

ПРИМЕНЕНИЕ БЛОЧНОГО ЧАСТОТНОГО ПРЕДСТАВЛЕНИЯ ИЗОБРАЖЕНИЙ С ЭЛЕМЕНТАМИ ИНВАРИАНТНОСТИ К ПОВОРОТУ ДЛЯ СРАВНЕНИЯ ТЕКСТУР¹

APPLICATION OF BLOCK FREQUENCY REPRESENTATION OF IMAGES WITH ROTATION INVARIANCE ELEMENTS FOR TEXTURE COMPARISON

**A. Shoberg
K. Shoberg**

Summary. Article describes block frequency transforms executions on example of the discrete cosine transform. Two execution variants lead to invariance to process direction. The proposed approach is used to store images in compressed form. At the same time, elements of invariance to rotations are preserved. Transform results are used to compare images including textures. The proposed approach can be used for other integral transforms.

Keywords: integral transforms, frequency domain, image compression, invariance, texture.

Шоберг Анатолий Германович

*К.т.н, доцент, Тихоокеанский государственный
университет, Хабаровск
shoberg@rambler.ru*

Шоберг Кирилл Анатольевич

*Аспирант, Тихоокеанский государственный
университет, Хабаровск
shoberg@bmail.ru*

Аннотация. В статье описываются варианты выполнения блочных частотных преобразований на примере дискретного косинусного преобразования. Два варианта выполнения приводят к инвариантности к направлению обработки. Предложенный подход используется для хранения изображений в сжатом виде. При этом сохраняются элементы инвариантности к поворотам. Результаты преобразований используются для сравнения изображений в том числе текстур. Предложенный подход может быть на другие интегральные преобразования.

Ключевые слова: интегральные преобразования, частотное представление, сжатие изображений, инвариантность, текстуры.

Введение

Публикации по применению частотных преобразований ограничиваются рассмотрением на основе традиционной схемы [1,2,3]. В эффективной системе сжатия данных JPEG применяется блочный подход к выполнению дискретного косинусного преобразования [2]. Сжатие изображения решает проблему уменьшения объема данных, необходимых для представления цифрового изображения. Это процесс, предназначенный для получения компактного представления изображения, тем самым снижая требования к хранению / передаче изображения. Целью сжатия является максимально возможное уменьшение количества бит при сохранении максимально возможного разрешения и визуального качества восстановленного изображения [4].

При сжатии часто используется промежуточное декоррелирующее преобразование (например, ДКП [2]).

Чтобы производить последующие операции, например, сравнение с шаблоном, требуется произвести обратное преобразование.

В ряде работ [5,6,7,8] предложены различные алгоритмы обработки данных непосредственно в частотном домене. При этом для вычисления частотных составляющих повернутого изображения используются дополнительные арифметические операции. Для ДКП можно использовать смену знака или умножение на -1 для ряда составляющих.

В настоящей статье рассматриваются вопросы формирования блочного частотного преобразования с элементами инвариантности к повороту за счет смены направления обработки для части блоков в кодере системы сжатия. При этом вычислительная сложность алгоритма остается прежней. База сжатых изображений на основе предлагаемой методики используется для сравнения текстур.

¹ Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-37-00181

Варианты выполнения двумерного блочное дискретного косинусного преобразование и его модификация

Основой двумерного дискретного косинусного преобразования является двумерное преобразование Фурье. Существует некоторое количество способов получения соотношений, которые позволяют производить расширения, в том числе и симметричные. Здесь возможны варианты с перекрытиями и без таковых. Наиболее востребованный путь заключается в том, что к изображению вплотную пристраивают его зеркальные отражения. Перекрытия отсутствуют. Таким образом, из первоначального массива, содержащего $N_1 \times N_2$ элементов получается массив из $2N_1 \times 2N_2$ элементов. На этой основе строится прямое ДКП [1]

$$X(v, u) = C(u)C(v) \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x(m, n) \cdot \cos\left(\frac{\pi u(2n+1)}{2N}\right) \cos\left(\frac{\pi v(2m+1)}{2N}\right)$$

где $X(m, n)$, $X(v, u)$ – входной и выходной наборы дискретных данных;

m, n, v, u – индексы во входном и выходном наборах;

$C(u)$, $C(v)$ – коэффициенты, зависящие от типа преобразования.

