

Поиск изображений. Синтез различных методов поиска при формировании результатов.

Наталья Васильева

Санкт-Петербургский
Государственный
Университет
natalia@ntc-it.ru

Александр Дольник

Санкт-Петербургский
Государственный
Университет
alexander.dolnik@gmail.com

Илья Марков

Санкт-Петербургский
Государственный
Университет
Ilya.Markov@gmail.com

Аннотация

В отчете представлены результаты исследования применимости методов синтеза данных к задаче поиска изображений.

При поиске изображений по содержанию распространенной практикой является независимый анализ различных низкоуровневых характеристик (цвет, текстура, форма). При таком подходе встает задача дальнейшего синтеза результатов поиска по каждой из характеристик.

Авторами предлагается новая функция синтеза WTGF, которая для ряда задач выигрывает у CombMNZ. Отличительной особенностью функции WTGF является возможность учитывать разные веса для входных источников.

В отчете представлены результаты экспериментов по использованию различных функций синтеза для решения задачи поиска по частично аннотированной базе изображений с текстовым запросом и классической задачи поиска изображений по содержанию. В качестве методов поиска по содержанию использованы методы поиска по цветовым и текстурным характеристикам изображений.

Также в данном отчете представлен ряд методов оценки эффективности методов синтеза и поиска изображений в целом.

1. Введение

За последние годы объем мультимедиа данных возрос в несколько раз. Однако вся эта информация была бы бесполезна без точного, удобного и быстрого поиска по ней. Эта работа посвящена задаче поиска изображений.

Все существующие на настоящий момент методы поиска графической информации можно разделить на два направления:

- § поиск по содержанию (content-based image retrieval – CBIR)¹;
- § поиск по текстовым аннотациям (description-based image retrieval – DBIR), который делится еще на три группы по способу построения аннотаций:
 - автоматические полученные аннотации,
 - полуавтоматические аннотации,
 - полученные вручную аннотации.

К аннотациям, полученным автоматически, относятся: взятые из "около картиночного" текста ключевые слова, различного рода другая мета информация (как время создания, географическое место создания) и прочее.

Под полуавтоматическими способами аннотирования понимаются методы, при которых часть аннотаций изображений описывается вручную, остальные автоматически получают аннотации изображений, близких по содержанию.

Заголовки изображений и написанные пользователем словесные аннотации – относят к полученным вручную аннотациям. Системы с ручным аннотированием относят к первому поколению систем текстового (по текстовому запросу) поиска изображений.

Существуют также разные схемы аннотации:

1. Полнотекстовые аннотации. Изображение полностью описывается словосочетанием или несколькими предложениями.
2. Аннотации ключевыми словами. Изображению приписывается несколько ключевых слов и создается словарь контрольных терминов.
3. Онтологии.

В отличие от систем DBIR, системы поиска изображений по содержанию не требуют наличия какой-либо дополнительной информации об изображении. Поиск в таких системах производится на основе анализа и сравнения низкоуровневых характеристик изображения, таких как цвет или текстура. Это позволяет уйти от трудоемкости и субъективности составленных вручную аннотаций, неточности аннотаций, полученных автоматически или полуавтоматически. Однако и у этого подхода есть свои минусы:

§ Типичным видом запроса для таких систем является изображение-образец, что менее удобно для пользователя, чем текстовый запрос.

§ Для поиска по коллекции разнородного содержания недостаточно анализа по какой-то одной из характеристик. Необходимо комбинировать методы поиска по различным характеристикам. Говоря о разнородном содержании, мы подразумеваем, что рассматриваемая коллекция изображений не

¹В англоязычной научной литературе по компьютерному зрению (computer vision) также можно встретить термин query by image content (QBIC) или термин content-based visual information retrieval (CBVIR).

принадлежит одной определенной тематике.

§ Система поиска изображений по содержанию сравнивает изображения на основе низкоуровневой информации, в то время как человек в большинстве случаев оценивает семантическую схожесть картинок. Проблема «семантического разрыва» – отсутствия связи между семантикой изображения и его низкоуровневыми характеристиками является одной из основных в области CBIR.

Использование текстового запроса позволяет решить проблему запроса-образца и сузить семантический разрыв. Но использование такого вида запроса невозможно без наличия в системе текстовой информации о изображениях. Предоставить аннотации для всех изображений базы - задача трудоемкая. В то же время, достаточно распространены частично аннотированные базы изображений. Любой фотохостинг предоставляет пользователям возможность комментировать свои или чужие фотографии, при этом наличие комментария не является обязательным. Данные комментарии можно использовать в качестве аннотаций для тех изображений, к которым они были написаны. Таким образом, задача поиска по частично аннотированной базе по текстовому запросу представляется достаточно актуальной. В данной работе предлагается подход к решению данной задачи с использованием метода комбинирования данных.

Также задача комбинирования крайне актуальна и применительно к «чистому» поиску по содержанию – выше была упомянута необходимость комбинирования результатов поиска по различным характеристикам изображения. В то же время, на сегодняшний день лишь небольшое количество работ посвящено исследованию методов комбинирования в данном контексте, хотя успешное использование данных методов может заметно улучшить качество поиска изображений по содержанию.

2. Обзор литературы

2.1 Поиск изображений по текстовому запросу

Для того чтобы был возможен поиск изображений по текстовому запросу, необходимо связать текстовую информацию об изображении с самим изображением. Первые системы поиска такого рода предусматривали ручное аннотирование изображений и поиск по данным аннотациям.

В дальнейшем, довольно широкое распространение получил подход, в основе которого лежит использование контекста изображений для получения аннотаций. В частности, аннотации выделялись из окружающего изображения текста. Примером может служить метод SRC, описанный в статье [13]. Однако, не всегда изображение индексируемой коллекции является иллюстрацией к тексту, и качество аннотаций, полученных таким способом, уступает аннотациям, полученным

вручную.

Некоторые исследователи уделяют особое внимание организации аннотаций. Например, Gertz с соавторами [10, 11] разработали модель графа аннотаций для нужд искусственного интеллекта (Human Brain Project - НБР). Узлы аннотаций служат для соединения так называемых узлов областей интереса изображения с концептуальными узлами, управляемыми словарем. Ребра графа обозначают отношения между узлами, например, аннотации или понятия. Они также разработали структуру для запросов к графу, основанную на путевых выражениях и предикатах. Аннотирование предполагалось производить вручную.

В последнее время все большее распространение получают методы автоматического аннотирования изображений, задачей которых является установление связи между особенностями низкоуровневых характеристик изображения или его сегмента с конкретными аннотациями [15, 19]. Например, в статье [15] приведен следующий метод автоматического аннотирования изображений. Для изображений в глобальной сети строятся кластеры похожих изображений по содержанию. Далее используется SRC метод для извлечения аннотаций из окружающего изображения текста, для тех из них, для которых это возможно. Снова используя кластеризацию, авторы выделяют наиболее вероятные ключевые слова для изображений одного кластера. Таким образом, аннотации получают все изображения кластера. Потенциально данный метод может работать с неограниченным словарем.

В работе [13] Herve Glotin и Sabrina Tollari предлагают использовать DIMATEX (дихотомический метод создания текстовых аннотаций). Они выделяют 13 различных уровней человеческого восприятия с двумя состояниями на каждом уровне (ярко или не ярко выражен признак). Далее метод сегментирует изображение и приписывает каждому сегменту ключевое слово, основываясь на тренировочном наборе данных. Таким образом, получают кластеры сегментов с аннотациями. Для аннотирования произвольной картинки, проверяется на каких ее сегментах и при каких кластерах достигается максимум.

Довольно распространено использование скрытой марковской модели (СММ) для решения задачи автоаннотирования. Например, данный подход использовался в работе [12]. Этот метод похож на описанный выше сегментированный подход, однако, сегменты в данном случае являются зависимыми друг от друга (совокупность сегментов образует состояние). Таким образом, увеличивается точность описания в целом.

2.2 Методы поиска изображений по содержанию

Методы поиска по содержанию сравнивают изображения на основе информации об их низкоуровневых характеристиках. Для коллекции

цветных изображений произвольной тематики наиболее значимыми характеристиками являются цвет и, в меньшей степени, текстура.

Самой простой и распространенной моделью для представления информации о цвете изображения являются цветовые гистограммы [1, 23, 25]. Для построения такой гистограммы цветовое пространство разбивается на области (близкие цвета попадают в одну область), и для изображения вычисляется количество пикселей, попавших в каждую из областей. В качестве метрики обычно используется сумма модулей разностей значений элементов гистограмм, относящихся к одной цветовой области.

В данной работе в качестве одного из методов поиска по цвету мы будем использовать модификацию цветовой гистограммы, учитывающую пространственное расположение цветов на изображении [1, 27].

Альтернативным методом представления цветовой характеристики изображения является статистическая модель. В рамках данной модели за основу берется статистическое распределение отдельных цветовых каналов [23, 24, 25]. Для оценки схожести двух изображений сравниваются параметры распределения их цветов (математическое ожидание, дисперсия и другие параметры статистического распределения). В качестве второго метода поиска по цвету, участвующего в синтезе, мы будем использовать статистический метод, предложенный M. Stricker и соавторами [25].

