

УДК 621.391:519.728

В.Г. Иванов, М.Г. Любарский, Ю.В. Ломоносов

СОКРАЩЕНИЕ СОДЕРЖАТЕЛЬНОЙ ИЗЫТОЧНОСТИ ИЗОБРАЖЕНИЙ НА ОСНОВЕ КЛАССИФИКАЦИИ ОБЪЕКТОВ И ФОНА

Введение

В течение многих лет в исследованиях проблемы сжатия данных и изображений преобладают методы обработки сигналов и методы теории информации, которые используют декоррелирующие свойства различных линейных преобразований, а также возможности квантования и энтропийного кодирования коэффициентов этих преобразований [1–5].

Развитие указанных 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 [6–9]. Повысить количественные и качественные характеристики этих форматов можно путем сокращения содержательной избыточности изображений. Суть такого подхода заключается в том, что в памяти хранится только часть изображения, представляющая интерес в данном конкретном применении.

Основной проблемой такого подхода к сжатию изображений остается выделение значимых (информативных) участков изображений (объектов) и участков менее информативных (фона). И хотя ее решению посвящено большое количество работ [10–16], в которых получены интересные результаты, в целом задача кодирования изображений на основе этих принципов еще далека от окончательного решения.

В данной работе элементы изображения разделяются на два класса (объект и фон) при помощи вейвлет-преобразования изображения и выделения контуров в пространстве коэффициентов высокочастотных областей многомасштабного анализа. Конечный результат выделения объектов на изображении — 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 представляет собой продолжение исследований авторов, основные результаты которых содержатся в публикациях [17–19].

1. Выделение объектов на плоскости изображений

Идея кодирования изображения с различной степенью качества реализована в графическом формате DjVu [20]. 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 определении границ резких перепадов яркости между областями с равномерным распределением ее 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 планов 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.

Ограничение приведенного метода выделения переднего плана заключается в том, что значимый (информативный) объект должен характеризоваться равномерным распределением 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 разделяет изображение и кодирует его с равной степенью качества, т.е. сводится к известному формату JPEG-2000.

Цель данной работы — 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> выбран файл `zelda.bmp`, который представлен на рис. 1. Параметры исходного изображения следующие: размер 256×256 , глубина цвета в градации серого 8 бит. Вейвлет-коэффициенты исходного изображения после одного шага преобразования представлены на рис. 2.



Рис. 1



Рис. 2

На практике вейвлет-преобразование проводится с помощью биортогональных вейвлет-базисов по формулам разложения

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

и синтеза

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

где $\{a_i\}$ — коэффициенты разложения остаточного члена, $\{b_i\}$ и $\{c_i\}$ — коэффициенты нового остаточного члена.

Для разложения и восстановления сигнала используют две разные пары квадратурных зеркальных фильтров: пару $\{\tilde{h}, \tilde{g}\}$ — при разложении и пару $\{h, g\}$ — 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

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

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

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

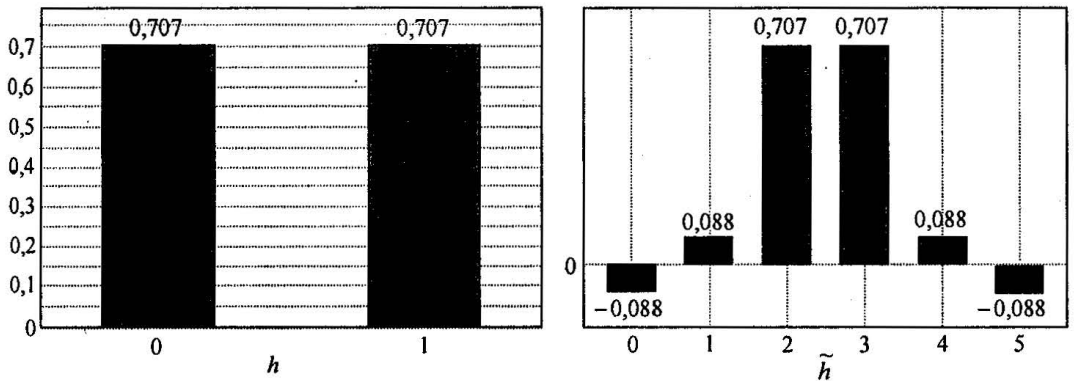


Рис. 3

При построении разделяющей битовой плоскости используются вейвлет-коэффициенты HG и GH областей (рис. 2), полученные в соответствии с выражением (1). После усредненного суммирования выбранных вейвлет-коэффициентов $(HG+GH)/2$ получаем изображение, представленное на рис. 4.