Данное выражение имеет преимущество по сравнению с преобразованием Фурье – в косинусном преобразовании присутствуют только действительные числа и сокращена избыточность спектральных составляющих. Обратное преобразование определяется близким выражением

$$x(m, n) = \sum_{v=0}^{N-1} \sum_{u=0}^{N-1} C(u)C(v)X(v, u) \cos\left(\frac{\pi u(2n+1)}{2N}\right) \cdot \cos\left(\frac{\pi v(2m+1)}{2N}\right)$$

Применяя блочный подход к выполнению преобразования с использованием двумерных базисных функций (частично рассмотренный в [9]), рассмотрим размещение соответствующих наборов.

Выделим из выражения преобразования базисную функцию $\Phi(v, u, m, n)$, представляющую собой произведение двух гармонических функций

$$\Phi(v, u, m, n) = \cos\left(\frac{\pi u(2n+1)}{2N}\right) \cos\left(\frac{\pi v(2m+1)}{2N}\right). \quad (1)$$

Для четного количества блоков L и K каждый размером $N \times N$ с учетом (1) можно записать прямое блочное ДКП в традиционной форме

$$X_T(v, u, l, k) = C(v)C(u) \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x(m, n, l, k) \Phi(v, u, m, n), \quad (2)$$

где $l = 0, 2, \dots, L$, $k = 0, 2, \dots, K$ – номера блоков по вертикали и горизонтали, соответственно;

$X_T(v, u, l, k)$ – результат блочного преобразования в традиционной форме для фиксированных l и k ;

$x(m, n, l, k)$ – исходные данные для блочного преобразования в традиционной форме для фиксированных l и k .

Выборка исходных данных в (2) для $x(m, n, l, k)$ будет осуществляться в соответствии с выражением

$$x(m, n, l, k) = \begin{cases} x(lN + m, kN + n) \\ x(lN + m, (k+1)N + n) \\ x((l+1)N + m, kN + n) \\ x((l+1)N + m, (k+1)N + n) \end{cases}. \quad (3)$$

Запись будет производиться в блок

$$X_T(v, u, l, k) = \begin{cases} X_T(lN + v, kN + u) \\ X_T(lN + v, (k+1)N + u) \\ X_T((l+1)N + v, kN + u) \\ X_T((l+1)N + v, (k+1)N + u) \end{cases}. \quad (4)$$

Восстановление осуществляется в соответствии с выражением

$$x(m, n, l, k) = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} C(v)C(u)X_T(v, u, l, k) \Phi(v, u, m, n), \quad (5)$$

где результат соответствует (3), а для вычислений используется (4).

Возможная блочная модификация из [10], распространенная на произвольное количество блоков для выражения аналогичного (2) может быть рассмотрена для двух видов.

Предлагаемое прямое блочное преобразование 1 типа для групп из четырех блоков будет рассчитываться в соответствии

$$x(m, n, l, k) = \begin{cases} x((l+1)N - (m+1), (k+1)N - (n+1)) \\ x((l+1)N - (m+1), (k+1)N + n) \\ x((l+1)N + m, (k+1)N - (n+1)) \\ x((l+1)N + m, (k+1)N + n) \end{cases}, \quad (6)$$

где размещение для (2) будет производиться подобными группами (X_T меняется на X_{S1})

