

Простые методы поиска изображений по содержанию

© Пименов В.Ю.

Санкт-Петербургский Государственный университет,
факультет Прикладной математики - процессов
управления
vitaly.pimenov@gmail.com

Аннотация

В статье описаны методы решения трех задач поиска изображений по содержанию: поиска нечетких дубликатов, поиска по образцу и аннотирования. Решение перечисленных задач основано на базовом математическом анализе свойств изображений. Для поиска нечетких дубликатов и аннотирования изображения представляются в виде векторов, составленных из значений яркости после применения преобразований Габора. Поиск по образцу осуществляется на основе нечетких цветовых гистограмм, вычисляемых в пространстве CIE $L^*a^*b^*$. Рассмотренные методы просты в реализации и обладают высокой скоростью. Экспериментальные результаты показывают, что предложенные методы обеспечивают качество решения задач сравнимое с альтернативными методами и, в некоторых случаях, превосходят их.

1. Введение

В настоящее время задачи поиска изображений по содержанию находятся в фокусе множества исследований. Трудности в формализации задачи обуславливают высокую сложность как разработки методов ее решения, так и разработки методик оценки качества полученных решений.

Цель настоящего исследования состоит в разработке простых для программной реализации методов решения задач поиска нечетких

дубликатов, поиска изображений по образцу и автоматического аннотирования изображений и экспериментальном анализе их эффективности. Подобные методы могут найти применения при создании систем решения задач поиска изображений по содержанию в качестве базы для построения процедур машинного обучения. Однако методы машинного обучения не рассматриваются в настоящей статье, так как требуют наличия существенных объемов данных для проведения обучения и, зачастую являются ресурсоемкими как с точки зрения их разработки, так и с точки зрения необходимости больших вычислительных мощностей.

Решению рассматриваемых в настоящей работе задач посвящено большое число исследований, заинтересованный читатель может обратиться к статьям [1-3], посвященным анализу современного состояния предметной области.

Содержание настоящей статьи организовано следующим образом. Раздел 2 посвящен описанию методу поиска нечетких дубликатов в коллекции изображений. Метод поиска изображений по образцу рассмотрен в Разделе 3. В Разделе 4 изложен метод аннотирования изображений. Раздел 6 содержит общие выводы.

2. Метод поиска нечетких дубликатов изображений

Предлагаемый метод поиска нечетких дубликатов в коллекции изображений предназначен для решения задачи кластеризации коллекции и основан на представлении изображения в виде вектора в евклидовом пространстве фиксированной размерности.

Векторная модель изображения строится на основе метода Gist [4] с использованием следующего алгоритма:

1. Приведение изображения к размеру 64×64 точки.
2. Построение пирамиды из трех изображений: 64×64 , 32×32 и 16×16 точек соответственно
3. На каждом уровне пирамиды для каждого цветового компонента изображения выполняется свертка с блоком фильтров Габора [5].
4. Из результатов свертки, приведенных к размеру 4×4 точки, последовательно составляется конечный вектор.

Данная операция выполняется для всех изображений коллекции, в результате чего производится переход от изображений к векторам в евклидовом пространстве.

Кластеризация набора векторов основана на модифицированном алгоритме одинарной связи, к которому добавлено ограничение на

отношение между расстоянием от анализируемого изображения до двух ближайших кластеров (nearest neighbor distance ratio) [6].