Рис. 4

Это изображение имеет более равномерное распределение яркости по сравнению с оригиналом (см. рис. 1) или остаточным членом разложения НН (см. рис. 2). Это 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) структуру (см. рис. 2). Области НГ и ГН имеют более гладкий рельеф яркости на всей плоскости, но при этом сохраняют самые существенные ее перепады. Выбор именно этих составляющих (НГ и ГН) обусловлен тем, что они получены в результате обработки исходного изображения одной парой фильтров h и \tilde{h} , но в обратной последовательности их применения. При этом область НГ содержит наиболее выраженные перепады яркости в вертикальном направлении, что обусловлено последовательностью операций свертки исходного изображения с соответствующими фильтрами h и \tilde{h} . Область ГН, наоборот, имеет резкие скачки яркости в горизонтальной плоскости.

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

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

где $X_{i,j}$ — точка усредненного изображения, $i, j = \overline{1, n}$, n — количество строк и столбцов в матрице изображения размерности $n \times n$.

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

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

где $i, j = \overline{1, n}$, n — количество строк и столбцов в матрице изображения размерности $n \times n$.

Пороговое значение привязывается к среднему значению яркости X_{mean} на всей плоскости изображения (см. рис. 4) и выбирается таким образом, чтобы исключить лишние случайных точек, т.е. тех точек, которые не имеют в своем ближайшем окружении себе подобных. В результате фильтруются одиночные точки, преодолевшие пороговое значение. На рис. 5, а наглядно показана 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 выполняется последовательное сканирование изображения по строкам и столбцам до выявления первой и последней ненулевой 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 (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 расстояние между точками разрыва, принадлежащими разным подмножествам $\{R_i\}$, т.е. между концом одного связного фрагмента границы и началом другого. После соединения граничных точек между собой получаем битовую плоскость, разделяющую объект и фон (рис. 5, в). Построение битовой плоскости — завершающий этап 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 (см. рис. 5, в).

В случае, когда сжатие изображений осуществляется при помощи JPEG-технологии на основе дискретного косинусного преобразования, плоскость изображения предварительно разбивается на фрагменты размером 8×8 . После одного шага вейвлет-преобразования плоскость изображения, представленная на рис. 4, уменьшилась в четыре раза по сравнению с плоскостью исходного изображения (см. рис. 1). Следовательно, размер фрагмента для битовой плоскости должен быть 4×4 , что полностью согласуется с разбиением исходного изображения на фрагменты 8×8 . Каждый фрагмент может принадлежать трем областям: объекта, фона и 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 .

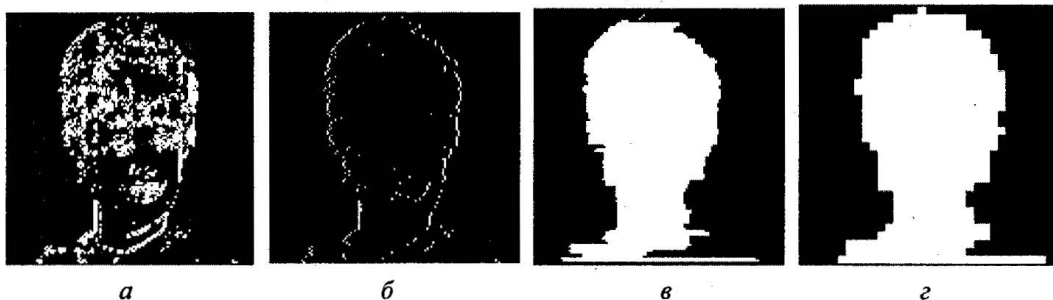


Рис. 5

В том случае, когда сжатие исходного изображения выполняется с использованием многомасштабного вейвлет-преобразования, сегментация битовой плоскости (см. рис. 5, в) не производится.

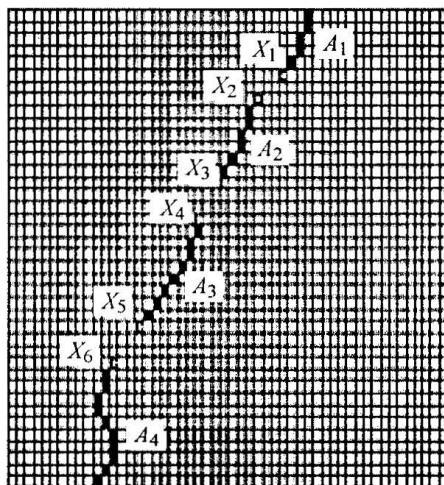


Рис. 6

2. JPEG- и вейвлет-компрессия объекта и фона изображений с различной степенью точности

Рассмотрим практические результаты кодирования изображений с различной степенью точности при использовании JPEG-технологий.

