

УДК 004.9

В.А. Ковалев

ВЛИЯНИЕ ИСКАЖЕНИЙ И ФРАГМЕНТАЦИИ ИЗОБРАЖЕНИЙ-ОБРАЗЦОВ НА КАЧЕСТВО ПОИСКА ЦВЕТНЫХ ИЗОБРАЖЕНИЙ ПО СОДЕРЖАНИЮ

Рассматривается проблема поиска по образцу цветных изображений в больших базах данных с использованием дескрипторов, построенных на матрицах совместной встречаемости цветов. Известно, что качество функционирования соответствующих систем поиска изображений зависит не только от используемых методов и алгоритмов, но и от широкого круга различных «шумовых» факторов. Приводятся результаты экспериментального исследования влияния на качество поиска таких факторов, как изменение размеров изображений и их пространственное искажение, а также использование фрагментов исходных изображений в качестве образца поиска.

Введение

Появившиеся в 1990-х гг. методы и программные средства поиска цветных изображений по содержанию стали достаточно эффективным ответом на вызовы наступающей «цифровой эры» в области съемки, накопления, хранения, поиска, обработки и анализа цифровых изображений и видеоданных [1, 2]. Известные трудности в формировании запроса на поиск в неиндексированных базах изображений без использования словесных описаний, ключевых слов и каких-либо количественных признаков привели к тому, что одной из популярных парадигм в технологии поиска стал так называемый поиск по образцу, когда запрос формулируется по принципу «найди изображения, наиболее похожие на это».

Целью данной работы является экспериментальное исследование влияния на качество поиска цветных изображений таких факторов, как изменение размеров изображений и пространственные искажения, а также использование фрагментов (частей) изображений в качестве изображений-образцов для поиска. Работа имеет практическую направленность и не претендует на полноту охвата всего спектра возможных причин и параметров, оказывающих влияние на результаты поиска цветных изображений в базах данных по содержанию. Однако предполагается, что полученные в результате проведенных вычислительных экспериментов зависимости и количественные оценки могут оказаться полезными для инженеров и программистов, разрабатывающих соответствующие прикладные программные комплексы.

1. Материалы и методы исследования

1.1. Тестовая база изображений

Технические характеристики. Тестовая база состояла из 10 000 цветных (RGB) изображений относительно небольших размеров, имеющих прямоугольную форму с горизонтальной или вертикальной ориентацией. Горизонтальные размеры варьировали в пределах от 222 до 512 пикселей, а вертикальные – в пределах от 164 до 382. Средний размер (площадь) изображений был эквивалентен изображению квадратной формы размером 350×350 пикселей.

Содержание изображений. Используемая база изображений создавалась сотрудниками Фраунhoferовского института компьютерной графики (Германия) с целью тестирования эффективности разрабатываемых методов, алгоритмов и программных средств поиска изображений по содержанию общего характера. Для обеспечения полноценного тестирования содержание изображений не ограничивалось какой-либо предметной областью или тематикой. Кроме того, тестовые изображения не имели никакой специфики относительно предпочтительной цветовой гаммы, наличия только реальных или только искусственных объектов, использования стандартизованных условий съемки, доминирования определенных текстурных свойств, объектов какой-либо определенной формы и т. п. Используя бытовую лексику, содержание тестовой базы может быть охарактеризовано как «изображения всего чего угодно».

