

УДК 004.932

С.Г. Удовенко, А.А. Шамраев, Е.О. Шамраева, Ю.Я. Чиженков

Харьковский национальный университет радиоэлектроники, Харьков

МЕТОД СЖАТИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ДИСКРЕТНОГО ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ

Усовершенствован алгоритм сжатия изображений с потерями для сжатия BMP-файлов. Результаты тестирования показали, что по степени сжатия при сопоставимых значениях PNSR модифицированный алгоритм превосходит существующие аналоги.

Ключевые слова: сжатие изображений, JPEG, дискретное вейвлет-преобразование, преобразование Хаара, локальная пороговая стратегия

Введение

На современном этапе развития компьютерных технологий все о́бщая часть информации представляется соо́мультимедиа. Цифровые видео и изображения требуют значительных расходов памяти на хранение. Задача эффективного сжатия графической информации становится все актуальнее.

В настоящее время существует достаточно много алгоритмов сжатия изображений. В их основе лежит выявление естественной избыточности исходной информации.

Сжатие изображений осуществляется ио́в пространственно, ио́в частотно́ областях. Наиболее яркими примерами пространственно сжатия изображений являются алгоритмы PCX, GIF, а частотно сжатия – JPEG.

В пространственно́ области нельзя добиться достаточно о́бщего коэффициента сжатия. Большинство алгоритмов сжатия работает в частотно́ области изображения, поскольку наиболее явно избыточность исходных данных проявляется при их спектральном анализе.

Поиск эффективного частотно́ представления изображения для сжатия является достаточно трудной задачей. С одной стороны, частотно́ преобразование довольно о́дательно вычислительно просто, а с другой – поучать хорошо «сжимаемые» частотные коэффициенты.

Вейвлет-преобразование о́дают множеством свойствами, по о́зными как при анализе, так и при сжатии изображений. Исследования показывают, что наилучшим преобразованием по соотношению «вычислительная сложность метода/размер сжатых данных» является преобразование Хаара, о́ разующее хорошо «сжимаемые» частотные коэффициенты.

Однако наряду с преимуществами методов, основанных на вейвлет-преобразовании, к их недостаткам можно отнести значительное искажение исходной информации при высоких степенях сжатия, а также отсутствие рекомендации по применению

тоо́и иио́о́азиса.

Поэтому актуальными при сжатии изображений на основе вейвлет-преобразования являются задачи оптимального выбора параметров вейвлет-коэффициентов и поиска метода оптимального кодирования вейвлет-коэффициентов. Решению этих задач посвящена настоящая статья.

Постановка задачи

Необходимо усовершенствовать алгоритм сжатия по цветным и по черно-белым (24 и 1) изображениям с регулируемым коэффициентом сжатия при допустимом уровне потерь.

Исходные изображения хранятся в формате Windows Bitmap (BMP). Алгоритм должен быть ориентирован на улучшение сжатия фотографий истинных изображений, т.е. изображений, полученных при помощи цифрового ио́анализа о́во фотоаппарата, сканера и др.

Средняя степень сжатия изображений и среднее время работы программы должны быть сопоставимы с существующими аналогами.

Вейвлет-преобразование

На основании понятия о векторном пространстве о́щим подходом к анализу сигнала $s(t)$ ставится их представление в виде взвешенной суммы простых составляющих – азисных функций $\Psi_k(t)$, помноженных на коэффициенты C_k :

$$s(t) = \sum_k C_k \Psi_k(t).$$

С помощью вейвлетов сигнал представляется совокупностью во новых пакетов, о́ разованных на основе некоторого исходного азисово́ функции $\Psi_0(t)$.

Число используемых вейвлетов при разложении сигнала задает уровень декомпозиции сигнала. За нулевой уровень декомпозиции часто принимают сам сигнал, а последующие уровни декомпозиции о́ разуют ниспадающее вейвлет-дерево.

Одна из основополагающих идей в представлении сигнала заключается в разбивке представления на две составляющих: непрерывную (аппроксимирующую) и периодическую (детализирующую), с последующим их уточнением итерационным методом.

Каждый шаг такого уточнения соответствует определению уровня декомпозиции и реставрации сигнала.