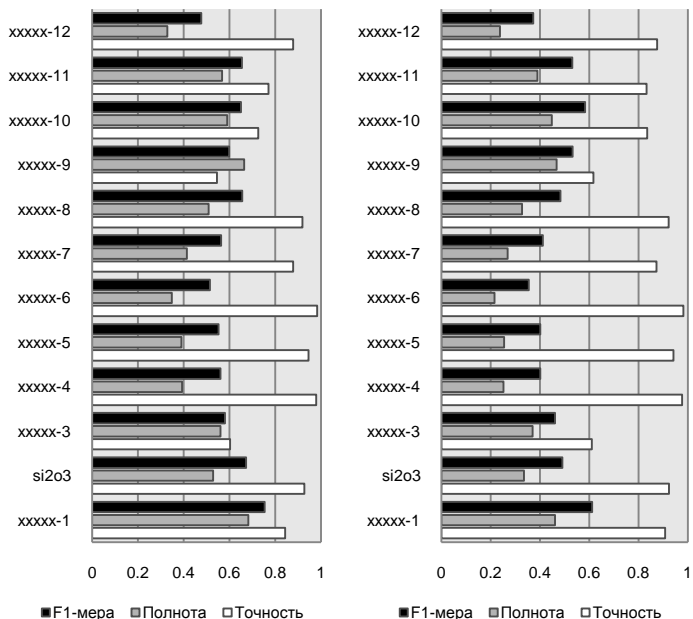


Рисунок 1. Гистограммы показателей качества решения задачи поиска нечетких дубликатов изображений для первого (слева) и второго (справа) ассессора.

На Рис.1 показаны гистограммы значений точности, полноты и F_1 -меры – показателей качества решения задачи поиска нечетких дубликатов. Предложенный в работе метод, обозначенный как «si2o3», имеет второе максимальное значение F_1 -меры среди двенадцати оценивавшихся методов по оценкам первого ассессора и пятое по оценкам второго ассессора. Это позволяет сделать вывод о том, что предложенный метод является конкурентноспособным.

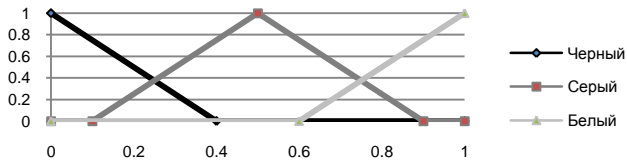
3. Метод поиска изображений по образцу

Для решения задачи поиска изображений коллекции по образцу предложен алгоритм, в основе которого лежит представление изображения в виде нечеткой цветовой гистограммы.

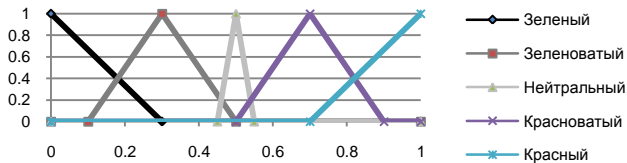
Метод расчета нечеткой цветовой гистограммы основан на методе, предложенном в [7] и включает следующие этапы:

1. Приведение исходного изображения к размеру, не превосходящему 100×100 точек. При этом производится сохранение пропорций исходного изображения и бикубическая интерполяция цветов.
2. Приведение полученного изображения в универсальное цветовое пространство CIE $L^*a^*b^*$ [5].
3. Вычисление значений функций принадлежности в каждой точке изображения.
4. Построение гистограммы значений функций принадлежности.

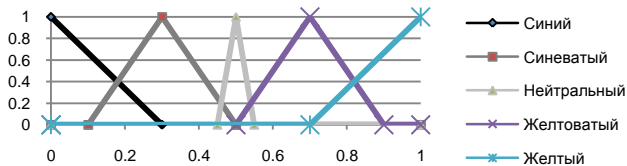
В работе, следуя оригинальному методу, рассматриваются отдельные функции принадлежности для каждого цветового канала, форма которых проиллюстрирована на Рис. 2.



(а)



(б)



(в)

Рисунок 2. Функции принадлежности: (а) канал L^* , (б) канал a^* , (в) канал b^* .

Отличительной чертой предлагаемого в настоящей работе метода является то, что гистограммы строятся без использования процедуры дефазификации. В каждой точке изображения строится

вектор 75-мерного пространства, компоненты которого являются средним арифметическим значений трех функций принадлежности. После чего производится усреднение данных векторов.

Расстояние между подобными векторами, описывающими изображение, вычисляется на основе функции пересечения гистограмм:

$$d(P, Q) = 1 - \frac{\sum_{i=1}^{75} \min(P_i, Q_i)}{\min(\sum_{i=1}^{75} P_i, \sum_{i=1}^{75} Q_i)},$$

где P, Q – сравниваемые вектора.

В отличие от базового метода, в рамках которого для изображения вычисляется единственная гистограмма, в предлагаемом методе гистограммы предложено вычислять шесть дополнительных гистограмм, которые строятся для участков изображения, показанных на Рис. 3. Это позволяет рассчитывать расстояние между изображениями с учетом простейшим геометрических преобразований, таких как отражение.