1.2. Способ описания содержания изображений

Для описания содержания изображений использовались матрицы совместной встречаемости цветов. Основываясь на общей концепции матриц совместной встречаемости, предложенной в работах [3, 4], и следуя методике конструирования конкретных их вариантов, изложенной в [4, 5], в качестве дескрипторов изображений были выбраны трехмерные матрицы совместной встречаемости пар цветов W_C типа «цвет – цвет – расстояние», которые могут быть определены как

$$W_C = // w_c(c(i), c(k), d_{i,k}) //,$$

$$w_c(c(i), c(k), d_{i,k}) = \text{card}\{ i, k \in R^2 \mid i \neq k, d_{i,k} = \text{round}(d(i, k)), d_{i,k} \leq D,$$

$$x_k = (x_i + \Delta x), y_k = (y_i + \Delta y), -D \leq \Delta x \leq D, 0 \leq \Delta y \leq D, \Delta y(D+1) + \Delta x > 0 \}.$$

В приведенном определении через (i, k) обозначена произвольная пара пикселей, идентифицируемая с помощью их индексов (номеров) i и k , которые расположены на расстоянии $d_{i,k}$ друг от друга. Параметры $c(i)$ и $c(k)$ соответствуют цветам этих пикселей, а w_c представляет собой количество (частоту встречаемости) пикселей с указанными цветами на расстоянии $d_{i,k}$. Весь диапазон RGB-цветов квантуется путем равномерного разбиения на заданное количество интервалов, а межпиксельное расстояние, округляемое до ближайшего целого значения, изменяется в пределах от 1 до D . Соответственно card обозначает количество, а round – операцию округления до ближайшего целого. Неравенства, представленные в последней строке определения, формализуют порядок перебора соседей k текущего пикселя i по принципу вперед и вниз от текущего. Дальнейшие детали об алгоритме перебора можно найти в работе [5].

В силу того что при традиционных восьми битах на каждый из RGB-каналов цветовое пространство имеет весьма большие размеры, на практике матрицы встречаемости цветов являются сильно разреженными, с большим количеством нулевых элементов. Это позволяет использовать достаточно грубые схемы квантизации с относительно небольшим количеством интервалов цветов. В данной работе использовалась схема с тремя битами на каждый из трех цветовых каналов, т. е. количество интервалов (квантов) цветов равнялось 512. Адаптивные схемы редуцирования цветового пространства не применялись, поскольку используемые изображения не имели доминирующих цветов. Таким образом, элементы матрицы w_c описывают «элементарные» цветовые сегменты, их границы и, опосредованно, пространственное расположение сегментов на изображении. На примере базы данных из 20 000 изображений было экспериментально показано [6], что такое представление является достаточно гибким и одинаково хорошо подходит для описания широкого класса сцен, начиная от простых комбинаций цветных объектов на однородном фоне и заканчивая высокочастотными цветными текстурами натурального происхождения.

Важно отметить, что при вычислении матриц совместной встречаемости рассматриваются все возможные пары в локальной окрестности каждого текущего пикселя. При этом пары с одними и теми же цветами, находящиеся на одних и тех же расстояниях, считаются идентичными независимо от порядка встречаемости цветов при переборе пикселей изображения. Фактически это означает, что для каждого из расстояний соответствующий слой результирующего трехмерного массива частот встречаемости представляет собой нижнюю треугольную матрицу. Завершающим шагом вычисления матриц является их нормализация, выполняемая путем деления всех элементов на их сумму. Перечисленные особенности алгоритма вычисления матриц приводят к тому, что дескрипторы совместной встречаемости обладают следующими важными свойствами:

- а) инвариантность по отношению к таким преобразованиям изображений, как поворот на угол, кратный 90° , и зеркальное отражение (из-за перебора всех возможных пар пикселей);
- б) сдвиг объектов изображений по однородному фону (из-за отсутствия координатной привязки);
- в) относительная независимость от размера, т. е. возможность сравнивать содержание изображений разного размера и формы (из-за нормализации матриц);
- г) возможность поиска по фрагментам изображений, что обусловлено такими факторами, как нормализация матриц и использование метрики $L1$, в которой вычисляется разность одноименных элементов матриц, т. е. относительных частот встречаемости некоторых цветов, представленных как на материнском изображении, так и на его фрагментах.

