

СПбГУ на РОМИП'2008: Синтез цветовых и текстурных признаков при поиске изображений по содержанию

© Наталья Васильева, Илья Марков

Санкт-Петербургский Государственный Университет
nvassilieva@hp.com, ilya.markov@gmail.com

Аннотация

В статье описаны методы поиска, которые использовались командой СПбГУ при участии в дорожке поиска изображений по содержанию семинара РОМИП'2008. Статья содержит описание визуальных признаков, а также методов синтеза результатов поиска по каждому из них, использовавшихся для формирования общей выдачи. В статье также рассматриваются полученные результаты и обсуждаются возможные причины неудачных результатов по некоторым из запросов.

1. Введение

Традиционным подходом к поиску изображений по содержанию является независимый анализ их визуальных характеристик (цвета, текстуры и формы объектов), каждая из которых описывается точкой в соответствующем пространстве признаков. Для получения общей выдачи на основе промежуточных результатов поиска по каждому из пространств признаков, используются методы синтеза данных (Data Fusion).

При поиске по разнородной коллекции изображений цвет и текстура являются наиболее значимыми характеристиками.

В данной работе мы приводим описание глобальных признаков цвета и текстуры, используемых нами для поиска изображений по

содержанию, а также методов синтеза, которые служат для построения общей выдачи по независимым результатам поиска с использованием каждого из признаков.

Участие в РОМИП'2008 позволило нам получить объективную оценку используемых методов поиска по разнородным коллекциям изображений, а также на основе анализа полученных результатов предположить возможные способы повышения точности работы наших методов.

Далее в разделе 2 приводится описание набора признаков и методов синтеза, которые были использованы для подготовки результатов по дорожке поиска изображений по содержанию семинара РОМИП'2008; раздел 3 содержит обсуждение полученных результатов; в разделе 4 сформулированы выводы, сделанные на основе данных результатов.

2. Описание методов поиска

2.1 Признаки

Стандартным способом представления цветовой характеристики изображения являются цветные гистограммы [8, 9, 1]: пространство всех цветов разбивается на подмножества с тем, чтобы похожие цвета, по возможности, попали в один элемент разбиения (цветовая область), и для каждой цветовой области подсчитывается количество пикселей, чей цвет принадлежит этой области. Для сравнения изображений на основании цветных гистограмм в качестве метрики используется сумма модулей разностей значений элементов гистограмм для каждой цветовой области. В данной работе в качестве одного из методов поиска по цвету мы будем использовать модификацию цветной гистограммы, учитывающую пространственное расположение цветов на изображении [12, 1].

Альтернативным методом представления цветовой характеристики является статистическая модель. В рамках данной модели за основу берется статистическое распределение отдельных цветных каналов [9]. Для оценки схожести двух изображений сравниваются параметры распределения их цветов (математическое ожидание, дисперсия и другие параметры статистического распределения). В [10] цветные каналы рассматриваются не как независимые распределения, а как части трехмерного распределения: помимо средних значений по каждому каналу для характеристики берутся также попарные ковариации каналов.

Именно такой статистический метод мы будем использовать в качестве второго метода поиска по цвету.

Для поиска по текстуре, мы будем использовать метод, основанный на анализе независимых компонент (Independent Component Analysis – ICA) [4]. С его помощью выделяются фильтры, которые призваны отражать основные направления текстуры для той базы изображений, на основе которой они строятся. В качестве характеристического вектора изображения берется набор его сверток с каждым из фильтров, а в качестве метрики используется дивергенция Kullback-Leibler. Особенностью фильтров, полученных методом ICA, является их естественный характер: они отражают основные направления текстуры для набора изображений, по которому были построены. В этом их преимущество по сравнению с широко применяемыми фильтрами Габора (Gabor), которые строятся на основе различных сечений функции Габора, то есть носят искусственный характер.

2.2 CombMNZColor

Одним из наиболее эффективных методов синтеза ранжированных списков, применяемых в текстовом поиске, является CombMNZ [7]. Вес элемента в результирующем списке вычисляется как сумма его весов во всех списках, подлежащих синтезу, умноженная на количество списков, где этот элемент присутствует с весом не равным 0. Если элемент отсутствует в каком-либо списке, подлежащем синтезу, его вес в этом списке также считается равным 0.

В [2, 5, 6] мы показали, что метод CombMNZ можно успешно применять в поиске изображений. В [2] также показано, что если списки, подлежащие синтезу, сильно пересекаются, CombMNZ возвращает результат лучший по сравнению с синтезом слабо пересекающихся списков. По этой причине, а так же по причине невозможности адекватной нормировки значений цветовых и текстурной метрик (см. раздел 3) было решено применить метод CombMNZ для синтеза результатов, полученных только по цветовой характеристике с помощью двух различных ее представлений, описанных выше.