Рисунок 3. Участки изображения, используемые для построения дополнительных цветных гистограмм¹.

Расстояние между изображениями P, Q , заданными гистограммами (P^0, \dots, P^6) и (Q^0, \dots, Q^6) , где индексом «0» обозначена гистограмма, построенная по полному изображению, вычисляется как взвешенное средние расстояний между гистограммами:

$$D(P, Q) = \sum_{i=0}^6 \sum_{j=0}^6 d(P^i, Q^j) \cdot \delta_{ij},$$

¹ Ссылка на изображение: <http://www.flickr.com/photos/ninhobomba/282410160/>.

где δ_{ij} – элемент матрицы, определяющей вес расстояния между гистограммами с индексами i и j . Данная матрица была задана в виде

$$\Delta = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Ранжирование изображений выполнялось в порядке возрастания расстояния от изображения-запроса до изображения-кандидата.

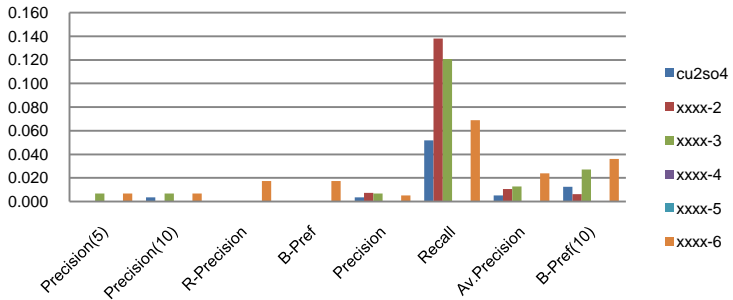
На Рис. 4 представлены гистограммы метрик качества, использованных при оценке методов поиска изображений по образцу для шести методов, представленных для оценки. Предложенный в настоящей работе метод обозначен как «su2so4». Описание метрик и методик оценки приведено в Приложении ? настоящего сборника.

Результаты эксперимента показали, что разработанный метод обеспечивает качество решения задачи, сравнимое с качеством, достигаемым при использовании альтернативных методов, и в некоторых случаях превосходит их. В частности, согласно методикам оценки «or weak» и «vote weak» рассматриваемый метод показал лучший результат по всем используемым метрикам качества. К достоинствам предложенного метода также стоит отнести простоту его реализации и вычислительную эффективность, так как вычисление расстояний между изображениями с помощью оценки пересечения гистограмм не требует операций умножения с плавающей точкой.

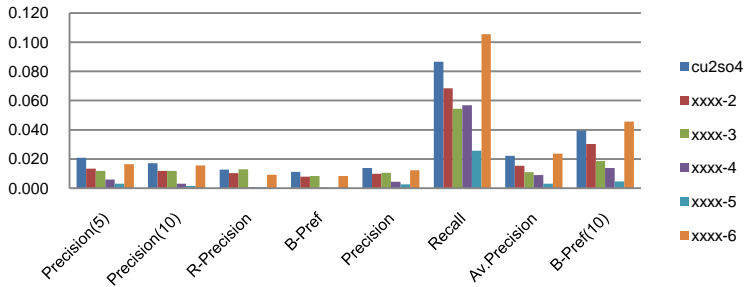
4. Метод аннотирования изображений

Предлагаемый метод аннотирования изображений принадлежит группе методов «грубой силы» (brute force) и основан на присвоении изображению аннотации согласно результатам поиска нечетких дубликатов данного изображения в заранее аннотированной коллекции большого объема. Данный метод сформулирован в предположении о том, что с возрастанием объема аннотированной коллекции до тысяч изображений на заданную текстовую метку, поиск нечетких дубликатов для заданного изображения позволит получить аннотации. При этом точность подобного метода будет определять точностью результатов поиска нечетких дубликатов.