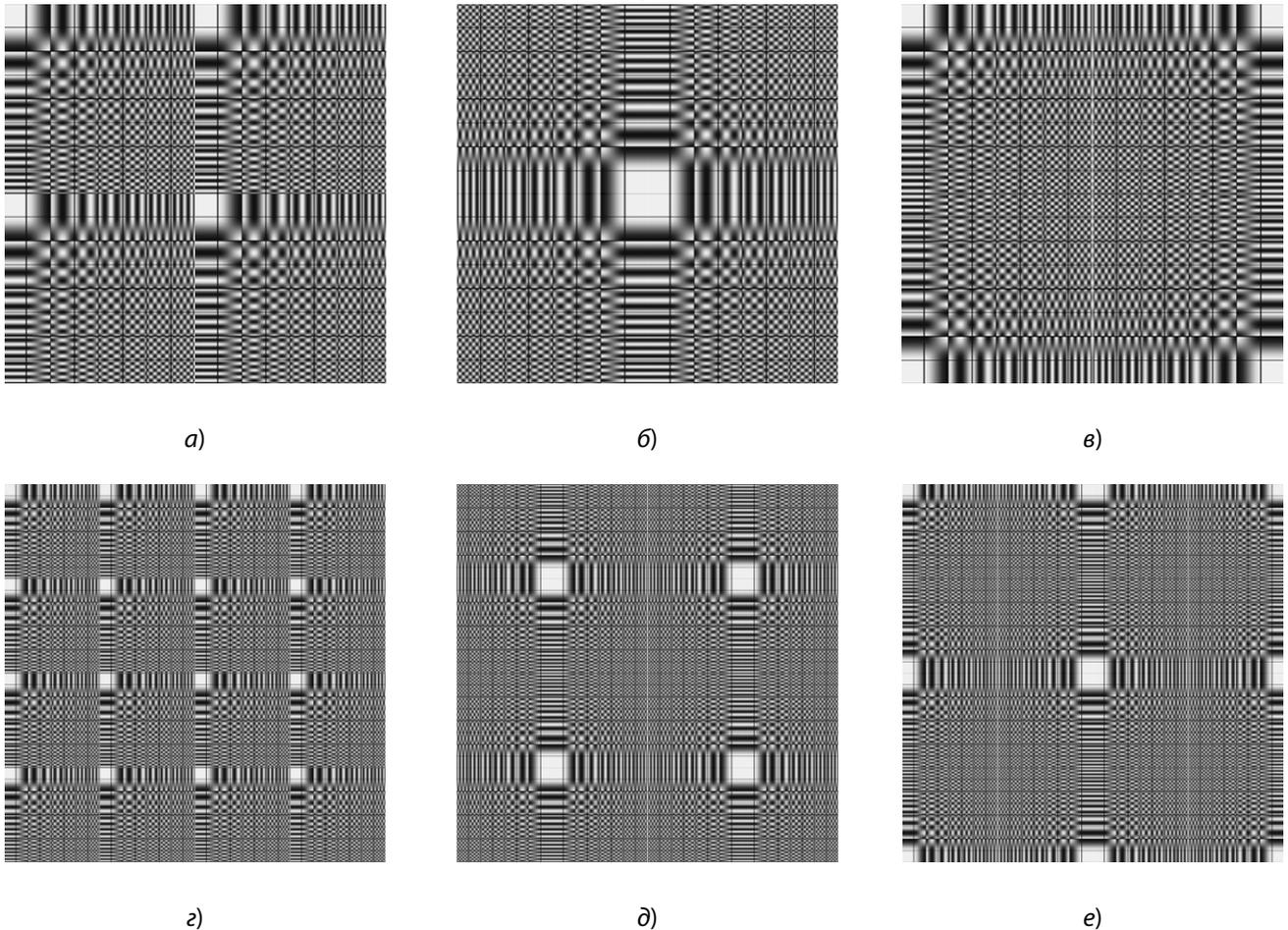


Рис. 1. Графическое представление наборов двумерных базисных функций ДКП на основе макроблоков

$$X_{S1}(v, u, l, k) = \begin{cases} X_{S1}((l+1)N - (v+1), (k+1)N - (u+1)) \\ X_{S1}((l+1)N - (v+1), (k+1)N + u) \\ X_{S1}((l+1)N + v, (k+1)N - (u+1)) \\ X_{S1}((l+1)N + v, (k+1)N + u) \end{cases} \quad (7)$$

Восстановление будет производиться по выражению (5) с такой же заменой.