Для описания текстуры изображения часто используются фильтры Габора (Gabor), которые строятся на основе различных сечений функции Габора. Для двух изображений сравниваются результаты свертки каждого из них с различными фильтрами. Альтернативой фильтрам Габора являются фильтры ИСА, которые строятся с помощью анализа независимых компонент по тренировочному набору изображений [3, 5]. Данные фильтры носят естественный характер и отражают основные направления текстуры изображений, по которым они строились. Более того, как показывают исследования [8], способ образования данных фильтров схож с процессом обучения зрительной системы человека. Это дает основание предположить, что результат сравнения изображений с использованием данных фильтров будет более соответствовать человеческому восприятию схожести изображений. В данной работе в качестве метода поиска по текстуре, так же участвующего в синтезе, мы будем использовать метод поиска, основанный на использовании фильтров ИСА.

2.3 Задачи слияния и комбинирования

Существует большое количество работ [14, 16, 17, 18, 21, 22], рассматривающих задачу синтеза или комбинирования данных. Общая идея представленных в них методов состоит в том,

чтобы, имея результаты работы некоторых поисковых алгоритмов или систем, можно было получить один общий результат, по тем или иным критериям превосходящий подаваемые на вход.

Ниже перечислены некоторые распространенные поисковые задачи, для решения которых используются алгоритмы синтеза.

§ Задача уточнения результатов поиска. Возникает в случае, когда совпадают области определения алгоритмов, но сами алгоритмы различны.

§ Задача дополнения результатов поиска. Возникает в случае, когда области определения алгоритмов не совпадают, а сами алгоритмы поиска при этом могут как совпадать, так и различаться.

§ Задача комбинированного поиска. Возникает в случае, когда запрос состоит из нескольких сущностей (комбинированный запрос). Например, одновременный поиск по изображению и уточняющему тексту. Зачастую такие задачи сводятся к тривиальному множественному пересечению, однако, иногда требуется не только выдать множество, но и расставить ранги его элементов, – в таком варианте задача усложняется.

Общая задача алгоритма синтеза может быть сформулирована следующим образом.

Пусть имеется некоторое множество объектов X и семейство алгоритмов $\{\mathcal{Z}_i\}$, заданных на некоторых подмножествах $U_i \subset X$. Алгоритмы заканчивают работу для всех входных данных из области определения. Результат работы алгоритма \mathcal{Z}_i – вектор некоторого метрического пространства с метрикой r .

Введем понятие ранга элемента $x \in U_i$ по отношению к элементу $x' \in U_i$ на множестве объектов из U_i с использованием алгоритма (или метода) \mathcal{Z}_i как значение метрики $r(\mathcal{Z}_i(x'), \mathcal{Z}_i(x))$.

Если применить алгоритм \mathcal{Z}_i ко всему множеству U_i и отсортировать результаты по рангу, полученный список можно рассматривать как результат работы поисковой системы \mathcal{Z}_i по тривиальному запросу x' .

На вход алгоритму синтеза подаются результаты работы нескольких алгоритмов ранжирования – ранжированные списки элементов из множества $H \cap X$, а на выходе получаем один ранжированный список, в котором учтены "мнения" всех ранжирующих алгоритмов.

В литературе рассмотрено множество алгоритмов синтеза (комбинирования) результатов, полученных различными методами поиска или при поиске по различным данным.

Каждый такой метод поиска в совокупности с данными, с которыми он работает, принято называть свидетелем. Разделяют гомогенных (однородных) и гетерогенных (неоднородных)

свидетелей.

В работе [20] (1983 год) авторы вводят модель p -нормы в качестве обобщенного способа обработки булевых запросов. Пусть для запроса, состоящего из n понятий $Q = \{t_1, t_2, \dots, t_n\}$, ранги понятий в документе d есть s_1, s_2, \dots, s_n . Тогда результирующее значение ранга может быть подсчитано по приведенной ниже формуле:

$$S(d) = \left(\frac{s_1^p + \dots + s_n^p}{n} \right)^{\frac{1}{p}} \quad (1)$$

В начале 1990-ых годов такая модель была признана не слишком эффективной, и после первой же конференции TREC (<http://trec.nist.gov/>), которая состоялась в ноябре 1992 года, появились работы Edward A. Fox [9], в которых были приведены результаты экспериментов по синтезу различных методов текстового поиска. Еще через год (1994) Edward A. Fox и Joseph A. Shaw представили статью [21], в которой описывались несколько функций синтеза:

1. CombMAX = MAX(частных ответов);
2. CombMIN = MIN(частных ответов);
3. CombSUM = SUM(частных ответов);
4. CombANZ = SUM(частных ответов) / (количество списков с ненулевыми ответами);
5. CombMNZ = SUM(individual similarities) * (количество списков с ненулевыми ответами).

В рамках данной работы Shaw & Fox комбинировали результаты, полученные с использованием различных схем поиска. Результаты говорили в пользу представленных ими новых методов слияния над p -нормами. Авторы выделили метод CombSUM, как наиболее эффективный. Однако, в 1995 году на новых данных TREC5 корейский исследователь Joop Ho Lee в своей статье [14] показал, что метод CombMNZ все-таки выигрывает у CombSUM. Для сравнения методов он предложил использовать два показателя: коэффициент изменения количества релевантных объектов ($R_{overlap}$) и коэффициент изменения нерелевантных объектов ($N_{overlap}$) в случае синтеза двух источников. Эти коэффициенты рассчитываются по приведенным ниже формулам.

$$R_{overlap} = \frac{R_{common} \times 2}{R_1 + R_2}; \quad (2)$$

$$N_{overlap} = \frac{N_{common} \times 2}{N_1 + N_2} \quad (3)$$

где $R_{common}(N_{common})$, $R_1(R_2)$, $N_1(N_2)$ – число релевантных (нерелевантных) ответов в результирующем списке, в первом исходном списке, во втором исходном списке.

В ряде работ рассматриваются функции синтеза, представляющие собой линейную комбинацию исходных данных. Например, в [22] рассматривается линейная функция синтеза HSC3D для случая гомогенных источников. Авторы вводят ряд дополнительных требований к функции синтеза.

Пусть f - искомая функция синтеза, S - вектор значений рангов входных списков (мнение различных свидетелей о ранге объекта). Функция синтеза должна обладать следующими свойствами.

§ Симметричность.

$S_1 = (s_{11}, s_{12}, \dots, s_{1m})$, S_2 - есть перестановка S_1 , тогда $f(S_1) = f(S_2)$.

§ V-монотонность.

$S_1 = (s_{11}, s_{12}, \dots, s_{1m})$, $S_2 = (s_{21}, s_{22}, \dots, s_{2m})$. Если $s_{1i} \leq s_{2i}$ для всех $i: 1 \leq i \leq m$, тогда $f(S_1) \leq f(S_2)$.

§ H-монотонность

$S_1 = (s_1, s_2, \dots, s_m)$, $S_2 = (s_1, s_2, \dots, s_{m+1})$. Если $s_{m+1} = 0$, тогда $f(S_1) = f(S_2)$, иначе $f(S_1) < f(S_2)$.

Предложенная авторами функция есть $f(S) = S(m) * s$, где m – количество исходных списков, $s_i = \frac{(K+1) * i}{K+i}$, $K = k_1 * (1-b) + b * |D| / avdl$, k_1, b - параметры, $|D|$ - длина документа (списка), $avdl$ - средняя длина документа (списка).

Примеры функций синтеза, основанных на вероятностном подходе, можно найти в работах [16, 17].

В работе [18] исследуется применимость методов синтеза, предложенных Edward A. Fox и соавторами для гетерогенных источников – видеопотоков. Экспериментально было показано, что функции синтеза CombSUM и CombMNZ подходят и для гетерогенных свидетелей. Данные функции работают в среднем лучше, чем при использовании взвешенного среднего или весовых коэффициентов, добавленных в формулы CombSUM и CombMNZ.

3. Постановка задачи

В данной работе рассматриваются задачи синтеза результатов различных методов поиска изображений с целью улучшения его качества. Все рассматриваемые задачи можно разделить на две группы. К первой относится поиск в частично аннотированной базе изображений по текстовому запросу. Ко второй – различные варианты синтеза результатов поиска по низкоуровневым характеристикам (в данной работе рассматриваются цвет и текстура).

3.1 Постановка задачи поиска изображений в частично аннотированной базе

Дана база изображений. Для некоторых из них имеется текстовая аннотация (в нашем случае – набор ключевых слов). Возможно, эти аннотации не совсем точны и не полностью описывают изображение. Требуется реализовать механизм поиска по всей базе по текстовому запросу.

В отличие от автоматического аннотирования, был выбран путь комбинирования полнотекстового поиска по аннотациям и поиска по содержанию. Предлагается использовать следующую схему.

1. Использовать полнотекстовый поиск по аннотациям.
2. Использовать изображения, полученные на

первом шаге, в качестве образцов для поиска по содержанию. Тем самым образуется множество списков и каждому списку можно приписать вес равный рангу объекта образца при полнотекстовом поиске.

