

УДК 681.325

Н.Н. Пономаренко, С.К. Абрамов, В.В. Лукин, А.С. Царан

Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков

ОПТИМИЗАЦИЯ ВЕСОВ МНОГОПАРАМЕТРОВОЙ МЕРЫ ПОДОБИЯ ДЛЯ ПОИСКА ИЗОБРАЖЕНИЙ

Статья посвящена задаче поиска изображений в Интернете и больших базах изображений. Описаны эксперименты по получению усредненных субъективных оценок подобия изображений, процедуры формирования банка элементарных расстояний между изображениями и оптимизации их весовых коэффициентов. Проанализирована эффективность полученной меры подобия, представляющей собой взвешенную сумму элементарных расстояний, приведены иллюстрации ее использования в практических ситуациях поиска.

меры подобия, поиск и индексация изображений в больших базах данных

Введение

Постановка проблемы. Задача поиска изображений в Интернете¹ по полному и частичному подобию становится все более актуальной вследствие недостаточной эффективности поиска изображений по ключевым словам. Как правило, по словарным запросам поисковая система вместе с нужными изображениями выдает пользователю очень много “мусора”. Одновременно значительная часть необходимых пользователю изображений может остаться за пределами поиска. Поэтому целесообразной является разработка эффективных методов поиска с комплексным использованием различных признаков.

Анализ литературы. Большинство описанных в научной литературе систем поиска изображений базируется на мерах подобия, представляющих собой комбинации каких-то элементарных расстояний между изображениями [1, 2]. В качестве простого примера такого расстояния можно привести разницу между средними уровнями яркости двух изображений. В целом же существует огромное число признаков изображений, использующихся в вычислении расстояний, которые можно условно разбить на три основных класса: признаки формы, цветовые и текстурные признаки [2 – 4].

Если рассматривать меру подобия как взвешенную сумму какого-то набора элементарных расстояний, то веса этих расстояний либо задаются экспертами, основываясь на своем опыте и каких-то разумных предположениях [5], или же вычисляются в ходе оптимизации [6]. В [6] нами пред-

лагается использовать для оптимизации весов расстояний усредненные субъективные оценки подобия изображений, полученные в ходе экспериментов. При этом одним из важных выводов, сделанных в [6], является целесообразность вычисления меры подобия изображений только на основании их пиктограмм (уменьшенных копий изображений), хранящихся в базе данных поисковой системы.

Цель исследования. В данной работе мы предлагаем метод экспериментальной оценки степени подобия изображений, а также процедуру оптимизации весов меры подобия на основе полученных экспериментальных данных. Раздел 1 посвящен описанию методики проведения экспериментов. В разделе 2 приводятся параметры банка расстояний и принципы его формирования. Раздел 3 посвящен оптимизации весов расстояний, а раздел 4 - анализу полученных результатов.

1. Описание экспериментов по получению усредненных субъективных оценок подобия изображений

В [6] нами был описан итеративный метод оптимизации весов многопараметровой меры подобия. Согласно этому методу в [6] была предложена упрощенная мера подобия, с помощью которой по большой базе изображений из 2,5 миллионов пиктограмм, предоставленной для проведения исследований поисковой системой Яндекс, были сформированы тестовые наборы изображений. При этом каждый набор представлял собой одно эталонное изображение и несколько изображений, в разной степени подобных эталону. Всего на первом этапе было сформировано 5000 тестовых наборов по 20 изо-

¹ Работа подготовлена при поддержке ООО "ЯНДЕКС" (www.yandex.ru)

бражений в каждом. После этого экспертом для проведения экспериментов было отобрано 500 тестовых наборов, в каждом из которых им было оставлено по 8 изображений.

Опишем проведенные эксперименты. Целью этих экспериментов являлось получение усредненных для большого числа людей оценок визуального подобия пар изображений, которые в дальнейшем могли бы использоваться для оптимизации весов меры подобия (рис. 1).