Как и преобразование Фурье, в вейвлет-преобразовании можно применять как к непрерывным сигналам – непрерывные вейвлет-преобразования, так и к цифровым сигналам – дискретные вейвлет-преобразования (ДВП).

ДВП может быть представлено в матричном виде и на основе аналогов Фурье.

Обозначим через вектор v^j последовательность конечно-дискретного сигнала $c_{j,n}$ на некотором уровне преобразования j . Этот вектор преобразуется в вектор, содержащий последовательности $c_{j+1,n}$ и $d_{j+1,n}$, каждая из которых имеет длину вдвое короче $c_{j,n}$.

Матричное преобразование может быть записано в виде матричного умножения $v^{j+1} = \hat{I}_j v^j$, где матрица \hat{I}_j – квадратная и состоит из нулевых и единичных элементов h_n , умноженных на $\sqrt{2}$.

В силу свойства h_n матрица \hat{I}_j является ортонормированной, и ортогональная матрица равна транспонированной.

В качестве иллюстрации рассмотрим следующую пример.

Возьмем фильтры $L=4$, последовательность длины $N=8$, а в качестве начального значения $j=0$.

Тогда операция матрично-векторного умножения:

$$\begin{bmatrix} c_{1,0} \\ c_{1,1} \\ c_{1,2} \\ c_{1,3} \\ d_{1,0} \\ d_{1,1} \\ d_{1,2} \\ d_{1,3} \end{bmatrix} = \sqrt{2} \begin{bmatrix} h_0 & h_1 & h_2 & h_3 & & & & \\ & & h_0 & h_1 & h_2 & h_3 & & \\ & & & h_0 & h_1 & h_2 & h_3 & \\ h_2 & h_3 & & & & & & \\ h_3 & -h_2 & h_1 & -h_0 & & & & \\ & & h_3 & -h_2 & h_1 & -h_0 & & \\ & & & h_3 & -h_2 & h_1 & -h_0 & \\ h_1 & -h_0 & & & & h_3 & -h_2 & \end{bmatrix} \begin{bmatrix} c_{0,0} \\ c_{0,1} \\ c_{0,2} \\ c_{0,3} \\ c_{0,4} \\ c_{0,5} \\ c_{0,6} \\ c_{0,7} \end{bmatrix}$$

Ортогональное преобразование есть умножение на ортогональную матрицу M_j^T :

$$\begin{bmatrix} c_{1,0} \\ c_{1,1} \\ c_{1,2} \\ c_{1,3} \\ d_{1,0} \\ d_{1,1} \\ d_{1,2} \\ d_{1,3} \end{bmatrix} = \sqrt{2} \begin{bmatrix} h_0 & & h_3 & h_3 & & & & h_1 \\ & h_1 & & -h_2 & h_3 & & & -h_0 \\ & & h_2 & & h_1 & -h_2 & & \\ h_3 & h_0 & & -h_0 & h_1 & & & \\ & h_1 & & & & -h_0 & h_3 & \\ & & h_2 & h_1 & & & -h_2 & \\ & & h_3 & h_2 & h_1 & & h_1 & h_3 \\ & & & h_3 & h_2 & & -h_0 & -h_2 \end{bmatrix} \begin{bmatrix} c_{0,0} \\ c_{0,1} \\ c_{0,2} \\ c_{0,3} \\ c_{0,4} \\ c_{0,5} \\ c_{0,6} \\ c_{0,7} \end{bmatrix}$$

Полное ДВП заключается в итеративном умножении верхних полей вектора v^{j+1} на квадратную матрицу \hat{I}_{j+1} , размер которой 2^{d-j} . Эта процедура может повторяться d раз, пока длина вектора не станет равной единице.

Целочисленные вейвлет-преобразования (ЦВП)

В основе целочисленных методов лежит модификация вейвлет-преобразования, позволяющая производить все вычисления в целочисленном виде. Полученное преобразование не является, строго говоря, вейвлет-преобразованием, но обладает всеми его свойствами. Теоретически при вейвлет-преобразовании потери информации не происходит. Однако при реализации возникают неизбежные ошибки округления вейвлет-коэффициентов. Вместе с тем, в некоторых приложениях ортогональность преобразования является важным свойством. Целочисленное вейвлет-преобразование позволяет достичь погрешности над точностью вычислений. Поэтому оно по-прежнему называется вейвлет-преобразованием. Кроме того, целочисленность вычислений ускоряет выполнение преобразования.