- Использовать метод синтеза ранжированных списков для комбинирования текстового поиска и поиска по содержанию.

В таком виде задача легко сводится к задаче синтеза.

Необходимо отметить, что данная схема поиска может быть расширена за счет использования более чем одного метода поиска по содержанию.

3.2 Задача синтеза результатов нескольких методов поиска по низкоуровневым характеристикам

Пусть имеется несколько независимых методов поиска по содержанию с использованием низкоуровневых характеристик. Задача заключается в том, чтобы изучить эффективность применения функций синтеза к результатам работы данных методов.

В нашем распоряжении имелись следующие методы поиска по низкоуровневым характеристикам:

- § метод поиска по цветовым характеристикам, основанный на статистическом представлении информации о цвете [25] – *ColorMoment*;
- § метод поиска по цветовым характеристикам, основанный на цветовых гистограммах [1] – *ColorHist*;
- § метод поиска по текстуре, использующий фильтры ICA [5] – *ICAHist*.

Ставилась задача произвести и проанализировать попарное смешивание результатов каждого из данных методов поиска. Предполагалось, что в среднем комбинированный поиск будет давать более удачные ответы на запрос, чем исходные методы поиска по отдельности.

4. Функция синтеза WTGF

В разделе 3 были поставлены задачи синтеза результатов поиска в частично аннотированной базе данных изображений (подраздел 3.1) и синтеза результатов нескольких поисковых методов по низкоуровневым характеристикам (подраздел 3.2).

В следующем разделе мы выделим основные свойства функции синтеза, подходящей для наших задач.

Для простоты изложения введем некоторые обозначения.

Ранжированный i -ый список будем обозначать символом α_i .

Ранжированный список состоит из пар вида $(x, r_x^{(ai)})$, где x – объект, а $r_x^{(ai)}$ – его ранг в списке α_i .

После синтеза ранг элемента x в результирующем списке будем обозначать как $r_x^{(0)}$.

$x \in \alpha_i$ означает, что элемент x принадлежит

какой-то паре из α_i и его ранг в этом списке не равен нулю.

4.1 Свойства функции синтеза

Метод синтеза будет эффективен только в том случае, когда он настроен под задачи целевой системы. Из всего множества предлагаемых в литературе способов синтеза мы не нашли метода, полностью удовлетворяющего всем условиям нашей задачи. Критерием отбора служило наличие или отсутствие следующих специфических ограничений на функцию синтеза, полученных из неформальных требований к нашей задаче:

- § для каждого метода поиска определено значение параметра, выражающее степень его значимости при формировании общего результата (иначе говоря, вес списка результатов поиска);
- § метод синтеза должен параметризоваться для смешивания как неоднородных (гетерогенных)², так и однородных (гомогенных) списков;
- § результат работы метода синтеза – ранжированный список – сам может являться объектом синтеза и участвовать в дальнейшем смешивании;
- § область задания объектов для всех списков одинакова ($x \in \alpha_i \Rightarrow x \in X$, где X – пространство изображений, и для $x \in X \Rightarrow \alpha_i, x \in \alpha_i \wedge \exists(x, 0) \in \alpha_i$);
- § каждый список рассматривается как особая точка зрения, которая дополняет общую картину, но не исключает мнения других источников (предполагается, однако, что имеется вес списка, который может учитываться как степень доверия той или иной точке зрения).

Приведенные выше неформальные требования были учтены в виде формальных свойств искомой функции.

Некоторым из требуемых свойств удовлетворяют и методы, рассматриваемые в литературе. Так, например, метод HSC3D удовлетворяет следующим свойствам.

- § Свойство симметричности

$R_1 = (r_x^{(a_1)}, r_x^{(a_2)}, \dots, r_x^{(a_m)})$, R_2 – есть перестановка R_1 , тогда $f(R_1) = f(R_2)$.

- § Свойство V-монотонности

$R_1 = (r_x^{(a_1)}, r_x^{(a_2)}, \dots, r_x^{(a_m)})$,

$R_2 = (r_y^{(a_1)}, r_y^{(a_2)}, \dots, r_y^{(a_m)})$, тогда из того что

$r_x^{(a_i)} \leq r_y^{(a_i)}$ для всех $i \in \{1 \dots m\} \Rightarrow f(R_1) \leq f(R_2)$

² Под неоднородностью в данном случае понимается то, что списки могут быть получены при помощи различных методов поиска по содержанию. Однако, при этом они должны быть сопоставимы между собой в соответствии с указанными весами, а также должны быть определены на всем множестве изображений (или для них должен существовать соответствующий проектор на рассматриваемое множество)

$f(R_2)$.

§ Свойство H-монотонности

$$R_1 = (r^{(a_1)}, r^{(a_2)}, \dots, r^{(a_m)}),$$

$$R_2 = (r^{(a_1)}, r^{(a_2)}, \dots, r^{(a_{m+1})}), \text{ тогда если}$$

$$r^{(a_{m+1})} \neq 0, \text{ то } f(R_1) < f(R_2)$$

Свойства симметричности и V-монотонности являются необходимыми свойствами и для метода синтеза, предлагаемого нами далее. Необходимыми, но не достаточными.

HSC3D имеет также некоторые особенности реализации, основанные на свойствах текста. Вследствие чего, функции типа HSC3D не применимы в чистом виде для решения поставленных нами задач.

Таким образом, необходимо найти функцию, удовлетворяющую следующему расширенному набору свойств. Она должна быть симметричной относительно своих переменных (списков), поскольку все переменные с одинаковым весом для нас равнозначны. Также должна быть возможность определить функцию от нескольких переменных (возможно, заранее не известно их точное количество) как суперпозицию вложенных функций от 2-ух переменных. Иначе говоря, вводя дополнительные ранжированные списки (результаты работы некоторых методов поиска) мы уточняем решение, а не строим его заново. Итак, первыми из необходимых свойств нашей функции, являются:

1. свойство симметричности:

$$R_1 = (r_x^{(a_1)}, r_x^{(a_2)}, \dots, r_x^{(a_m)}), R_2 - \text{ есть}$$

перестановка R_1 , тогда $f(R_1) = f(R_2)$;

2. свойство композиции:

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

На основании вышесказанного дальнейшие условия будем вводить, используя 2 списка. Обозначим их, как α_1 и α_2 с рангами $r_x^{(a1)}$ и $r_x^{(a2)}$ соответственно. Будем считать, что объект x присутствует в обоих списках (если это не так, его всегда можно добавить в список с нулевым рангом).

Введем следующие естественные условия.

3. Будем считать, что для определения итогового ранга элемента x необходимо получить на вход набор весов списков и значения рангов x в каждом из них. Иначе говоря, результирующее значение ранга объекта x не зависит от рангов других объектов.

4. Пусть имеется N списков: $\alpha_1, \alpha_2, \dots, \alpha_N$. Элементу x в них сопоставляется следующий набор рангов:

$$r_x^{(a_1)}, r_x^{(a_2)}, \dots, r_x^{(a_N)}.$$

Тогда после синтеза ранг $r_x^{(0)}$ должен удовлетворять условию:

$$\min\{r_x^{(a_1)}, r_x^{(a_2)}, \dots, r_x^{(a_N)}\} \leq r_x^{(0)} \leq$$

$$\max\{r_x^{(a_1)}, r_x^{(a_2)}, \dots, r_x^{(a_N)}\}$$

Взаиморасположение объектов в результирующем списке определяется после того, как определен ранг каждого элемента.

4.2 Функция стабилизации высокоранговых элементов и правила конусов

Перед тем, как продолжить формализацию требований, необходимо ввести свойство взвешенной стабилизации высокоранговых элементов.

Пусть имеется некоторый список α . Будем рассматривать его вес $w(a)$ как показатель уверенности в том, насколько значения, приведенные в данном списке, соответствуют запросу.

Для каждого элемента $x \in \alpha$ можно рассмотреть функцию, зависящую от веса списка $w(a)$ и от ранга элемента в списке $r_x^{(a)}$: $g(r_x^{(a)}, w(a))$.

Эту функцию мы назвали функцией стабилизации высокоранговых элементов или гравитационной функцией. Она определена в области $[0..1, 0..1]$, непрерывна и монотонна, убывает по обоим параметрам. Ее задача заключается в выработке компромисса между мнениями различных источников с учетом как рангов элементов, так и весов самих источников.

Для того, чтобы связать абстрактную функцию стабилизации высокоранговых элементов с функцией синтеза введем следующие правила (мы назвали их "правилами конусов"³).

§ Выполнение граничных условий.

- Если есть список с весом, близким к единице, и в нем есть объект с рангом, почти равным единице, то вероятность того, что ранг изменится после синтеза, очень мала;
- При смешивании двух списков, один из которых обладает весом, почти равным 0, таковой почти не вносит вклада в результат (за исключением добавления новых элементов). При этом 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 ($r_x^{(a_1)} = r_y^{(a_2)}$), то в результирующем списке их ранги должны быть различны, причем $r_x^{(0)} > r_y^{(0)}$.