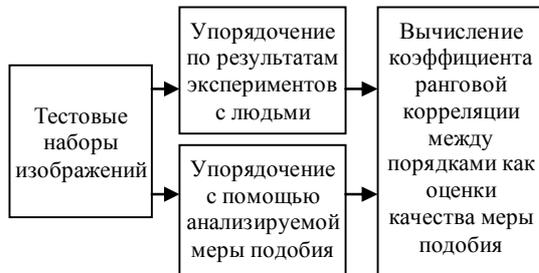


Рис. 1. Использование результатов экспериментов для оценки качества меры подобия

Всего в экспериментах участвовало 25 человек. Каждый участник должен был отсортировать каждый из 500 тестовых наборов по степени подобия эталону. На рис. 2 приведен пример окна программы проведения эксперимента.

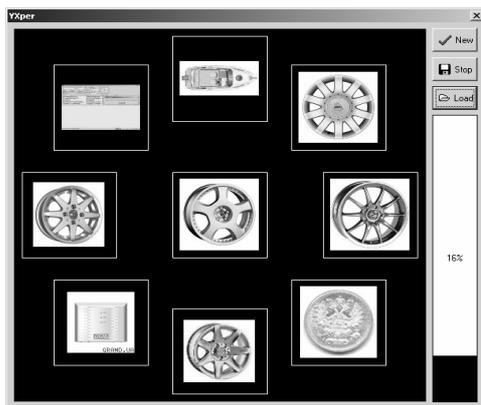


Рис. 2. Программа проведения экспериментов

В центре экрана (рис. 2) находилось эталонное изображение, а по его периметру – 8 изображений,



Рис. 4. Примеры упорядоченных по результатам экспериментов тестовых наборов изображений

которые участник эксперимента должен был упорядочить по степени подобия эталону. Для этого он должен был указать мышкой самое похожее на эталон изображение, затем – самое похожее из оставшихся, и т.д. После нажатия на изображение оно убиралось с экрана.

На рис. 3 приведено окно программы эксперимента для этого же тестового набора после выбора участником эксперимента трех наиболее подобных эталону изображений.

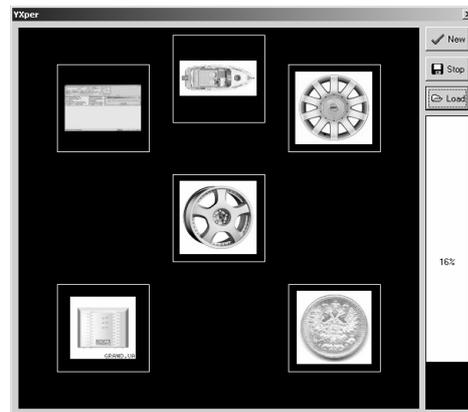


Рис. 3. Результат выбора участником эксперимента трех наиболее похожих на эталон изображений

Следует отметить, что в каждом эксперименте восемь тестовых изображений располагались вокруг эталона всегда в случайном порядке, чтобы исключить возможные зависимости порядка их выбора от порядка расположения на экране.

Каждый эксперимент длился от полутора до двух часов. 25 участников экспериментов упорядочили по 500 наборов из 8 изображений, осуществив, таким образом, около 100000 оценок визуального подобия изображений.

Полученные порядки изображений в наборах были усреднены по всем участникам экспериментов. На рис. 4 приведено три примера таких упорядоченных в усредненном порядке тестовых наборов изображений. Крайнее изображение слева - эталон, а остальные располагаются в порядке убывания визуального подобия эталону.

Так как результаты экспериментов представляют собой порядки изображений в соответствующих тестовых наборах, то для нахождения корреляции между результатами отдельных участников экспериментов нами использовались ранговые коэффициенты корреляции Спирмэна и Кэндалла [7]. В табл. 1 и 2 приведены значения этих коэффициентов корреляции, как между данными разных участников, так и между данными одного и того же участника (для оценки повторяемости результатов три человека выполнили эксперимент дважды).