Пусть $\{c_n^0\}_{n=0}^{N-1}$ – исходный сигнал, где верхний индекс показывает уровень разложения (0), нижний – конкретную точку сигнала. Пусть $\{c_n^1\}_{n=0}^{N_1-1}$ и $\{d_n^1\}_{n=0}^{N_1-1}$ – составляющие разложения на первом уровне (низкочастотная – аппроксимирующая и высокочастотная – детализирующая части, соответственно). Здесь $N_1 = N/2$.

Преобразование Хаара использует следующие фильтры декомпозиции:

$$\tilde{h}_n = \left\{ \frac{1}{2}, \frac{1}{2} \right\},$$

$$\tilde{g}_n = \left\{ \frac{1}{2}, -\frac{1}{2} \right\}$$

где \tilde{h}_n – низкочастотный фильтр, а \tilde{g}_n – высокочастотный фильтр.

Целочисленное **вейвлет-преобразование (2, 2)** эквивалентно вейвлет-преобразованию Хаара, используемому с ведущими фильтрами декомпозиции:

$$\tilde{h}_0 = \tilde{h}_1 = \tilde{g}_0 = 1/2.$$

Вычисление ведется с ведущим образом:

$$\begin{aligned} d_k^1 &= c_{2k}^0 - c_{2k+1}^0, k = 0, \dots, N_1 - 1, \\ c_k^1 &= \text{int} \left(\frac{d_k^1}{2} \right) + c_{2k+1}^0, k = 0, \dots, N_1 - 1, \\ c_{N_1-1}^1 &= \text{int} \left(\frac{d_{N_1-1}^1}{2} + c_{N_1-1}^0 \right). \end{aligned}$$

Здесь int означает операцию округления. Таким образом, все элементы $\{c_n^1\}_{n=0}^{N_1-1}$ и $\{d_n^1\}_{n=0}^{N_1-1}$ будут целыми числами. Алгоритм реконструкции имеет вид:

$$\begin{aligned} c_{2k+1}^0 &= c_k^1 - \text{int} \left(\frac{d_k^1}{2} \right), k = 0, \dots, N_1 - 1, \\ c_{2k}^0 &= d_k^1 + c_{k+1}^0, k = 0, \dots, N_1 - 1. \end{aligned}$$

Вейвлет-преобразование Лэйзи заключается в простом разложении исходного сигнала на четную и нечетную части.

На этапах декомпозиции и реконструкции используются одни и те же формулы:

$$\begin{aligned} c_k^0 &= c_{2k}^1, k = 0, \dots, N_1 - 1, \\ d_k^0 &= c_{2k+1}^1, k = 0, \dots, N_1 - 1, \end{aligned}$$

Следует отметить интересное свойство данного преобразования. Если пиксели изображения представляются некоторым числом i и t , то такое же число i и t может быть использовано для представления значений вейвлет-коэффициентов. Данное свойство преобразования поучительно называть свойством сохранения точности.

Целочисленное **вейвлет преобразование (1,3)** – это разновидность неинтерполяционного преобразования, использующее ортонормальную пару фильтров: $\tilde{h}_n = \{1, 0, 0\}$ и $\tilde{g}_n = \{1/4, -1/2, 1/4\}$. Вычисления начинаются с вейвлет-Лэйзи с последующим изменением высокочастотных коэффициентов

$$\begin{cases} d_k^1 = \text{int} \left(\frac{c_k^1 + c_{k+1}^1}{2} \right) - d_k^{1,0}, k = 0, \dots, N_1 - 2, \\ d_{N_1-1}^1 = c_{N_1-1}^1 - d_{N_1-1}^{1,0}. \end{cases}$$

Реконструкция происходит с ведущим образом:

$$\begin{cases} \tilde{h}_{2k}^0 = c_k^1, k = 0, \dots, N_1 - 1, \\ \tilde{h}_{2k+1}^0 = \text{int} \left(\frac{c_{2k}^0 + c_{2k+2}^0}{2} \right) - d_k^1, k = 0, \dots, N_1 - 2, \\ \tilde{h}_{N_1-1}^0 = c_{N_1-2}^1 - d_{N_1-1}^1. \end{cases}$$

Целочисленное **вейвлет-преобразование (2, 6)** эквивалентно использованию ведущих фильтров аналога:

$$\tilde{h}_n = \{0, 0, 1/2, 1/2, 0, 0\},$$

$$\tilde{g}_n = \{-1/16, -1/16, 1/2, -1/2, 1/16, 1/16\}.$$

Декомпозиция выполняется аналогично преобразованию Хаара.