§ Степень свободы низкоранговых элементов. Чем меньше ранг элемента, тем больше

³ Идея почерпнута из правила метода HSC3D, названного авторами гравитационным правилом в 3D пространстве. В чистом виде данное правило не применимо для нашей задачи, однако, в названии предлагаемого нами метода от него сохранена буква G – gravitation.

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.

Теперь мы можем добавить новое условие к списку свойств функции синтеза.

5. Для функции синтеза должны выполняться "правила конусов".

Пункты (1)-(4) не учитывают веса самих списков. Пункт (5) устраняет этот недочет.

Рассмотрим следующую функцию:

$$r_x^{(0)} = \max\{r_x^{(a_1)}, r_x^{(a_2)}, \dots, r_x^{(a_N)}\} \quad (4)$$

Она удовлетворяет условиям симметричности и композиции. Также, естественным образом соблюдаются условия (3, 4). Однако условие (5) явно нарушается, так как веса списков никак не учитываются.

Рассмотрим функцию взвешенного среднего для вектора, которая удовлетворяет свойствам (1)-(4):

$$f(R) = \frac{\sum_i r_x^{(a_i)} \cdot w(a_i)}{\sum w(a_i)}, \quad (5)$$

где $R = (r_x^{(a_1)}, r_x^{(a_2)}, \dots, r_x^{(a_m)})$, $w(a_i)$ – вес списка α_i .

Для выполнения свойства (5) заменим вес элемента $w(a_i)$ на значение функции стабилизации высокоранговых элементов $g(r_x^{(a_i)}, w(a_i))$.

$$f(R) = \frac{\sum_i r_x^{(a_i)} \cdot g(r_x^{(a_i)}, w(a_i))}{\sum g(r_x^{(a_i)}, w(a_i))} \quad (6)$$

В качестве функции g возьмем следующую функцию:

$$g(r_x^{(a_i)}, w(a_i)) = w^2(a_i) \cdot \left(r_x^{(a_i)} + \frac{1}{12} \right)^4 \quad (7)$$

Такая форма объясняется условиями нашей задачи:

§ зависимость ранга и веса должна быть мультипликативной;

§ мы не должны учитывать значение списка с нулевым весом ($w \equiv 0$), элемент сколь высокого ранга в нем бы не содержался;

§ ситуация, когда $r_x^{(a_i)} \equiv 0$, то есть объект x не присутствует в списке α_i , тоже должна оказывать на результат некоторое воздействие, а именно, уменьшать значение результирующего ранга элемента, если вес списка не равен нулю;

§ функция должна быть монотонна и непрерывна по обоим параметрам.

Константы степеней (2,4) и сдвига ранга ($\frac{1}{12}$) определяются экспериментальным образом, исходя из особенности данных и конкретных применяемых методов.

Получившейся функции была присвоена аббревиатура *WTGF* (weighted total with gravitation function – взвешенное среднее с гравитационной функцией). Она разрешает приведенные ниже проблемы.

4.3 Реализация вычислений

Синтез можно осуществлять двумя способами.

Вариант 1. Последовательно использовать стандартную формулу для одинаковых элементов. Ниже данный вариант представлен схематически.



Рис. 1. Последовательное смешивание

Вариант 2. Смешать списки попарно, затем повторить ту же процедуру для получившихся результатов и т.д. Схема данного подхода представлена ниже.

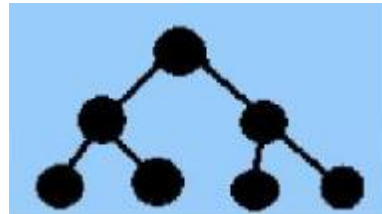


Рис. 2. Попарное смешивание. Круги – смешиваемые списки.

У обоих вариантов есть свои плюсы и минусы, оба поддаются распараллеливанию. Основным минус первого варианта заключается в том, что поиск одинаковых объектов в разных списках работает медленно и сложно реализуем, если реализовывать его оптимально. Основным минус второго – хранение "лишних" переменных.

На данный момент нами используется синтез на основе слияния отсортированных списков (вариант 2'). Основные шаги алгоритма следующие:

1. производится сортировка элементов списков по идентификатору;
2. в конец добавляется ограничивающий (максимальный) элемент;
3. производится сортировка слиянием путем соединения двух списков по приведенным выше формулам.

Такая реализация добавляет к методу *WTGF* суффикс *MT* (Merge by Two).

В целях распараллеливания вычислений используемую формулу (6) можно преобразовать к виду последовательного смешивания двух списков.

$$r_x^{(0)} = \frac{g(r_x^{a_1}, w(a_1)) \cdot r_x^{a_1} + g(r_x^{a_2}, w(a_2)) \cdot r_x^{a_2}}{g(r_x^{a_1}, w(a_1)) + g(r_x^{a_2}, w(a_2))}. \quad (8)$$

При этом вес результирующего списка будет вычисляться по формуле:

$$w_0 = (w_{a_1}^2 + w_{a_2}^2)^{\frac{1}{2}} \quad (9)$$

Вышеописанные алгоритмы были реализованы с использованием продуктов Microsoft: база данных была размещена на MSSQL Server 2005 (предоставлен MSDAA), алгоритмы написаны на языке C#. Для интеграции с MSSQL Server 2005 использовались возможности среды .NET. Код разрабатывался на тестовой версии Visual Studio 2005.

5. Методы оценки

После анализа существующей литературы были предложены следующие способы оценки методов поиска изображений, использующих алгоритмы синтеза независимых оценок.

5.1 Relevance/Nonrelevance overlap evaluation

Метод оценки с использованием коэффициентов перекрытия релевантных/нерелевантных объектов [14], благодаря своей простоте и достаточной показательности, применяется для оценки эффективности функций синтеза уже давно.

Для того чтобы вычислить коэффициент перекрытия релевантных изображений $R_{overlap}$ (формула 2, разд. 2.3) и коэффициент перекрытия нерелевантных изображений $N_{overlap}$ (формула 3, разд. 2.3) необходимо сосчитать количество изображений, релевантных и не релевантных запросу, во входных и результирующем списках.

Для получения этих данных мы предлагаем использовать следующий эксперимент. Все списки, как до слияния, так и после, объединяются в теоретико-множественном смысле. Полученное единое множество предлагается ассессору для оценки. На основании этой оценки вычисляются коэффициенты $R_{overlap}$ и $N_{overlap}$.

5.2 Mark Recapture evaluation

Помимо основной аннотации изображениям присваиваются дополнительные аннотации следующими возможными способами.

§ Ассессор вводит текстовый запрос и оценивает множество изображений, полученное в результате поиска. На основании этой оценки тем изображениям, которые, по мнению ассессора, наиболее соответствуют запросу, в качестве дополнительной аннотации присваивается введенный ассессором текст.

§ Случайным образом из базы выбирается несколько изображений, и наиболее характерные из них аннотируются вручную. Надо отметить, что такой способ довольно трудоемок и не дает такой кластеризации изображений по

дополнительным аннотациям, которая получается при использовании первого способа.

Предположим теперь, что в базе имеется некоторое количество изображений с одной и той же дополнительной аннотацией. Обозначим это множество через A . Одно из изображений, принадлежащих A , выбирается в качестве запроса для поиска с последующим синтезом по схемам, рассматриваемым в разделе 3. В случае полнотекстового поиска по аннотациям важно, чтобы дополнительная аннотация выступала только в роли запроса и не участвовала в самом поиске.

Качество каждого поиска в отдельности можно теперь оценить, подсчитав полноту и точность для изображений из A , попавших в результат. А произведя смешивание и вычислив для тех же изображений характеристики, указанные в предыдущем разделе, можно оценить качество алгоритма смешивания.

5.3 Fuzzy set evaluation

Пусть на вход методу синтеза подается N списков $\{\alpha_1, \alpha_2, \dots, \alpha_N\}$ с набором весов $\{w_1, w_2, \dots, w_N\}$. Каждый из них по очереди показывается ассессору в паре со списком-результатом. Для каждой пары ему предлагается ответить на вопрос: как изменилось качество и порядок изображений в результирующем списке по сравнению со входным? В качестве ответа нужно выбрать один из трех вариантов:

- § стало лучше;
- § не изменилось;
- § стало хуже.

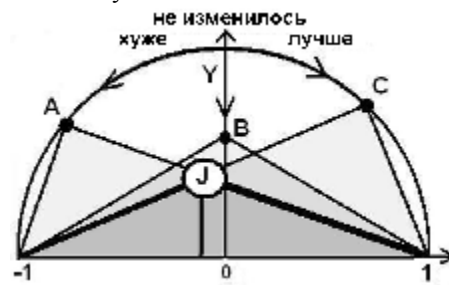


Рис. 3. Общий вид нечетких множеств

На основе ответов ассессора для каждого входного списка на полуокружности (см. рис. 3) выбирается одна из трех точек A , B или C по следующим правилам. Если ответ был "стало хуже", выбирается точка A , которая отстоит от верхней точки полусферы на $w_i \cdot K \cdot \frac{\rho}{2}$ радиан (радиус окружности = 1) в направлении против часовой стрелки, где $K = \frac{1}{\sum w_i}$. Для ответа "стало лучше" аналогично выбирается точка C , но в противоположном направлении. Если ответ был "не изменилось", то точка B располагается на расстоянии $w_i \cdot K$ от верхней точки полусферы по направлению к центру окружности.

