

В.Г. ИВАНОВ, канд. техн. наук,
М.Г. ЛЮБАРСКИЙ, д-р физ.-мат. наук,
Ю.В. ЛОМОНОСОВ, канд. техн. наук (г. Харьков)

JPEG И ВЕЙВЛЕТ-КОМПРЕССИЯ ОБЪЕКТА И ФОНА ИЗОБРАЖЕНИЯ С РАЗЛИЧНОЙ СТЕПЕНЬЮ КАЧЕСТВА

У даній роботі приводяться практичні результати комп'ютерного моделювання стиску деяких класів зображень. Зображення перед кодуванням попередньо розбивається на дві частини – об'єкт і тло. Поділ зображення здійснюється в площині вейвлет-коефіцієнтів. Кодування кожної складової зображення робиться окремо з різним ступенем точності на основі JPEG-технологій з використанням узагальненого перетворення Фур'є (дискретне косинусне перетворення – DCT), а також із застосуванням багатомасштабної обробки вейвлет-коефіцієнтів.

In the given operation the practical results of computer simulation of contraction of some classes of the maps are resulted. The map, before encoding, is previously broken down into two parts - plant and background. The separation of the map is carried out in a plain of wavelet-coefficients. The encoding by each component the maps is made separately with different degree of accuracy on the basis of JPEG-know-hows with usage of a generalized Fourier transform (discretic cosine transforming - DCT), and also with applying of multiscale treating of wavelet-coefficients.

Введение. В течение многих лет доминирующей тенденцией в исследованиях проблемы сжатия данных и изображений являются методы обработки сигналов и методы теории информации, которые используют декоррелирующие свойства различных линейных преобразований и возможности квантования и энтропийного кодирования коэффициентов этих преобразований [1].

Развитие этих методов привело к созданию известных JPEG-форматов, которые используют субоптимальные косинусные преобразования Фурье и методы вейвлет-преобразований, позволяющие в 20 и более раз сократить объем графических данных, сохранив при этом хорошее визуальное качество [2, 3]. Повысить количественные и качественные характеристики этих форматов можно в результате сокращения содержательной избыточности изображений. Суть такого подхода заключается в том, что в памяти хранится только часть изображения, представляющая интерес в данном конкретном применении.

Постановка задачи и её решение. Основной проблемой при сокращении содержательной избыточности изображений остается выделение значимых (информативных) участков изображений (объектов) и участков менее информативных (фона). И хотя решению этой проблемы посвящено большое количество работ [4 – 6], в которых получены интересные результаты, в целом задача кодирования изображений на основе этих принципов еще далека от своего окончательного решения.

В данной работе разделение элементов изображения на два класса (объект и фон) производится при помощи вейвлет-преобразования изображения, и выделения контуров в пространстве коэффициентов высокочастотных областей многомасштабного анализа. Конечным результатом выделения объектов на изображении является построение битовой плоскости, которая содержит информацию о расположении объекта и фона. Дальнейшая обработка выделенных областей осуществляется при помощи JPEG-технологий на базе дискретного косинусного преобразования, а также с помощью многомасштабного вейвлет-анализа. Эта работа является продолжением исследований авторов, основные результаты которых были опубликованы в [7 – 9].

Выделение объектов на плоскости изображений. Идея кодирования изображения с различной степенью качества реализована в графическом формате DjVu [10]. В этом формате изображение разделяется на передний план, задний план и битовую плоскость соответствия, которая содержит информацию о взаимном размещении переднего и заднего плана на плоскости изображения. Разделение изображения на передний план (в рассматриваемом случае – аналог объекта) и задний план (фон) основано на нахождении границ резких перепадов яркости между областями с равномерным распределением её значений. Подобный алгоритм выделения значимых областей изображений позволяет сохранить эти области с лучшим качеством, а степень сжатия повысить за счет большей компрессии заднего плана. Кодирование переднего и заднего плана изображения производится на основе вейвлет-преобразования данных с различным разрешением, причем разрешение заднего плана в несколько раз меньше, чем разрешение переднего плана. Таким образом, обрабатываемое изображение в формате DjVu кодируется с различной степенью качества, что по сравнению с форматом JPEG-2000 (JP2), где также используется вейвлет-преобразование для всей плоскости изображения, имеет некоторое преимущество в сжатии для определенных классов изображений.