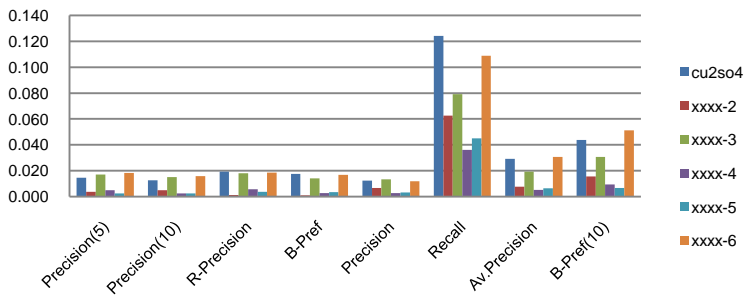
And strong



And weak



Or strong



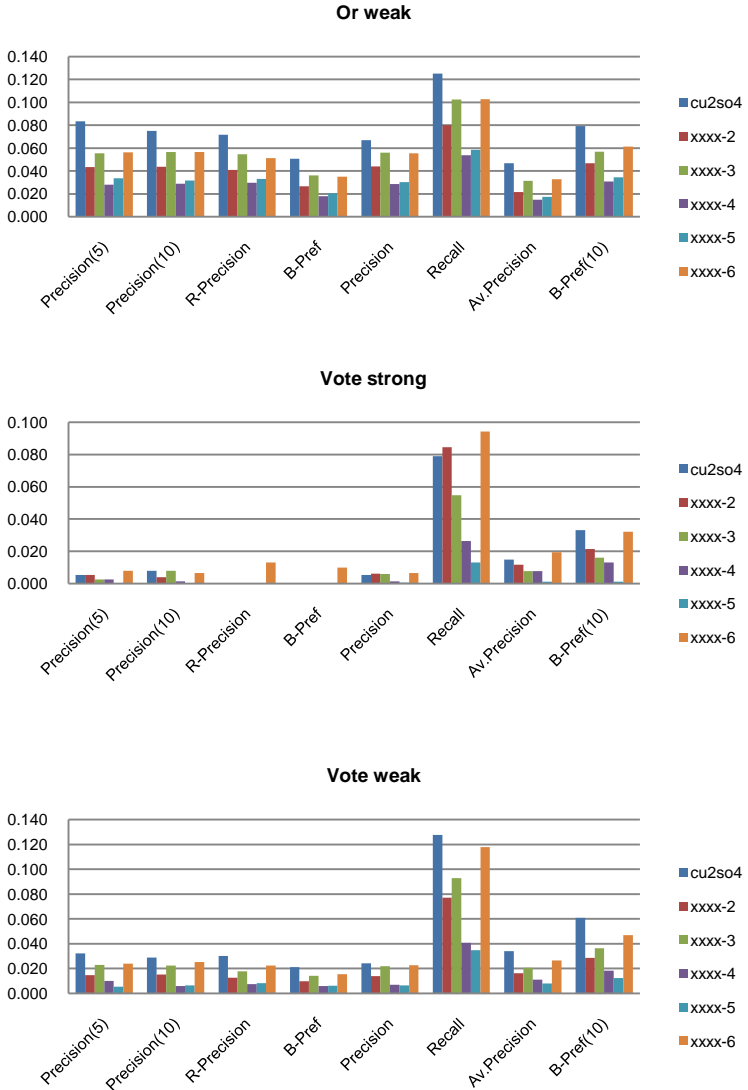


Рисунок 4. Гистограммы значений метрик качества поиска изображений по содержанию. Сверху вниз: методики «and strong», «and weak», «or strong», «or weak», «vote strong», «vote weak».

Таблица 1. Текстовые метки, использованные в задаче аннотирования изображений.

Текстовая метка	Количество размеченных изображений	Количество изображений, получивших метку
Растение	3000	445
Группа людей	1660	218
Здание	1421	96
Небо	1246	54
Кошка	1485	19
Собака	1603	11
Дорога	1288	4
Торт	1233	2
Птица	2126	2
Море	2382	1
Горы	2029	1

В качестве заранее аннотированной коллекции в настоящем исследовании была использована коллекция изображений, составленная на основе сверхбольшой коллекции ImageNet², для изображений которой создана англоязычная разметка на основе семантической сети WordNet. В качестве меток-кандидатов для последующего назначения анализируемым изображениями были выбраны словосочетания русского языка, для англоязычных аналогов которых в коллекции ImageNet доступны более 1000 изображений. Для данных наборов изображений были рассчитаны векторные модели, согласно методу, описанному в разделе 3.

