

## МЕТОДЫ И АЛГОРИТМЫ ПОИСКА ИЗОБРАЖЕНИЙ

Огневой Г.Д.

*This article describes the concepts search of images in large image libraries. Described most effective algorithms comparison images. Nominate the concept of optimal search engine*

В настоящее время оцифровка и хранение больших объемов визуальных материалов - не проблема с технической точки зрения. Актуальной проблемой сегодня является обеспечение эффективного содержательного доступа к релевантной информации в электронных коллекциях изображений.

До последнего времени традиционным являлся поиск визуальной информации, опирающийся на индексирование текстовых описаний, ассоциированных с изображением. При очевидной необходимости организации доступа к визуальной коллекции посредством поиска по текстовой информации, ассоциированной с изображениями, такой подход представляется недостаточным. Неоднозначность соответствия между визуальным содержанием и текстовым описанием снижает показатели точности и полноты поиска. Некоторые изображения вообще трудно описать словами (очевидный пример - абстрактные картины).

В связи с этим возникает проблема организации доступа к современным электронным коллекциям изображений с использованием комплекса средств - как текстовых описаний, так и характеристик визуального содержания, простейших типа цветовой гаммы, и более сложных, связанных с распознаванием образов. Текстовое описание и визуальная поисковая информация дополняют друг друга, обеспечивая возможность разностороннего поиска. Поиск может выполняться итеративно: сначала поиск на основе ключевых слов, как более быстрый способ, затем среди отобранного множества материалов - более трудоемкий поиск с использованием визуальных характеристик.

В процессе поиска изображений по визуальному содержанию, процесс сравнения должен осуществляться не по самим исходным изображениям, а по данным, полученным в результате предварительной обработки этих изображений, т.к. сравнение изображений непосредственно друг с другом занимает слишком много процессорного времени, и в случае использования даже на минимальных хранилищах, где количество изображений немногим больше сотни - абсолютно неэффективно. В случае записи данных в строковом формате, эффективность метода зависит от алгоритмов используемых для формирования строки. Одним из наиболее эффективных и универсальных является использование перцептивного хэш-алгоритма.

Перцептивные хэш-алгоритмы описывают класс функций для генерации сравнимых хэшей. Характеристики изображения используются для генерации индивидуального (но не уникального) отпечатка, и эти отпечатки можно сравнивать друг с другом.

Перцептивные хэши — это другая концепция по сравнению с криптографическими хэш-функциями (MD5 и SHA1). В криптографии каждый хэш является случайным. Данные, которые используются для генерации хэша, выполняют роль источника случайных чисел, так что одинаковые данные дадут одинаковый результат, а разные данные — разный результат. Из сравнения двух хэшей SHA1 на самом деле можно сделать только два вывода. Если хэши отличаются, значит, данные разные. Если хэши совпадают, то и данные, скорее всего, одинаковые (поскольку существует вероятность коллизий, то одинаковые хэши не гарантируют совпадения данных). В отличие от них, перцептивные хэши можно сравнивать между собой и делать вывод о степени различия двух наборов данных.

Алгоритмы вычисления перцептивного хэша обладают одинаковыми базовыми

свойствами: картинку можно изменять в размере, менять соотношение сторон и даже слегка менять цветовые характеристики (яркость, контраст и т.д.), но они всё равно совпадают по хэшу. Для получения хэша имеется несколько распространённых алгоритмов. Один из простейших хэш-алгоритмов отображает среднее значение низких частот. В изображениях высокие частоты обеспечивают детализацию, а низкие частоты показывают структуру. Большая, детализированная фотография содержит много высоких частот. В очень маленькой картинке нет деталей, так что она целиком состоит из низких частот. Самый быстрый способ избавиться от высоких частот — уменьшить изображение. Для сравнения разных хэшей, подсчитывается количество разных битов (расстояние Хэмминга). Нулевое расстояние означает, что это, скорее всего, одинаковые изображения (или вариации одного изображения). Дистанция 5 означает, что изображения в чём-то отличаются, но в целом всё равно довольно близки друг к другу. Если дистанция 10 или больше, то это, вероятно, совершенно разные изображения.