Ограничением приведенного метода выделения переднего плана есть то, что значимый (информативный) объект должен иметь равномерное распределение яркости, быть замкнутым по форме и иметь четкие границы перепада яркости по сравнению с задним планом изображения (символьные данные). В случае, когда наиболее информативная область реалистичного изображения не имеет замкнутой формы с постоянным значением яркости, приведенный алгоритм не производит разделения изображения и кодирует его с равной степенью качества, т.е. сводится к известному формату JPEG-2000.

Целью данной работы является разработка и анализ эффективного алгоритма кодирования на основе выделения наиболее информативных областей изображения, не имеющих равномерного распределения яркости внутри себя и замкнутых границ, очерчивающих объект. В качестве такого изображения, из библиотеки стандартных изображений

<http://www.icsl.ucla.edu>, был выбран файл lena.bmp, который представлен на рис. 1. Параметры исходного изображения: размер 256×256, глубина цвета в градации серого 8 бит. Вейвлет коэффициенты исходного изображения, после одного шага преобразования, представлены на рис. 2.

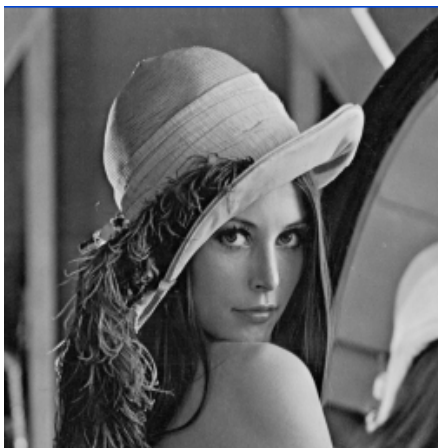


Рис.1. Исходное изображение

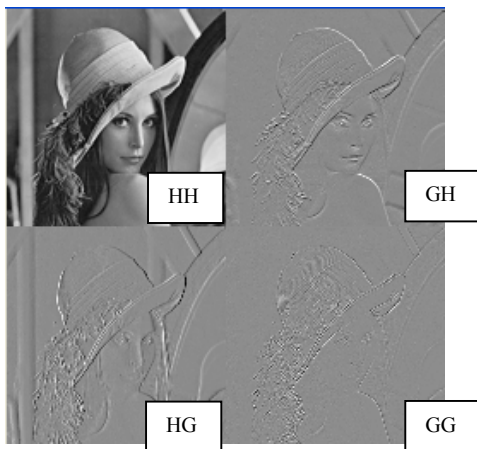


Рис. 2. Вейвлет коэффициенты после одного шага преобразования

На практике вейвлет-преобразование проводится с помощью биортогональных вейвлет-базисов по формулам разложения (1) и синтеза (2). Если $\{a_i\}$ – коэффициенты разложения остаточного члена, то $\{b_i\}$ и $\{c_i\}$ –

коэффициенты, соответственно нового остаточного члена в следующем масштабе представления

$$b_i = \sum_s h_s a_{2i+s} \text{ и } c_i = \sum_s g_s a_{2i+s} . \quad (1)$$

Аналогичным образом, если известны коэффициенты $\{b_i\}$ и $\{c_i\}$, то коэффициенты $\{a_i\}$ можно восстановить в соответствии с выражением

$$a_i = \sum_s h_s b_{i-s} + \sum_s g_s c_{i-s} . \quad (2)$$

Для разложения и восстановления сигнала используют две разные пары квадратурных зеркальных фильтров: пара $\{\tilde{h}, \tilde{g}\}$ при разложении и пара $\{h, g\}$ при восстановлении. Основное уравнение, в данном случае имеет вид

$$\overline{h(\omega)\tilde{h}(\omega)} + \overline{h(\omega + \pi)\tilde{h}(\omega + \pi)} = 2 .$$

Для этого равенства существует семейство решений:

$$\begin{aligned} \tilde{h}_{n,m}(\omega) &= \sqrt{2} \left(\frac{1 + e^{i\omega}}{2} \right)^n ; \\ h_{n,m}(\omega) &= \sqrt{2} \left(\frac{1 + e^{i\omega}}{2} \right)^2 P_m \left(\sin^2 \frac{\omega}{2} \right) e^{-im\omega} , \end{aligned} \quad (3)$$

где $P_m(x) = \sum_{s=0}^{m-1} C_{m-1+s}^s x^s$ – многочлен Дебеша.