На заключительном этапе вычисления дескриптора каждого изображения набор ненулевых элементов соответствующей матрицы совместной встречаемости представляется в виде списка (<номер элемента матрицы>, <значение>) и записывается в файл дескрипторов в порядке возрастания номеров элементов, что обеспечивает их быстрый перебор при сравнении. Следует отметить, что на практике обычно игнорируются не только элементы матрицы, которые строго равны нулю, но и все «шумовые» элементы, значение которых меньше некоторого (достаточно малого) порога. Данный прием позволяет существенно сократить количество элементов, включаемых в дескриптор, что значительно уменьшает его размер и повышает скорость работы системы поиска изображений при условии обеспечения практически тех же самых результатов поиска. Так, например, при установке минимального значения порога, эквивалентного 0,2 % от площади изображения, среднее значение количества элементов матрицы, включаемых в дескриптор, равно 66 [6].

1.3. Сравнение дескрипторов изображений

Дескрипторы изображений, т. е. множества ненулевых элементов дескриптора изображения-образца поиска и каждого из дескрипторов базы данных, сравнивались с помощью метрики L1 [7], т. е. степень различия между изображениями R вычислялась как нормализованная сумма абсолютных значений разностей однотипных элементов сравниваемых дескрипторов. При отсутствии какого-либо элемента в дескрипторе его значение полагалось равным нулю. Соответственно степень близости изображений в пространстве признаков (степень сходства S) вычислялась как $S = 1 - R$. Следуя сложившейся традиции, для количественной оценки результатов поиска выводились и использовались только первые $N = 20$ изображений, наиболее близких к заданному образцу.

2. Поиск изображений при различных условиях

2.1. Типичные результаты поиска

На рис. 1 показаны примеры образцов поиска и $N = 3$ наиболее близких результатов, найденных в тестовой базе из 10 000 изображений. Последняя колонка иллюстрирует инвариантность результатов поиска по отношению к таким преобразованиям, как зеркальное отражение и поворот.



Рис. 1. Изображения-образцы (а) и $N = 3$ наиболее похожих результата поиска для каждого из них вместе с соответствующей количественной мерой сходства S (б)

Следует особо подчеркнуть, что при визуальной оценке результатов поиска похожих изображений в базах данных пользователи подобных сервисов интуитивно ожидают увидеть среди N наиболее близких изображений действительно «очень похожие» (в соответствии с их индивидуальными и, надо сказать, весьма различными представлениями о «похожести»). Вместе с тем очевидно, что в результате поиска в N наиболее близких будут включены только изображения из числа фактически присутствующих в используемой базе изображений. Соответственно в общем случае (и особенно в случае небольших баз изображений) совсем не обязательно, что они будут действительно «очень похожи» на заданный образец и будут отвечать ожиданиям пользователя.

2.2. Влияние пространственных искажений на результаты поиска

В качестве типичного примера пространственных искажений изображений в соответствующей серии экспериментов была использована операция их редуцирования к некоторому «стандартизованному» размеру: до небольших, в нашем случае 128×128 пикселей, миниатюр (preview), которые часто используются в интерфейсных целях в современных компьютеризированных сервисах (рис. 2). При этом информация об исходной портретной или пейзажной форме исходных изображений во внимание не принималась. Кроме приведенного варианта пространственных искажений в литературе описываются и такие важные случаи, как искажения изображений, связанные с различиями в освещении сцены [8], и некоторые другие. Однако, по мнению автора, в данной работе рассматривается гораздо более массовый и понятный способ пространственного искажения изображений из числа встречающихся на практике.



Рис. 2. Примеры исходных изображений и их искаженных (уменьшенных до 128×128 пикселей) версий

Схема вычислительных экспериментов по оценке влияния пространственных искажений на результаты поиска представлена на рис. 3, а их результаты – на рис. 4. Процесс тестирования включал серию из трех независимых экспериментов, описанных ниже.