Таблица 1

Значения корреляции Спирмэна

| Корреляция | Мин. | Макс. | Среднее |
|--|------|-------|---------|
| Между результатами отдельного участника и усредненными | 0,57 | 0,83 | 0,75 |
| Между результатами одного человека (повторяемость) | 0,58 | 0,79 | 0,70 |
| Между результатами разных участников экспериментов | 0,53 | 0,73 | 0,63 |

Таблица 2

Значения корреляции Кэндалла

| Корреляция | Мин. | Макс. | Среднее |
|--|------|-------|---------|
| Между результатами отдельного участника и усредненными | 0,44 | 0,72 | 0,63 |
| Между результатами одного человека (повторяемость) | 0,46 | 0,67 | 0,57 |
| Между результатами разных участников экспериментов | 0,37 | 0,62 | 0,51 |

Данные табл. 1 и 2 свидетельствуют о высоком качестве использованных в экспериментах тестовых наборов изображений. Действительно, если бы эти наборы были слишком простыми, то корреляция между отдельными людьми была бы близка к единице, как и повторяемость результатов. В то же время, если бы наборы были слишком сложными, то есть изображения в них были бы в одинаковой степени подобны эталону, то корреляция между отдельными людьми была бы близка к нулю, как и повторяемость результатов.

2. Описание банка расстояний

В банк расстояний нами были включены самые различные расстояния между параметрами (признаками) изображений, характеризующими их форму, цвет, текстуру, гистограммы и пр., в том числе и предложенные нами в [6, 8]. Всего в банк, параметры которого приведены в табл. 3, вошло 309 рас-

стояний. Крестик в таблице означает, что данный тип параметров изображений учитывается в данной группе расстояний.

Таблица 3

Параметры банка расстояний

| № группы расстояний | Общее число расстояний | Тип учитываемых параметров | | | | | | | Общая информативность |
|---------------------|------------------------|----------------------------|------------------|----------|-------|---------|----------|-----------|-----------------------|
| | | Статистические | Пространственные | Цветовые | Формы | Яркости | Размеров | Контраста | |
| 1 | 87 | + | + | + | + | + | | | 0,2100 |
| 2 | 1 | + | + | | + | + | | + | 0,0419 |
| 3 | 95 | + | + | + | | + | | | 0,2576 |
| 4 | 1 | + | + | | + | + | | | 0,0425 |
| 5 | 4 | + | | + | | + | | + | 0,0304 |
| 6 | 111 | + | | + | | + | | | 0,2720 |
| 7 | 6 | + | | + | | | | | 0,0661 |
| 8 | 1 | + | | | | + | | | 0,0239 |
| 9 | 1 | + | | | | | | + | 0,0022 |
| 10 | 2 | | | | | | + | | 0,0535 |

Для последующего выбора стартовой точки для оптимизации весов расстояний важно, чтобы их масштабы были соотносимы между собой. Выполнения этого условия мы достигали следующим образом. Для одного миллиона случайно выбранных пар пиктограмм для каждого расстояния d_i вычислялись и запоминались значения среднего \bar{d}_i и дисперсия этих значений σ_i^2 . В дальнейшем, при вычислении значения этого расстояния оно нормировалось на σ_i .

Информативность групп расстояний, приведенная в табл. 3, оценивалась по результатам оптимизации, описанной в следующем разделе, как сумма весов нормированных расстояний, входящих в эти группы.

Каждое из расстояний банка было включено в интегральную меру подобия в двух вариантах (таким образом, в нее вошло 618 расстояний). Во втором варианте каждое расстояние d_i было заменено на $\min(d_i, \bar{d}_i/10)$. Такое ограничение по максимальному значению придало разрабатываемой мере подобия некоторую нелинейность. Веса для обоих вариантов каждого расстояния оптимизировались независимо друг от друга.

3. Процедура оптимизации весов расстояний

В нашем случае оптимизация заключалась в нахождении весов расстояний, при которых значение корреляции Спирмэна между усредненными данными экспериментов и порядками изображений, полученными с помощью оптимизируемой меры подобия, было бы как можно большим (рис. 1). При этом веса ограничивались только положительными

значениями, чтобы исключить ситуации, когда мера подобия могла бы оценивать отличие двух изображений отрицательным числом (значение меры подобия должно равняться нулю для полностью идентичных изображений).

Поскольку число оптимизируемых весов велико, многократный расчет целевой функции занимает существенное (даже для современных вычислительных средств) время. В связи с этим для нахождения максимума целевой функции целесообразным оказывается применение метода координатного подъема, при котором оптимизация ведется раздельно по каждому весовому коэффициенту одним из методов одномерной оптимизации. Такой подход позволяет существенно упростить процедуру оптимизации и сократить время расчета целевой функции за счет сокращения количества необходимых операций при изменении значения только одной переменной.