Для вейвлет-преобразования исходного изображения, рис. 1, использовалась пара квадратурных зеркальных фильтров h и \tilde{h} с параметрами $n = 1$ и $m = 2$ выражения (3), общий вид, которых представлен на рис. 3.

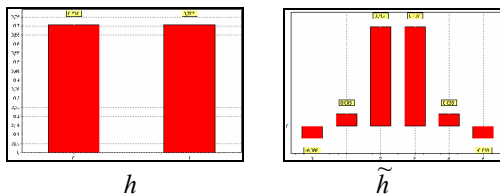


Рис. 3. Пара квадратурных зеркальных фильтров с параметрами $n = 1$ и $m = 2$

При построении разделяющей битовой плоскости, использовались вейвлет коэффициенты HG и GH областей (рис. 2), которые были получены в соответствии с выражением (1). После усредненного суммирования выбранных вейвлет-коэффициентов $(HG+GH)/2$ получим изображение, которое представлено на рис. 4.



Рис. 4. Изображение полученное в результате усредненного суммирования вейвлет коэффициентов HG и GH областей

Полученное изображение имеет более равномерное распределение яркости по сравнению с оригиналом (рис. 1), или остаточным членом разложения НН (рис. 2). Это качество достигается за счет применения соответствующей пары фильтров h и \tilde{h} при вейвлет преобразовании исследуемого изображения. Исходное изображение раскладывается на “грубую” (НН) и ”тонкую” (GG) структуры (рис. 2). Области HG и GH имеют более гладкий рельеф яркости на всей плоскости, но при этом сохраняют наиболее существенные ее перепады. Выбор именно этих составляющих (HG и GH) обусловлен тем, что они получены в результате обработки исходного изображения одной парой фильтров h и \tilde{h} , но только в обратной последовательности их применения. При этом область HG содержит наиболее выраженные перепады яркости в вертикальном направлении, что обусловлено последовательностью операций свертки исходного изображения с соответствующими фильтрами h и \tilde{h} . Область GH, наоборот, имеет резкие скачки яркости в горизонтальной плоскости.

Усредненное суммирование HG и GH плоскостей позволяет сохранить резкие изменения яркости, которые обязательно присутствуют в обоих направлениях. Усредненное суммирование HG и GH областей формирует плоскость изображения для пороговой обработки (рис. 4) по формуле

$$X_{i,j} = (HG_{i,j} + GH_{i,j}) / 2, \quad (4)$$

где $X_{i,j}$ – точка усредненного изображения, $i, j = 1, \dots, n$; n – количество строк и столбцов в матрице изображения размерностью $n \times n$.

Применение пороговой обработки к изображению, представленному на рис. 4, позволяет получить совокупность точек, между которыми значения разности яркости превышает пороговое значение ($|X_i - X_{i+1}| \geq KX_{cp}$), (рис. 5, а). Формирование плоскости $Z_{i,j}$, которая содержит точки превышения порогового значения, (рис. 5, а) согласно правилу

$$Z_{i,j} = \begin{cases} 1, & \text{если } |X_{i,j} - X_{i,j+1}| \geq KX_{\text{cp}}, \text{ или } |X_{i,j} - X_{i+1,j}| \geq KX_{\text{cp}}; \\ 0, & \text{если } |X_{i,j} - X_{i,j+1}| < KX_{\text{cp}}, \text{ или } |X_{i,j} - X_{i+1,j}| < KX_{\text{cp}}. \end{cases} \quad (5)$$

где $i, j = 1, \dots, n - 1$, n – количество строк и столбцов в матрице изображения размерностью $n \times n$.

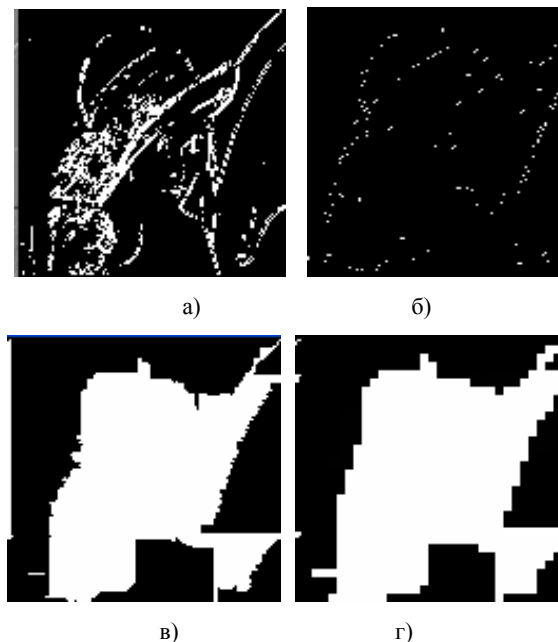


Рис. 5 а) совокупность точек, превысивших разностный порог; б) граничные точки разделения объекта и фона; в) битовая плоскость разделения; г) битовая плоскость после сегментации

Пороговое значение привязывалось к среднему значению яркости X_{cp} на всей плоскости изображения (рис. 4), и выбиралось таким образом, чтобы исключить наличие случайных точек, т.е. тех точек, которые не имеют в своем ближайшем окружении себе подобных. Таким образом, производится фильтрация одиночных точек преодолевших пороговое значение. На рис. 5, а наглядно видны область с высокой плотностью перепадов яркости и область, которая практически не содержит смежных точек превышения порога. Пороговая обработка изображения, представленного на рис. 4, проводилась по строкам и столбцам матрицы изображения.

Затем, для отделения объекта от фона, производится последовательное сканирование изображения по строкам и столбцам до выявления первой и последней не нулевой точки. После разделения изображения на области

объекта и фона, необходимо создать непрерывную и замкнутую границу между этими областями. На рис. 5,б представлены граничные точки, которые разделяют области объекта и фона. Эта граница раздела имеет разрывы и не является замкнутой линией. Для формирования непрерывного и замкнутого контура требуется соединить точки разрыва границы между связными фрагментами контура (см. рис. 6). Точкой разрыва линии границы считается любая точка, имеющая только одну смежную точку, т. е. одного "соседа". Для соединения точек разрыва линии границы X_j необходимо найти минимальное расстояния между точками разрыва, принадлежащими разным подмножествам $\{R_i\}$, т.е. концом связного фрагмента границы и началом другого. После соединения граничных точек между собой получим битовую плоскость, разделяющую объект и фон, рис. 5, в. Построение битовой плоскости является завершающим этапом выделения объекта на плоскости изображения. Точки внутри контура считаются принадлежащими объекту и им присваивается значение 1, точки, лежащие вне контура, считаются фоном и им присваивается значение 0, рис. 5, в.

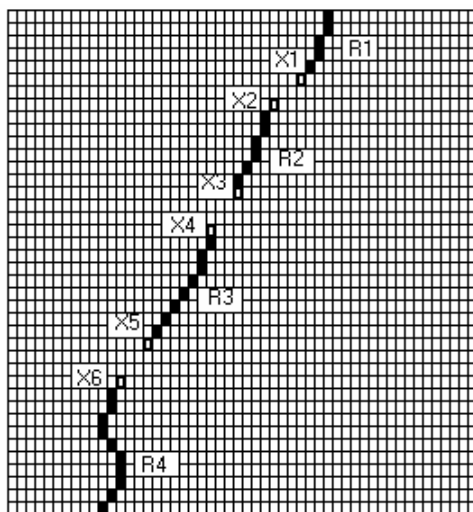


Рис. 6. Связные фрагменты R_i и точки разрыва линии границы X_j между областями объекта и фона

В случае, когда компрессия изображений осуществляется при помощи JPEG-технологии на основе дискретного косинусного преобразования, то предварительно осуществляется разбиение плоскости изображения на фрагменты размером 8×8 [JPEG]. После одного шага вейвлет-преобразования плоскость изображения, представленная на рис. 4, уменьшается в 4 раза по

сравнению с плоскостью исходного изображения, рис. 1. Следовательно, размер фрагмента, для битовой плоскости, должен быть 4×4 , что полностью согласуется с разбиением исходного изображения (рис. 1) на фрагменты 8×8 . Каждый фрагмент может принадлежать трем областям: область объекта, область фона, область границы битовой плоскости. Таким образом, фрагментами области объекта можно считать те фрагменты, которые принадлежат, как области самого объекта, так и области границы битовой плоскости. Все остальные сегменты плоскости изображения принадлежат области фона.

На рис. 5, г представлена битовая плоскость разделения изображения после сегментного разбиения на фрагменты 4×4 .

В том случае, когда сжатие исходного изображения (рис. 1) производится с использованием многомасштабного вейвлет преобразования, то семантика битовой плоскости (рис. 5, в) не производится.

JPEG- и вейвлет-компрессия объекта и фона изображений с различной степенью точности. Сначала рассмотрим практические результаты кодирования изображений с различной степенью точности при использовании JPEG-технологий.

Исходное изображение (рис. 1) необходимо разбить на сегменты размером 8×8 и осуществить кодирование объекта и фона изображения по методу JPEG-технологии [5 – 7].

Согласно метода JPEG-компрессии ядро одномерного косинусного преобразования имеет вид:

$$L_x(0) = \frac{1}{\sqrt{N}} \sum X(m); \quad L_x(R) = \sqrt{\frac{2}{N}} \sum_{m=0}^{N-1} X(m) \cos \frac{(2m+1)k\pi}{2N}. \quad (7)$$

В этом выражении $X(m)$ – значение отчетов в строках соответствующих рабочих матриц, полученных на втором шаге JPEG-технологий, $L_x(k)$ – коэффициенты дискретного косинусного преобразования, а значения m и k изменяются соответственно от 0 до $N-1$ и от 1 до $N-1$ с шагом один.

Различная степень точности кодирования объекта и фона определяется количеством косинусных коэффициентов, которые используются в рабочих матрицах объекта и фона. Для фона выберем минимальное значение -1 (имеется ввиду первый коэффициент рабочей матрицы 8×8 в каждом сегменте изображения), а для объекта максимальное – 64 коэффициента. Таким образом, сжатие изображения будет осуществляться только за счет сегментов, которые входят в область объекта на битовой плоскости. Область фона будет представлена только одним коэффициентом косинусного преобразования в каждом сегменте области фона. Такой подход позволяет обеспечить максимальное сжатие области фона изображения, а в области объекта обеспечить такую степень компрессии и качества изображения, которая являлась бы приемлемой для решения поставленной задачи.

На рис. 7 представлена зависимость общего коэффициента сжатия всей плоскости изображения (Ксж) от среднеквадратической ошибки (СКО) на плоскости объекта.

На рис. 8, представлено исходное изображение после кодирования методом JPEG компрессии. Рис. 8, а показывает восстановленное изображение после раздельного кодирования объекта и фона с различной степенью качества, а рис. (8, б) после кодирования всей плоскости изображения с одинаковой степенью качества. СКО на плоскости объекта в обоих вариантах составляет 12%, а увеличение коэффициента сжатия с 19 до 24 обусловлено большей степенью компрессии области фона (в данном случае Ксж фона составляет 64 раза, так как только один косинусный коэффициент рабочей матрицы 8×8 представляет каждый сегмент области фона). При JPEG кодировании изображений битовую плоскость разделения необходимо кодировать энтропийными методами и хранить или передавать вместе с самим изображением в сжатом виде, что в незначительной мере влияет на степень компрессии самого изображения.