Для выбора наилучшего вейвлет-преобразования рассмотренные ЦВП и применены к растровому файлу Lena.bmp. В выходные файлы записываются коэффициенты всех субопос, кроме диагональных. Полученные файлы сжимаются при помощи энтропийного метода арифметического сжатия. Анализ полученных результатов показывает, что преобразование Хаара обладает наименьшей вычислительной сложностью и формирует хорошо сжимаемые частотные коэффициенты. Данное преобразование можно применять как к цветным изображениям, независимо от используемого цветового пространства RGB и YCbCr.

Усовершенствование метода сжатия изображений на основе ЦВП

Процесс сжатия изображений можно разбить на четыре этапа:

- применение ортонормальных дискретных ортонормальных преобразований к изображению;
- выделение наиболее значимых частотных коэффициентов (декомпозиция частотных коэффициентов);
- квантование полученных частотных коэффициентов;
- сжатие выделенных значимых коэффициентов энтропийным и усовершенствованным методом сжатия.

Поскольку рассматриваемый метод Хаара относится к методам с потерями, то важным вопросом является **оценка качества восстановленного изображения**. В работе для сравнения исходного $f(x, y)$ и восстановленного $\hat{f}(x, y)$ изображения размера $m \times n$ используется отношение пиковой сигнал/шум (PSNR):

$$d(x, y) = 10 \log_{10} \frac{255^2 * m * n}{\sum_{x,y} (f(x, y) - \hat{f}(x, y))^2}.$$

Эффективность выделения значимых коэффициентов зависит от **способа обхода матрицы** преобразования. Можно выделить различные методы обхода: по строкам и столбцам, по спирали, зигзагообразное сканирование.

Применение зигзагообразного сканирования для декомпозиции изображения с помощью ЦВП является приоритетным, т.к. позволяет сосредоточить большее количество значимых коэффициентов вначале преобразованного вектора, что позволяет эффективно их закодировать при сжатии.