1. *Раздельный поиск.* Каждое изображение исходной базы данных подавалось на вход программы в качестве образца поиска, и идентификаторы получаемых $N = 20$ наиболее близких результатов поиска запоминались в качестве «правильных». Та же самая процедура независимо выполнялась над искаженными версиями изображений, и результаты поиска по «искаженной» базе данных сравнивались с результатами, полученными при поиске исходных изображений. Количественная оценка воспроизводимости результатов среди ближайших 4, 8, 12, 16 и 20 наиболее похожих ответов показана на рис. 4 с помощью кривой синего цвета.

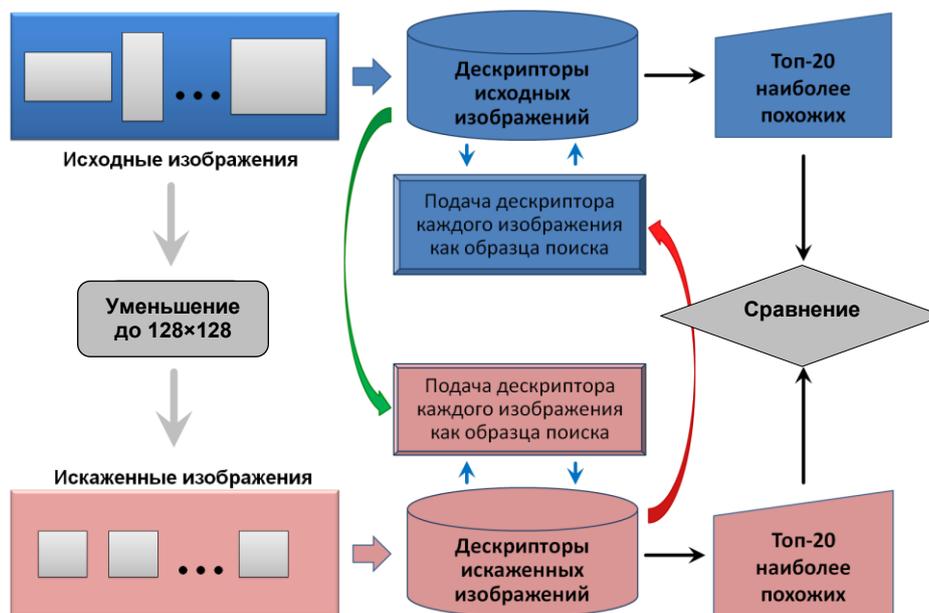


Рис. 3. Общая схема экспериментов по оценке влияния пространственных искажений на качество поиска

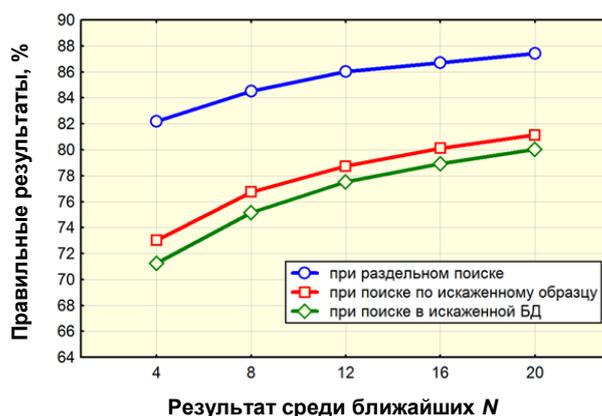


Рис. 4. Графики правильных результатов поиска (в процентах) среди $N = 4, 8, 12, 16$ и 20 изображений, ближайших к заданному изображению-образцу

2. *Поиск по искаженному образцу.* В качестве образца поиска использовался дескриптор искаженного изображения, в то время как поиск наиболее похожих осуществлялся среди исходных, неискаженных изображений (см. красную стрелку на рис. 3). Результаты эксперимента показаны на рис. 4 кривой красного цвета.