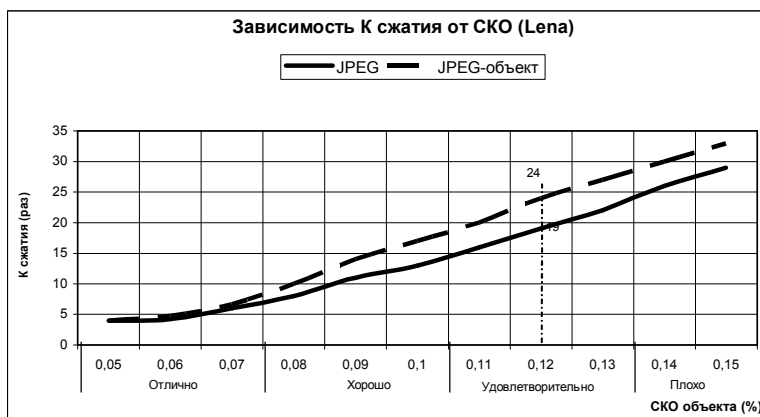


Рис. 7. Зависимость коэффициента сжатия всей плоскости изображения (Ксж) от среднеквадратической ошибки (СКО) на плоскости объекта.

Кодирование изображений на основе вейвлет преобразований с предварительным выделением объекта имеет некоторые отличия от JPEG кодирования области объекта и фона.

Выделение объекта на плоскости изображения производится таким же образом, как и в случае преобразования в JPEG формат. Однако, после формирования битовой плоскости разделения изображения на объект и фон, (рис. 5, г) применить её необходимо только к вейвлет-коэффициентам, представленным в областях HG, GH и GG на каждом шаге разложения.

В результате, все вейвлет-коэффициенты, которые принадлежат области объекта на всех этапах разложения, сохраняются, а коэффициенты, которые принадлежат области фона в соответствующих масштабах, игнорируются. При таком подходе, общий коэффициент сжатия будет более высоким, но на восстановленном изображении текстура фона будет полностью отсутствовать или заменена на абстрактную текстуру, которая не имеет никакого отношения к обрабатываемому изображению.



а) Ксж = 24, СКО объекта = 12% б) Ксж = 19, СКО объекта = 12%

Рис. 8. Изображение после раздельного кодирования методом JPEG-технологий

Использование вейвлет преобразования при кодировании изображений с разной степенью точности не требует отдельного кодирования и хранения битовой плоскости вместе с изображением. Битовая плоскость используется только на этапе разложения изображения, при восстановлении вейвлет-коэффициенты будут распределены в обратном порядке Z-сканирования плоскости разложения. Этот факт можно отнести к преимуществу вейвлет кодирования изображений по сравнению с методом JPEG.

На рис. 9 представлена зависимость коэффициента сжатия от СКО при кодировании изображения с использованием вейвлет преобразования.

Для отметки СКО = 12% степень сжатия при использовании вейвлет преобразования соответствует 23 (кривая – Вейвлет). При использовании тех же фильтров разложения и восстановления, но с разделением изображения на объект и фон степень сжатия увеличилась до 26 при том же значении СКО = 12% (кривая – Вейвлет-объект). Следует заметить, что при аналогичном значении ошибки восстановления (СКО = 12%) степень сжатия при JPEG кодировании изображения с выделением объекта равна 24 (рис. 7), что превышает соответствующий показатель при обычном вейвлет анализе (рис. 9, кривая – Вейвлет).