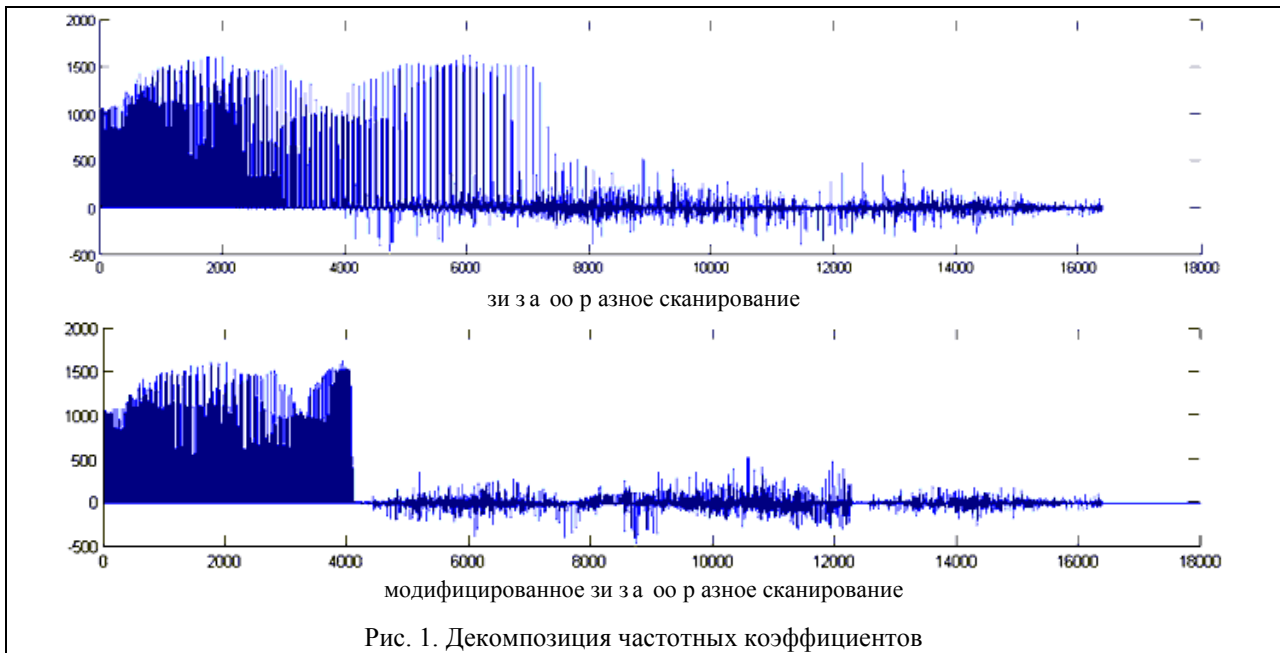


Рис. 1. Декомпозиция частотных коэффициентов

Возможны два варианта формирования вектора коэффициентов из множества вейвлет-коэффициентов. В первом случае происходит перемешивание коэффициентов разных ветвей вейвлет-дерева, находящихся на различных уровнях преобразования, а во втором варианте коэффициенты различных уровней преобразования записываются единой группой в результирующий вектор. Первый вариант группирования вейвлет-коэффициентов предпочтительнее второго, поскольку коэффициенты одного подуровня и масштаба располагаются рядом. Эффективность по сравнению с обычным зигзаговым сканированием проиллюстрирована на рис. 1. Из рисунка видно, что модифицированный метод зигзагового сканирования обеспечивает меньшую дисперсию частотных коэффициентов и, следовательно, является более эффективным.

В отличие от JPEG кодирования [1, 2], при **квантовании вейвлет-коэффициентов** используется однородное квантование, при котором все коэффициенты из всех субполос квантуются одним и тем же коэффициентом. Таким образом, некоторая часть информации теряется и мы учаем матрицы детализации с большим количеством нулей, что приводит к увеличению степени сжатия. В [3] был предложен метод выбора порога квантования с помощью кумулятивных профилей энергии. Бы рассмотрен одномерный сигнал, к которому применено преобразование Хаара с оптимальным порогом, а полученные частотные коэффициенты упорядочены в порядке убывания абсолютных значений:

$$L_1 \geq L_2 \geq L_3 \geq L_n.$$

Для одинаковых коэффициентов их положение в последовательности не изменяется. Поскольку преобразование Хаара сохраняет энергию исходного

сигнала, то энергия сигнала и сумма квадратов коэффициентов может быть вычислена как сумма квадратов коэффициентов:

$$E_f = \sum_{k=1}^{n \cdot m} L_k^2,$$

где n и m — размеры сигнала.

Кумулятивные профили энергии определяются следующим образом:

$$\left(\frac{L_1^2}{E_f}, \frac{L_1^2 + L_2^2}{E_f}, \frac{L_1^2 + L_2^2 + L_3^2}{E_f}, \dots, 1 \right). \quad (1)$$

При выборе порога квантования задается допустимая энергия с потерями сигнала ϵ_{rq} . Значимые коэффициенты определяются с помощью следующего выражения:

$$\frac{L_1^2 + L_2^2 + \dots + L_k^2}{E_f} = \epsilon_{rq}.$$

Коэффициенты накапливаются до тех пор, пока не будет достигнуто значение ϵ_{rq} . Все коэффициенты меньшие L_k являются незначимыми и обнуляются.

Выражение (1) описывает общую пороговую технику, которая была адаптирована для вычисления локальных порогов в двумерных сигналах. В этом случае профили энергии вычисляются для каждого отдельного субполоса детализации и пороги выражаются таким образом, чтобы допустимая энергия сжатого сигнала была не меньше ϵ_{rq} .