Аналогичный подход используется для предлагаемого прямого и обратного блочного преобразования 2 типа. Используется выражение (2) для групп из четырех блоков в соответствии

$$x(m, n, l, k) = \begin{cases} x(lN + m, kN + n) \\ x(lN + m, (k+2)N - n) \\ x((l+2)N - m, kN + n) \\ x((l+2)N - m, (k+2)N - n) \end{cases} \quad (8)$$

запись будет производиться следующим образом

$$X_{S2}(v, u, l, k) = \begin{cases} X_{S2}(lN + v, kN + u) \\ X_{S2}(lN + v, (k+2)N - u) \\ X_{S2}((l+2)N - v, kN + u) \\ X_{S2}((l+2)N - v, (k+2)N - u) \end{cases} \quad (9)$$

На рис. 1 представлен ряд вариантов размещения наборов матриц двумерных базисных функций ДКП при увеличении их общего количества. На рис. 1, а представлен набор из четырех блоков для традиционного блочного преобразования (2), (3) и (4). На рис. 1, б и рис. 1, в наборы из 4-х блоков составлены в виде групп с элементами центральной симметрии. При этом рис. 1, б формируется на основе (6) и (7), а макроблок (группа из 4-х блоков) рис. 1 в, формируется на основе (8) и (9).

Увеличение количества макроблоков по вертикали и горизонтали для каждого из рассматриваемых вариантов (X_T , X_{S1} и X_{S2}) не влияет на симметрии в предлагаемых матрицах. Для традиционного блочного преобразования элементы центральной симметрии отсутствуют

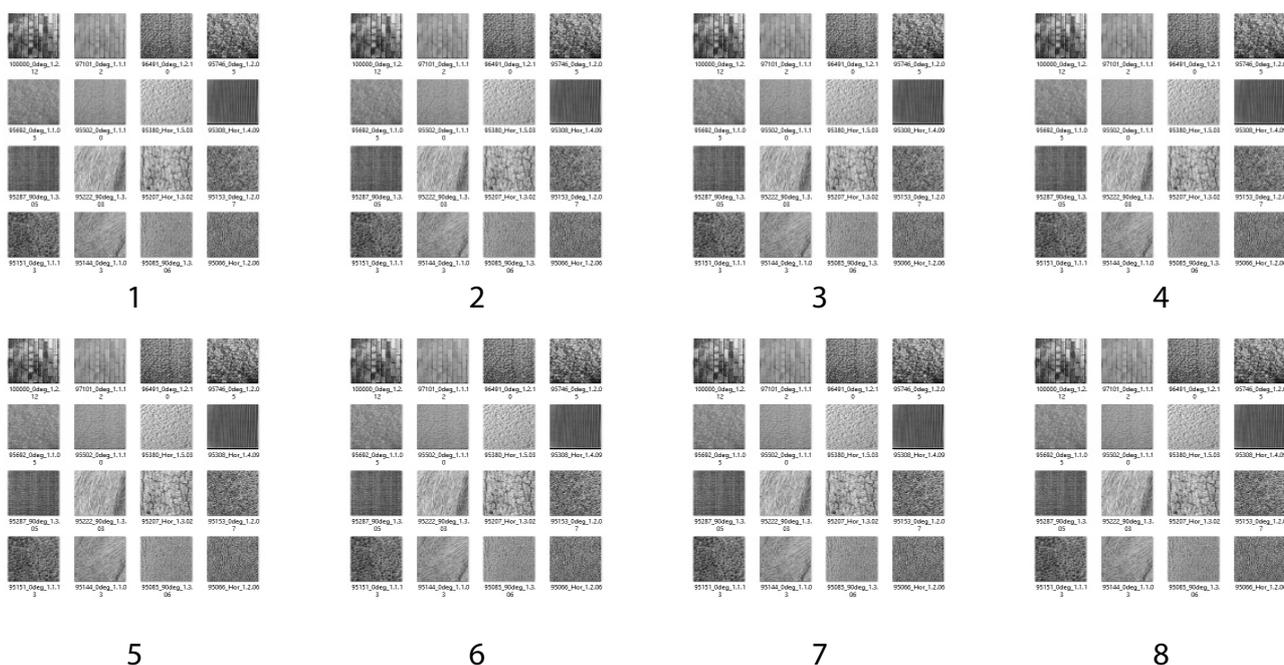


Рис. 2. Результат ранжирования по базе текстур (постоянная составляющая)

(рис. 1, г). Для вариантов на основе (6), (7) (рис. 1, д) и (8), (9) (рис. 1, е) элементы центральной симметрии по-прежнему присутствуют. Следует заметить, что на рис. 1, г повторяемость макроблоков составляет 1 блок, а для рис. 1, д и рис. 1, е повторение осуществляется через 2 блока.

Выражения (6), (7) и (8), (9) показывают, что они связаны между собой и отличаются смещением друг от друга на один блок по горизонтали и вертикали. Макроблок (рис. 1, б) присутствует в блочной матрице (рис. 1, е) по центру (смещение в блоках 1, 1), а макроблок (рис. 1, в) присутствует в центре рис. 1, д.

В рассматриваемых случаях для предлагаемых схем выполнения преобразования в целом увеличивается количество групп наборов матриц с элементами центральной симметрии. Изучение наборов элементов с наличием симметрии производится на основе «Теории групп». В минимальном случае следует отметить следующие положения.

В двумерном случае группы симметрий делятся на ряд классов, среди которых можно выделить:

- ◆ циклические группы, состоящие из поворотов вокруг неподвижной точки на углы, кратные $360^\circ/n$;
- ◆ бесконечные группы, которые получаются из одномерных групп симметрий добавлением переносов;
- ◆ диэдральные группы и мн. др.

Если рассматривать группы симметрий на примере квадрата, то он имеет четыре оси симметрии и одну центральную при которых при движении он отображается сам на себя. При этом некоторые вершины поменяются местами, а некоторые останутся неподвижными [11]. Таких движений 8, включающих 4 осевых симметрии и 4 поворота вокруг центра (0° , 90° , 180° , 270°). Эти восемь движений образуют некоммутлируемую группу симметрий квадрата. Четыре поворота вокруг центра образуют циклическую группу, которая является подгруппой группы симметрий квадрата. Эта группа коммутативна. Эти положения применимы при рассмотрении предлагаемой модифицированной схемы.

Так как для выражений (2) — (9) используются различные варианты размещения (местоположения результатов для различных частот не совпадают), для сравнения результатов можно использовать различные виды выполнения блочного преобразования с необходимым размещением результатов. Так блочное преобразование в традиционной форме с размещением результата в виде предлагаемого блочного преобразования 1 типа X_{TM1} для (2) следует использовать (3) с записью соответствующей (7). Для получения X_{TM2} (аналогично описанному) следует использовать (3) с записью соответствующей (9). В другом случае блочное преобразование 1 типа в предлагаемой форме с размещением результата в виде традиционного блочного преобразования X_{S1T} для (2) следует использовать (6) с записью соответствующей (4). Для получения X_{S2T} следует использовать (8)

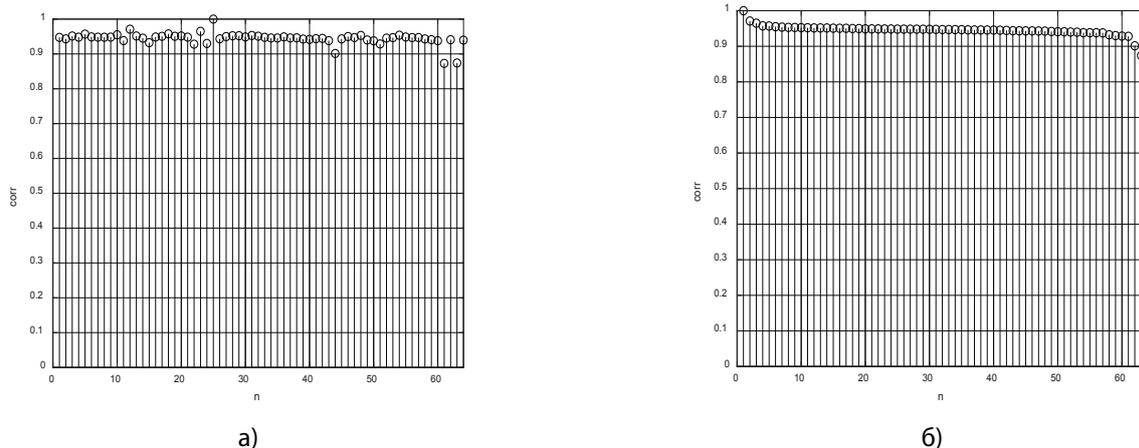


Рис. 3. Коэффициенты корреляции между образцом и изображениями в базе для постоянной составляющей

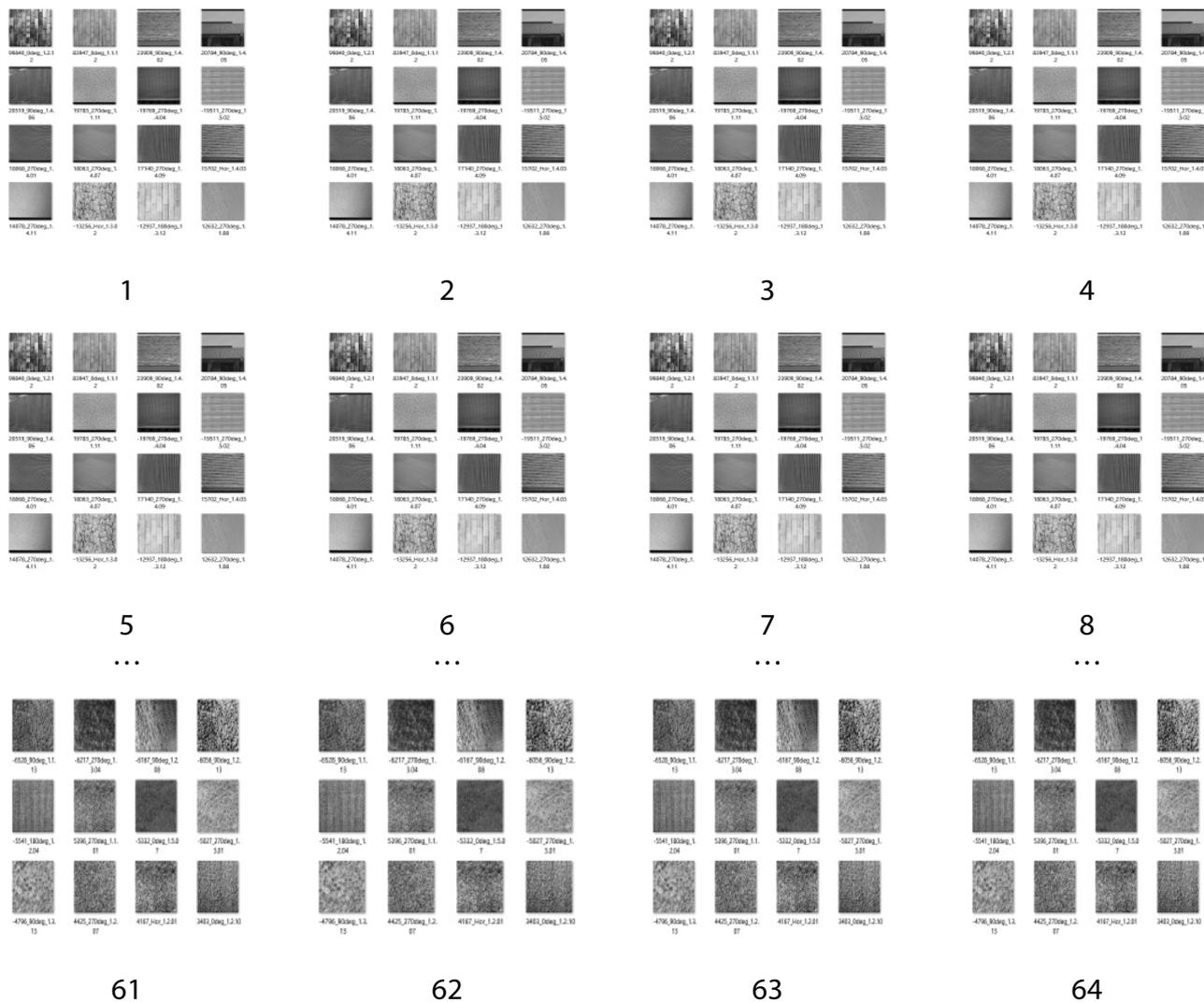


Рис. 4. Результат ранжирования по базе текстур (1 частотная составляющая)

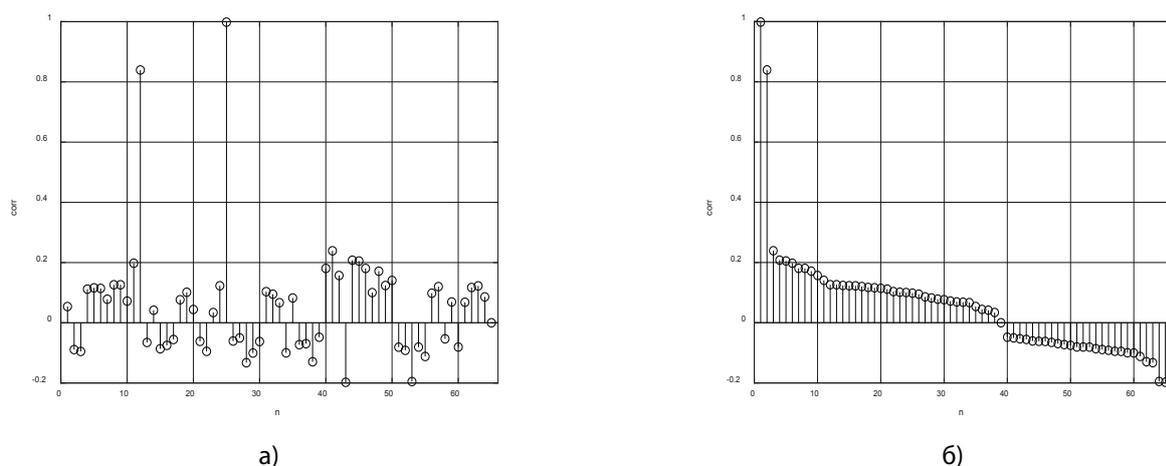


Рис. 5. Коэффициенты корреляции между образцом и изображениями в базе для постоянной — а, б и 1-й частотной составляющей — в, г

с записью соответствующей (4). Кроме того возможны виды $XS1S2$ — чтение (6), запись (9) и $XS2S1$ — чтение (8), запись (7).

Выражения (2) — (9) могут быть использованы для произвольных изображений или двумерных функций.

Результаты моделирования

На основе выражений, предложенных выше, производилось ДКП изображений блоками 8×8 (аналогично JPEG [2]) с выборкой и записью (6), (7). Производилось последовательное сравнение (на основе корреляции) образца с частотными представлениями изображений в базе данных. Использовались повороты на углы кратные 90° . Сравнения производились как по отдельным составляющим, так и по наборам частотных составляющих. В качестве исследуемых использовались в частности текстуры, как образцы с достаточно близкими характеристиками, и соответственно распознавание, которых достаточно затруднено.

Количество текстур в базе — 64 шт. Изображения представлены в 256 градациях яркости.

На рис. 2 представлены результаты сравнения с образцом по постоянной составляющей по всему изображению (представлены 8 текстур, имеющих максимальные значения корреляции с шаблоном). Образец для сравнения изображен на рис. 2, 1. Первые два изображения имеют максимальное соответствие. Текстура (рис. 2, 2) отличается от (рис. 2, 1) сниженным значением яркости. Образцы, представленные на рис. 2, 3–2, 7, могут рассматриваться как ошибочные. Изображение (рис. 2, 8) может рассматриваться, как близкое по ха-

рактеристикам (большое количество параллельных линий).

Зависимости, представленные на рис. 3, отражают значения коэффициента корреляции между образцом и текстурами из базы данных.

График на рис. 3, а показывает значения коэффициентов корреляции относительно положения изображения в базе данных, а график на рис. 3, б представляет те же значения в отсортированном по амплитуде виде. Для постоянной составляющей динамический диапазон изменения коэффициента корреляции составляет от 1 до 0.87 и составляет 0.13. Он имеет достаточно гладкий вид (рис. 3, б) без выбросов с высокими амплитудными характеристиками.

На рис. 4 показаны результаты ранжирования текстур, полученные в результате поиска по образцу с использованием 1-й частотной составляющей. В качестве образца использовалось тестовое изображение рис. 1, 1.