Получившиеся треугольники (нечеткие

множества) всех списков пересекаются (пересечение обязательно должно быть по построению, учитывая, что нет списков с нулевыми весами), образуя новый треугольник. Аргумент его вершины – и есть требуемая оценка качества смешивания.

5.4 Competition statistics

Ассессору одновременно показываются результаты работы различных методов синтеза (для одних и тех же входных данных) и предлагается сравнить их между собой, присваивая им "места". После накопления некоторой статистики можно определить метод–лидер.

6. Описание экспериментов

6.1 Поиск по частично аннотированной базе

Для оценки эффективности различных функций синтеза применительно к решению задачи поиска по частично аннотированной базе по схеме, описанной в разделе 3.1, использовался метод *Relevance/Nonrelevance overlap evaluation*. Данный метод оценки прост в реализации, опробован ранее и, в то же время, наиболее эффективен для оценки именно метода синтеза, а не поиска в целом. Результат поиска зависит не только и не столько от метода синтеза, сколько от данных, подаваемых ему на вход.

В качестве экспериментальной базы использовалась база изображений, полученных с сайта фотогалереи flickr (<http://www.flickr.com>). В ней присутствуют как изображения, аннотированные ключевыми словами, так и изображения без аннотаций. Аннотирование производилось пользователями, загружающими изображения на сайт.

Поиск по содержанию производился с помощью метода поиска по цветовым характеристикам, предложенного в [1]. Для реализации полнотекстового поиска использовался компонент MSFTS корпорации Microsoft, включенный в поставку набора MSSQL Server 2005.

В ходе эксперимента была создана тестовая база однословных запросов. В качестве тестовых запросов было выбрано 100 наиболее часто встречающихся в базе ключевых слов.

В ходе эксперимента участнику показывалась страница, на которой был расположен случайный тестовый запрос и множество изображений, полученное в результате объединения (в теоретико-множественном смысле) первых N элементов из списков, подававшихся на вход методам синтеза, и первых N элементов результирующих списков для каждого из методов. Параметр N выбирался таким образом, чтобы ассессору показывалось не более 20 изображений. Изображения показывались в произвольном порядке.

Далее участнику эксперимента предлагалось выбрать один из вариантов ответа для каждого изображения:

- § не соответствует запросу (-1),
- § частично соответствует запросу (0),
- § соответствует запросу (1).

Внешний вид экспериментального стенда представлен на рис.4.



Рис. 4. Внешний вид экспериментального стенда

6.2 Синтез методов поиска по содержанию

Целью данного эксперимента было сравнить эффективность функций WTGF и CombMNZ применительно к синтезу результатов различных методов поиска по содержанию. В качестве таких методов были взяты методы *ColorMoment*, *ColorHist* и *ICAHist* (см. раздел 3.2).

Для оценки эффективности попарного смешивания результатов каждого из этих методов, была использована накопленная в ходе предыдущих экспериментов статистика по базе из 285 изображений, представляющих собой подмножество Corel Photo Set. Для изображений данной базы у нас имелись оценки ассессоров об их попарном сходстве.

7. Анализ результатов

7.1 Анализ активности ассессоров

На основе данных, полученных в эксперименте по оценке метода WTGF применительно к задаче поиска по частично аннотированной базе, была собрана статистика по ответам участников, позволяющая проследить интересные зависимости их поведения от времени суток. Результаты анализа данной статистики могут быть учтены во время проведения подобных экспериментов.

Эксперимент продолжался 2 недели (14 дней). За это время на сайте зарегистрировалось 136 человек. Внесли свой вклад в эксперимент (хотя бы один раз прошли тест) 98 человек. Всего было поставлено 484 опыта. Таким образом, в среднем, каждый участник дал примерно 5 ответов (точнее 4.93877551).

Распределение количества ассессоров, работавших с тестовым стендом, по времени суток и датам почти равномерно. В основном, опрашиваемые являлись русскоговорящими пользователями интернет. Возрастная категория опрашиваемых пользователей – от 18 до 35 лет.

Сутки были разбиты на 3 временных интервала по 8 часов:

- § 07:00 – 15:00 (утреннее и дообеденное время);
- § 15:00 – 23:00 (послеобеденное время и вечер);
- § 23:00-07:00 (ночь).

Как показал анализ оценок ассессоров, результаты эксперимента зависят от времени суток. Утром оценки более «пессимистичны», чем в послеобеденное время, а ночью самые «оптимистичные». Ночью участники скорее укажут «частично соответствует» или не укажут значение вообще, чем поставят отрицательную оценку.

Об этом можно судить по сводной гистограмме, представленной ниже на рис. 5. Гистограмма составлена следующим образом:

- § выбирались лишь те значения эксперимента, которые соответствовали интересующему временному интервалу;
- § для каждого результата эксперимента (один запрос, оцененный одним человеком), считался средний ответ ассессора $ans \in [-1...1]$ по всем имеющимся изображениям, подаваемым на оценку;
- § полученные коэффициенты ans использовались для построения гистограммы (количество экспериментов со значением ans).

Для гистограмм были построены аппроксимирующие нормальные распределения. Заметно, что значения сильно расходятся в зависимости от времени суток.

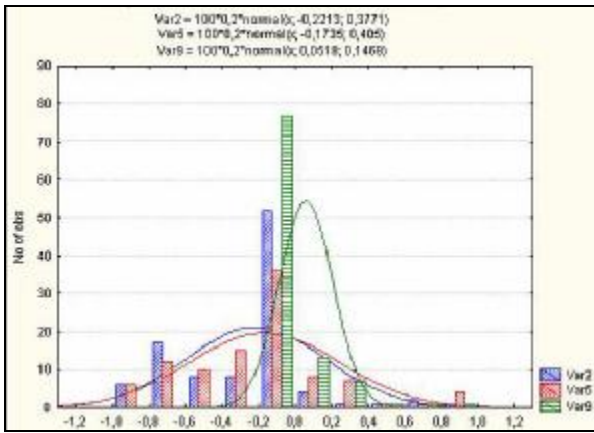


Рис. 5. Количество и средние значения ответов пользователей для трех временных интервалов.

Ниже приводится расшифровка обозначений гистограммы:

- Var2 – утреннее и дообеденное время;
- Var5 – послеобеденное время и вечер;
- Var9 – ночь.

7.2 Метод оценки алгоритмов синтеза

Для вычисления коэффициента перекрытия ($R_{overlap}$) изображений, соответствующих запросу, использовалась следующая формула:

$$R_{overlap}(x) = \frac{M \cdot R^{(0)}(x)}{\sum_i R^{(a_i)}(x)}, \quad (10)$$

где $R^{(a_i)}(x)$ – количество объектов, соответствующих запросу, в первых x элементах списка α_i , усредненное по всем различным объектам; $R^{(0)}(x)$ – среднее количество объектов, соответствующих запросу, в первых x элементах результирующего списка, также усредненное по всем различным объектам.

Если этот коэффициент больше единицы, значит, в результирующем списке больше релевантных объектов, чем в среднем в тех списках, которые подавали на вход функции синтеза. Чем больше данный коэффициент, тем лучше.

Второй коэффициент, который использовался для оценки качества методов синтеза, – коэффициент перекрытия изображений, нерелевантных запросу ($N_{overlap}$). Он вычислялся по формуле, аналогичной формуле (10).

$$N_{overlap}(x) = \frac{M \cdot N^{(0)}(x)}{\sum_i N^{(a_i)}(x)}, \quad (11)$$

где $N^{(a_i)}(x)$ – количество объектов, несоответствующих запросу, в первых x элементах списка α_i , усредненное по всем различным объектам; $N^{(0)}(x)$ – количество объектов, несоответствующих запросу, в первых x элементах результирующего списка, также усредненное по всем различным объектам. Таким образом чем меньше коэффициент $N_{overlap}(x)$, тем лучше работает метод.

Однако, относительно данных коэффициентов стоит сделать одну существенную оговорку. Поскольку ассессору для оценки предлагалось лишь ограниченное число изображений, то точно вычислить коэффициенты $N^{(a_i)}(x)$ и $R^{(a_i)}(x)$ возможно было только для нескольких первых элементов списка. Согласно описанию эксперимента, участнику показывались объединенные в теоретико-множественном смысле результаты сравниваемых методов синтеза и первые N элементов от Z-обхода⁴ списков, подававшихся методам на вход.

Таким образом, например, при смешивании 10 списков в среднем имелись данные только о первых двух элементах в исходных списках. Остальные элементы оставались вне поля зрения пользователя.

Тем не менее, следует отметить, что в большинстве случаев удавалось получить полную оценку по всем 20 первым элементам, так как результирующие и исходные списки сильно перекрывались.

Следует также учитывать, что кроме оценок «соответствует запросу» и «не соответствует запросу», была введена оценка «частично соответствует запросу».