Обозначив через ϵ_A допустимую энергию и низкочастотную аппроксимирующую субполосу сигнала, а через $\epsilon_1 \dots \epsilon_{3n}$ — допустимые энергии отдельных высокочастотных субполос детализации от декомпозиции $1 \dots n$ уровней, можем записать:

$$1 = \epsilon_A + \epsilon_1 + \epsilon_2 + \dots + \epsilon_{3n}.$$

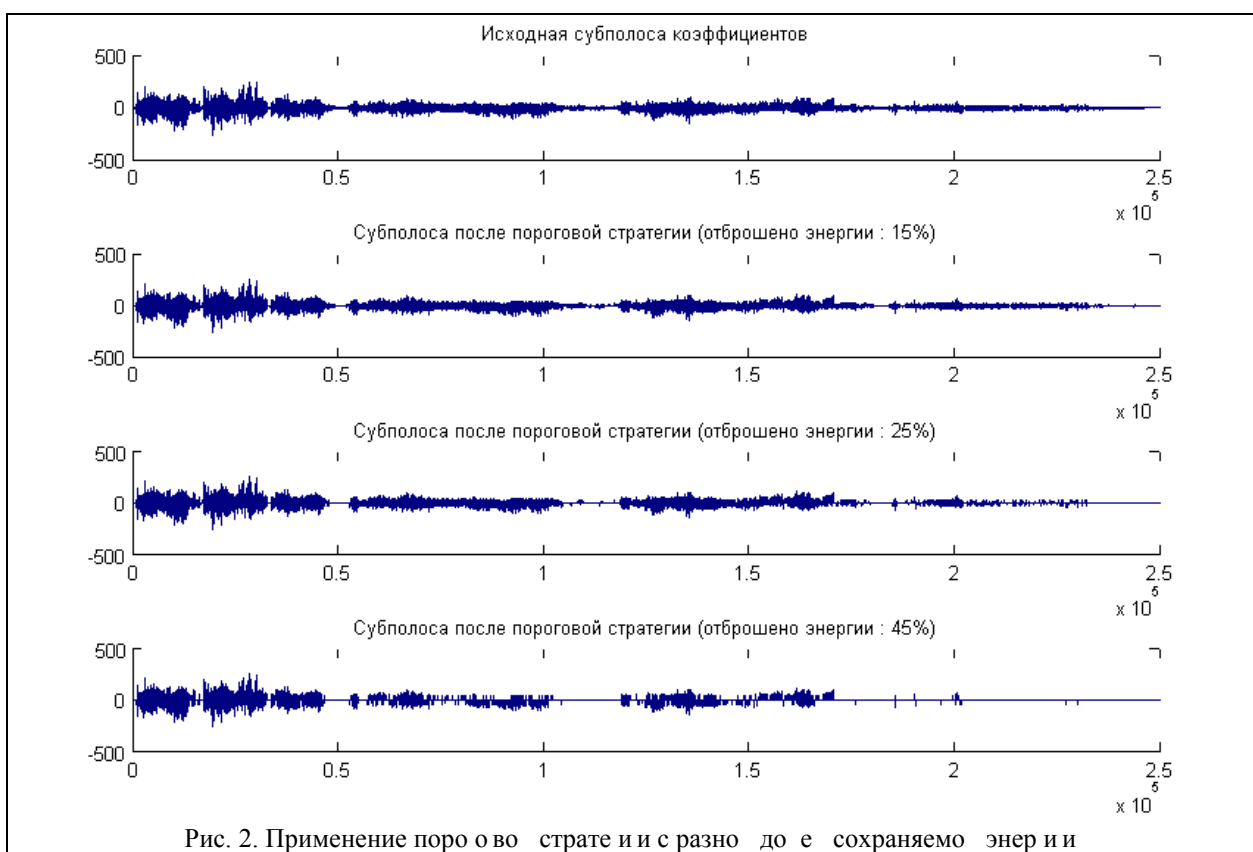


Рис. 2. Применение поро во страте и с разно до е сохраняемо энер и и

До я энер ии γ д я каждо су по осы дета и зации может ы ть вычис ена по форму е:

$$\gamma = \frac{\epsilon_{rq} - \epsilon_A}{\epsilon_1 + \epsilon_2 + \dots + \epsilon_{3n}}$$

Учитывая, что $1 - \epsilon_A = +\epsilon_1 + \epsilon_2 + \dots + \epsilon_{3n}$ по учим

$$\gamma = \frac{\epsilon_{rq} - \epsilon_A}{1 - \epsilon_A},$$

де $\epsilon_A < \epsilon_{rq} < 1$.

