других изображений в БД, а критерием поиска является заданная степень различия образца поиска и изображений, хранящихся в БД. Для сравнения изображений используется комплексная оценка – взвешенная сумма примитивов. Предусмотрена возможность управления преобразованием комбинации примитивов в число, выражающее подобие данного изображения образцу запроса: каждому примитиву приписывается вес (от 0.0- незначительный примитив до 1.0- наиболее существенный), что позволяет определить, какие примитивы наиболее важны для оценки подобия изображений. При сравнении изображений используется одно или более «расстояний подобия» для пары вычисленных примитивов. Вначале такое расстояние вычисляется для каждого примитива в отдельности, после чего вычисленные значения комбинируются для формирования окончательной оценки.

При разработке систем использовались разнообразные инструментальные средства. Система QVIC написана на языке C, для создания интерфейса пользователя использован X/Motif. При создании системы VisualSEEK использованы Java, HTML и C++. Интерфейс пользователя реализован на Java, для представления результатов запросов используется HTML, выполнение запроса и сопутствующие программы написаны на C++. В Virage для реализации интерфейса также используется HTML. Информация о том, какие инструментальные средства использованы при разработке процессора запросов, фирмой- разработчиком не предоставляется.

### **Заключение**

Из вышесказанного можно сделать следующий вывод: архитектуры рассмотренных систем во многом сходны и являются частными случаями схемы, изображенной на рис.1, а следовательно, похожи также средства хранения и поиска

изображений. Различия заключаются в методах индексирования и способах определения подобия изображений, организации интерфейса с пользователем.

Функциональные возможности рассмотренных систем приведены в табл. 1, где они сравниваются по ряду критериев.

Таблица 1 Сравнение систем контекстного поиска изображений с точки зрения их функциональных возможностей

Название системы	Функциональные характеристики							
	Поиск В WWW	Возможность построения образца поиска	Использование ключевых слов при выполнении запроса	Задание параметров поиска изображений	Гибкая комбинация параметров запроса	Выбор пользователем критерия сходства изображений	Индексирование по глобальным характеристикам	Индексирование по локальным характеристикам
QBIC	+	-	+	+	-	-	+	+
Visual SEEK	+	+	+	+	-	+	+	+
Virage	-	-	-	+	+	-	+	-

С точки зрения функциональных возможностей систем лучшей из рассмотренных является система VisualSeek, предоставляющая пользователю возможность выделения большого количества визуальных характеристик. В этой системе реализован более мощный механизм поиска, использующий закрашенные области и их пространственное расположение.

Можно выделить следующие направления дальнейших исследований в области поиска изображений по образцу:

- Новые модели поиска и индексирования;
- Учет локальных характеристик изображения, в частности, композиции
- Уменьшение вычислительных затрат при выделении характеристик изображения и при сравнении изображений по этим характеристикам

### Литература

1. J. R. Smith and S.-F. Chang, "Searching for Images and Videos on the World-Wide Web," to appear in IEEE Multimedia Magazine, Summer, 1997. (also Columbia University CU/CTR Technical Report #459-96-25). Demo: <http://www.ctr.columbia.edu/webseek>
2. J. R. Smith and S.-F. Chang, "VisualSEEK: A Fully Automated Content-Based Image Query System," ACM Multimedia Conference, Boston, MA, Nov. 1996. Demo: <http://www.ctr.columbia.edu/VisualSEEK>
3. Amarnath Gupta. Visual Information Retrieval Technology- A Virage Perspective
4. IBM Almaden Research Center. Query by Image and Video Content: the QBIC System. Computer, September 1995, 23-31.
5. John R. Smith. Shih-Fu Chang. VisualSEEK: a Content-Based Image/Video Retrieval System. System Report and User's Manual, version 1.0 beta. December 1995.
6. John R. Smith. Shih-Fu Chang. Automated Image Retrieval Using color and Texture.
7. John R. Smith and Shih-Fu Chang. Tools and Techniques for Color Image Retrieval ACM Multimedia 1996: Boston, MA
8. Colour Science Glossary: <http://ziggy.derby.ac.uk/colour/info/glossary/>
9. Chee- Keong Lau. The Effect of Color Quantization in Difference Color Models: <http://www.erc.msstate.edu/~rjm/ColorQuant>