⁴ Под Z-обходом подразумевается обход сперва всех первых элементов всех списков, далее всех вторых, потом всех третьих и т.д.

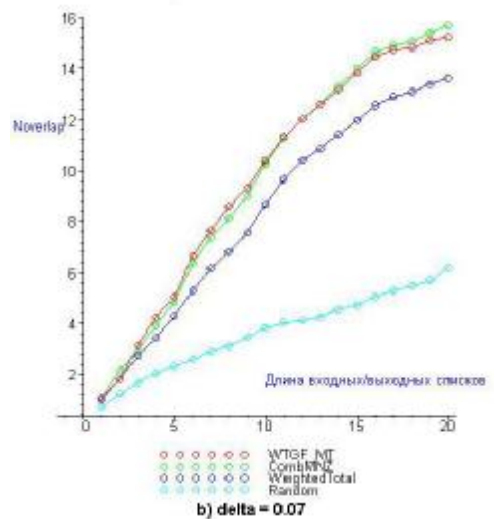
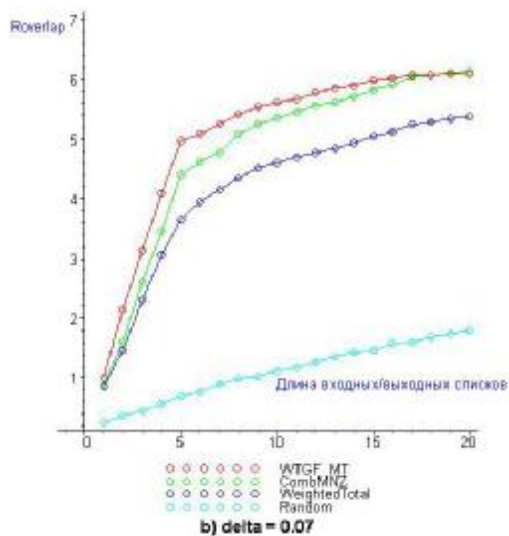
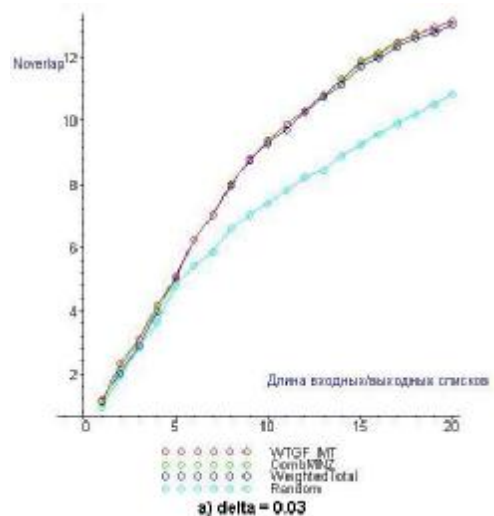
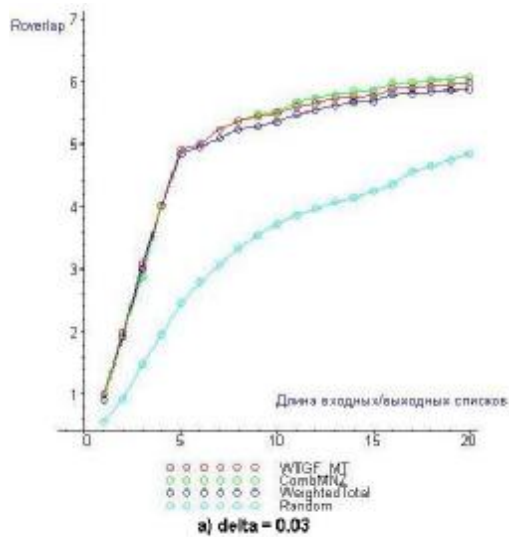


Рис. 6. Графики значений $R_{overlap}$ для различных функций синтеза в зависимости от размера списков (для 5 входных списков).

Рис. 7. Графики значений $N_{overlap}$ для различных функций синтеза в зависимости от размера списков (для 5 входных списков).

7.3 Результаты эксперимента: поиск по частично аннотированной базе

Прежде чем приступить к анализу результатов данного эксперимента, необходимо пояснить смысл коэффициента δ , параметризующего поиск по цветовым характеристикам. Он означает допустимое отступление от изображения-образца по используемой метрике. Например, если $\delta=0.03$, для эксперимента выбираются изображения, которые по расстоянию в заданной метрике отличаются от образца не более чем на 0.03.

На рис. 6, 8 представлены графики значений коэффициента $R_{overlap}$ для различных функций смешивания в зависимости от количества рассматриваемых элементов от начала списков для 5 и 10 смешиваемых списков соответственно. На графиках 7, 9 – зависимость коэффициента $N_{overlap}$ от количества рассматриваемых элементов для 5 и 10 списков.

На рис. 6(a), 7(a), 8(a), 9(a) представлены результаты эксперимента при смешивании списков с параметром $\delta = 0.03$. На рис. 6(b), 7(b), 8(b), 9(b) представлены результаты эксперимента при смешивании списков с параметром $\delta = 0.07$.

Как можно видеть, при 10 смешиваемых списках предложенный метод *WTGF_MT* выигрывает у методов: *Random*, *WeightedTotal*, *CombMNZ*. Однако, можно заметить, что при различных конфигурациях системы отрыв меняется. Чем больше коэффициент δ , тем лучше работает поиск, что говорит в среднем о небольшом количестве действительно близких изображений для запросов, участвовавших в эксперименте.

Графики на рис. 6, 8 ($R_{overlap}$) показывают общее превосходство метода *WTGF_MT* над методами *CombMNZ*, *WeightedTotal*, *Random*, *WeightedTotal*.

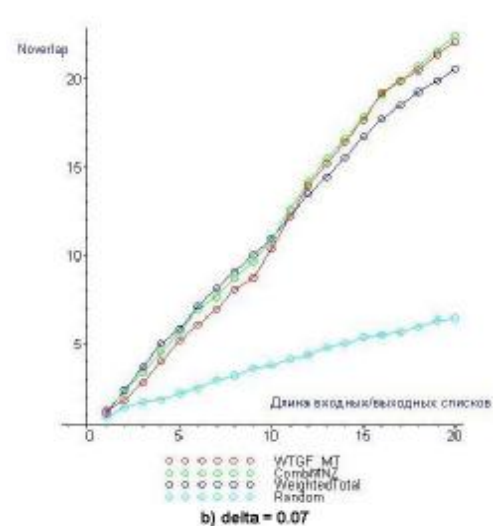
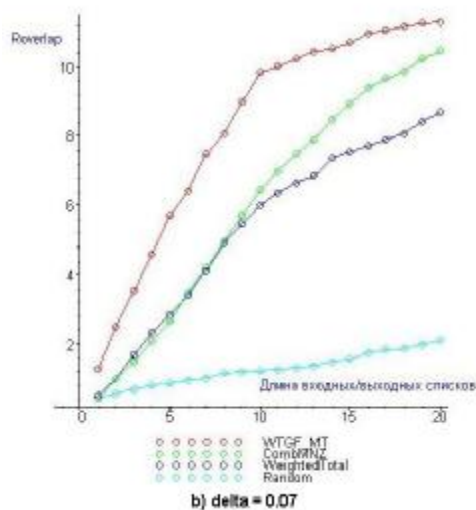
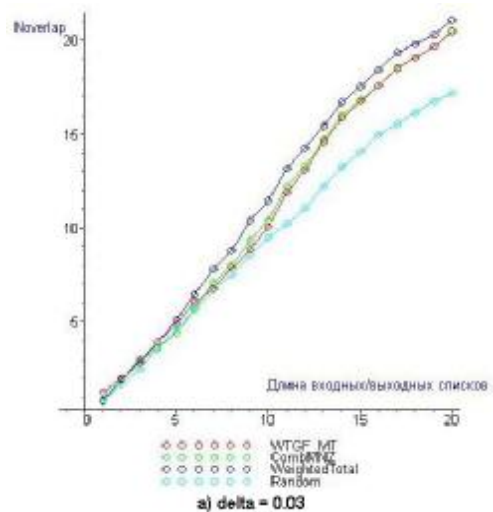
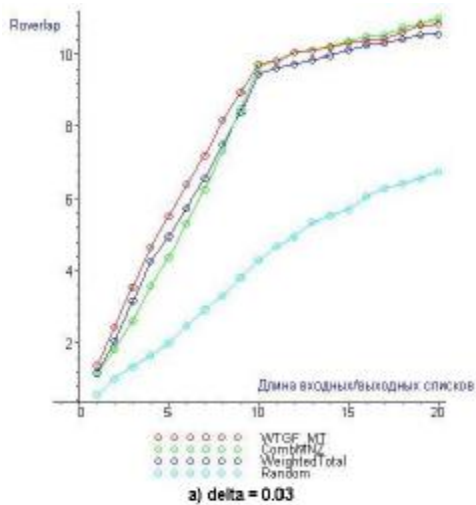


Рис. 8. Графики значений $R_{overlap}$ для различных функций синтеза в зависимости от размера списков (для 10 входных списков).