2.3 MixedMetrics

Метод синтеза результатов цветового и текстурного поисков по одному изображению-запросу на основе смешанных метрик

рассматривается в [3, 5, 6]. Согласно данному методу, итоговая оценка схожести изображений с запросом вычисляется по следующей формуле: $a \cdot C + (1 - a) \cdot T$, где C – метрика по цветовой характеристике, T – метрика по текстурной характеристике, a – параметр из отрезка $[0, 1]$, зависящий от изображения-запроса. В той же работе было показано, что для всех изображений-запросов оптимальное значение параметра a вычисляется однозначно, а для схожих запросов оно одно и то же.

В данной работе мы используем группы семантически близких изображений и оптимальные значения параметра a для них, определенные в [3] и [5], со скидкой на особенности нормировки, которые обсуждаются в разделе 3. В качестве результата поиска по цвету используется результат CombMNZColor, а для получения результата текстурного поиска используются вектора, построенные с помощью фильтров ICA.

3. Эксперименты на РОМИП'2008

В дорожке поиска изображений по содержанию были представлены результаты шести прогонов от трех участников. Мы представили к оценке результаты двух прогонов:

- **CombMNZColor** – поиск по цветовым моментам и цветовым гистограммам; в качестве метода синтеза использовался алгоритм CombMNZ;
- **MixedMetrics** – поиск с использованием алгоритма MixedMetrics; в качестве результатов поиска по цвету брался результат первого прогона (синтез цветовых моментов и цветовых гистограмм с помощью CombMNZ), в качестве текстурных признаков использовались вектора, построенные с помощью фильтров ICA.

Во время индексирования коллекции РОМИП'2008 мы столкнулись со следующими сложностями. Предварительные эксперименты с текстурными признаками на основе фильтров ICA показали, что они крайне неустойчивы на большой разнородной коллекции изображений. Данные признаки показывают приемлимые результаты только на запросах с тривиальной текстурой – очень гладкой (как, например, изображения неба или воды) или интенсивной и регулярной (например, изображения городов с высоты птичьего полета).

Второй проблемой являлся вопрос нормировки значений метрик по цветовым и текстурным признакам. Ранее мы использовали нормировку по среднему и дисперсии:

$$\tilde{x}_i = \frac{x_i - \bar{x}}{\sigma},$$

где x_i - исходное значение метрики, \bar{x} - среднее значение метрики (математическое ожидание), σ - среднеквадратичное отклонение значений метрики. Такая нормировка гарантирует для результатов по всем метрикам нулевое среднее значение и одинаковое среднеквадратичное отклонение.

После применения такой нормировки к значениям цветовой и текстурной метрик на коллекции РОМИП'2008 мы обнаружили, что минимальные значения текстурной метрики для каждого из запросов на порядок меньше минимальных значений цветовой метрики. Это означает, что при попытке синтеза результатов поиска по цвету и текстуре текстурная метрика автоматически имеет существенный приоритет.

Линейная нормировка значений, задаваемая выражением

$$\tilde{x}_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}},$$

где x_{\min} и x_{\max} - минимальное и максимальное значение метрики соответственно, приводила к схожему результату.

Чтобы уменьшить влияние текстурной метрики, которая в среднем дает худшие результаты, чем цветовая, было принято решение использовать небольшие веса для текстурной метрики при синтезе результатов поиска по цвету и текстуре с помощью MixedMetrics.

Несмотря на перечисленные выше сложности, ожидалось, что результаты прогона MixedMetrics будут лучше результатов CombMNZColor, поскольку первый использовал большее число признаков. Был получен результат, в целом оправдавший наши ожидания.

2.3 Результаты экспериментов

Для оценки результатов дорожки поиска изображений по содержанию на РОМИП'2008 использовались следующие метрики: Precision(10), Bpref-10, Bpref, Recall, Average precision, Precision, R-precision, Precision(5). Для оценки использовалась трехуровневая шкала релевантности ответов систем: релевантен, частично

релевантен, нерелевантен. Оценка производилась по схемам AND (сильные требования к релевантности – ассессоры единогласны в своей оценке) и OR (слабые требования к релевантности – хотя бы один ассессор имеет такое мнение). Итого было получено четыре набора оценок:

- AND-strong – релевантными считаются только те документы, которые были отмечены всеми ассессорами, как релевантные;
- OR-strong – релевантными считаются документы, отмеченные хотя бы одним ассессором, как релевантные;
- AND-weak – релевантными считаются документы, отмеченные всеми ассессорами как релевантные или частично релевантные;
- OR-weak – релевантными считаются документы, отмеченные хотя бы одним ассессором, как релевантные или частично релевантные.