3. *Поиск в искаженной базе.* Данный эксперимент являлся противоположным эксперименту 2, т. е. в качестве образца поиска использовался дескриптор исходного изображения, а поиск осуществлялся среди искаженных изображений (см. стрелку зеленого цвета на рис. 3 и зеленую кривую на рис. 4, представляющую точность получаемых результатов).

Проведенные эксперименты показали, что поиск по искаженному образцу среди искаженных же изображений обеспечивает более высокую воспроизводимость результатов (в среднем на 7–9 % лучше) по сравнению с гибридными схемами, использующими исходные и искаженные изображения одновременно. В частности, для 10 000 запросов на поиск по образцу среднее количество корректных результатов среди $N = 20$ наиболее похожих изображений достигает 87,4 %.

2.3. Поиск по фрагментам изображений

На практике довольно часто встречаются ситуации, когда используется не полное исходное цифровое изображение, а тот или иной его фрагмент. Более того, в сети Интернет нередко

можно наблюдать случаи, когда некоторая удачная иллюстрация или ее часть заимствуются различными авторами без ссылки и без разрешения ее автора (правообладателя). Поэтому возможность автоматического поиска изображений по их фрагментам представляется весьма полезной вне зависимости от того, идет ли речь об обнаружении некорректных заимствований или об обычном использовании свободных мультимедийных ресурсов.

Следует подчеркнуть, что в отличие от известных работ по проблеме поиска изображений по фрагментам (см., например, [9, 10]), которые посвящены повышению качества поиска за счет использования некоторых специфических фрагментов изображений, в данной работе проблема выбора каких бы то ни было «характерных» областей (участков, фрагментов) не рассматривается. Вместо этого фрагменты различного размера выбираются вслепую, случайным образом, без оценки их информативности и важности относительно других.

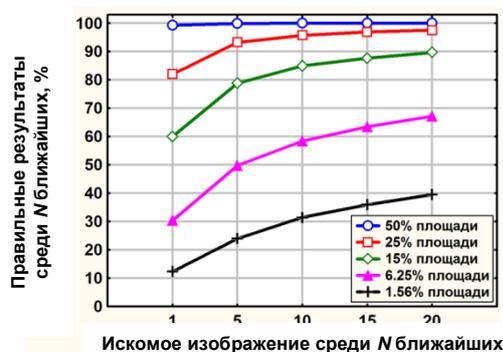
Таким образом, данный подраздел посвящен экспериментальной оценке эффективности решения задачи поиска исходных изображений по образцу на основе дескрипторов совместной встречаемости в условиях, когда в качестве образца поиска выступает некоторый фрагмент исходного изображения. Естественно, что при этом дается только базовая количественная оценка потенциальной возможности решения подобных задач без рассмотрения всей технологии поиска дубликатов в архивах изображений или в сети Интернет, а также других родственных задач.

В рассматриваемом случае экспериментальное исследование потребовало довольно значительных вычислительных затрат и состояло из следующих основных этапов:

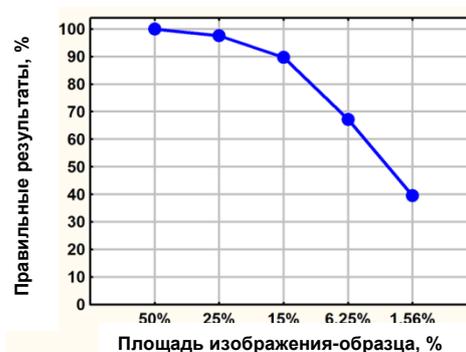
1. *Подготовка фрагментов изображений.* Размеры фрагментов изображений варьировались в широком диапазоне от 50 до 1,56 % от площади исходного изображения, включая следующие пять значений: 50; 25; 15; 6,25 и 1,56 %. Для каждого размера и каждого изображения базы данных в позиции, выбираемой на изображении случайным образом, вырезались 10 фрагментов, т. е. брались всего 50 фрагментов с каждого изображения или в общей сложности 500 000 фрагментов из 10 000 тестовых изображений используемой базы данных. Типичные примеры фрагментов показаны на рис. 5, а.