Использовались два метода оптимизации: простой метод перебора с последовательным уменьшением шага и более сложный метод, учитывающий возможную многомодальность целевой функции. Остановимся на нем более подробно.

Большинство методов оптимизации [9] рассчитаны на поиск лишь одного экстремума в заданной области. Поэтому для поиска глобального максимума необходимо вначале провести поиск областей локализации локальных максимумов, далее в каждой из этих областей найти их точное положение и величину, а затем среди найденных локальных максимумов определить глобальный.

Для решения первой задачи предлагается использовать метод кусочно-линейной аппроксимации [9], а для решения второй – метод квадратичной интерполяции-экстраполяции [10]. Третья задача решается простым сравнением величин локальных максимумов.

Использование этих методов в паре позволяет с одной стороны обеспечить высокую скорость вычислений из-за относительной простоты этих методов. С другой стороны – достаточную устойчивость при нахождении глобального максимума, поскольку даже при некоторой ошибке линейной аппроксимации и неправильном определении области локализации используемый метод квадратичной интерполяции-экстраполяции позволяет находить решение даже вне заданной области.

В результате оптимизации весов коэффициент корреляции Спирмэна удалось повысить с 0,06 для упрощенной меры подобия [6] до 0,6 для итоговой меры подобия, а коэффициент Кендалла – с 0,04 до 0,49. При этом анализ различных методов оптимизации показал, что, во-первых, целевая функция имеет много локальных максимумов, и, во-вторых, результат оптимизации мало зависит от используемого метода оптимизации. Так, для табличного ме-

тода удается достичь корреляции Спирмэна, равной 0,602, в то время как для более сложного метода оптимизации это значение больше лишь на 0,002.

В то же время результат оптимизации в существенной степени определяется суммарной информативностью расстояний в банке расстояний. На рис. 5 приведены графики оптимизации весов двух вариантов меры подобия. В первую вошли все расстояния из банка расстояний (это собственно наша итоговая мера подобия), а во вторую вошли все расстояния, за исключением расстояний между параметрами, характеризующими форму изображений.

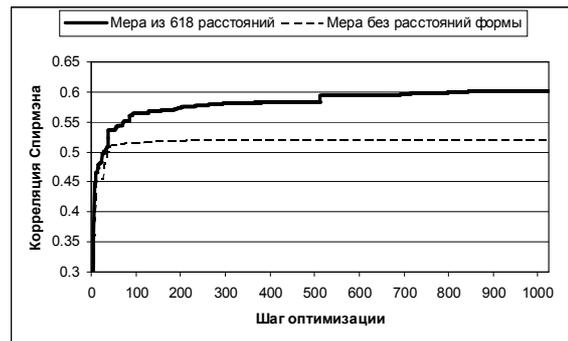


Рис.5. Результаты оптимизации с учетом и без учета расстояний, характеризующих форму изображений

Как видим, мера подобия без расстояний формы обеспечивает корреляцию с восприятием человека лишь на уровне 0,52. Также на графике можно наблюдать несколько локальных максимумов, из которых с трудом выбиралась процедура оптимизации.

4. Анализ результатов и примеры поиска изображений

На основании анализа полученных данных можно сделать два основных вывода. Во-первых, достигнуто достаточно хорошее соответствие итоговой меры подобия восприятию человека, так как корреляция 0,6 превышает минимальную границу для человека 0,57 (см. табл. 1). Второй же вывод состоит в том, что качество меры подобия напрямую зависит от состава банка расстояний и почти не зависит от способа нахождения весов этих расстояний. Отметим, что изображения на рис. 6, как и на рис. 7, цветные, хотя представлены здесь в оттенках серого цвета.

Приведем два примера поиска изображений с помощью полученной меры подобия. На рис. 6 приведен пример простого поиска, когда по изображению в плохом разрешении (в данном случае 100×80 пикселей) ищутся его варианты в лучшем разрешении.