Таким о р а з о м, д я т о о ч т о ы н а т и п о р о д я каждо су по осы, нео х одимо рассчитать куму я тивны профи ь энер ии, а затем вы р ать поро таким, что ы до я сохраненно энер ии равня ась γ :

$$\frac{L_1^2 + L_2^2 + \dots + L_k^2}{E_S} = \gamma,$$

де L_1, \dots, L_k – частотные коэффициенты в су по о се, E_S – о щая энер ия в су по о се, а поро до жен ы ть вы р а н меньше чем L_k . Чем о ь ше энер ии содержится в низкочастотно аппроксимирующе су по о се, тем меньше значение γ , и тем меньшее ко ичество энер ии до жно ы ть сохранено высокочастотными коэффициентами. Пред о женны метод дает учшие резу ьтаты, так как поро квантования вы ирается в зависимости от распространения энер ии в су по осах дета и зации, а не просто как до я от максима ьно возможно о коэффициентен

та.

С едует отметить, что до я сохраняемо энер ии ϵ_{rq} д я каждо о изо р ажения вы ирается экспериментально и я в яется одними из задаваемых параметров разра о танно о а о ритма сжатия. Разница между су по осами коэффициентов дета и зации до и пос е применения поро во страте ии на ядно продемонстрирована на рисунке 2.

Тестирование степени сжатия производи ось на на о ре фотореа истичных тестовых bitmap-изо р ажении (рисунок 3). Усовершенствованны а о ритм сравнива ся с JPEG и азо вым ве в ет прео р азованием. Чис енные резу ьтаты тестов приведены в та . 1.

Та и ца 1

Резу ьтаты сжатия тестово о на о ра изо р ажении

	Размер исходно о фа а, кБ	JPEG, кБ	Базовое ЦВП, кБ	ЦВП с о ка ьно поро во страте и е , кБ
lena	257	34,9	36,2	31,5
barbara	257	29,3	29,5	27,7
man	257	37,4	38,5	35,4
cameraman	257	30,6	30,8	27,9
goldhill	257	42,5	44,1	41,1

Как видно из резу ьтатов тестирования, модифицированные а о ритм и е о реа и зация о есп ечивают учшую степень сжатия среди тестируемых методов.

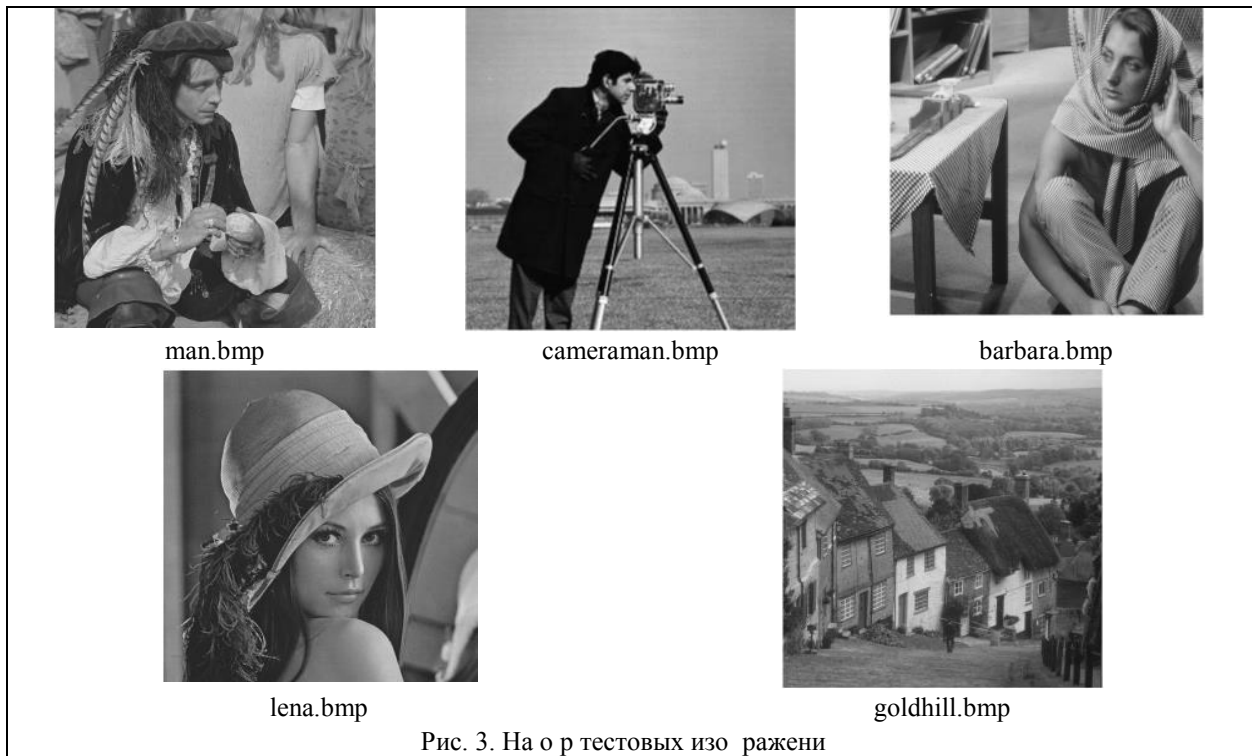


Рис. 3. На о р тестових изо ра жени

Выводы

В рамках данно о исс едования ы усовершенствован а ритм сжатия изо ра жени с потерями на основе це очис енн о ве в ет-прео р азования. Бы а пред ожена модификация метода вы ора поро а квантования, представ яющая со о о ка ьную поро о вую страте ию, о адающую о ьше эффективностью по сравнению с о а ьно стратегией. К ючевым моментом яв яется то, что она позво яет сократить о ьшее ко ичество коэффициентов при одинаковых значениях показате я PSNR. Сокращение происходит с учетом распространения енер ии в су по осах, поэтому при меньше сохраняемо енер ии матрицами дета изац ии эффективность их кодирования повышается.

Бы а также пред ожена модификация метода зи за о о р азно о сканирования матриц дета изац ии, что позво и о существенно сократить из ы точность за счет то о, что коэффициенты не то ько одно о подуровня, а и масшта а р аспо а аются рядом.

Применение рассмотренных усовершенствовани позво и о повысить эффективность арифметическо о кодирования, примененно о к под отов енн о пос едовате ьности коэффициентов, на 6 – 13%, по сравнению с азо вым ве в ет-прео р азованием и показа о свою конкурентоспос о ность по отношению к методу сжатия JPEG.

Список литературы

1. Solomon D. *Data Compression* / D. Solomon // NY.: Springer-Verlag, 2004. – 898p.
2. Сэломон Д. *Сжатие данных, изображений и звука* / Д. Сэломон // М.: Техносфера, 2004. – 368 с.
3. Walker J.S. *A Primer on Wavelets and Their Scientific Applications* / Walker J.S. – Second Edition. – Chapman & Hall/CRC., 2008. – 320 p.

Поступила в редколлегию 10.07.2012

Рецензент: д-р техн. наук, проф. Е.В. Бодянки, Харьковский национа ьны университет радиоэ ектроники, Харьков.

МЕТОД СТИСНЕННЯ ЗОБРАЖЕНЬ НА ОСНОВІ ДИСКРЕТНОГО ВЕЙВЛЕТ-ПЕРЕТВОРЕННЯ

С.Г. Удовенко, А.А. Шамраев, Е.О. Шамраева, Ю.Я. Чиженов

Удосконалено алгоритм стиснення зображень з втратами для стиснення BMP-файлів. Результати тестування показали, що за ступенем стиснення при порівняних значеннях PSNR модифікований алгоритм перевершує надіснуючі аналоги.

Ключові слова: стиснення зображень, JPEG, дискретне вейвлет-перетворення, перетворення Хаара, локальна порогова стратегія.

COMPRESSION METHOD OF IMAGES BASED ON DISCRETE WAVELET TRANSFORM

S.G. Udovenko, A.A. Shamraev, E.O. Shamraeva, U.Y. Chigenkov

The algorithm of lossy image compression to compress BMP-files. Test results showed that the improved is superior over existing counterparts in compression degree with comparable values of PSNR.

Key words: image compression, JPEG, discrete wavelet transform, Haar, a local threshold strategy.