Исходное изображение (см. рис. 1) необходимо разбить на сегменты размером 8×8 и осуществить кодирование объекта и фона изображения по методу JPEG-технологии [7, 15, 17].

Согласно методу JPEG-компрессии, ядро одномерного косинусного преобразования имеет вид:

$$L_x(0) = \frac{1}{\sqrt{N}} \sum X(m); \quad L_x(R) = \sqrt{\frac{2}{N}} \sum_{m=0}^{N-1} X(m) \cos \frac{(2m+1)k\pi}{2N}. \quad (6)$$

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; значения m и k изменяются соответственно от 0 до $N-1$ и от 1 до $N-1$ с шагом, равным 1.

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, т.е. 64 коэффициента. Таким образом, сжатие изображения осуществляется только за счет сегментов, которые входят в область объекта на битовой плоскости. Область фона представлена лишь 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 сжатия и качества изображения, которая была бы приемлема для решения поставленной задачи.

На рис. 7 представлена зависимость общего коэффициента сжатия всей плоскости изображения ($K_{\text{сжмг}}$) от среднеквадратической ошибки (СКО) на плоскости объекта.

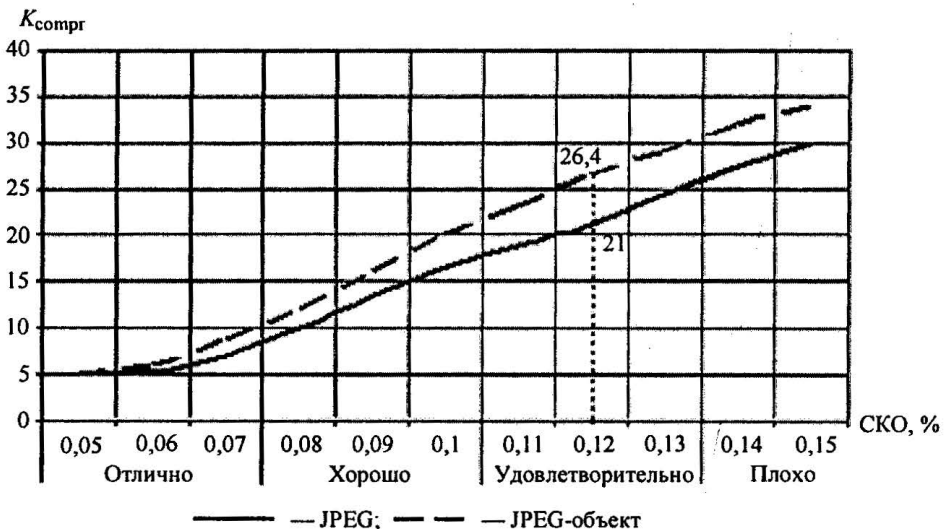


Рис. 7.

На рис. 8 дано исходное изображение после кодирования методом JPEG-компрессии. Рис. 8, а показывает восстановленное изображение после отдельного 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 21 раза до 26,4 обусловлено большей степенью сжатия области фона (в данном случае $K_{\text{сomp}}$ фона равен 64, 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 незначительно влияет на степень сжатия самого изображения.

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



Рис. 8

Сама битовая плоскость при каждом шаге разложения изображения подвергается соответствующему масштабированию, которое должно быть согласовано с каждым шагом вейвлет-преобразования. Таким образом, битовая плоскость в соответствующем масштабе представления применяется к областям HG, GH и GG на каждом шаге вейвлет-преобразования, что позволяет обеспечить сохранность вейвлет-коэффициентов области объекта на участках HG, GH и GG, которые, в свою очередь, отображают высокочастотную составляющую исходного изображения. В результате все вейвлет-коэффициенты, принадлежащие области объекта на всех этапах разложения, сохраняются, а коэффициенты, которые принадлежат области фона в соответствующих масштабах, игнорируются. Необходимо заметить, что использование битовой плоскости на остаточном члене разложения (область HH) приводит к потере низкочастотной составляющей области фона, вследствие чего из исходного изображения полностью удаляется текстура фона. При таком подходе общий коэффициент сжатия более высокий, но на восстановленном изображении текстура фона полностью отсутствует или заменяется абстрактной текстурой, не имеющей никакого отношения к обрабатываемому изображению. Использование вейвлет-преобразования при кодировании изображений с разной степенью точности не требует отдельного кодирования и хранения битовых плоскостей.