Рис. 9. Графики зависимости значений $N_{overlap}$ для различных функций синтеза в зависимости от размера списков (для 10 входных списков).

Ситуация по коэффициенту $N_{overlap}$, которую можно увидеть на графиках на рис. 7 и рис. 9, несколько хуже для метода *WTGF_MT*. Самым оптимальным по показателю $N_{overlap}$ является график функции *Random*, однако, эти данные нельзя считать правдоподобными. Такое положение графика обуславливается тем, что для случайного метода оказывалось слишком мало оцененных элементов в выходном списке.

Также была собрана статистика по количеству релевантных и нерелевантных изображений в зависимости от позиции в тексте. На рис. 10 представлен график релевантных изображений при 10 смешиваемых списках. На графике видно, что сначала лидирует метод *WTGF_MT*. Таким образом, для метода *WTGF_MT* свойственна тенденция "перемещать" релевантные изображения к началу списка в большей степени, чем всем остальным методам.

Таким образом, по различным показателям метод *WTGF_MT* действительно можно считать эффективным.

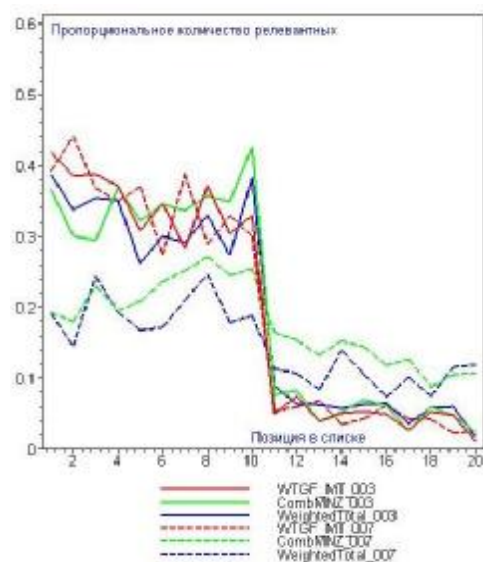


Рис. 10. Графики зависимости количества релевантных изображений от позиции при 10 входных списках

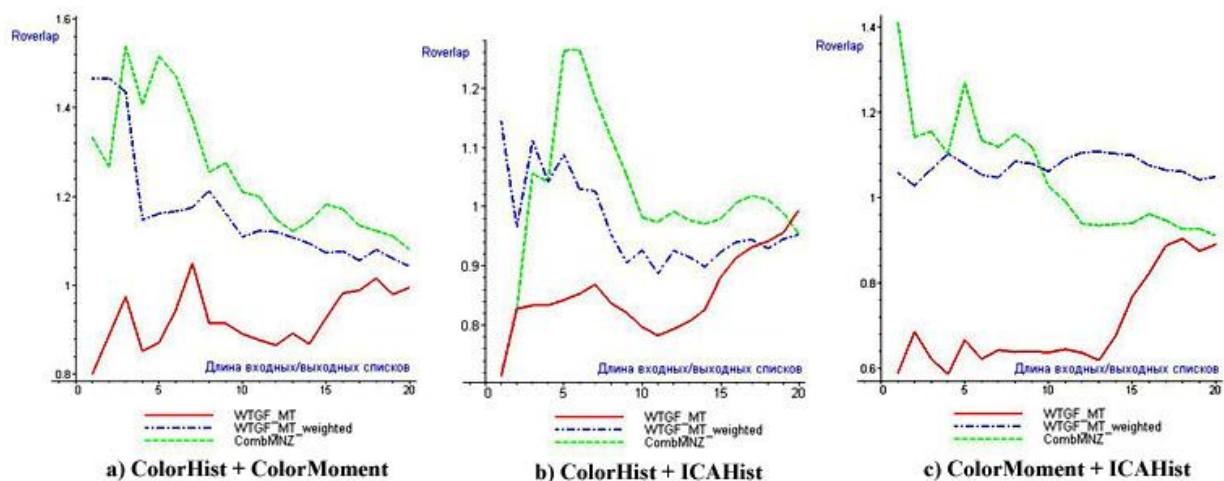


Рис. 11 Графики зависимости значений $R_{overlap}$ от размера списков для различных функций синтеза применительно к различным методам поиска по содержанию: а) синтез ColorHist и ColorMoment; б) ColorHist и ICAHist; в) ColorMoment и ICAHist.

7.4 Результаты эксперимента: синтез методов поиска по содержанию

Для оценки эффективности синтеза методов поиска по содержанию с использованием функций WTGF и CombMNZ также использовались коэффициенты $R_{overlap}$ и $N_{overlap}$. На вход функциям синтеза подавалось два ранжированных списка – результаты работы двух методов поиска по одному и тому же изображению-запросу, случайным образом выбранному из базы. Далее, на основе имеющихся данных о релевантности изображений базы друг другу, вычислялось количество изображений, релевантных (нерелевантных) запросу в исходных списках и в результирующем списке, полученном с помощью функции синтеза. Усредненные значения этих параметров использовались для вычисления коэффициентов $R_{overlap}$ и $N_{overlap}$.

Функция WTGF_MT участвовала в каждом эксперименте (синтезе каждой пары методов) дважды. Один эксперимент проводился с равными весами для обоих смешиваемых методов, в рамках второго – для участвующих в эксперименте методов поиска экспериментальным образом подбирались веса, улучшающие результат синтеза.

Зависимости коэффициента $R_{overlap}$ от размера смешиваемых списков представлены на графиках на рис. 11. Здесь релевантными запросу считались изображения, отмеченные ассессорами в среднем как «соответствующие» или «частично соответствующие» запросу. Веса для функции WTGF_MT_weighted представлены в таблице ниже.

Табл.1. Веса для WTGF_MT_weighted.

ColorHist + ColorMoment	0.01	0.19
ColorHist + ICAHist	0.02	0.25
ColorMoment + ICAHist	0.612	0.012

На графиках видно существенное превосходство метода CombMNZ над методом WTGF без использования весов. Использование весов существенно повышает эффективность функции WTGF, однако не делает ее безусловным лидером. Данный результат можно объяснить характером данных, на которых производился эксперимент, особенностями функций CombMNZ и WTGF и методов поиска по содержанию. Скачкообразный характер графиков и постепенное убывание коэффициента $R_{overlap}$ объясняется тем, что порядок элементов во входных списках, полученных с помощью методов поиска по содержанию, не соответствует степени релевантности элементов запросу. Методы поиска по содержанию работают на основе низкоуровневых характеристик, что позволяет выбирать изображения, похожие по цвету или текстуре на запрос, но это не гарантирует семантическую схожесть ответа с запросом. В то же время человеку свойственно оценивать схожесть изображений именно на основе семантики. Нельзя сказать, что методы поиска по содержанию не работают вовсе – в среднем среди элементов, близких к запросу по содержанию, больше элементов близких к запросу и по семантике. Но соотношение рангов, полученных с помощью методов поиска по содержанию, может не соответствовать степени релевантности запросу. В то же время, изображения, оказавшиеся в обоих смешиваемых списках, с большей вероятностью окажутся релевантными запросу. Так как для функции CombMNZ характерно существенное повышение ранга элементов, присутствующих во всех входных списках, то она показывает лучший результат по сравнению с WTGF_MT без весов.

Для большинства изображений базы, на которой ставился эксперимент, цветовая характеристика играет более существенную роль, чем текстура. Об этом говорят результаты работы

[2], в рамках которой и была собрана статистика по оценкам ассессоров. Этим можно объяснить улучшение результата работы WTGF_MT_weighted по сравнению с WTGF_MT.

На основе вышесказанного, можно предположить, что в условиях возможной недостоверности источников результат функций синтеза будет тем лучше, чем больше исходных списков, при условии их пересечения. В таком случае, преимущество получают те элементы, которые попали в большее число исходных списков. Действительно, если вернуться к результатам первого эксперимента, то видно, что смешивание 10 списков дает лучший результат, по сравнению со смешиванием 5. В данном эксперименте на вход функциям синтеза подавалось всего 2 списка.

В то же время, нельзя сказать, что использование функций синтеза в данной ситуации не улучшило результат поиска. На графиках видно, что при подсчете коэффициента для первых N элементов, где $N \leq 10$, значение $R_{overlap}$ для функций CombMNZ и WTGF_MT_weighted больше единицы. Это говорит о том, что все же функциям синтеза удалось переместить релевантные элементы ближе к началу результирующего списка. Лучший результат наблюдается при синтезе результатов двух методов по цвету. Это объясняется большой степенью перекрытия исходных списков.

7.5 Обсуждение результатов

Эксперименты показали превосходство функции WTGF над функциями CombMNZ, CombSUM и взвешенным средним при решении задачи поиска изображений по частично аннотированной базе. Для решения данной задачи использовался синтез нескольких источников. С увеличением числа источников метод поиска давал лучшие результаты и увеличивался отрыв функции WTGF от прочих функций синтеза. Среди особенностей входных данных для функций синтеза в этом эксперименте можно отметить невысокую степень перекрытия различных исходных списков и наличие в каждом из них одного элемента с достаточно достоверным высоким рангом (результат текстового поиска по аннотациям, выступавший в роли запроса при формировании соответствующего списка).