а)



б)



в)

Рис. 5. Фрагменты изображений различных размеров (а); количество случаев (в процентах), когда искомое изображение было найдено среди N ближайших (б) либо среди всех $N = 20$ (в)

2. *Вычисление дескрипторов изображений.* Как уже упоминалось выше, все дескрипторы (трехмерные гистограммы частоты совместной встречаемости) нормализовывались, чтобы избежать зависимости от размеров изображений.

3. *Собственно проведение поиска по фрагментам.* Данный этап выполнялся в виде пяти серий отдельных экспериментов, по одной серии для каждого размера фрагментов. Каждый из 100 000 фрагментов подавался в качестве образца поиска и ближайшие $N = 20$ результатов запоминались. Таким образом, в каждой из пяти серий задача сравнения дескриптора фрагмента-образца с дескрипторами 10 000 исходных изображений решалась 100 000 раз путем выполнения 10^9 сравнений.

4. *Анализ результатов.* Исходное изображение считалось найденным по заданному фрагменту, если оно появлялось на каком-либо месте среди $N = 20$ изображений, наиболее близких к заданному образцу. Графики, представленные на рис. 5, б, иллюстрируют процент случаев, когда искомое исходное изображение находилось на первом месте, а также среди ближайших 5, 10, 15 или 20 изображений, являющихся результатом поиска. Кроме того, для удобства процент корректных результатов поиска среди $N = 20$ ближайших изображений как функция размера фрагмента-образца показан отдельно на рис. 5, в. Следует отметить, что в представленных данных вероятность случайного появления искомого изображения среди наиболее близких результатов поиска не учитывалась в силу ее малости.

Как и следовало ожидать, доля корректных результатов поиска исходного изображения зависит от размера фрагмента, используемого в качестве образца. Так, например, при размере фрагмента, составляющем 50 % площади материнского изображения, оно находится в 100 % случаев, причем в подавляющем большинстве из них материнское изображение оказывается на первом месте (см. графики синего цвета на рис. 5). С другой стороны, что довольно неожиданно, использование в качестве образца фрагментов весьма малого размера, составляющих лишь 1,56 % от площади исходного изображения, тем не менее позволяет найти исходное изображение среди $N = 20$ результатов в 39,5 % случаев. При этом в 12 % случаев оно оказывается на первом месте (см. кривую черного цвета на рис. 5).

Заключение

Экспериментальные результаты, полученные на тестовой базе из 10 000 цветных изображений общего характера, позволяют сделать следующие выводы.

Пространственное искажение цветных изображений путем приведения их к миниатюрам стандартизованного размера в 128×128 пикселей приводит к тому, что результирующий набор из $N = 20$ изображений, ближайших к заданному образцу поиска, может содержать порядка 13–20 % ошибок. В частности, наихудший результат с 20 % ошибок (соответственно точность 80 %) получается при «смешанной схеме», т. е. при поиске по неискаженному образцу в искаженной базе изображений. Противоположная смешанная схема, т. е. поиск по искаженному образцу в исходной (неискаженной) базе изображений, занимает среднее, второе место с показателем качества 18,9 % ошибок (точность 81,1 %). Наилучшие результаты (12,6 % ошибок и, соответственно, точность 87,4 %) получаются при «однородной» схеме, т. е. при поиске по искаженному изображению-образцу в искаженной базе изображений.

При поиске по фрагменту изображения его оригинала результаты в существенной степени зависят от размера используемого фрагмента. При размере фрагмента, составляющем 50 % от площади исходного изображения, материнское изображение находится среди $N = 20$ наиболее похожих всегда (100 %), причем в подавляющем большинстве случаев (99,3 %) оно оказывается на первом месте. При уменьшении фрагмента-образца поиска точность поиска падает. Однако даже при использовании фрагментов, составляющих менее 2 % от площади исходного изображения, рассматриваемый метод обеспечивает автоматический поиск исходного изображения почти в 40 % случаев.

Несмотря на то что рассматриваемые дескрипторы в принципе могут использоваться для описания структуры и поиска в базах данных полутоновых изображений, качество их поиска будет сравнительно невысоким. Причина этого заключается в том, что для покрытия всего цветового пространства рассматриваемые матрицы совместной встречаемости используют достаточно

грубую схему его разбивки на $8*8*8 = 512$ цветов. Учитывая, что гамме серого цвета в пространстве RGB соответствуют цвета с равными цветовыми компонентами $r = g = b$, максимальное количество различных уровней серого будет равно восьми. Этого может быть недостаточно для целого ряда задач хранения и поиска полутоновых изображений по содержанию. Более подробные сведения о дескрипторах полутоновых изображений можно найти в [4].

Автор выражает глубокую благодарность доктору Штефану Волмеру и профессору Георгиусу Сакасу за предоставленную базу изображений и плодотворные дискуссии по проблеме поиска изображений по содержанию.

Список литературы

1. Content-based image retrieval at the end of the early years / A.W.M. Smeulders [et al.] // IEEE Transactions on PAMI. – 2000. – Vol. 22(2). – P. 1349–1380.
2. A survey of content-based image retrieval with high-level semantics / Y. Liu [et al.] // Pattern Recognition. – 2007. – Vol. 40(1). – P. 262–282.
3. Kovalev, V.A. Feature extraction and visualization methods based on image class comparison / V.A. Kovalev // Medical Imaging 1994 Intern. Symp., Image Processing. – 1994. – Vol. 2167. – P. 691–701.
4. Kovalev, V.A. Multidimensional co-occurrence matrices for object recognition and matching / V.A. Kovalev, M. Petrou // Graphical Models and Image Processing. – 1996. – Vol. 58, № 3. – P. 187–197.
5. Ковалев, В.А. Анализ текстуры трехмерных медицинских изображений / В.А. Ковалев. – Минск : Белорус. наука, 2008. – 264 с.
6. Kovalev, V.A. Color co-occurrence descriptors for querying-by-example / V.A. Kovalev, S. Volmer // Intern. Conf. on Multimedia Modelling. – Lausanne, 1998. – P. 32–38.
7. Ковалев, В.А. Влияние мер близости в пространстве признаков на качество поиска медицинских изображений по содержанию / В.А. Ковалев, А.А. Дмитрук // Информатика. – 2011. – № 2(30). – С. 5–11.
8. Liao, C.-J. Complementary retrieval for distorted images / C.-J. Liao, S.-Y. Chen // Pattern Recognition. – 2002. – Vol. 35. – P. 1705–1722.
9. A scheme of fragment-based faceted image search / T. Komamizu [et al.] // Database and Expert Systems Applications. – 2012. – Vol. 7447. – P. 450–457.
10. Li, J. IRM: Integrated region matching for image retrieval / J. Li, J.Z. Wang, G. Wiederhold // Proc. of the 2000 ACM Multimedia Conf. – Los Angeles, 2000. – P. 83–86.

Поступила 08.04.2015

*Объединенный институт проблем
информатики НАН Беларуси,
Минск, Сурганова, 6
e-mail: vassili.kovalev@gmail.com*

V.A. Kovalev

THE EFFECT OF DISTORTIONS AND FRAGMENTATION ON RESULTS OF CONTENT-BASED RETRIEVAL OF COLOR IMAGES

This paper is dealing with the problem of content-based image retrieval using color co-occurrence matrices. It is well-known that there are a number of factors influencing the results of image retrieval. The specific contribution of this paper lies on quantitative assessment of the effect of spatial image distortions as well as on estimating the accuracy of retrieving original images using its fragments of different size as query examples. The presented study is capitalized on database containing 10 000 color RGB images of very different content. The necessary experimental results are provided.