вой плоскости вместе с изображением. Битовая плоскость используется только на этапе разложения изображения, при восстановлении вейвлет-коэффициенты распределяются 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.

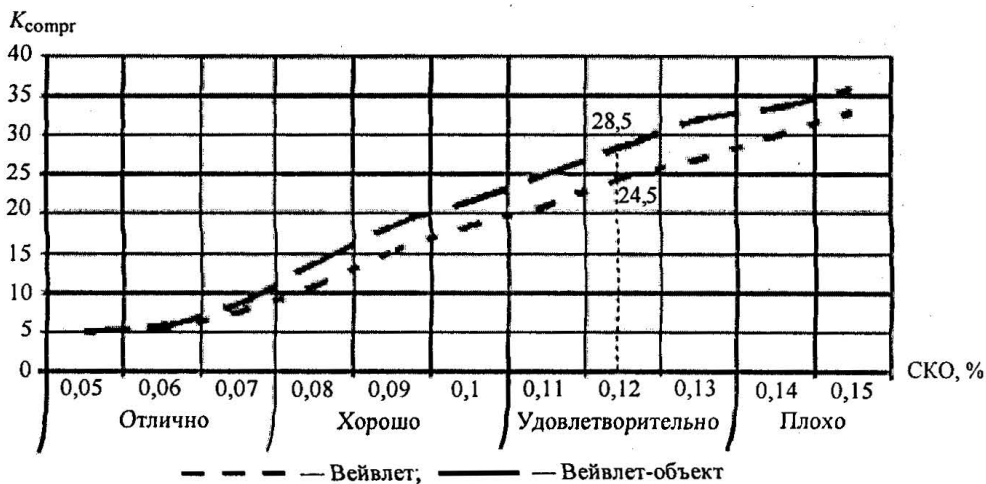
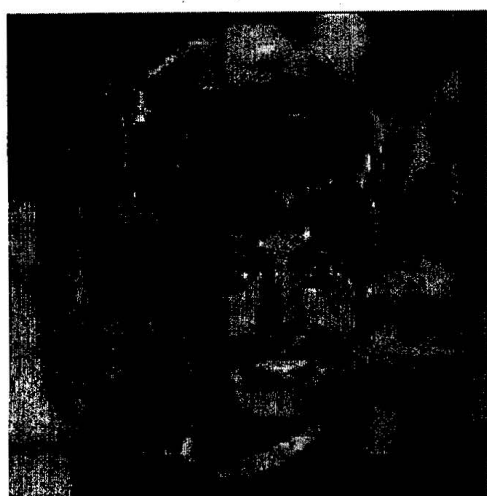


Рис. 9

Для СКО = 12 % при использовании вейвлет-преобразования происходит сжатие в 24,5 раза. При использовании тех же фильтров разложения и восстановления, но с разделением изображения на объект и фон, степень сжатия увеличивается и составляет 28,5 раза при том же значении СКО = 12 %. Следует заметить, что при таком же значении ошибки восстановления (СКО = 12 %) в случае JPEG-кодирования происходит сжатие изображения в 26,4 раза с выделением объекта (см. рис. 7), что превышает соответствующий показатель при обычном вейвлет-анализе на 8 % (см. рис. 9, вейвлет).

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



а



б

Рис. 10

Заклучение

Предложенный метод разделения изображения на объект и фон позволяет кодировать их отдельно различными методами с различной степенью качества. 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 большим сжатием, но с худшим качеством, и таким образом повысить общий коэффициент сжатия для всего изображения.

Предложенный подход позволяет 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 использование метода выделения объекта в схеме классического JPEG-формата дает более высокое повышение эффективности кодирования, чем выделение объекта в схеме многомасштабного вейвлет-преобразования изображения (JPEG-2000). Это объясняется естественной локальной природой многомасштабного анализа, схема преобразований которого уже априори приспособлена для разделения изображений.

V.G. Ivanov, M.G. Lyubarskiy, Yu.V. Lomonosov

СКРОЧЕННЯ ЗМІСТОВНОЇ НАДЛИШКОВОСТІ ЗОБРАЖЕНЬ НА ОСНОВІ КЛАСИФІКАЦІЇ ОБ'ЄКТІВ І ФОНУ

Наведено практичні результати комп'ютерного моделювання стиснення деяких класів зображень. Перед кодуванням зображення розбивається на дві частини — об'єкт і фон. Поділ зображення здійснюється в площині вейвлет-коефіцієнтів. Кодування кожної складової зображення виконується окремо з різним ступенем точності на основі JPEG-технологій з використанням узагальненого перетворення Фур'є (дискретного косинусного перетворення), а також із застосуванням багатомасштабної обробки вейвлет-коефіцієнтів.