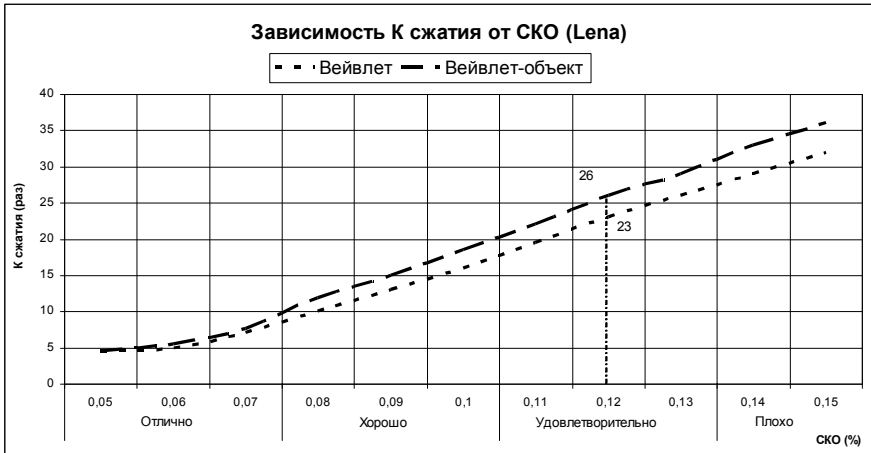


Рис. 9. Зависимость коэффициента сжатия (Ксж) от среднеквадратической ошибки (СКО) при вейвлет анализе изображения

На рис. 10 (а, б) представлены восстановленные изображения, полученные после вейвлет кодирования всей плоскости изображения, (рис. 10, а), а так же после раздельного кодирования области фона и объекта, (рис. 10, б), при одинаковом значении ошибки восстановления, СКО = 12%.



а) Ксж = 23, СКО = 12%
(без выделения объекта)



б) Ксж = 26, СКО = 12%,
(с выделением объекта)

Рис. 10. Восстановленные изображения после выделения объекта и вейвлет-кодирования

Выводы. Предложенный метод разделения изображения на объект и фон, позволяет кодировать их отдельно, как с различной степенью качества, так и с использованием различных методов. Это дает возможность представлять

наиболее информативные области изображения с более высоким качеством, чем остальные его составляющие. Благодаря такому подходу можно сохранить высокое качество информативных областей (объектов), менее информативные области (фон) представить с большей компрессией, но с худшим качеством и таким образом повысить общий коэффициент сжатия для всего изображения.

Предложенный подход позволил выявить резервы классической схемы JPEG-кодирования и получить лучшие количественные характеристики по сравнению с вейвлет-кодированием на основе формата JPEG 2000.

Также следует отметить, что относительное повышение эффективности кодирования предложенного метода выше в классическом JPEG-формате, чем в JPEG 2000. Это объясняется естественной локальной природой вейвлет-преобразований.

Список литературы: 1. *Претт У.* Цифровая обработка изображений. – М.: Мир, 1982. – Кн. 2. – 480 с. 2. *Уэлстид С.* Фракталы и вейвлеты для сжатия изображений в действии. Учеб. пособ. – М.: Изд-во "Триумф", 2003. – 336 с. 3. *Сэломон Д.* Сжатие данных, изображений и звука. – М.: Техносфера, 2004. – 368 с. 4. *Оброблення сигналів і зображень та розпізнавання образів // Праці Сьомої Всеукраїнської міжнародної конференції УкрОБРАЗ'2004.* – К.: Видання Міжнародного науково-навчального центру інформаційних технологій та систем, 2004. – 306 с. 5. *Гонсалес Р., Вудс Р.* Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072 с. 6. *Оброблення сигналів і зображень та розпізнавання образів / Праці Восьмої Всеукраїнської міжнародної конференції УкрОБРАЗ'2006.* – К.: Видання Міжнародного науково-навчального центру інформаційних технологій та систем, 2006. – 212 с. 7. *Иванов В.Г., Ломоносов Ю.В., Любарский М.Г.* Фурье и вейвлет анализ изображений в плоскости JPEG технологий // Проблемы управления и информатики. – 2004. – №5. – С. 111–124. 8. *Иванов В.Г., Любарский М.Г., Ломоносов Ю.В.* Сжатие изображений на основе компенсации контуров при вейвлет преобразовании // Проблемы управления и информатики. – 2006. – № 3. – С. 89–101. 9. *Иванов В.Г., Любарский М.Г., Ломоносов Ю.В.* Сжатие изображений на основе выделения и кодирования объектов с различным визуальным качеством / Праці Восьмої Всеукраїнської міжнародної конференції "Оброблення сигналів і зображень та розпізнавання образів" (УкрОБРАЗ'2006). – К.: Кібернетичний центр НАН України, 2006. – С. 159–163. 10. <http://www.dejavu.research.att.com/djvu/sci/djvuspec>.

Поступила в редакцію 02.04.2007