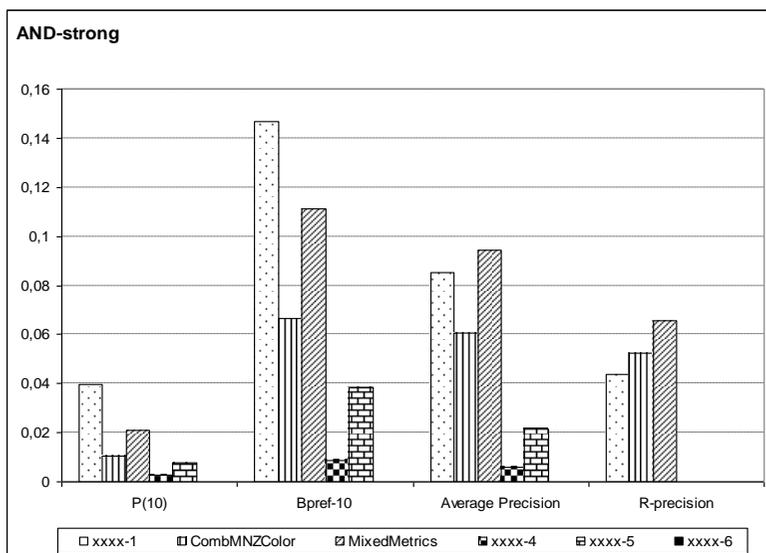


Рис. 1. Результаты участников дорожки поиска изображений по содержанию, оценка AND-strong.

Необходимо отметить, что между полученными результатами по различным метрикам и с использованием различных схем оценки наблюдается значительная корреляция. В связи с этим, в данном отчете для иллюстрации мы приводим результаты только по некоторым, наиболее интересным на наш взгляд метрикам для схем AND-strong (рис. 1) и OR-weak (рис. 2).

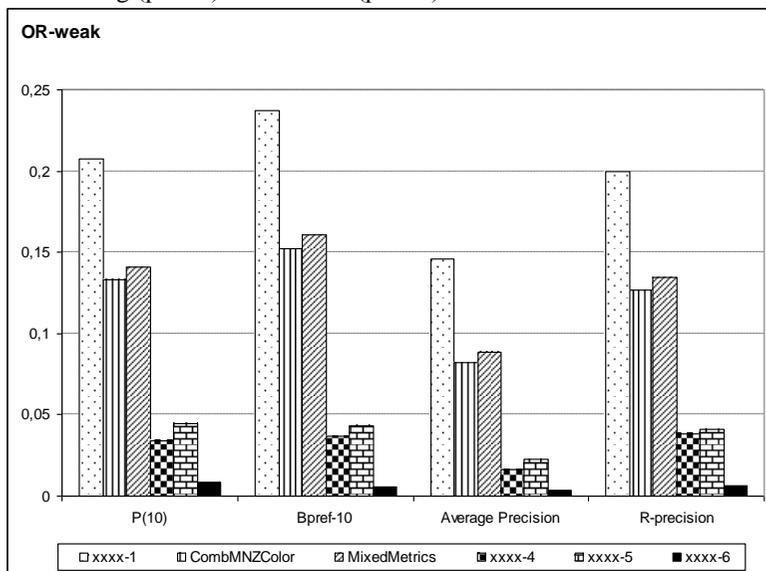


Рис. 2. Результаты участников дорожки поиска изображений по содержанию, оценка OR-weak.

2.3 Анализ результатов

На графиках 1, 2 представлены результаты участников дорожки поиска изображений по содержанию для оценок AND-strong и OR-weak. Видно, что участник xxxx-1 является явным лидером практически по всем метрикам и способам оценки. MixedMetrics и CombMNZColor показали второй и третий результаты соответственно. Результаты по остальным метрикам и способам оценки, не включенные в данный отчет, демонстрируют аналогичное поведение сравниваемых систем.

С целью выявления слабых мест наших методов поиска мы проанализировали некоторые запросы, по которым были получены плохие результаты по сравнению с остальными участниками. Ниже на рис. 3 приведено несколько таких запросов.



Рис. 3. Примеры запросов, на которых CombMNZColor и MixedMetrics показали плохие результаты. Номера запросов слева направо: 1590, 1124, 1463, 978.

Во время анализа промежуточных результатов поиска (по каждому из использовавшихся признаков, до применения методов синтеза), мы сделали следующие наблюдения.

Для части проблемных запросов пересечение промежуточных результатов поиска очень мало (например, для запросов 1590 и 1463). При этом в результатах поиска по каждому из признаков может присутствовать некоторое количество изображений, релевантных запросу. Но множества релевантных изображений для различных признаков не пересекаются. В результате синтеза все релевантные изображения получают низкие ранги и не попадают в общий ответ системы.