Сравнение перцептивных хэшэй достаточно производительно, но в некоторых случаях недостаточно эффективно поскольку не учитывает цветовую гамму и наличие на изображении объектов, представляющих в некоторых случаях основную информацию которую содержит это изображения. Важность для человека цветового восприятия изображения объясняет значение методов поиска визуальной информации на основании схожести с образцом по цветовым характеристикам. Возможно также использование таких показателей, как средний или основной цвета, а также множества цветов; эти характеристики имеет смысл использовать для локального индексирования областей изображения.

Идея метода цветowych гистограмм для индексирования и сравнения изображений сводится к следующему. Все множество цветов разбивается на набор непересекающихся, полностью покрывающих его подмножеств. Для изображения формируется гистограмма, отражающая долю каждого подмножества цветов в цветовой гамме изображения. Для сравнения гистограмм вводится понятие расстояния между ними. Известны различные способы построения и сравнения цветowych гистограмм, отличающиеся между собой изначальной цветовой схемой, размерностью гистограммы и определением расстояния между гистограммами.

При разбиении RGB-цветов по яркости вычисляется интенсивность каждого цвета на основании его красной, синей и зеленой составляющих. Полученное значение, заключенное между числами 0 и 255, попадает в один из 16 интервалов, на которые разбивается диапазон возможных значений. В качестве расстояния между гистограммами используется сумма модулей разности соответствующих элементов гистограмм; некоторое усовершенствование метода достигается при вычислении расстояния на основании поэлементного сравнения гистограмм с учетом соседних элементов. Этот метод наиболее эффективен для черно-белых полутоновых изображений. Для цветных RGB-изображений лучшие результаты дает другой способ - разбиение RGB-цветов по прямоугольным параллелепипедам.

Цветовое RGB-пространство рассматривается как трехмерный куб, каждая ось которого соответствует одному из трех основных цветов (красному, зеленому или синему), деления на осях пронумерованы от 0 до 255 (большее значение соответствует большей интенсивности цвета). При таком рассмотрении любой цвет RGB-изображения может быть представлен точкой куба. Для построения цветовой гистограммы каждая сторона делится на 4 равных интервала, соответственно RGB-куб делится на 64 прямоугольных параллелепипеда. Гистограмма изображения отражает распределение точек RGB-пространства, соответствующих цветам пикселей изображения, по параллелепипедам. В качестве расстояния между гистограммами используется покомпонентная сумма модулей разности между ними. Несмотря на предельную простоту подхода, он показывает довольно стабильные результаты. Например, если необходима картинка в ярко-красных тонах, можно взять за образец фотографию алой розы и получить из базы в качестве результата поиска фотографии, на которых крупным планом представлен красный флаг, черепичная крыша, женщина в красном платье и т.д. Более точное сравнение изображений достигается с помощью техники квадродеревьев, когда методы вычисления и сравнения цветowych гистограмм применяются не ко всему изображению, а к его четверти

(одной шестнадцатой и т. д.). Сравнение изображений основывается на расстоянии, определенном как Евклидово в пространстве расстояний между гистограммами их частей. Этот метод дает результат, семантически отличный от других вариантов: изображения, различающиеся только по взаимному расположению идентичных по цвету объектов, считаются непохожими, в то время как могли быть определены как близкие без использования этой техники. Целесообразность ее применения определяется значением для пользователя расположения на картинке-образце определенных цветовых областей. В качестве примера можно привести задачу поиска фотографий солнечного заката на море.

Особый интерес представляют методы индексирования, применяемые не для изображения в целом, а для отдельных его областей, так как не все части изображения одинаково информативны. Пространственное сегментирование изображения может осуществляться автоматически, когда выделяются области с некоторыми общими свойствами - одинаковыми или сильно схожими значениями того или иного примитива.