Максимальные соответствия по-прежнему имеют образец и его менее яркая копия (рис. 4, 1 и рис. 4, 2). Из первых 8 выбросом можно рассматривать только текстуру рис. 4, 6. Остальные изображения имеют достаточно сходную структуру (наличие регулярных прямых линий). При этом найдены также повернутые текстуры (рис. 4, 3 и рис. 4, 8). Текстуры, имеющие наименьшее сходство (нерегулярная структура), вынесены в конец списка (рис. 4, 61 — рис. 4, 64)

Графики на рис. 5, отражают значения коэффициента корреляции между образцом и текстурами из базы данных. Зависимость на рис. 5, а показывает значения

коэффициентов корреляции относительно положения изображения в базе данных, а график на рис. 5, 6 представляет те же значения в отсортированном по амплитуде виде.

Для 1-й частотной составляющей динамический диапазон изменения коэффициента корреляции значительно шире — от 1 до -0.2 и имеет резко выраженный нестационарный вид (рисунок 5, 6). По сравнению с постоянной составляющей динамический диапазон изменения коэффициента корреляции шире в 9.2 раза.

Заключение

Для блочных частотных преобразований при учете направления обработки возможны варианты выполнения, приводящие к различным результатам. Только часть из них имеет элементы инвариантности к направлению обработки. Их реализации бывают двух основных типов,

отличающиеся друг от друга смещением в размещении блоков базисных функций. Описанный подход может использоваться для произвольного четного количества блоков и для больших размерностей.

Предложенные выражения (6) — (9) показывают, что вычислительная сложность при реализации преобразований не изменяется, как для стандартных алгоритмов — $O(n^2)$, так и для их быстрых реализаций — $O(n \log(n))$. Это происходит вследствие того, что для каждого блока расчеты производятся независимо. Количество умножений и сложений остается прежним, изменяется только порядок выборки и записи элементов.

Описанные выражения могут быть использованы для компактного хранения данных различных размерностей, ускоренным поиском в частотной области с наличием повернутых экземпляров, как с использованием отдельных составляющих, так и их взвешенных наборов.

ЛИТЕРАТУРА

1. Блейхут Р. Быстрые алгоритмы цифровой обработки сигналов. — М.: Мир, 1989. — С. 448.
2. Гонсалес Р., Вудс Р. Цифровая обработка изображений. — М.: Техносфера, 2005. — С.1072.
3. Шапиро Л., Стокман Дж., Компьютерное зрение. М.: Бином. Лаборатория знаний, 2006. — С.721.
4. Ватолин Д., Ратушняк А., Смирнов М., Юкин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. М.: Диалог-МИФИ, 2002. — С. 384.
5. Chang T. and Kuo C.-C. J. Texture Analysis and Classification with Tree-Structured Wavelet Transform IEEE Transactions on Image Processing, Vol. 2, No. 4, Oct., 1993. — PP. 429–441
6. Hsu Y.S., Prum S., Kagel J. H., Andrews A. C., Pattern recognition experiments in the Mandala/Cosine domain IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-5, No. 5, Sep. 1983. — PP. 521–529
7. Ngo C.-W., Pong T.-C., Chin R. T. Exploiting image indexing techniques in DCT domain Pattern Recognition 34 (2001), May, 2000. — PP. 1841–1851
8. Park, Y.S.; Park H. W. Design and analysis of an image resizing filter in the block-DCT domain. IEEE T. Circ. Syst. Vid. 2004, 14. — PP. 274–279.
9. Шоберг А. Г. Современные методы обработки изображений: мо-дифицированное вейвлет-преобразование. — Хабаровск: Изд-во Тихоокеан. гос. ун-та. 2014. С. — 125.
10. Шоберг А. Г. Повышение устойчивости к повороту при разделимом частотном преобразовании изображений // Вопросы радиоэлектроники, сер. Техника телевидения, 2015, № 6(23). — С. 75–80.
11. Каргополов М.И., Мерзляков Ю.И. Основы теории групп. М: Наука, Физматлит, 1996. — С. 287.

© Шоберг Анатолий Германович (shoberg@rambler.ru), Шоберг Кирилл Анатольевич (shoberg@mail.ru).

Журнал «Современная наука: актуальные проблемы теории и практики»