Представленные на РОМИП'2008 методы поиска работают на основе глобальных признаков цвета и текстуры и не различают фон и объекты изображения. В случаях, когда фоновая область занимает существенную часть изображения, а также когда фон хорошо описывается глобальными признаками и в базе есть изображения с подобным фоном, поиск производится в основном именно по признакам фона, а не объектов, которые, возможно, представляют больший интерес для пользователя. Примером такого запроса является запрос 978.

Для некоторых запросов недостаточно использования глобальных признаков цвета и текстуры (например, запрос 1124). Судя по таблицам релевантности, ассессоры судили о релевантности запроса 1124 и ответов систем на данный запрос по семантическому подобию. В числе релевантных ответов оказались изображения кошек и собак (животные, как и курица на запросе); человек с открытым ртом (у курицы на изображении-образце также открыт клюв). Однако для определения семантического подобия между изображениями недостаточно использования низкоуровневых визуальных признаков (как глобальных, так и локальных). На сегодняшний день задача определения семантического подобия изображений остается открытой.



Рис. 4. Примеры релевантных изображений по оценке OR-weak на запрос 1124.

4. Выводы

Проведенные эксперименты в рамках семинара РОМИП'2008 показали, что комбинирование небольшого набора глобальных цветовых и текстурных признаков вполне оправданно при поиске по разнородной коллекции изображений. В то же время, анализ результатов поиска по проблемным запросам для описанных методов позволил нам сделать следующие предположения о возможных способах повышения точности поиска:

- совершенствование алгоритмов синтеза в случае когда пересечение промежуточных результатов поиска мало;
- выделение фоновой области на изображении и понижение веса признаков фона при поиске;

Литература

- [1] Васильева Н., Б. Новиков. Построение соответствий между низкоуровневыми характеристиками и семантикой статических изображений . *Труды 7-ой Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» - RCDL'2005*, – Ярославль, Россия, 2005.
- [2] Васильева Н., Дольник А., Марков И. Поиск изображений. Синтез различных методов поиска при формировании результатов. *Сборник «Интернет-математика 2007»*.
- [3] Марков И., Васильева Н., Яремчук А. Поиск изображений. Выбор оптимальных весов для слияния метрик по цвету и текстуре в зависимости от запроса-образца. *Труды 9-ой Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» - RCDL'2007*, – Переславль-Залесский, Россия, 2007.

- [4] Borgne H., Guerin-Dugue, A. Antoniadis. Representation of images for classification with independent features. In *Pattern Recognition Letters*, volume 25, pages 141-154, 2004.
- [5] Markov I., Vassilieva N. (2008) Image retrieval. Color and texture combining based on query-image. In *Proceedings of International Conference on Image and Signal Processing (ICISP 2008)*, July 1–3, Cherbourg-Octeville, France, pp. 430–438.
- [6] Markov I., Vassilieva N. (2008) Building Up Low-Level Centroids for Groups of Perceptually Similar Images. In *Proceedings of 8th International Baltic Conference on Databases and Information Systems (Baltic DB&IS 2008)*, June 2–5, Tallinn, Estonia, pp. 341–348.
- [7] Shaw J. A., Fox E. A. Combination of multiple searches [Electronic resource]. – 1994. Mode of access: <http://citeseer.ist.psu.edu/fox94combination.html>
- [8] Stricker M., Orengo M. Similarity of Color Images. In *Proceedings of the SPIE Conference*. – 1995. – Vol. 2420. – P. 381-392.
- [9] Stricker M., Dimai A. Color Indexing with Weak Spatial Constraints. In *Proceedings of the SPIE Conference*, 1996.
- [10] Stricker M., Dimai A. Spectral Covariance and Fuzzy Regions for Image Indexing. In *Machine Vision and Applications*, volume 10, pages 66-73, 1997.
- [11] Swain M.J., Ballard D.H. Color Indexing. In *International Journal of Computer Vision*, volume 7(1), pages 11-32, 1991.
- [12] Vassilieva N., Novikov B. A similarity retrieval algorithm for natural images. In *Proceedings of the Baltic DB&IS'2004*. – Riga, Latvia: Scientific Papers University of Latvia, 2004. – Vol. 672.

SPbGU at ROMIP'2008.
Fusing Color and Texture Features in
Content-Based Image Retrieval

Natalia Vassilieva, Ilya Markov

The article describes retrieval methods were used by SPbGU team for participation in content-based image retrieval track at ROMIP'2008. Visual features are described as well as data fusion methods used to combine the retrieval results of each independent feature. The experimental results are presented and possible causes of poor results for some queries are discussed.