Задача выявления границ объекта решается путем локализации на изображении резких перепадов яркости цвета. С этой целью вычисляется градиент функции интенсивности в каждой точке изображения, после чего подавляются значения меньше установленного порога. Выбор порога является одним из ключевых вопросов выделения перепадов.

В результате обработки получается бинарная матрица, где единицам соответствуют точки со значительным перепадом яркости, нулям – все остальные. С целью борьбы с шумом и ликвидации возможных разрывов в контурах применяются морфологические операции, после чего в бинарной матрице единицами представлены точки, принадлежащие искусственно утолщенным на предыдущем этапе границам объектов. Для выделения точек внешнего контура используется обход полученной области по внешней ее стороне. Небольшие объекты при этом исключаются из рассмотрения. Для выделенных объектов могут быть определены и включены в индекс такие параметры, как координаты на изображении, размеры, характеристики цвета, измерения текстуры и формы. Существует успешная практика использования формы объектов для индексирования изображений с целью их дальнейшего сравнения.

Пользователю изображений должна быть предоставлена возможность строить запросы с использованием различных визуальных средств - в терминах не только визуальных примитивов, но и высокоуровневых объектов. Для этого в поисковом образе должен отражаться факт присутствия на изображении объектов, а также их размеры и расположение на кадре. Задача нахождения на изображении объектов в настоящее время не ставится глобально. Как правило, речь идет об объектах определенного класса, особенно интересных для рассматриваемой предметной области. В большинстве случаев, когда нужно простое сравнение двух достаточно похожих фрагментов изображения его реализуют через их ковариацию. Берётся образец и передвигается по изображению по X и Y в поисках точки, где отличие образца (J) от изображения (I):

$$\sum_{i < W, j < H} |I(x + i, y + j) - J(i, j)|$$

достигает своего минимума.

Этот способ очень быстр в реализации, интуитивен и досконально известен.

Недостатки:

- Неустойчивость при смене освещения
- Неустойчивость при изменении масштаба или повороте изображения
- Неустойчивость, если часть изображения — изменяющийся фон

- Маленькая скорость работы — если нужно обнаружить область  $n \times n$  на изображении  $m \times m$ , то количество операций будет пропорционально  $n^2 \cdot (m-n)^2$ .

Каждый из перечисленных методов имеет свои достоинства и недостатки. Для разработки приложения по поиску изображений следует применять комплекс методов при индексировании изображений: нахождение перцептивного хэша, метод гистограмм, поиск объектов. При этом возможность выделения характерных и важных объектов, составление текстовых описаний также следует предоставить пользователю. Только при таком комплексном подходе, а также возможности выбора методов поиска (в случае безрезультатного поиска по объектам, к примеру, использовать поиск по перцептивному хэшу) программа поиска изображений будет наиболее эффективна.

## ЛИТЕРАТУРА

1. Н.С. Байгарова, Ю.А. Бухштаб, А.А. Воробьев, А.А. Горный  
Организация управления базами визуальных данных  
Препринт Института прикладной математики им. М.В. Келдыша РАН, 2000, N 6
2. Н.С. Байгарова, Ю.А. Бухштаб, А.А. Горный  
Методы индексирования и поиска визуальных данных  
Препринт Института прикладной математики им. М.В. Келдыша РАН, 2000, N 7
3. Baron, J. L., Fleet, D. J., and Beauchemin, S. S., Performances of optical flow techniques.  
Int. Journal of Computer Vision, 12:1, pp.43—77, 1994
4. Jain, R. and Gupta, A., Visual Information Retrieval,  
Communications of the ACM, 1997, vol. 40, no. 5.
5. Jain, R., Pentland, A.P., Petkovic, D.,  
Workshop Report: NSF – ASPA Workshop on Visual Information Management Systems, 1995.  
<http://www.virage.com/vim/vimsreport95.html>
6. Looney, C.G., Pattern Recognition Using Neural Networks,  
Theory and Algorithms for Engineers and Scientists, Oxford University Press, 1997.
7. Sobel, I., An isotropic image gradient operator.  
Machine Vision for Three-Dimensional Scenes, pp.376-379. Academic Press, 1990