

Устойчивый поиск изображений по полному и тематическому подобию с использованием многопараметровой классификации

Н.Н. Пономаренко

В.В. Лукин

С.К. Абрамов

Национальный аэрокосмический университет им. Н.Е. Жуковского "ХАИ", Харьков, Украина

Аннотация

Исследуется задача поиска изображений по подобию в больших базах изображений с комплексным использованием различных признаков. Предложен итеративный подход определения весов признаков меры подобия, представляющей собой взвешенную сумму большого числа расстояний между признаками изображений. Проведены эксперименты по получению усредненных субъективных оценок подобия изображений (всего в них было осуществлено 100 тысяч оценок подобия), использованных в дальнейшем для оптимизации весов меры подобия. Сформирован банк расстояний между признаками изображений (более 600 расстояний), которые предложено вычислять только по пиктограммам (уменьшенным копиям) изображений. Предложен ряд параметров для сортировки изображений. Получены оптимизированные веса меры подобия, высокая эффективность которой продемонстрирована на примерах поиска и в ходе анализа коэффициентов ранговых корреляций с экспериментальными данными.

1. Введение и обзор работ по исследуемой тематике

Количество изображений, размещаемых в Интернете и индексируемых поисковыми системами, растет в геометрической прогрессии. В базе данных поисковой системы Яндекс на конец 2005 года было около 60 млн. изображений, а на конец 2006 года их число достигло уже 133 млн. Все больше людей пользуются поиском изображений в Интернете для оформления своих презентаций, публикаций и просто для удовлетворения любопытства. При этом классифицированной является лишь незначительная часть этих изображений (в основном, это касается платных баз изображений). Для остальных же доступен только поиск по ключевым словам в сопутствующем тексте. Недостаточная эффективность такого подхода [32,33] делает актуальной разработку альтернативных методов поиска изображений, в частности, поиска по полному или частичному подобию.

Задачи разработки мер визуального подобия изображений и тесно связанные с ними задачи создания адекватных математических моделей психо-визуального восприятия человека в последнее время являются областью интенсивных исследований [14,33,36]. При этом все разрабатываемые меры подобия можно условно разбить на меры "близкого" подобия и меры "частичного" подобия. К мерам "близкого" подобия

относятся меры, предназначенные, в основном, для верификации различных методов обработки изображений и видео - сжатия, фильтрации, восстановления и др. [13,31,36]. К мерам "частичного" подобия относятся более сложные меры, предназначенные для поиска изображений и их индексации в больших базах изображений [8,32,33]. Основной особенностью мер "частичного" подобия является, как правило, высокая степень различия между изображениями, которые могут иметь неодинаковые размеры, пропорции, цветовую гамму, контраст и другие характеристики.

Возможные ситуации поиска можно классифицировать по их сложности и учитываемым типам подобия на три больших класса (см. рис. 1).

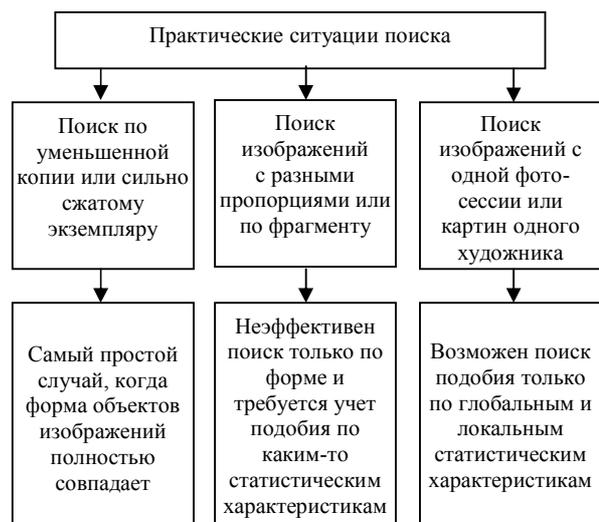


Рис. 1. Классификация ситуаций поиска по учитываемым типам подобия

Основными требованиями к системам поиска изображений являются: 1) высокая скорость вычисления меры подобия; 2) простота модификации системы поиска; 3) учет особенностей психовизуального восприятия человека [38]; 4) робастность к небольшим изменениям пропорций, яркости, контраста и др.

Все современные наиболее развитые системы поиска изображений основаны на каком-либо объединении в одной мере подобия большого количества учитываемых признаков [8,28,32,33] изображений (см. рис. 2), основными классами которых являются: признаки формы (наиболее проблемный) [19,35], цветовые признаки [22,33] и текстурные признаки [12,37].



Рис. 2. Обобщенная структурная схема вычисления меры подобия для поисковой системы

Соответственно системы поиска изображений можно классифицировать по: 1) способу объединения расстояний между различными признаками в одну меру подобия; 2) на обучаемые (с обратной связью) и не обучаемые. Чаще всего встречается построение меры подобия изображений как взвешенной суммы отдельных элементарных расстояний между признаками изображений [9,32,33], хотя иногда используются и различные нелинейные комбинации расстояний [15], и даже нейроструктуры [26,27]. Веса расстояний могут задаваться экспертами [9] либо вычисляться в ходе обучения. При этом используются самые различные методы обучения, от генетических алгоритмов [34] до самоорганизующихся карт [20]. Часто обучение продолжается и на этапе обработки поисковых запросов пользователя и их итеративного уточнения [17,20,39].

В данной работе мы рассмотрим построение меры подобия как взвешенной суммы большого числа расстояний с оптимизацией их весов по результатам экспериментов. Такая мера подобия, кроме простоты вычисления, обладает простотой модификации (наращиваемости), а также позволяет осуществлять поиск по частичному подобию (для этого достаточно обнулить часть весовых коэффициентов). Кроме того, на ее основе достаточно просто может быть осуществлена древовидная кластеризация пространства признаков, что важно для построения пригодной к практическому использованию системы поиска изображений.

Работа организована следующим образом. В разделе 2 описывается предлагаемый подход к построению меры подобия. Раздел 3 посвящен описанию упрощенной меры подобия для формирования тестовых наборов изображений, а раздел 4 - экспериментам по получению усредненных субъективных оценок подобия изображений. В разделе 5 описывается разработанный банк расстояний, а в разделе 6 - оптимизация их весов. В разделе 7 приводятся некоторые примеры поиска с использованием полученной меры подобия и осуществляется их анализ. В разделе 8 предлагается несколько параметров для сортировки найденных изображений. И, наконец, в разделе 9 обсуждаются аспекты построения практической системы поиска изображений на основании предложенного подхода.

2. Идея исследования

Прежде, чем описать предлагаемый метод построения меры подобия, вернемся к ограничениям, накладываемым на пригодную к практическому использованию систему поиска. При добавлении изображения в базу данных поисковой системой сохраняется его уменьшенная копия (thumbnail, пиктограмма) с размерами от 100x100 (поисковая система Яндекс) до 128x128 (поисковая система Google) пикселей. В формате JPEG каждая такая пиктограмма занимает примерно 2-3 килобайта памяти.

Очевидно, что для обеспечения высокой скорости поиска все признаки каждого изображения должны предварительно рассчитываться и храниться в этой же базе данных вместе с его пиктограммой. При этом разумным представляется выделение объема памяти для хранения вектора признаков изображения меньшего, чем требуется для хранения пиктограммы, например, не более 512 байт на изображение.

Для обеспечения простоты внедрения и низкой стоимости сопровождения (наращиваемости) разрабатываемой системы поиска изображений все признаки, используемые при вычислении меры подобия, желательно рассчитывать только на основании сохраненных в базе данных пиктограмм изображений. В противном случае при любом изменении алгоритма вычисления меры подобия может потребоваться скачивание оригинальных версий всех индексированных в базе данных изображений. При среднем размере изображения 110 Кб (данные Яндекс) и общем числе изображений в сотни миллионов, объем скачиваемой информации может составить десятки терабайт, что неприемлемо из-за больших временных затрат на скачивание и обработку данных.

Таким образом, при разработке нашей меры подобия, мы будем вычислять все признаки изображений только по их пиктограммам с размером банка признаков не более 512 байт. При этом на этапе разработки поисковой системы банк признаков может содержать большее число признаков, а по окончании ее разработки в нем могут быть оставлены только наиболее информативные из них.

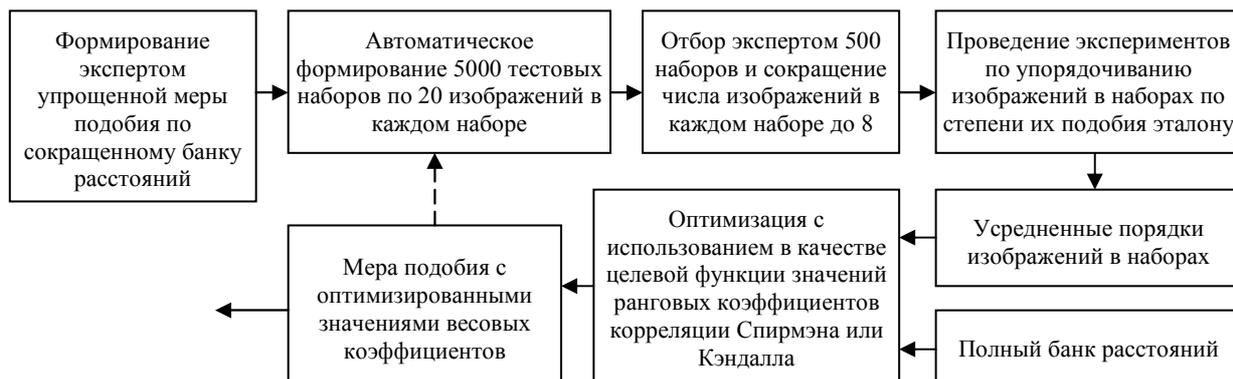


Рис. 4. Структурная схема предлагаемого метода итеративной оптимизации весов меры подобия

Пусть $M = \sum_{n=1}^N a_n d_n$ - мера подобия, вычисляемая как

взвешенная сумма n расстояний, где d_n - значение n -го расстояния, а a_n - его вес. Наиболее сложной частью разработки такой меры подобия, особенно, когда N - большое число, является нахождение (оптимизация) весов $\{a_n\}$.

Формализовать и автоматизировать процесс оптимизации $\{a_n\}$ можно введением в процесс разработки экспериментальной части по субъективной оценке степени подобия изображений. В ходе этих экспериментов для большого количества пар изображений (3000-5000) должны быть получены оценки визуального подобия $\{S_k\}$, где k - номер пары изображений, причем каждая такая оценка должна быть получена усреднением субъективных оценок большого числа участников экспериментов (20-25 человек). Таким образом, для хорошей статистической достоверности должно быть произведено не менее 100000 оценок субъективного подобия пар изображений. Данные, полученные в ходе экспериментов, в дальнейшем могут использоваться для вычисления целевой функции при оптимизации весов меры подобия, как это показано на рис. 3.

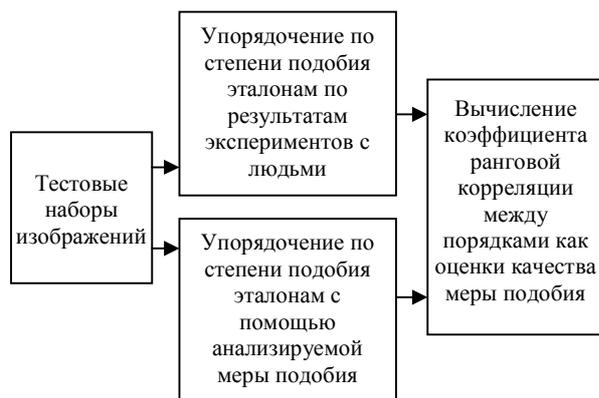


Рис. 3. Использование результатов экспериментов для оценки качества меры подобия

При таком подходе качество оптимизации весов меры подобия напрямую зависит от состава сформированных тестовых наборов изображений. Каждый тестовый набор (одно эталонное изображение и 8-10 изображений в различной

степени подобных эталону) должен содержать изображения как с большим подобием эталону, так и с низким подобием эталону, причем желательно, чтобы для определения подобия участнику эксперимента приходилось учитывать все группы признаков. Если учесть большое число необходимых тестовых наборов (как минимум, несколько сотен), то их формирование должно осуществляться в автоматическом или полуавтоматическом режиме с использованием какой-то меры подобия. То есть для разработки эффективной меры подобия нам требуется уже имеющаяся в наличии достаточно эффективная мера подобия. Получается замкнутый круг.

В качестве одного из приемлемых решений этой проблемы мы предлагаем итеративный метод оптимизации весов меры подобия [4] с полуавтоматическим формированием тестовых наборов, иллюстрируемый схемой на рис. 4.

Эксперт по небольшому банку расстояний (5-10 расстояний) формирует упрощенную меру подобия, которая и используется для формирования тестовых наборов изображений. При этом веса расстояний выбираются экспертом, исходя из его опыта. Такая мера подобия будет далека от совершенства, поэтому автоматически формируются 5000 наборов вместо требуемых 500, с 20-ю тестовыми изображениями в каждом наборе вместо требуемых 8. Затем эксперт отбирает 500 наиболее информативных на его взгляд наборов, оставляя в каждом по 8 изображений из 20, частично компенсируя таким образом недостатки упрощенной меры подобия.

Пунктирная линия на схеме означает то, что весь процесс оптимизации можно повторять итеративно, на каждой последующей итерации используя для формирования тестовых наборов полученную на предыдущей итерации меру подобия. Однако нет никакой гарантии сходимости процесса, так как многое зависит от стартовой точки - упрощенной меры подобия, сформированной экспертом.

Помочь эксперту в процессе формирования упрощенной меры подобия могли бы оценки взаимного масштаба, избыточности и, особенно, шумности расстояний. Если их избыточность и масштаб можно эффективно оценить, имея большую базу изображений (а Яндекс для

проведения исследований предоставил нам фрагмент своей базы данных из 2,5 миллионов пиктограмм), то оценка шумности расстояний является сложной задачей. Особенно это касается расстояний между признаками формы изображений, что иллюстрирует пример на рис. 5, взятый из [35].

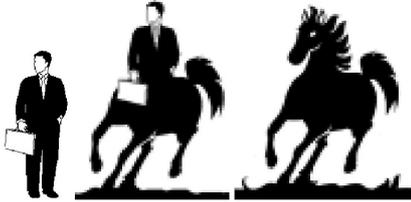


Рис. 5. Иллюстрация к сложности оценки шумности признаков формы

Одним из критериев высокой шумности признаков и расстояний может служить ситуация, когда для трех признаков П1, П2 и П3 расстояние между признаками П1 и П2 невелико и равно расстоянию между признаками П2 и П3, в то время как расстояние между признаками П1 и П3 очень велико. Однако для признаков формы этот критерий неэффективен. Изображение человека на рис. 6 по форме достаточно сильно подобно изображению кентавра. Изображение кентавра по форме, в свою очередь, достаточно сильно подобно изображению коня. В то же время изображения человека и коня абсолютно различаются по форме, что, однако, не свидетельствует о высокой шумности признаков формы.

3. Упрощенная мера подобия

В соответствии с вышеописанным методом нами была разработана упрощенная мера подобия. В нее были включены расстояния, характеризующие отличия изображений по форме, контрасту, цвету, яркости, гистограммам, пропорциям (мы также учли тот факт, что сравниваемые изображения могут быть фрагментами друг друга или третьего изображения), детализированности, процентному содержанию фона и объектов с определенными характеристиками. Опишем эти расстояния подробнее.

3.1. Отличие по форме, контрасту и яркости

Для вычисления робастных к изменениям яркости и контраста расстояний между двумя одинаковыми по размерам двумерными массивами (черно-белыми изображениями) будем использовать следующее выражение:

$$D_{rob}(A, B) = \sqrt{\frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M (A_{ij} - \bar{A} - (B_{ij} - \bar{B})K)^2}, \quad (1)$$

где A - эталонное изображение, B - изображение, для которого вычисляется расстояние, N, M - размеры изображения, \bar{A} , \bar{B} - средние значения яркости на A и B , K - масштабный коэффициент:

$$K = \max \left(\frac{\sum_{i=1}^N \sum_{j=1}^M (A_{ij} - \bar{A})(B_{ij} - \bar{B})}{\sum_{i=1}^N \sum_{j=1}^M (A_{ij} - \bar{A})^2}, 0 \right). \quad (2)$$

Использование K , вычисленного в соответствии с (2), минимизирует D_{rob} для A и B [29]. Ограничение K положительными числами исключает из поиска инвертированные изображения и существенно уменьшает вероятность ошибок при нахождении подобий. Если $K=0$, B отличается от A больше, чем изображение, залитое одним уровнем яркости.

K характеризует разницу в контрастах изображений, которую можно вычислять как

$$\Delta_k = \begin{cases} 1-K, & K < 1 \\ 1-1/K, & K > 1 \end{cases}. \quad (3)$$

Для вычисления расстояний между формой изображений D_S , контрастом изображений D_C и яркостью изображений D_B предварительно будем переводить изображения из цветового пространства RGB в YUV [10] и уменьшать до размеров 15x15, сохраняя только компоненту Y. При этом для хранения компоненты Y требуется всего 225 байт, и она представляет собой низкочастотную (наиболее информативную) часть изображения, характеризующая его форму и контраст. Подобным уменьшением достигается и некоторая робастность к отличию в пропорциях изображений (до 10%). Далее для компоненты Y вычисляются следующие расстояния:

$D1$ - изображение B уменьшается таким образом, чтобы быть фрагментом, вписанным в изображение A , после чего для всех возможных положений фрагмента вычисляются расстояния (2) и из них выбирается минимальное;

$D2$ - аналогично $D1$, только до фрагмента изображения B уменьшается изображение A ;

$D3$ и $D4$ - значения Δ_k (3) для положений фрагмента, соответствующих вычислению $D1$ и $D2$;

$D5$ и $D6$ - значения выражения $|\bar{A} - \bar{B}|$ для положений фрагмента, соответствующих вычислению $D1$ и $D2$;

После этого вычисляются значения D_S , D_C и D_B :

$$\begin{aligned} D_S &= \min(D1, D2, 40), \\ D_C &= \min(D3, D4, 1), \\ D_B &= \min(D5, D6, 50). \end{aligned} \quad (4)$$

Поясним, какими соображениями мы руководствовались, вводя константы 40, 1, 50 в вычисление расстояний D_S , D_C и D_B . Когда форма объектов на изображениях отличается не очень сильно, человек пытается оценить степень этого отличия (рис. 6). Если же отличие очень большое, то его оценка становится некоей константой, соответствующей полному несовпадению в субъективном восприятии человека (рис. 7). Назовем ее порогом полного несовпадения (ППН).



Рис. 6. Форма изображения б) более соответствует форме эталонного изображения а), чем форма изображения в), и это отличие можно оценить



Рис. 7. Форма изображений б) и в) одинаково "очень сильно" отличается от изображения а)

3.2. Отличие по цвету изображений

Отличие по цвету будем вычислять в цветовом пространстве xyz [10]. Для каждого изображения будем находить и хранить средние значения x и y (с диапазоном $0...255$). Тогда расстояние между цветом изображений A и B может быть вычислено как:

$$D_L = \min(\sqrt{(x_A - x_B)^2 + (y_A - y_B)^2}, 12), \quad (5)$$

где $\underline{x}_A, y_A, \underline{x}_B, y_B$ - соответственно средние значения x и y для изображений A и B , 12 - ППН.

3.3. Отличие по форме гистограмм

Обозначим как $\{H_A^Y\}, \{H_A^U\}, \{H_A^V\}$ соответственно гистограммы компонент Y, U и V изображения A , где $l=1...L$. Для вычисления интегральных характеристик гистограмм будем использовать $L=256$, а для хранения в базе данных и последующего вычисления расстояний между двумя гистограммами будем использовать $L=16$ (сохраненные нормализованные 8-битные данные по $\{H_A^Y\}, \{H_A^U\}$ и $\{H_A^V\}$ будут занимать всего 48 байт памяти).

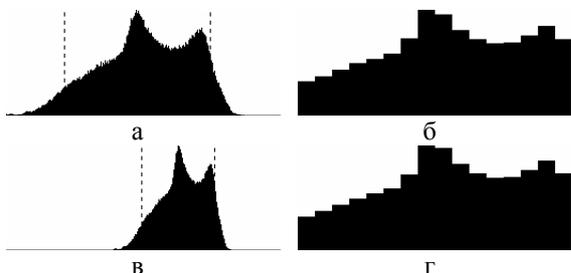


Рис. 8. а) исходная гистограмма (пунктирными линиями обозначены нижняя и верхняя границы), б) сохраненная 16-элементная гистограмма, в) гистограмма этого же изображения с измененными контрастом и яркостью, г) сохраненная 16-элементная гистограмма для гистограммы в)

Для обеспечения робастности по отношению к

изменениям контраста для каждой гистограммы предлагается выполнять следующие действия. Находятся нижняя и верхняя границы (рис. 8, а), отсекающие по 5% от общего числа пикселей. Затем часть гистограммы, находящаяся между этими границами, сжимается (растягивается) до L элементов (рис. 8, б) и нормализуется (с умножением на 255 и округлением для обеспечения 8-битности). Полученный массив (16 байт) сохраняется в базе данных.

При вычислении расстояния между формой гистограмм рассматриваются все их возможные сдвиги друг относительно друга и выбирается сдвиг, обеспечивающий минимальное расстояние. При этом гистограммы по необходимости на краях дополняются нулями.

Расстояние между формой двух гистограмм $D_H(HX, HY)$ будем вычислять как

$$D_H = \sqrt{\frac{1}{L} \sum_{l=1}^{L_0} (HX_l - HY_l)^2},$$

где L_0 - размерность обеих гистограмм после дополнения нулями.

Интегральное расстояние между гистограммами изображений A и B будем вычислять как взвешенную сумму расстояний между формами их гистограмм для компонент Y, U и V :

$$D_{HW} = (4D_H(H_A^Y, H_B^Y) + D_H(H_A^U, H_B^U) + D_H(H_A^V, H_B^V))/6.$$

Расстояние D_H между гистограммами изображений A и B будем вычислять как (90 - ППН):

$$D_H = \min(D_{HW}, 90). \quad (6)$$

3.4. Отличие в пропорциях изображений

Расстояние между пропорциями изображений A и B будем вычислять в соответствии со следующим выражением:

$$D_w = \min(1 - \min(\frac{X_S Y_S}{X_S Y_S}, \frac{X_S Y_S}{X_S Y_S}), 0,35), \quad (7)$$

где X_A, Y_A и X_B, Y_B - размеры изображений A и B , 0,35 - ППН.

3.5. Отличие по степени детализированности

Введем понятие степени детализированности изображения, оцениваемой по его яркостной компоненте Y . Для этого вычислим значение средней локальной дисперсии на изображении σ_{loc}^2 , рассчитываемое в скользящем окне 5×5 пикселей. Получим робастную оценку диапазона значений пикселей $Diap = E_{max} - E_{min} + 1$, где E_{max} и E_{min} - верхняя и нижняя границы гистограммы, отсекающие по 5% общего числа пикселей (см. рис. 9). Оценка детализированности изображения, робастная к изменению контраста, может быть найдена как $Det = \sigma_{loc} / Diap$.

Расстояние между степенями детализированности изображений A и B будем вычислять как (0,09 - ППН):

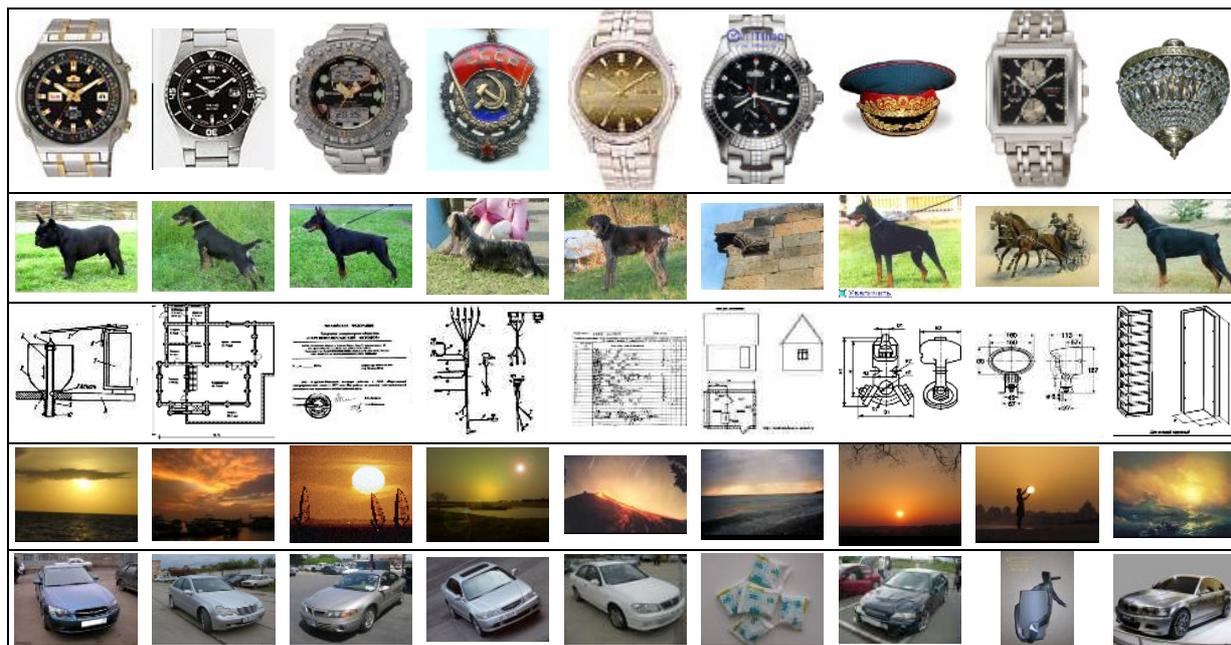


Рис. 9. Примеры сформированных тестовых наборов

$$D_D = \min(|Det_A - Det_B|, 0,09). \quad (8)$$

3.6. Отличие по содержанию фона

Вычислим выход медианного фильтра с окном 5x5 пикселей для компонент Y, U, V изображения и построим трехмерную гистограмму для этого выхода. Перед вычислением гистограммы для повышения скорости вычислений и обеспечения робастности к флуктуациям уровня фона осуществим равномерное квантование значений пикселей с шагом квантования 8. Число уровней трехмерной гистограммы будет равно 32768. Введем в качестве меры процентного содержания фона на изображении $Pr(A)$ отношение значения максимального пика полученной гистограммы к общему числу точек. Тогда расстояние между соответствующими мерами двух изображений может быть вычислено как (0,11 - ППН):

$$D_P = \min(|Pr(A) - Pr(B)|, 0,11). \quad (9)$$

3.7. Отличие по процентному содержанию объектов определенного цвета

Так как большую долю изображений в Интернете занимают фотографии людей, нам представляется важным включить в вычисление базовой меры подобия параметр, характеризующий процентное содержание на изображении пикселей, цвет которых соответствует цвету человеческой кожи. Вычисление этого параметра удобно осуществлять в цветовом пространстве xyu .

Опытным путем нами были найдены три опорные точки в плоскости xy , приблизительно соответствующие телесным оттенкам цвета: $(x=0,35, y=0,36)$, $(x=0,38, y=0,37)$, $(x=0,42, y=0,41)$. Для каждого изображения A подсчитывается доля числа пикселей $Prop(A)$, расстояние от которых в плоскости xy хотя бы до одной из приведенных

выше опорных точек меньше 0,02. Соответствующее расстояние между изображениями вычисляется как (0,08 - ППН):

$$D_G = \min(|Prop(A) - Prop(B)|, 0,08). \quad (10)$$

3.8. Интегральная мера подобия

Будем вычислять интегральную меру подобия как взвешенную сумму расстояний, описанных в подразделах 3.1-3.7:

$$D_I(A, B) = a_1 D_S + a_2 D_C + a_3 D_B + a_4 D_L + a_5 D_H + a_6 D_W + a_7 D_D + a_8 D_P + a_9 D_G. \quad (11)$$

Результаты предварительных исследований показали, что хорошее соответствие субъективной оценке подобия изображений обеспечивают значения весовых коэффициентов $a_1 \dots a_9$, равные соответственно $a_1=1/100$, $a_2=1/5$, $a_3=1/250$, $a_4=1/100$, $a_5=1/50$, $a_6=1/4$, $a_7=1$, $a_8=1$, $a_9=1/2$.

3.9. Формирование тестовых наборов изображений

Для проведения исследований нам был предоставлен фрагмент базы изображений Яндекс размером в 2,5 миллиона пиктограмм. ППН для каждого расстояния задавался как 50% квантиль гистограммы его распределения, построенной на 1 миллионе случайно выбранных пар изображений [4]. Примеры тестовых наборов, которые были сформированы в соответствии со схемой на рис. 5, приведены на рис. 9.

4. Эксперименты по получению усредненных субъективных оценок подобия изображений

Опишем проведенные эксперименты. Всего в них приняло участие 25 человек. Каждый участник должен был отсортировать каждый из 500 тестовых наборов по степени подобия эталону.

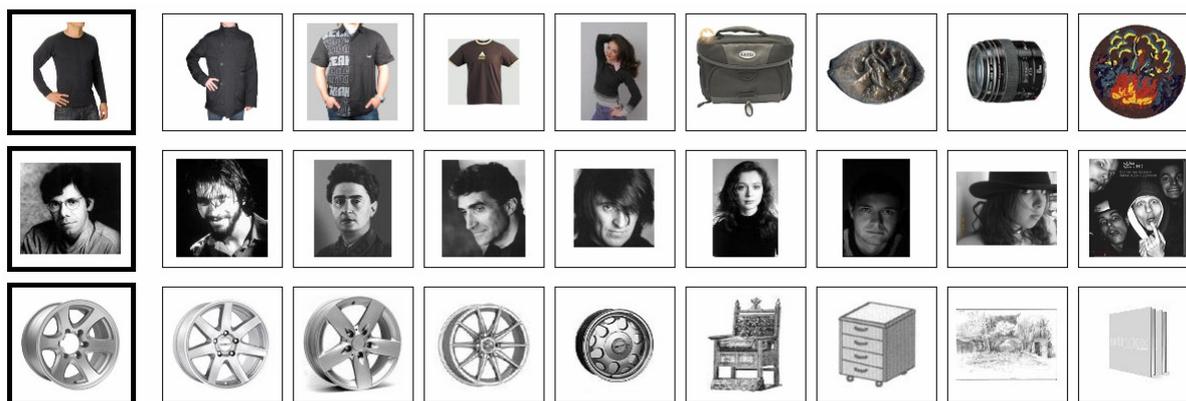
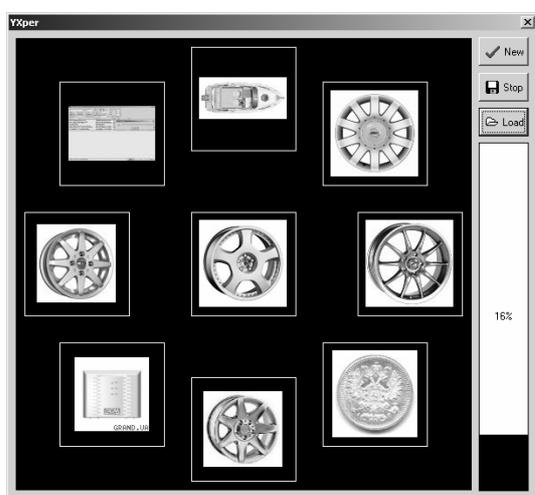
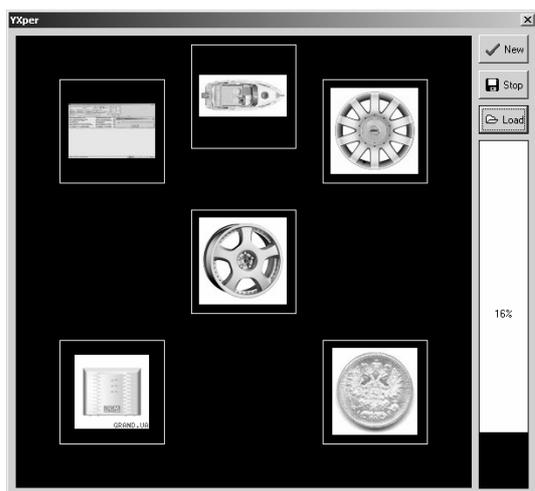


Рис. 11. Примеры упорядоченных по результатам экспериментов тестовых наборов изображений

На рис. 10, а) приведен пример окна программы проведения эксперимента.



а)



б)

Рис. 10. Окно программы проведения экспериментов

В центре экрана (рис. 10, а) находилось эталонное изображение, а по его периметру - 8 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. 10, б) приведено окно программы эксперимента для этого же тестового набора после 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. 11 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 [21]. 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.

5. Описание банка расстояний

В банк расстояний нами были включены самые различные расстояния между признаками изображений (подробности по банку признаков изображений, занимающих по 459 байт на изображение, доступны на [5]), характеризующими их форму, цвет, текстуру, гистограммы и пр., в том числе и предложенные нами в [1,4,6]. Всего в банк, параметры которого приведены в Таблице 3, вошло 309 расстояний (подробности доступны на [5]). 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 1-го 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 $\min(d_i, \bar{d}_i/10)$. Использование такого ППН, ограничения по максимальному значению, придало разрабатываемой мере подобия некоторую нелинейность. Веса для обоих вариантов каждого расстояния оптимизировались независимо друг от друга. Как показал дальнейший анализ, использование ППН оказалось вполне оправданным, так как общая информативность всех расстояний с ППН (с 310-го по 618-е) составила 0,39 (суммарная информативность всех расстояний равна единице).

6. 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 (см. рис. 1) [3]. 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 [2] 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 [2], а для решения второй – метод квадратичной интерполяции-экстраполяции [7]. 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 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 метода 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.

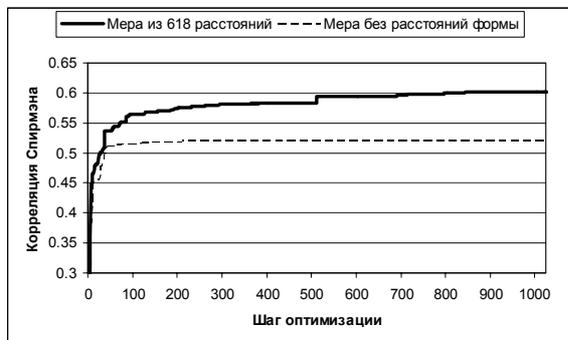


Рис. 12. Графики оптимизации с учетом и без учета расстояний, характеризующих форму изображений

На рис. 12 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.

Как видим, мера подобия без расстояний формы обеспечивает корреляцию с восприятием человека лишь на уровне 0,52.

7. 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). Тем не менее, отметим, что вряд ли когда-либо программа сумеет учесть ряд признаков, которые используются человеком при оценке степени подобия и сортировке изображений (например, признаки пола людей на изображении). 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.

Приведем три примера поиска изображений с помощью полученной меры подобия. На рис. 13 приведен пример простого поиска, когда по изображению в плохом разрешении (в данном случае 100x80 пикселей) ищутся его варианты в лучшем разрешении.



Рис. 13. Пример результатов поиска (образец для поиска выделен рамкой)

В данном случае удалось не только найти несколько вариантов данного изображения (на рис. 14 представлены не все из них), среди которых было изображение с размерами 640x547 пикселей. Было найдено и изображение-прототип (два последних изображения на рис. 13), из которого, видимо, было получено изображение-образец.

На рис. 14 приведен более сложный пример поиска портретных изображений. Интересно, что среди первых 270 найденных изображений (на рис. 15 показано первые 11) в данном случае всего 16 изображений не являлись портретами (ошибка классификации менее 6%). На [5] можно увидеть все 270 изображений.



Рис. 14. Пример поиска портретных изображений (образец выделен рамкой)

И, наконец, на рис. 15 приведен результат самого сложного поиска, иллюстрирующий эффективность используемого подхода в ситуациях, когда не удастся найти изображения с хорошим подобием по форме, и приходится учитывать более «тонкие» подобию в локальных характеристиках изображений. В данном случае по заданному изображению иконы найдены изображения других икон, хотя формы объектов на всех изображениях сильно отличаются друг от друга.



Рис. 15. Пример результатов поиска изображений, подобных по стилю (образец выделен рамкой)

На [5] доступны еще несколько интересных примеров поиска с помощью предложенной меры подобия, иллюстрирующих ее работоспособность в различных практических ситуациях.

8. Параметры для сортировки найденных изображений

Для сортировки найденных изображений в современных поисковых системах используется только признак размера. Так в поисковой системе Google можно выбрать, выводить на экран только большие, средние, маленькие или же все найденные изображения. В то же время сортировать изображения можно и по другим параметрам. Мы предлагаем сортировать их:

- 1) По средней цветовой температуре;
- 2) По цветовой насыщенности;
- 3) По степени детализированности;
- 4) По наличию однородных участков;
- 5) По процентному содержанию точек телесного цвета.

цвета.

На [5] доступны примеры случайных выборок изображений, отсортированных по вышеперечисленным параметрам, а также выборки изображений, обладающих наибольшим и наименьшим значением параметра в имеющейся базе из 2,5 миллионов изображений. Остановимся несколько подробнее на вычислении этих параметров.

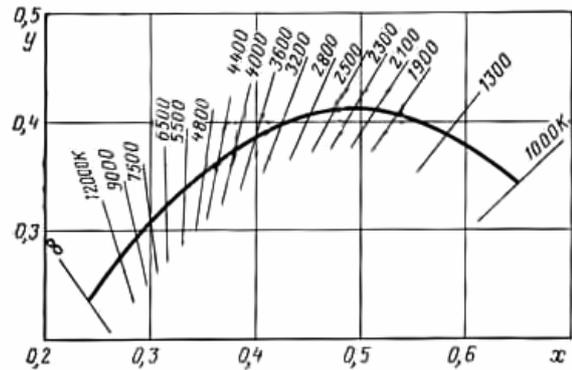


Рис. 16. Вычисление цветовой температуры в цветовом пространстве xuz

На рис. 16 приведен график, по которому вычисляется цветная температура изображения. В цветовом пространстве xuz вычисляются средние значения x и u пикселей изображения, и затем для полученной точки в плоскости xu находится ближайшая точка графика. На рис. 17 приведены примеры изображений с высоким и низким значением цветовой температуры.



Рис. 17. Изображения с теплой (верхняя строка) и холодной (нижняя строка) цветовой температурой

Отметим, что сортировать изображения можно как по значению цветовой температуры изображения, так и по его отклонению от значения цветовой температуры изображения-образца.

В качестве цветовой насыщенности изображения мы предлагаем использовать значение меры энтропии изображения [1], вычисляемой в цветовом пространстве YUV (см. рис. 18).

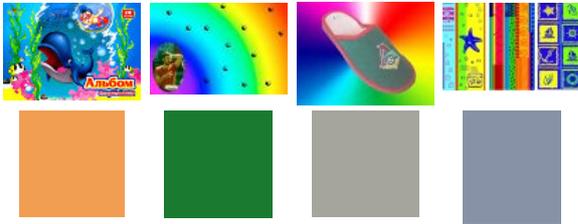


Рис. 18. Изображения с высокой (верхняя строка) и низкой (нижняя строка) цветовой насыщенностью

В качестве степени детализации изображения мы предлагаем использовать значение робастной оценки детализированности Det (см. подраздел 3.5), помноженную на цветовую насыщенность изображения (см. рис. 19).



Рис. 19. Изображения с высокой (верхняя строка) и низкой (нижняя строка) детализацией

Параметр, характеризующий наличие на изображении объекта на однородном фоне, мы предлагаем вычислять как разницу между цветовой насыщенностью в центре изображения и цветовой насыщенностью на краях изображения (при этом изображение делится по площади на центр и края примерно пополам, см. рис. 20).



Рис. 20. Изображения на однородном фоне (верхняя строка) и их противоположности (нижняя строка)

И, наконец, для вычисления процентного содержания точек телесного цвета мы предлагаем использовать признак изображения, описанный в подразделе 3.7 (примеры не приводятся из-за большого числа в них эротических изображений).

Заметим, что предложенные параметры сортировки - лишь небольшая часть параметров, которые могли бы использоваться для упрощения пользователю анализа найденных изображений. Эта задача заслуживает отдельной проработки на этапе создания практической системы поиска изображений.

9. Построение практической системы поиска изображений

В этом разделе мы кратко остановимся на аспектах создания системы поиска изображений по их подобию, пригодной для практического использования.

Очевидно, любая эффективная мера подобия, которая только может быть разработана, не будет пригодна для вычисления в реальном времени степени подобия заданного пользователем изображения и нескольких сотен миллионов изображений, проиндексированных поисковой системой. А ведь популярная система поиска должна иметь возможность обрабатывать сотни таких поисковых запросов одновременно и выдавать результаты поиска с задержкой не более единиц секунд [16,18].

Нам целесообразным представляется использование разработанной меры подобия для предварительной кластеризации [11] пространства признаков, как это показано на рис. 21.

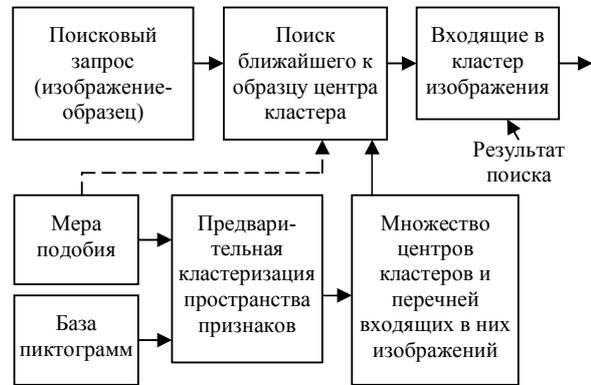


Рис. 21. Использование меры подобия в построении практической системы поиска изображений

При этом поиск подобия при поступлении запроса осуществляется поисковой системой только для центров кластеров, число которых может быть на несколько порядков меньше числа изображений в базе данных.

Наиболее же быстрый поиск может быть обеспечен при использовании кластеризации с построением древовидных структур [23], исследованной в наших работах [24,25,30]. В этом случае поиск ближайшего к образцу центра кластера осуществляется вообще без использования меры подобия и представляет собой последовательность шагов, на каждом из которых происходит сравнение одного из признаков изображения с заданным порогом, и в зависимости

от результата сравнения, происходит переход на одну из ветвей дерева. В этом случае число кластеров может быть очень большим, так как для нахождения ближайшего к образцу кластера требуется число сравнений, приблизительно равное двоичному логарифму от числа кластеров.

10. Выводы

В результате проведенных исследований:

1) Разработана многопараметровая мера подобия, состоящая из 618 элементарных расстояний, и оптимизированы ее веса, что позволило достичь значения ранговой корреляции Спирмэна 0,6 и Кэндалла 0,49 по сравнению с усредненными результатами сортировки людьми наборов изображений по степени их подобия эталонам. При этом нижняя граница этих корреляций между отдельными участниками экспериментов и усредненными данными была соответственно 0,57 и 0,44;

2) Предложено пять новых параметров для сортировки изображений, которые позволяют сортировать изображения, как по значению параметра, так и по его отличию от значения этого параметра у заданного изображения.

Особенностями проведенных исследований и предложенного подхода являются:

1) Все признаки, используемые в вычислении меры подобия, рассчитываются только по пиктограммам изображений;

2) Подход к оптимизации весов меры подобия позволяет использовать неограниченный по размерам банк элементарных расстояний и оценивать их информативность;

3) При определении целевой функции для оптимизации весов меры подобия предложено использовать усредненные данные экспериментов по получению субъективных оценок степени визуального подобия изображений;

4) В ходе экспериментов было осуществлено около 100000 оценок степени визуального подобия изображений;

5) Введено понятие порога предельного несоответствия для вычисления расстояний между признаками изображений, которое позволило придать некоторую нелинейность разработанной мере подобия. Показано, что использование рекомендованных значений ППН позволяет существенно повысить общую информативность банка расстояний;

6) Поисковой системой Яндекс для проведения исследований была предоставлена уникальная база данных, содержащая 2,5 миллионов пиктограмм. Использование этой базы позволило автоматизировать процессы формирования тестовых наборов изображений, а также расчета и анализа статистических характеристик признаков, что позволило оценить их масштабы и избыточность и, таким образом, точнее выбрать стартовые точки для оптимизации их весов.

По результатам исследований было опубликовано 4 научные статьи [1,3,4,6]. Дальнейшие исследования могут быть связаны с расширением банков признаков и расстояний, а также с вопросами автоматизации оценки шумности признаков изображений по большим базам изображений.

11. Литература

- [1] 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. - N2 (21). - С. 24-28.
- [2] Батищев Д.И. Методы оптимального проектирования: Учеб. пособие для вузов. – М.: Радио и связь, 1984. – 248 с.
- [3] Пономаренко Н.Н., Абрамов С.К., Лукин В.В., Царан А.С. Оптимизация весов многопараметровой меры подобия для поиска изображений // принята в журнал Системы обработки информации. - 2007.
- [4] Пономаренко Н.Н., Лукин В.В., Абрамов С.К., Зеленский А.А. Меры подобия изображений для поисковых систем // Радиоэлектронные и компьютерные системы. - 2007. - N2 (21). - С. 49-58.
- [5] Пономаренко Н.Н., Лукин В.В., Абрамов С.К. Информационная страница проекта “Устойчивый поиск изображений по полному и тематическому подобию с использованием многопараметровой классификации”. [Электрон. ресурс] - 2007. - Режим доступа: <http://k504.xai.edu.ua/html/nauka/Yandex/Yandex2007.htm>
- [6] Роечко А.А., Февралев Д.В., Пономаренко Н.Н., Лукин В.В. Применение устойчивых оценок параметров выборок данных при обработке изображений // Восточноевропейский журнал передовых технологий. - 2007. - N3/2 (27). - С. 21-31.
- [7] Щуп Т. Решение инженерных задач на ЭВМ: Пер. с англ. – М.: Мир, 1982. – 238 с.
- [8] 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.
- [9] Bach J.R., Fuller C., Gupta A., Hampapur A., Horowitz B., Humphrey R., Jam R., Shu C.F. The Virage Search Engine: An Open Framework for Image Management // In Proc. of the SPIE Storage and Retrieval for Still Image and Video Databases IV. - 1996. - Vol. 2670. - P. 77-87.
- [10] Berns R.S. Principles of Color Technology. - New York: John Wiley & Sons, 2000. - 272 p.
- [11] Castelli V., Thomasian A., Li C.-S. CSVD: Clustering and singular value decomposition for approximate similarity search in high-dimensional spaces // IEEE Trans. on Knowledge and Data Engineering. - 2003. - Vol. 15. - Issue 3. - P. 671-685.
- [12] Do M.N., Vetterli M. Rotation invariant texture characterization and retrieval using steerable wavelet domain hidden Markov models // IEEE Transactions on Multimedia. - 2002. - Vol. 4. - Issue 4. - P. 517-527.
- [13] Egiazarian K., Astola J., Ponomarenko N., Lukin V., Battisti F., Carli M. New full-reference quality metrics based on HVS // CD-ROM Proceedings of the Second

- International Workshop on Video Processing and Quality Metrics. - Scottsdale, USA, 2006. - 4 p.
- [14] Furht B., Marques O. The handbook of video databases: Design and Applications. - CRC Press, 2003. - 1232 p.
- [15] Grauman K., Darrell T. Efficient image matching with distributions of local invariant features // Proceedings of Computer Vision and Pattern Recognition. - 2005. - Vol. 2. - P. 627-634.
- [16] Guldogan E., Guldogan O., Gabbouj M. Visual media retrieval using transform based layered query scheme // Proceedings of ICIP. - 2005. - Vol. 1. - P. 521-524.
- [17] Guo G.-D., Jain A.K., Ma W.-Y., Zhang H.-J. Learning similarity measure for natural image retrieval with relevance feedback // IEEE Transactions on Neural Networks. - 2002. - Vol. 13. - Issue 4. - P. 811-820.
- [18] Hacid H., Zighed D.A. Content-based retrieval in large image databases // Proc. of International Conference on Granular Computing. - 2006. - P. 498-501.
- [19] Kiraniyaz S., Ferreira M., Gabbouj M. Automatic object extraction over multiscale edge field for multimedia retrieval // IEEE Trans. on Image Processing. - 2006. - Vol. 15. - Issue 12. - P. 3759-3772.
- [20] Koskela M. Interactive Image Retrieval Using Self-Organising Maps // Ph.D. Thesis in Helsinki University of Technology. - 2003.
- [21] Kruskal W.H. Ordinal measures of association // J. Amer. Statist. Assoc. - 1958. - Vol. 53. P. 814-861.
- [22] Ling H., Okada K. Diffusion distance for histogram comparison // Proceedings of Computer Vision and Pattern Recognition. - 2006. - P. 246-253.
- [23] Liu G. Techniques and data structures for efficient multimedia retrieval based on similarity // IEEE Trans. on Multimedia. - 2002. - Vol. 4. - Issue 3. - P. 372-384.
- [24] Lukin V., Ponomarenko N., Zelensky A., Astola J., Egiazarian K. Automatic Design of Locally Adaptive Filters for Pre-processing of Images Subject to Further Interpretation // Proc. of 2006 IEEE Southwest Symp. on Image Analysis and Interpretation. - 2006. - P. 41-45.
- [25] Lukin V.V., Ponomarenko N.N., Zelensky A.A., Astola J.T., Egiazarian K.O. Object classification, segmentation, and parameter estimation in multichannel images by classifier learning with clustering of local parameters // Proceedings of the SPIE conference Image Processing: Algorithms and Systems V. - 2007. - SPIE Vol. 6497. - Art. 64970L.
- [26] Ma W.-Y., Manjunath B.S. Texture features and learning similarity // Proceedings of Computer Vision and Pattern Recognition. -1996. - P. 425-430.
- [27] Mauro C., Gori M., Maggini M. An adaptive visual information retrieval system // Proceedings of International Conference on Document Analysis and Recognition. - 2001. - P. 898-902.
- [28] Missaoui R., Sarifuddin M., Vaillancourt J. Similarity measures for efficient content-based image retrieval // IEE Proceedings on Visual, Image and Signal Processing. - 2005. - Vol. 152. - P. 875-887.
- [29] Ponomarenko N., Egiazarian K., Lukin V. Acceleration of fractal image compression by correlation trees // Interdisciplinary Applications of Fractal and Chaos Theory. - Editura Academiei Romane, Editors R. Dobrescu, C. Valilescu. - Romania, 2004. P. 75-84.
- [30] Ponomarenko N.N., Lukin V.V., Zelensky A.A., Egiazarian K.O., Astola J.T. Locally adaptive image filtering based on learning with clustering // Proc. of Image Proc.: Algorithms and Syst. IV. - 2005. - P. 94-105.
- [31] Ponomarenko N., Silvestri F., Egiazarian K., Carli M., Astola J., Lukin V. On between-coefficient contrast masking of DCT basis functions // CD-ROM Proc. of the Third International Workshop on Video Processing and Quality Metrics. - USA, 2007. - 4 p.
- [32] 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.
- [33] Santini S. Exploratory Image Databases.- Academic Press, 2001. - 512 p.
- [34] Tran K.D. Content-Based Retrieval Using a Multi-Objective Genetic Algorithm // Proceedings of IEEE SoutheastCon. - 2005. - P. 561-569.
- [35] Veltkamp R.C. Shape Matching: Similarity Measures and Algorithms // In proceedings of the International Conference on Shape Modeling and Applications. - Genova, Italy, 2001. - P. 188-197.
- [36] Wang Z., Bovik A., Sheikh H., Simoncelli E. Image quality assessment: from error visibility to structural similarity // IEEE Transactions on Image Processing. - 2004. - Vol. 13, issue 4. - P. 600-612.
- [37] Yuan H., Zhang X.-P. Texture image retrieval based on a Gaussian mixture model and similarity measure using a Kullback divergence // Proceedings of ICME. - 2004. Vol. 3, P. 1867-1870.
- [38] Zeng W., Daly S., Lei S. An overview of the visual optimization tools in JPEG 2000 // Signal Processing: Image Communication. - 2002. - Vol.17. - P. 85-104.
- [39] Zhou Z.-H., Dai H.-B. Query-sensitive similarity measure for content-based image retrieval // Proc. of Int. conf. on Data Mining. - 2006. - P. 1211-1215.

Robust search of images by full and partial similarity using multi-parameter classification

N.N.Ponomarenko V.V.Lukin S.K.Abramov

A task of image search in large databases by their similarity with complex use of different features is considered. An iterative approach to determination of weights for features of similarity measure which is a sum of a large number of weighted distances between image features is proposed. Experiments for obtaining averaged subjective estimates of image similarity (totally 100 000 estimates of similarity have been accomplished) have been carried out. They have been further used for weight optimization of similarity measure. A bank of distances between image features (more than 600 distances) has been formed; these features have been determined only for thumbnails of images. Several original parameters for image ranking have been proposed. Optimized weights of similarity measure have been obtained, their high effectiveness has been demonstrated on search examples and by quantitative analysis of rank correlation factors for experimental data.