Второй эксперимент по проверке эффективности методов синтеза для комбинирования результатов работы методов поиска по содержанию показал важность использования весов входных списков для функции WTGF. В данном эксперименте можно отметить лидерование метода CombMNZ. Особенностью входных данных в этом случае является невысокая достоверность информации о рангах элементов во входных списках, а также в случае синтеза двух методов поиска по цвету большая степень перекрытия исходных списков.

Именно для этого случая методам синтеза в большей степени удалось улучшить результаты поиска.

Таким образом, можно сделать вывод, что функция WTGF выигрывает у CombMNZ при условии достоверности информации о рангах элементов во входных списках и небольшом перекрытии входных списков. Использование CombMNZ дает лучший результат при достаточном перекрытии входных списков и возможной недостоверности рангов элементов.

В целом, использование методов синтеза для решения поставленных нами задач оправдало себя – это позволило улучшить результаты поиска.

8. Заключение

Мы рассмотрели две схемы использования методов синтеза применительно к задаче поиска изображений: использование синтеза для реализации поиска по частично аннотированной базе с текстовым запросом и синтез методов поиска по цветовым и текстурным характеристикам.

Нами были сформулированы требования к функции синтеза, подходящей для решения задачи поиска изображений. Анализ наиболее распространенных на сегодняшний день методов синтеза показал, что они не удовлетворяют всем сформулированным требованиям. В частности, они не учитывают веса исходных множеств, которые можно трактовать как степень доверия к тому или иному источнику информации. Нами была предложена новая функция WTGF (weighted total with gravitation function), удовлетворяющая данным требованиям.

С помощью экспериментов мы показали, в каких случаях более эффективна новая функция WTGF (по показателю $R_{overlap}$), а когда лучший результат дает классическая функция CombMNZ. В целом, можно отметить, что использование методов синтеза позволило улучшить результаты поиска по сравнению с результатами отдельно взятых источников.

Также в работе приводится обсуждение нескольких методов оценки эффективности синтеза применительно к задаче поиска изображений.

В дальнейшем мы планируем повторить эксперимент по синтезу результатов методов поиска по содержанию на базе фотографий Flickr. Возможно, изменение параметров функции WTGF позволит улучшить полученный результат.

9. Литература

- [1] Васильева Н. Построение соответствий между низкоуровневыми характеристиками и семантической статических изображений / Н. Васильева, Б. Новиков // Труды 7-ой Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные

- коллекции» - RCDL'2005, – Ярославль, Россия, 2005.
- [2] Марков И. Поиск изображений. Выбор оптимальных весов для слияния метрик по цвету и текстуре в зависимости от запроса-образца / И. Марков, Н. Васильева, А. Яремчук // Принято к публикации в трудах RCDL'2007
- [3] Bell A. J. The 'Independent Components' of natural scenes are edge filters / Antony J. Bell, Terrence J. Sejnowsky // *Vision Research*. – 1997. – No 37. – P. 3327 - 3338
- [4] Belkin N. J. Combining the evidence of multiple query representations for information retrieval / N. J. Belkin, P. Kantor, E. A. Fox, J. A. Shaw // *Inf. Process. Manage.* – 1995. – Vol. 31. – No 3. – P. 431-448.
- [5] Borgne H. Representation of images for classification with independent features / Borgne, A. Guerin-Dugue, A. Antoniadis // *Pattern Recognition Letters*. – 2004. – Vol. 25. – P. 141-154.
- [6] Chowdhury A. Analyses of multiple-evidence combinations for retrieval strategies / A. Chowdhury, O. Frieder, D. Grossman, C McCabe // *SI-GIR '01: Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*. – New York, NY, USA: ACM Press, 2001. – P. 394-395.
- [7] Deer P.J., Eklund P.W. On the Fusion of Image Features Eklund. [Electronic resource]. – Mode of access: <http://citeseer.ist.psu.edu/162546.html>
- [8] Field D. J. Relations between the statistics of natural images and the response properties of cortical cells / D. J. Field // *J Opt. Soc. Am.* – 1987. – No 4(12). – P. 2370 - 2393
- [9] Fox E. A. Combining evidence from multiple searches / E. A. Fox, M. P. Koushik, J. A. Shaw et al. // *TREC*. – 1992. – P. 319-328.
- [10] Gertz, M. Annotating scientific images: A concept-based approach [Electronic resource]. – Mode of access: <http://citeseer.ist.psu.edu/528338.html>
- [11] Gertz, M. Integrating scientific data through external, concept-based annotations, citeseer.ist.psu.edu/526357.html.
- [12] Ghoshal, A. Hidden markov models for automatic annotation and content-based retrieval of images and video / A. Ghoshal, P. Irc-ing, S. Khudanpur // *SIGIR '05: Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*. "— New York, NY, USA: ACM Press, 2005. "— Pp. 544-551.
- [13] Glotin, H. Fast image auto-annotation with visual vector approximation clusters / H. Glotin, S. Tollari // *Proc. of Fourth International Workshop on Content-Based Multimedia Indexing (CBMI2005)*. - 2005. - June.
- [14] Lee J. H. Analyses of multiple evidence combination / J. H. Lee // *SIGIR '97: Proceedings of the 20th annual international ACM SIGIR conference on Research and development in information retrieval*. –New York, NY, USA: ACM Press, 1997. – P. 267-276.
- [15] Li X. Image annotation by large-scale content-based image retrieval / X. Li, L. Chen, L. Zhang et al. // *MULTIMEDIA '06: Proceedings of the 14th annual ACM international conference on Multimedia*. – New York, NY, USA: ACM Press, 2006. – P. 607-610.
- [16] Lillis D. Probability-based fusion of information retrieval result sets / D. Lillis, F. Toolan, A. Mur et al. // *Proceedings of the 16th Irish Conference on Artificial Intelligence and Cognitive Science*. – University of Ulster: AICS 2005, 2005. – P. 147-156.
- [17] Lillis D. Probfuse: a probabilistic approach to data fusion / D. Lillis, F. Toolan, R. Collier, J. Dunning // *SIGIR '06: Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*. – New York, NY, USA: ACM Press, 2006. – P. 139-146.
- [18] McDonald K. A comparison of score, rank and probability-based fusion methods for video shot retrieval / K. McDonald, A. F. Smeaton // *CIVR*. – 2005. – P. 61-70.
- [19] Mojsilovic A. Capturing image semantics with low-level descriptors / A. Mojsilovic, B. Rogowitz // *Proceedings of International Conference on Image Processing, ICIP 2001*. – Thessaloniki, Greece, September 2001.
- [20] Salton G. Extended boolean information retrieval / G. Salton, E. A. Fox, H. Wu // *Commun. ACM*. – 1983. – Vol. 26. – No 11. – P. 1022-1036.
- [21] Shaw J. A., Fox E. A. Combination of multiple searches [Electronic resource]. – 1994. Mode of access: <http://citeseer.ist.psu.edu/fox94combination.html>
- [22] Shuming Shi. Latent additivity: Combining homogeneous evidence / Shuming Shi, Ruihua Song, Ji-Rong Wen // *Tech. rep. Microsoft Research*. – August 2006.
- [23] Stricker M. Similarity of Color Images / M. Stricker, M. Orengo // *Proceedings of the SPIE Conference*. – 1995. – Vol. 2420. – P. 381-392.
- [24] Stricker M. Color Indexing with Weak Spatial Constraints / M. Stricker, A. Dimai // *Proceedings of the SPIE Conference*. – 1996.
- [25] Stricker M. Spectral Covariance and Fuzzy Regions for Image Indexing / M. Stricker, A. Dimai // *Machine Vision and Applications*. – 1997. – Vol. 10. – P. 66-73.
- [26] Swain M.J. Color Indexing / M.J. Swain, D.H. Ballard // *International Journal of Computer Vision*. – 1991. – Vol. 7(1). – P. 11-32.
- [27] Vassilieva N. A similarity retrieval algorithm for natural images / N. Vassilieva, B. Novikov // *Proceedings of the Baltic BD-BIS'2004*. – Riga, Latvia: Scientific Papers University of Latvia, 2004. – Vol. 672.
- [28] Zeng, H. Learning to cluster web search results [Electronic resource]. – Mode of access: <http://citeseer.ist.psu.edu/701862.html>

Image Retrieval. Combining multiple search methods' results.

N. Vassilieva, A. Dolnik, I. Markov

Fusion techniques applied to content-based image retrieval domain are discussed. Two kinds of retrieval tasks (text query search in partly annotated database and classical content-based search) are stated. A set of specific conditions (including "cone conditions") are derived from stated tasks. New fusion function WTGF based on those conditions is proposed. The main advantage of the function is that it takes into account merged lists' weights.

Experimental results for both retrieval tasks are given. Experiments show that WTGF outperforms CombMNZ for some tasks. Color- and texture-based techniques are used for the experiments.

A set of techniques for fusion functions and search methods evaluation are also described. Relevance/nonrelevance overlap evaluation, mark recapture evaluation, fuzzy set evaluation are among them.