В данном случае удалось не только найти несколько вариантов данного изображения (на рис. 6 представлены не все из них), среди которых было

изображение с размерами 640 × 547 пикселей. Было найдено и изображение-прототип (два последних изображения на рис. 6), из которого, видимо, было получено изображение-образец.

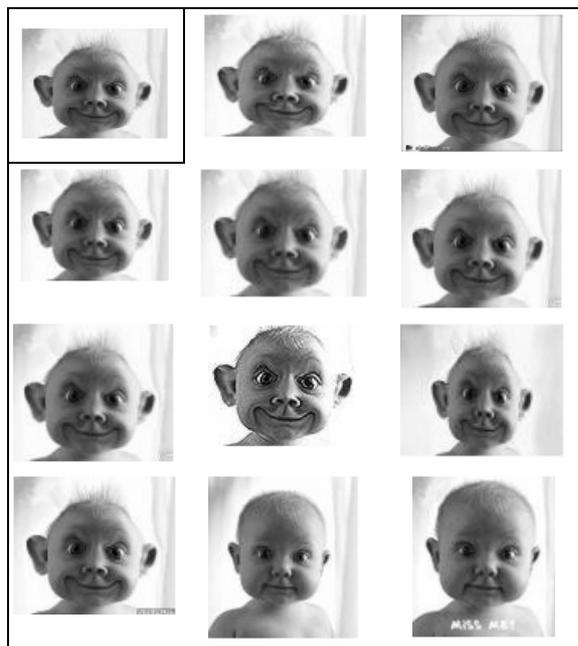


Рис. 6. Пример результатов поиска (образец для поиска выделен рамкой)



Рис. 7. Пример результатов поиска изображений, подобных по стилю (образец для поиска выделен рамкой)

На рис. 7 приведен результат более сложного поиска, иллюстрирующий эффективность используемого подхода в ситуациях, когда не удается найти изображения с хорошим подобием по форме, и приходится учитывать более тонкие подобию в локальных статистических характеристиках изображений.

В данном случае по заданному изображению иконы найдены изображения других икон, хотя формы объектов на всех изображениях сильно отличаются друг от друга.

Список литературы

1. Santini S. *Exploratory Image Databases*. – Academic Press, 2001. – 512 p.
2. Rui Y., Huang T.S., Chang S.F. *Image retrieval: current techniques, promising directions and open issues // Journal of visual communication and image representation*. – 1999. – No 10. – P. 39-62.
3. Aslandogan Y., Yu C.T. *Techniques and Systems for Image and Video Retrieval // IEEE Transactions on Knowledge and Data Engineering*. – 1999. – Vol. 11, No 1. – P. 56-63.
4. Veltkamp R.C. *Shape Matching: Similarity Measures and Algorithms // Proceedings of the International Conference on Shape Modeling and Applications*. – Genova, Italy, 2001. – P. 188-197.
5. Bach J.R., Fuller C., Gupta A., Hampapur A., Horowitz B., Humphrey R., Jam R., Shu C.F. *The Vi-rage Search Engine: An Open Framework for Image Management // Proceedings of the SPIE Storage and Retrieval for Still Image and Video Databases IV*. – 1996. – Vol. 2670. – P. 77-87.
6. Пономаренко Н.Н., Лукин В.В., Абрамов С.К., Зеленский А.А. *Меры подобия изображений для поисковых систем // Радиоэлектронные и компьютерные системы*. – 2007. – № 2 (21). – С. 49-58.
7. Kruskal W.H. *Ordinal measures of association // Journal Amer. Statist. Assoc.* – 1958. – Vol. 53. – P. 814-861.
8. Абрамов С.К., Лукин В.В., Пономаренко Н.Н. *Мера содержания фона на основе энтропии для поиска и сортировки изображений в базах данных // Радиоэлектронные и компьютерные системы*. – 2007. – № 2 (21). – С. 24-28.
9. Батищев Д.И. *Методы оптимального проектирования: Учебн. пособие для вузов*. – М.: Радио и связь, 1984. – 248 с.
10. Щуп Т. *Решение инженерных задач на ЭВМ: Пер. с англ.* – М.: Мир, 1982. – 238 с.

Поступила в редколлегию 1.06.2007

Рецензент: д-р техн. наук, проф. В.К. Волосюк, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков.