

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ  
УНИВЕРСИТЕТ

На правах рукописи

Васильева Наталья Сергеевна

ПОСТРОЕНИЕ И КОМБИНИРОВАНИЕ ПРИЗНАКОВ В ЗАДАЧЕ  
ПОИСКА ИЗОБРАЖЕНИЙ ПО СОДЕРЖАНИЮ

05.13.11 — Математическое и программное обеспечение  
вычислительных машин, комплексов и компьютерных сетей

АВТОРЕФЕРАТ

диссертации на соискание ученой степени  
кандидата физико-математических наук

Санкт-Петербург  
2010

Работа выполнена на кафедре системного программирования математико-механического факультета Санкт-Петербургского государственного университета.

Научный руководитель: доктор физико-математических наук,  
проф. НОВИКОВ Борис Асенович

Официальные оппоненты: доктор физико-математических наук,  
проф. КАРКИЩЕНКО Александр Николаевич  
(ОАО “НИИАС”, Москва)

кандидат физико-математических наук,  
ОКРУГИН Михаил Борисович  
(Quest Software, Inc., Санкт-Петербург)

Ведущая организация: Институт системного программирования РАН  
(Москва)

Защита диссертации состоится “\_\_\_” \_\_\_\_\_ 2010 года в \_\_ часов на заседании совета Д212.232.51 по защите докторских и кандидатских диссертаций при Санкт-Петербургском государственном университете по адресу: 198504, Санкт-Петербург, Петродворец, Университетский пр., д. 28, математико-механический факультет, ауд. 405.

С диссертацией можно ознакомиться в Научной библиотеке Санкт-Петербургского государственного университета по адресу: 199034, Санкт-Петербург, Университетская наб., д. 7/9.

Автореферат разослан “\_\_\_” \_\_\_\_\_ 2010 года.

Ученый секретарь  
диссертационного совета  
доктор физико-математических наук,  
профессор

Даугавет И. К.

## Общая характеристика работы

**Актуальность темы.** Исследования в области поиска изображений по содержанию стали актуальны в последние десятилетия в связи с ростом емкости доступных накопителей данных и широким распространением цифровой фотографии, и, как следствие, ростом числа и объемов коллекций изображений.

Поиск по содержанию (Content Based Image Retrieval, CBIR) является приоритетным подходом к решению задачи поиска изображений. Методы поиска по содержанию работают на основе анализа численных характеристик составляющих изображение пикселей и не требуют наличия текстовых аннотаций или другой дополнительной информации об изображении. Это позволяет избежать трудоемкости и субъективности составленных вручную аннотаций, неточности аннотаций, полученных автоматически или полуавтоматически. Однако на сегодняшний день эффективность систем поиска по содержанию значительно уступает эффективности поиска по аннотациям. Основной проблемой поиска по содержанию большинство исследователей признают так называемый “*семантический разрыв*”. Человек, сравнивая два изображения, в первую очередь сравнивает их смысловое наполнение – семантику, в то время как оценка системы основывается на сравнении низкоуровневых признаков, описывающих визуальные характеристики изображения (цвет, текстуру, форму объектов). Задачи уменьшения семантического разрыва и повышения качества результатов поиска по содержанию являются актуальными в области поиска изображений и информационного поиска в целом.

**Цель работы.** Основной целью работы является создание высокоэффективных с точки зрения качества результата методов поиска по содержанию для коллекций изображений произвольной тематики, позволяющих уменьшить семантический разрыв. Повышение качества результатов поиска возможно за счет построения более эффективных признаков, а также за счет синтеза различных методов поиска.

Для достижения поставленной цели были решены следующие задачи.

- Разработка эффективного цветового признака в соответствии с особенностями зрительного восприятия человека.
- Формулирование требований к методам синтеза в контексте задачи поиска изображений.
- Разработка эффективных методов синтеза для комбинирования результатов поиска по различным пространствам признаков.

**Основные результаты.** В работе получены следующие основные результаты.

1. Новый цветовой признак на основе цветовой гистограммы, учитывающий пространственное расположение цветов, и функция расстояния для соответствующего пространства признаков.
2. Рекомендации по выбору оптимальной схемы квантования цветового пространства при построении цветовой гистограммы в зависимости от используемой метрики и размеров коллекции, полученные по результатам экспериментального исследования.
3. Требования к универсальным (не зависящим от изображения-запроса) методам синтеза для комбинирования результатов поиска по различным пространствам признаков. Метод синтеза с использованием среднего взвешенного с гравитационной функцией (WTGF - Weighted Total with Gravitation Function), удовлетворяющий сформулированным требованиям.
4. Алгоритм поиска в частично аннотированной коллекции изображений по текстовому запросу, не требующий предварительного автоаннотирования всей коллекции. В основе алгоритма лежит идея использования методов синтеза.
5. Адаптивный метод синтеза результатов поиска по цветовым и текстурным признакам в зависимости от изображения-запроса, центроидный метод классификации запроса.

**Научная новизна.** В работе предложен новый цветовой признак изображений на основе гистограммы, отражающий пространственное расположение цветов, и функция расстояния для соответствующего пространства признаков. Сформулированы требования к методам синтеза результатов поиска применительно к задаче поиска изображений, а также предложено два новых метода синтеза: взвешенное среднее с гравитационной функцией и адаптивный синтез в зависимости от изображения-запроса. Предложен центроидный метод классификации изображения-запроса в рамках решения задачи адаптивного синтеза. Предложен алгоритм поиска в частично аннотированной коллекции изображений по текстовому запросу, не требующий предварительного автоаннотирования всей коллекции.

**Теоретическая ценность и практическая значимость.** Теоретическую ценность работы составляют предложенное расширение классической цветовой гистограммы вкуче с функцией расстояния для соответствующих векторов признаков, гипотеза о существовании единой оптимальной

смешанной метрики для групп семантически и визуально подобных изображений, лежащая в основе метода адаптивного синтеза, требования к универсальной функции синтеза в контексте задачи поиска изображений.

Предложенные в работе методы могут быть эффективно использованы в таких современных практических задачах, как навигация и поиск в частных архивах цифровых фотографий и в сети Интернет, защита авторских прав (поиск копий авторской фотографии), и другие. Эффективные методы поиска по содержанию помогут значительно дополнить и повысить качество результатов поиска существующих поисковых систем в Интернет. Большинство таких систем предлагают пользователям поиск изображений по ключевым словам, построенным по контексту изображения, и не используют анализ содержания изображений. Предложенный алгоритм поиска по текстовому запросу в частично аннотированных коллекциях изображений позволит использовать текстовый запрос без предварительного автоаннотирования всех изображений коллекции.

**Апробация работы.** Основные результаты диссертации докладывались на Российском семинаре по Оценке Методов Информационного Поиска, РОМИП 2008 (Дубна, Россия, октябрь 2008); на Второй Российской летней школе по информационному поиску “Russian Summer School in Information Retrieval” (Таганрог, сентябрь 2008); на Международной Конференции по Обработке Изображений и Сигналов “International Conference on Image and Signal Processing” (Шербург-Октевиль, Франция, июль 2008); на Шестой и Восьмой Международных Балтийских Конференциях по Базам Данных и Информационным Системам “Baltic Conference on Databases and Information Systems” (Рига, Латвия, июнь 2004 и Таллинн, Эстония, июнь 2008 ); на семинаре Московской Секции ACM SIGMOD (Москва, Россия, ноябрь 2007); на семинаре победителей конкурса научных проектов “Интернет-Математика 2007” (Переславль-Залесский, октябрь 2007); на Седьмой и Девятой Всероссийских научных конференциях “Электронные библиотеки: перспективные методы и технологии, электронные коллекции” (Ярославль, Россия, октябрь 2005 и Переславль-Залесский, Россия, октябрь 2007); на семинарах группы исследования методов организации информации при лаборатории исследования операций НИИМ.

Результаты диссертации были частично использованы в работах по гранту РФФИ 07-07-00268а, по гранту компании “Яндекс” для победителей конкурса научных проектов “Интернет-Математика 2007”, во внутренних проектах лаборатории Hewlett-Packard.

**Публикации.** Основные результаты представлены в работах [1-10]. Статьи [1-2] опубликованы в журнале, входящем в перечень ВАК. Статьи [3-10] написаны в соавторстве. В работах [3-6] автору принадлежат метод центроидной классификации запроса, включая алгоритм находжде-

ния центроидов, общая постановка экспериментов, реализация цветовых и текстурных признаков, анализ результатов экспериментов, Маркову И. Е. – реализация описываемых методов, создание экспериментальных стендов, сбор результатов экспериментов. В статье [7] автору принадлежат схема поиска по текстовому запросу в частично-аннотированной коллекции, требования к универсальной функции синтеза, анализ результатов экспериментов и анализ применимости предложенных методов синтеза, соавторам – определение функции WTGF, создание экспериментальных стендов, сбор результатов экспериментов. В статье [8] автору принадлежит гипотеза о существовании единой оптимальной смешанной метрики для групп семантически и визуально подобных изображений, определение метода адаптивного синтеза, реализация цветовых и текстурных признаков, соавторам – реализация метода адаптивного синтеза, постановка и сбор результатов экспериментов. В работах [9-10] автору принадлежат алгоритм поиска по подобию с использованием текстового запроса, алгоритм построения соответствий между низкоуровневыми признаками и семантикой изображений, определение цветовой гистограммы, учитывающей пространственное расположение цветов, реализация цветовых и текстурных признаков, постановка и анализ результатов экспериментов, Новиков Б. А. оказывал помощь в формулировании задач и при редактировании текста статей.

**Структура и объем диссертации.** Диссертация состоит из введения, 3 глав, заключения и списка литературы. Общий объем диссертации составляет 164 страницы машинописного текста. Библиография содержит 149 наименований. Рисунки и таблицы нумеруются по главам.

## Содержание работы

Во **введении** содержится предварительная информация о предмете исследования, обосновывается актуальность тематики диссертационной работы и кратко излагаются ее основные результаты.

В **первой главе** “Методы поиска изображений по содержанию” представлены направления и задачи области исследования (Content-Based Image Retrieval, CBIR), рассматриваются основные алгоритмы построения векторов признаков и метрики для соответствующих пространств, обсуждаются известные на сегодняшний день подходы к комбинированию различных признаков.

В п. 1.1 представлена традиционная архитектура систем поиска изображений по содержанию и выделены основные направления исследований в области CBIR: построение пространств признаков, многомерное индексирование и проектирование систем поиска. Диссертация решает задачи первого из направлений.

В п. 1.2 перечислены основные проблемы области СВІR: семантический разрыв, неудобство запроса-образца, необходимость обеспечения высокой скорости работы систем поиска на больших объемах данных.

В п. 1.3 приведена общая классификация подходов к построению векторов признаков. Выделены классы цветовых, текстурных и признаков формы. В пп. 1.4 – 1.6 описаны основные алгоритмы построения векторов признаков для цвета, текстуры и формы объектов. Для каждого из классов приведена более подробная классификация в соответствующем разделе.

П. 1.7 содержит описание известных на сегодняшний день методов синтеза, используемых для формирования общей выдачи на основе результатов нескольких независимых поисковых алгоритмов. Рассматриваются в том числе и методы синтеза, используемые в текстовом поиске. Обсуждается их применимость в задаче поиска изображений.

В п. 1.8 рассматриваются некоторые из наиболее известных на сегодняшний день систем поиска изображений по содержанию с тем, чтобы получить представление какие из многочисленных характеристик и их признаков используются в реальных системах поиска.

Во **второй главе** “Поиск по цвету” предлагается новый цветовой признак на основе цветовой гистограммы, учитывающий пространственное расположение цветов, а также описывается экспериментальное исследование по выбору оптимальной схемы квантования цветового пространства для построения цветовой гистограммы.

Цвет является наиболее значимой характеристикой при поиске по коллекции цветных изображений произвольной тематики. Цветовые гистограммы – наиболее широко используемые цветовые признаки в системах поиска изображений по содержанию. Эффективность поиска по цветовым гистограммам во многом зависит от выбора цветового пространства и схемы его квантования.

П. 2.1 описывает рассмотренные в работе *палитры* – наборы цветовых диапазонов, полученных путем равномерного квантования некоторого цветового пространства с определенными шагами для каждой из координатных осей. Для обозначения палитры будем использовать запись вида  $ColorSpace\langle K \times L \times M \rangle$ , где  $ColorSpace$  – обозначение цветового пространства;  $K, L, M$  – число уровней квантования для трех координатных осей пространства.

Формально палитра определяется следующим образом:

$$ColorSpace\langle K \times L \times M \rangle = \{\Delta_{k,l,m}\}_{k=1..K,l=1..L,m=1..M}, \quad (1)$$

$$\Delta_{k,l,m} = \{(x, y, z) \in ColorSpace | x \in (x_k, x_{k+1}], y \in (y_l, y_{l+1}], z \in (z_m, z_{m+1}]\},$$

$$x_j = j\delta_x, \quad y_j = j\delta_y, \quad z_j = j\delta_z.$$

Здесь  $\delta_x, \delta_y, \delta_z$  – шаги квантования по осям пространства *ColorSpace*.

В работе использовались цветовые пространства HSI\*, nHSI\*, HCL (Hue, Chroma and Luminance), nHCL, CIELab, nCIELab и nRGB. Описание пространств и формальные правила преобразования RGB  $\rightarrow$  HSI\*, RGB  $\rightarrow$  HCL, RGB  $\rightarrow$  CIELab, RGB  $\rightarrow$  nRGB приведены в п. 2.1.1. Пространства nHSI\*, nHCL, nCIELab были получены из пространства nRGB с помощью тех же преобразований, что и пространства HSI\*, HCL, CIELab из RGB соответственно.

В п. 2.1.2 обосновывается выбор шагов при равномерном квантовании для перечисленных выше пространств. Однако равномерное квантование пространств HSI\* и HCL не учитывает такие особенности системы зрительного восприятия человека, как меньшая чувствительность к изменению оттенка цвета при недостаточной или чрезмерной яркости, недостаточной насыщенности цвета. Чтобы учесть эти особенности, введем для пространств семейства HSV еще одну схему квантования *ColorSpace* $\langle K \times L \times M \rangle_{th}$  – равномерное квантование с граничными условиями:

$$\begin{aligned}
ColorSpace\langle K \times L \times M \rangle_{th} &= \Delta^B \cup \Delta^W \cup \\
&\quad \{\Delta_m^G\}_{m=1..M} \cup \{\Delta_{k,l,m}\}_{k=1..K, l=1..L, m=1..M}, \\
\Delta^B &= \{(x, y, z) \in ColorSpace | z \leq I^B\}, \\
\Delta^W &= \{(x, y, z) \in ColorSpace | z \geq I^W\}, \\
\Delta_m^G &= \{(x, y, z) \in ColorSpace | y \leq S^G, z \in (z_m, z_{m+1}]\}, \\
\Delta_{k,l,m} &= \{(x, y, z) \in ColorSpace | x \in (x_k, x_{k+1}], y \in (y_l, y_{l+1}], z \in (z_m, z_{m+1}]\}, \\
x_j &= j\delta_x, \quad y_j = S^G + j\delta_y, \quad z_j = I^B + j\delta_z.
\end{aligned} \tag{2}$$

Здесь  $I^B$  – пороговое значение яркости, такое что все цвета с меньшей яркостью не отличимы от черного;  $I^W$  – пороговое значение яркости, такое что все цвета с большей яркостью не отличимы от белого;  $S^G$  – пороговое значение насыщенности, такое что все цвета с меньшей насыщенностью не отличимы от серого;  $\delta_x, \delta_y, \delta_z$  – шаги квантования по осям пространства *ColorSpace*.

Полный перечень палитр, рассмотренных в работе, представлен в п. 2.3.1, в таблице 2.1.

В п. 2.2. вводится новый цветовой признак на основе цветовой гистограммы. Для каждого ненулевого элемента гистограммы предлагается вычислять центр масс пикселей соответствующего цвета. Таким образом цветовой признак изображения  $I$  представляет собой набор векторов

$$HistSP(I) = \{\mathbf{c}_i | \mathbf{c}_i = (p_i, x_i, y_i)\}_{i=1..N}, \tag{3}$$

где  $N$  – число цветовых диапазонов;  $p_i$  – отношение количества пиксе-



лей, принадлежащих  $i$ -му цветовому диапазону, к общему числу пикселей в изображении;  $x_i, y_i$  – нормированные относительно размера изображения координаты центра масс пикселей  $i$ -го цветового диапазона.

Для определения расстояния между изображениями  $I^{(1)}$  и  $I^{(2)}$  по соответствующим им признакам

$$HistSP(I^{(1)}) = \{\mathbf{c}_i^{(1)} | \mathbf{c}_i^{(1)} = (p_i^{(1)}, x_i^{(1)}, y_i^{(1)})\}_{i=1\dots N}$$

и

$$HistSP(I^{(2)}) = \{\mathbf{c}_i^{(2)} | \mathbf{c}_i^{(2)} = (p_i^{(2)}, x_i^{(2)}, y_i^{(2)})\}_{i=1\dots N}$$

мы предлагаем использовать следующую функцию:

$$D_{HistSP}(I^{(1)}, I^{(2)}) = \sum_{i=1}^N \left( \left( D_p(\mathbf{c}_i^{(1)}, \mathbf{c}_i^{(2)}) + \alpha \right) \left( D_{xy}(\mathbf{c}_i^{(1)}, \mathbf{c}_i^{(2)}) + \beta \right) - \alpha\beta \right), \quad (4)$$

где

$$D_p(\mathbf{c}_i^{(1)}, \mathbf{c}_i^{(2)}) = |p_i^{(1)} - p_i^{(2)}|, \quad (5)$$

$$D_{xy}(\mathbf{c}_i^{(1)}, \mathbf{c}_i^{(2)}) = \begin{cases} \frac{\sqrt{(x_i^{(1)} - x_i^{(2)})^2 + (y_i^{(1)} - y_i^{(2)})^2}}{max_{xy}}, & p_i^{(1)} > 0 \wedge p_i^{(2)} > 0 \\ 1, & p_i^{(1)} = 0 \vee p_i^{(2)} = 0 \end{cases}, \quad (6)$$

$max_{xy}$  – максимальное возможное расстояние между центрами масс цветowych пятен в нормированных координатах (длина диагонали изображения),  $\alpha$  и  $\beta$  – параметры.

Функция (4) не является метрикой, так как не удовлетворяет неравенству треугольника. Выбор данной функции расстояния обосновывается в п. 2.2.

П. 2.3 содержит описание и анализ результатов экспериментов по оценке эффективности поиска по классическим цветовым гистограммам и признакам HistSP в цветовых пространствах HSI\*, nHSI\*, HCL, nHCL, CIELab, nCIELab, nRGB с использованием равномерного квантования и равномерного квантования с граничными условиями с различными шагами.

Согласно полученным результатам, добавление информации о пространственном расположении цветов в гистограмму позволило существенно повысить точность (до 28%) и полноту (до 15%) результатов поиска вне зависимости от выбора цветового пространства и схемы квантования. Экспериментальное исследование зависимости эффективности поиска по цветовым гистограммам от выбора шага квантования цветового пространства доказало существование пороговых значений, таких что выбор меньших шагов квантования (увеличение числа цветовых диапазонов) не приводит к улучшению результатов поиска. При равномерном квантовании для всех цветовых пространств и функций подобию, участвовавших в экспериментах,

оптимальным оказалось использование порядка 500 цветовых диапазонов. Добавление граничных условий позволило повысить показатели точности и полноты для схем с небольшим числом цветовых диапазонов. По данным показателям схема квантования  $\langle 6, 2, 3 \rangle_{th}$  (41 цветовой диапазон) для пространств семейства  $HSV$  в большинстве случаев превзошла схемы равномерного квантования со значительно большим числом диапазонов или лишь незначительно уступила им.

В **третьей главе** “Синтез методов поиска при формировании результатов” обсуждаются подходы к комбинированию независимых признаков для формирования единого результата поиска. Рассматриваются различные схемы применения функций синтеза в решении задачи поиска изображений: использование синтеза для реализации поиска в частично аннотированной базе по текстовому запросу, универсальный (независящий от запроса) синтез методов поиска по визуальным признакам и адаптивный синтез методов поиска по цвету и текстуре.

В п. 3.1 рассматривается общая задача синтеза взвешенных ранжированных списков.

Для простоты дальнейшего изложения введем некоторые обозначения.

Ранжированный  $i$ -ый список будем обозначать символом  $\alpha_i$ . Ранжированный список состоит из пар вида  $(x, r_i(x))$ , где  $x$  – объект, а  $r_i(x)$  – его ранг в списке  $\alpha_i$ . Обозначим набор всех рангов объекта  $x$  из различных ранжированных списков как вектор  $R_x = (r_1(x), r_2(x), \dots, r_N(x))$ , где  $N$  – число ранжированных списков. В контексте задачи информационного поиска ранг объекта является значением функции подобия, определенной на некотором пространстве признаков, и отражает степень близости данного объекта запросу в этом пространстве признаков.

Будем рассматривать вес  $w_i$  ранжированного списка  $\alpha_i$  как показатель уверенности в том, насколько ранги объектов в данном списке соответствуют их степени подобия запросу. Обозначим набор весов всех ранжированных списков как вектор  $W = (w_1, w_2, \dots, w_N)$ .

Обозначим как  $\alpha_0$  результирующий ранжированный список,  $r_0(x)$  – ранг объекта  $x$  в результирующем списке. Задача синтеза сводится к поиску подходящей функции синтеза  $f(R_x, W)$ , с помощью которой можно вычислить ранг в результирующем списке каждого из элементов исходных списков:

$$\forall x, i : (x, r_i(x)) \in \alpha_i \quad r_0(x) = f(R_x, W).$$

П. 3.1.1 содержит описание практических задач, которые сводятся к задаче синтеза взвешенных ранжированных списков. В п. 3.1.2 определены основные свойства функции синтеза, подходящей для решения поставленных задач.

**Свойство 1** Симметричность.

$R_1 = (r_1(x), r_2(x), \dots, r_N(x))$ ,  $W_1 = (w_1, w_2, \dots, w_N)$ ,  $R_2, W_2$  – перестановки  $R_1$  и  $W_1$  соответственно, тогда  $f(R_1, W_1) = f(R_2, W_2)$ .

**Свойство 2** Монотонность по каждому из аргументов.

$R_1 = (r_1(x), r_2(x), \dots, r_N(x))$ ,  $R_2 = (r_1(y), r_2(y), \dots, r_N(y))$ ,  
 $W_1 = (w_{11}, w_{12}, \dots, w_{1N})$ ,  $W_2 = (w_{21}, w_{22}, \dots, w_{2N})$ ,  
тогда  $r_i(x) \leq r_i(y) \quad \forall i \in \{1, \dots, N\} \Rightarrow f(R_1, W_1) \leq f(R_2, W_1)$ ,  
 $w_{1i} \leq w_{2i} \quad \forall i \in \{1, \dots, N\} \Rightarrow f(R_1, W_1) \leq f(R_1, W_2)$ .

**Свойство 3** Суперпозиция.

Результат работы функции синтеза есть ранжированный список, который может использоваться в качестве входных данных для функции синтеза.

**Свойство 4** Ограниченность (MinMax условие).

Пусть имеется  $N$  списков:  $\alpha_1, \alpha_2, \dots, \alpha_N$ . Элементу  $x$  в них сопоставляется следующий набор рангов:  $R_x = (r_1(x), r_2(x), \dots, r_N(x))$ . Тогда после синтеза ранг  $r_0(x)$  должен удовлетворять условию:  
 $\min(r_1(x), r_2(x), \dots, r_N(x)) \leq r_0(x) \leq \max(r_1(x), r_2(x), \dots, r_N(x))$ .

**Свойство 5** Для функции синтеза должны выполняться следующие "правила конусов":

- Выполнение граничных условий.
  - Вероятность изменения ранга объекта, присутствующего с высоким рангом в списке с большим весом, мала.
  - При синтезе двух списков, один из которых обладает весом, близким к нулю, таковой почти не вносит вклада в результат. Высока вероятность того, что его элементы изменят свой ранг.
- Влияние весов списков. Если есть два элемента ( $x \in \alpha_1$  и  $y \in \alpha_2$ ) из списков с различными весами ( $w_1 > w_2$ ), но с одинаковыми рангами ( $r_1(x) = r_2(y) > 0$ ), то в результирующем списке их ранги должны быть различны, причем  $r_0(x) > r_0(y)$ .
- Степень свободы низкоранговых элементов. Чем меньше ранг элемента, тем больше степень свободы изменения его ранга в результирующем списке.
- Степень свободы высокоранговых элементов в списках с высоким весом. Чем больше ранг элемента и выше вес списка, в котором он встречается с высоким рангом, тем меньше степень свободы изменения его ранга в результирующем списке.

- **Согласованность.** Незначительные изменения веса списка или ранга элемента не должны повлечь значительных изменений результирующего ранга элемента.

В п. 3.1.2 определяется новая функция синтеза WTGF (Weighted Total with Gravitation Function) – взвешенное среднее с гравитационной функцией, удовлетворяющая сформулированным требованиям. Данная функция задана следующим образом:

$$f(R_x, W) = \frac{\sum_i r_i(x) \cdot g(r_i(x), w_i)}{\sum_i g(r_i(x), w_i)}, \quad (7)$$

где  $g$  – функция стабилизации высокоранговых элементов или гравитационная функция

$$g(r_i(x), w_i) = w_i^2 \cdot \left( r_i(x) + \frac{1}{12} \right)^4. \quad (8)$$

П. 3.1.2 содержит обоснование выбора вида функции WTGF, а также описание реализации вычислений данной функции. В п. 3.1.3 и п. 3.1.4 приводятся описание и результаты экспериментального сравнения функции WTGF с известными функциями синтеза CombMNZ и взвешенным средним при решении поставленных выше практических задач.

Согласно полученным результатам, функция WTGF более эффективна при условии достоверности информации о весах источников синтеза и их небольшом пересечении. Использование CombMNZ дает лучший результат при достаточном пересечении источников синтеза. В целом, использование методов синтеза для решения поставленных нами задач оправдало себя – это позволило улучшить результаты поиска.

В п. 3.2 рассматривается новый метод адаптивного синтеза результатов поиска по цветовым и текстурным признакам, зависящий от запроса-образца. Данный метод основан на подтверждаемой экспериментами гипотезе о том, что при использовании линейной комбинации в качестве функции синтеза можно определить оптимальные веса для изображений-запросов из разных семантических классов. Классифицировав запрос по заданным ранее классам и тем самым определив для него оптимальные веса, возможно повысить точность результатов поиска.

В п. 3.2.1 вводится понятие *смешанных метрик*  $D_{mixed}$ , которые являются линейной комбинацией цветовой и текстурной метрик с определенными весами.

$$D_{mixed}(I, Q) = \alpha \cdot D_{color}(I, Q) + (1 - \alpha) \cdot D_{texture}(I, Q), \quad (9)$$

где  $I$  и  $Q$  – изображения,  $D_{color}$  – цветовая метрика,  $D_{texture}$  – текстурная метрика,  $\alpha$  – параметр из отрезка  $[0, 1]$ . Метод адаптивного синтеза заклю-

чается в применении на этапе поиска смешанной метрики с оптимальным значением параметра  $\alpha$  для заданного запроса.

В п. 3.2.2 приводятся описание и результаты эксперимента, подтверждающего сформулированную гипотезу, в п. 3.2.3 – эксперимента по сравнению эффективности метода адаптивного синтеза с известным методом синтеза CombMNZ. Согласно полученным результатам, средняя точность поиска по всем изображениям тестового множества для оптимальных смешанных метрик составила 42.76%, для метода CombMNZ – 39.68%, что говорит о превосходстве метода адаптивного синтеза.

В п. 3.2.4 исследуется применимость известных методов классификации (методы Байеса и опорных векторов, SVM) для классификации запроса в контексте метода адаптивного синтеза, предлагается центроидный классификатор. Согласно результатам эксперимента, классификаторы Байеса и SVM не вполне подходят для решения поставленной задачи. Основные ограничения данных методов связаны с недостаточной представительностью и несбалансированностью обучающего множества. Такие условия обучения классификаторов приводят к их ненадежной работе на этапе классификации. В то же время, центроидный классификатор, будучи крайне простым в обучении и быстрым в принятии решения, дает приемлемые результаты классификации в сравнении с методами Байеса и опорных векторов в контексте решаемой нами задачи.

П. 3.3 содержит основные выводы по исследованию методов синтеза применительно к комбинированию независимых признаков в задаче поиска изображений.

**Заключение** содержит список основных результатов, полученных в работе.

## Работы автора по теме диссертации

### Статьи в журналах, рекомендованных ВАК:

- [1] Васильева Н.С. Выбор шага квантования при построении цветовой гистограммы в задаче поиска изображений // Вестник СПбГУ. Сер. 10: Прикладная математика, информатика, процессы управления. – Изд-во СПбГУ, 2009. – Вып. 2. – С. 155–164.
- [2] Васильева Н.С. Методы поиска изображений по содержанию // Программирование. – 2009. – № 3. – С. 1–30.

## Другие публикации:

- [3] Васильева Н.С., Марков И.Е. СПбГУ на РОМИП'2008: Синтез цветовых и текстурных признаков при поиске изображений по содержанию // Труды Российского семинара по Оценке Методов Информационного Поиска РОМИП 2007-2008. – 2008. – С. 135–144.
- [4] Markov I., Vassilieva N. Building up low-level centroids for groups of perceptually similar images // Proceedings of the 8th International Baltic Conference on Databases and Information Systems. – Tallinn, Estonia, 2008. – P. 341–348.
- [5] Markov I., Vassilieva N. Image retrieval. Color and texture combining based on query-image // Proceedings of the International Conference on Image and Signal Processing. – Cherbourg-Octeville, France, 2008. – P. 430–438.
- [6] Markov I., Vassilieva N. Query classification in content-based image retrieval // Databases and Information Systems V - Selected Papers from the Eighth International Baltic Conference, DB&IS 2008 / Ed. by H.-M. Naav, A. Kalja. – IOS Press, 2008. – Vol. 187 of Frontiers in Artificial Intelligence and Applications. – P. 281–288.
- [7] Васильева Н.С., Дольник А.С., Марков И.Е. Поиск изображений. Синтез различных методов поиска при формировании результатов // Интернет-Математика 2007: Сборник работ участников конкурса / Под ред. П. Браславского. – Екатеринбург: Изд-во Урал. ун-та, 2007. – С. 46–55.
- [8] Марков И.Е., Васильева Н.С., 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”. – 2007. – С. 195–200.
- [9] Васильева Н.С., Новиков Б.А. Построение соответствий между низкоуровневыми характеристиками и семантикой статических изображений // Труды 7-ой Всероссийской научной конференции “Электронные библиотеки: перспективные методы и технологии, электронные коллекции”. – 2005. – С. 236–240.
- [10] Vassilieva N., Novikov B. A similarity retrieval algorithm for natural images // Proceedings of the 6th International Baltic Conference on Databases and Information Systems (Doctoral Consortium), Vol. 673. – Riga, Latvia, 2004. – P. 151–154.