После этого для изображений, входивших в задание по аннотированию были вычислены расстояния до каждого из наборов. Полученные величины сравнивались с эмпирически определенным пороговым значением и, в случае, если порог не был превышен, изображению присваивалась текстовая метка соответствующего набора. В Табл. 1 приведен список текстовых меток, каждая из которых была присвоена хотя бы одному изображению (ряд меток-кандидатов не был присвоен ни одному из изображений, входивших в задание).

Значения качественных показателей, полученные по результатам экспертной оценки приведены в Табл. 2. Исходя из абсолютных значений средней точности при микроусреднении, можно сделать вывод о том, что применение предложенного метода позволяет строить корректные текстовые аннотации. Низкое значение средней

² <http://www.image-net.org/>

точности при макроусреднении объясняется тем, что для большого числа изображений не было найдено ни одной аннотации.

Значение полноты при микроусреднении равно единице, так как в оценке участвовал единственный метод в результате чего множество меток, в соответствии с которым производилось вычисление значения полноты совпадало с множеством меток, полученных рассматриваемым методом.

Таблица 2. Значения показателей качества решения задачи построения текстовых аннотаций.

Показатель	Методика «And»	Методика «Or»
Micro Precision	0,704	0,739
Macro Precision	0,275	0,289
Micro Recall	1	1
Macro Recall	0,285	0,298
AvgRelevantPerTask	0,300	0,315

Экспериментальная оценка предложенного в работе метода построения текстовых аннотаций показала, что использование метода «грубой силы» является перспективным подходом, однако эффективности его применения существенно зависит от богатства заранее размеченной коллекции.

5. Заключение

В статье предложены методы решения трех задач поиска изображений по содержанию: поиска нечетких дубликатов в коллекции изображений, поиска изображений по образцу, построения текстовых аннотаций. Разработанные методы, в основе которых лежат векторные модели изображения, отличаются простотой реализации и не требуют использования высокопроизводительных вычислительных систем. Вместе с тем, результаты экспериментов, проведенных в рамках семинара РОМИП подтверждают конкурентноспособность данных методов.

Основным направлением дальнейшего развития полученных результатов является синтез разработанных методов и методов машинного обучения с целью установления достижимых пределов эффективности исходных методов.

Литература

- [1] B. Thomee, M. Huiskes, E. Bakker, M. Lew. Large Scale Image Copy Detection Evaluation. *Proc. 1st ACM Int. Conf. on Multimedia Information Retrieval*. 2008. P. 59-66.
- [2] L. Chen, F. Stentiford. Comparison of Near-Duplicate Image Matching. *Proc. 3rd Eur. Conf. on Visual Media Production*. 2006. P. 38-42.
- [3] W. Zhao, C. Ngo, H. Tan, X. Wu. Near-Duplicate Keypoint Identification with Interest Point Matching and Pattern Learning. In *IEEE Trans. on Multimedia*. V. 9. № 5. 2007. P. 1037-1048.
- [4] A. Oliva, A. Torralba. Modeling the shape of the scene: a holistic representation of the spatial envelope. *Int. J. Comput. Vis.* V. 42. № 3. 2001. P. 145-175.
- [5] Д. Форсайт, Ж. Понс. Компьютерное зрение. Современный подход. Вильямс. 2004.
- [6] D. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vis.* V. 60. 2004. P. 91-110.
- [7] K. Konstantinidis, A. Gasteratos, I. Andreadis. Image retrieval based on fuzzy color histogram processing. *Optics Communications*. V. 248. № 4-6. 2005. P. 375-386.

Simple methods for content-based image retrieval

Vitaly Pimenov

Saint-Petersburg State University, Faculty of Applied Mathematics
and Control Processes
vitaly.pimenov@gmail.com

The paper introduces simple yet efficient methods designed for following problems in content-based image retrieval: near-duplicate detection, search by image query and automatic annotation. Image is modeled as a vector of Gabor filter responses to detect near-duplicates and provide textual annotations. Search is implemented on top of fuzzy color histograms built in CIE $L^*a^*b^*$ color space. Proposed methods are easy to implement and do not require much computational resources. At the same time experimental evaluation shows that these methods allow quality of results comparable to alternative solutions and in several cases allow better quality.