V.G. Ivanov, M.G. Lyubarskiy, Yu.V. Lomonosov

CUTTING OF CONTENT REDUNDANCY OF THE IMAGIES ON THE BASIS OF CLASSIFICATION OF OBJECTS AND BACKGROUND

In the given operation the practical results of computer simulation of compression of some classes of the imagies are presented. The image, before encoding, is previously separated into two parts — object and background. The separation of the image is carried out in a plain of wavelet-coefficients. The encoding of each component of the image is made separately with different degree of accuracy on the basis of JPEG-technologies with usage of a generalized Fourier transform (discrete cosine transform — DCT), and also with applying of multiscale processing wavelet-coefficients.

1. Лебедев Д.С., Цуккерман И.И. Телевидение и теория информации. — М.: Энергия, 1965. — 219 с.
2. Прат У., Кейн Д., Эндрюс Х. Кодирование изображений посредством преобразования Адамара // Труды Института инженеров по электротехнике и радиоэлектронике. — 1969. — 57, № 1. — С. 66–67.

3. *Ахмед Н., Рао К.Р.* Ортогональные преобразования при обработке цифровых сигналов. — М. : Связь, 1980. — 248 с.
4. *Джайн А.К.* Сжатие видеoinформации : Обзор // Труды Института инженеров по электротехнике и радиоэлектронике. — 1981. — 69, № 4. — С. 31–58.
5. *Прэтт У.* Цифровая обработка изображений. Кн. 2 / Пер. с англ. — М. : Мир, 1982. — 480 с.
6. *Ватолин Д., Ратушняк А., Смирнов М., Юкин В.* Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. — М. : Диалог-Мифи, 2002. — 384 с.
7. *Миано Дж.* Форматы и алгоритмы сжатия изображений в действии : Учеб. пособ. — М. : Триумф, 2003. — 336 с.
8. *Уэлстид С.* Фракталы и вейвлеты для сжатия изображений в действии : Учеб. пособ. — М. : Триумф, 2003. — 320 с.
9. *Сэлмон Д.* Сжатие данных, изображений и звука. — М. : Техносфера, 2004. — 368 с.
10. *Дуда Р., Харт П.* Распознавание образов и анализ сцен / Пер. с англ. — М. : Мир, 1976. — 368 с.
11. *Денисов Д.А., Низовкин В.А.* Сегментация изображений на ЭВМ // Зарубежная радиоэлектроника. — 1985. — № 10. — С. 3–30.
12. *Кунт М., Икономопулос А., Кошер М.* Методы кодирования изображений второго поколения // Труды Института инженеров по электротехнике и радиоэлектронике. — 1985. — 73, № 4. — С. 59–86.
13. *Системы технического зрения (принципиальные основы, аппаратное и математическое обеспечение) / А.Н. Писаревский, А.Ф. Чернявский, Г.К. Афанасьев и др. / Под общ. ред. А.Н. Писаревского, А.Ф. Чернявского. — Л. : Машиностроение. Ленингр. отд., 1988. — 424 с.*
14. *Оброблення сигналів і зображень та розпізнавання образів // Пр. VII Всеукраїнської міжнародної конференції «УкрОБРАЗ'2004». — Київ : Міжнародний науково-навчальний центр інформаційних технологій та систем НАН України і Міністерства освіти та науки України, 2004. — 306 с.*
15. *Гонсалес Р., Вудс Р.* Цифровая обработка изображений. — М. : Техносфера, 2005. — 1072 с.
16. *Оброблення сигналів і зображень та розпізнавання образів // Пр. VIII Всеукраїнської міжнародної конференції «УкрОБРАЗ'2006». — Київ : Міжнародний науково-навчальний центр інформаційних технологій та систем НАН України і Міністерства освіти та науки України, 2006. — 212 с.*
17. *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.
18. *Иванов В.Г., Любарский М.Г., Ломоносов Ю.В.* Сжатие изображений на основе компенсации контуров при вейвлет-преобразовании // Там же. — 2006. — № 3. — С. 89–101.
19. *Иванов В.Г., Любарский М.Г., Ломоносов Ю.В.* Сжатие изображений на основе выделения и кодирования объектов с различным визуальным качеством // Пр. VIII Всеукраїнської міжнародної конференції «УкрОБРАЗ'2006». — Київ : Кібернетичний центр НАН України, 2006. — С. 159–163.
20. <http://www.dejavu.research.att.com/djvu/sci/djvuspec>.

Получено 25.12.2006