

УДК 621.391

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

ФУРЬЕ- И ВЕЙВЛЕТ-АНАЛИЗ ИЗОБРАЖЕНИЙ В ПЛОСКОСТИ JPEG-ТЕХНОЛОГИЙ

На сегодняшний день с файловым форматом JPEG связано практически все аппаратное и программное обеспечение, так или иначе касающееся изображений. Однако есть все основания говорить, что ему на смену идет новый формат JPEG 2000. Вместо дискретного косинусного преобразования, базового для JPEG, JPEG 2000 использует технологию вейвлет-преобразований (всплесков), которая основывается на представлении сигнала в виде суперпозиции специальных функций — волновых пакетов [1–3]. В результате при той же компрессии изображение получается более гладким и четким.

Фрактальные алгоритмы сжатия изображений также обеспечивают достаточно высокую степень сжатия, но требуют огромных вычислительных мощностей при архивации. К тому же фрактальный формат FIF не поддерживается операционными системами Windows. Поэтому говорить о широком практическом использовании этого формата пока еще рано.

В данной работе приводятся результаты компьютерного моделирования сжатия некоторых классов изображений с использованием обобщенных преобразований Фурье (дискретное косинусное преобразование, преобразование Хаара) и вейвлет-преобразований на базе JPEG-технологий. Эта работа — логическое продолжение исследований авторов (см. [4]).

Постановка задачи и ее решение. Идея замены одноцветного изображения как непосредственного объекта кодирования коэффициентами его двумерного преобразования Фурье известна давно [5, 6]. Изображение подвергается унитарному преобразованию, и полученные в результате коэффициенты преобразования квантуются, кодируются и передаются по каналу связи или записываются в архив. Сегодня этот метод утвердился как эффективное и практически удобное средство кодирования одноцветных, цветных и спектрональных изображений, в том числе и в телевизионных системах, действующих в реальном масштабе времени.

Преимущества обработки сообщений с использованием ортогональных преобразований вытекают из особенностей распределения энергии среди коэффициентов преобразования.

Наличие корреляции между компонентами X_i и X_k ($i, k = \overline{1, M}$) исходного вектора сигнала X , дисперсии σ_i^2 ($i = \overline{1, M}$) которых, как правило, распределены равномерно, приводит к существенно неравномерному распределению дисперсий по элементам вектора Y , полученного с помощью матрицы P оператора преобра-

зования с системой базисных векторов p_i ($i = \overline{1, M}$). При этом, как показали экспериментальные результаты, большие дисперсии сосредоточены в малом количестве элементов вектора Y , а дисперсии остальных его элементов сравнительно малы [7–9]. Следовательно, в большинстве случаев такие элементы можно отбросить или отвести на их кодирование малое количество двоичных единиц при незначительной потере качества исходного сообщения X . Таким образом, показателем эффективности использования того или иного преобразования может служить мера неравномерного распределения дисперсий по элементам Y . В качестве такой меры можно использовать теоретико-информационное понятие энтропии [10], но примененное уже не к распределению вероятностей дискретного по значениям сообщения X , а к распределению дисперсий σ_j^2 ($j = \overline{1, M}$) по элементам вектора Y .

Если ортогональному преобразованию подвергаются блоки по m отсчетов, то энергия распределения дисперсий определяется уравнением [11, 12]

$$E(P, M, m) = -\frac{M}{m} \sum_{j=1}^m \sigma_j^2 \log_2 \sigma_j^2, \quad m = \overline{1, M}. \quad (1)$$

Из этого выражения следует, что чем неравномернее распределены дисперсии σ_j^2 вектора Y , тем меньше $E(P, M, m)$. Преобразование с матрицей Y , минимизирующее (1), наиболее эффективно в смысле компактного представления вектора X . Таково преобразование Карунена–Лозва — разложение изображения по собственным векторам [13, 14]. Однако это преобразование характеризуется очень большим объемом вычислений, примерно равным M^4 , быстрого вычислительного алгоритма не существует. Кроме того, требуется предварительное знание статистических характеристик ансамбля передаваемых изображений. Пока что оно интересно лишь с теоретической точки зрения для оценки потенциальных возможностей ортогональных преобразований.

Сложность преобразований существенно уменьшается при использовании разделимых ортогональных (в общем случае унитарных) преобразований, имеющих быстрые вычислительные алгоритмы. К ним относятся преобразования Фурье, Хаара и косинусное, которые наиболее интересны для целей сжатия данных.

Основное свойство дискретного косинусного преобразования состоит в том, что его базисные векторы очень хорошо аппроксимируют собственные векторы теплицевых матриц. По своим декоррелирующим свойствам оно приближается к преобразованию Карунена–Лозва, сохраняя при этом возможности алгоритмов быстрых вычислений Фурье. Этим, собственно, и объясняется использование этого преобразования в JPEG-форматах представления и передачи изображений.

Наиболее эффективна, с точки зрения вычислительных затрат, приходящихся на одну ординату преобразования, ортонормированная система кусочно-постоянных функций базиса Хаара [15]. Кроме этого, базис Хаара обладает весьма интересным свойством локальности, которое лежит в основе современной теории вейвлет-преобразований.

Кодирование цветных изображений на основе обобщенных преобразований Фурье в терминах JPEG-технологий можно представить в виде блок-схемы, изображенной на рис. 1.

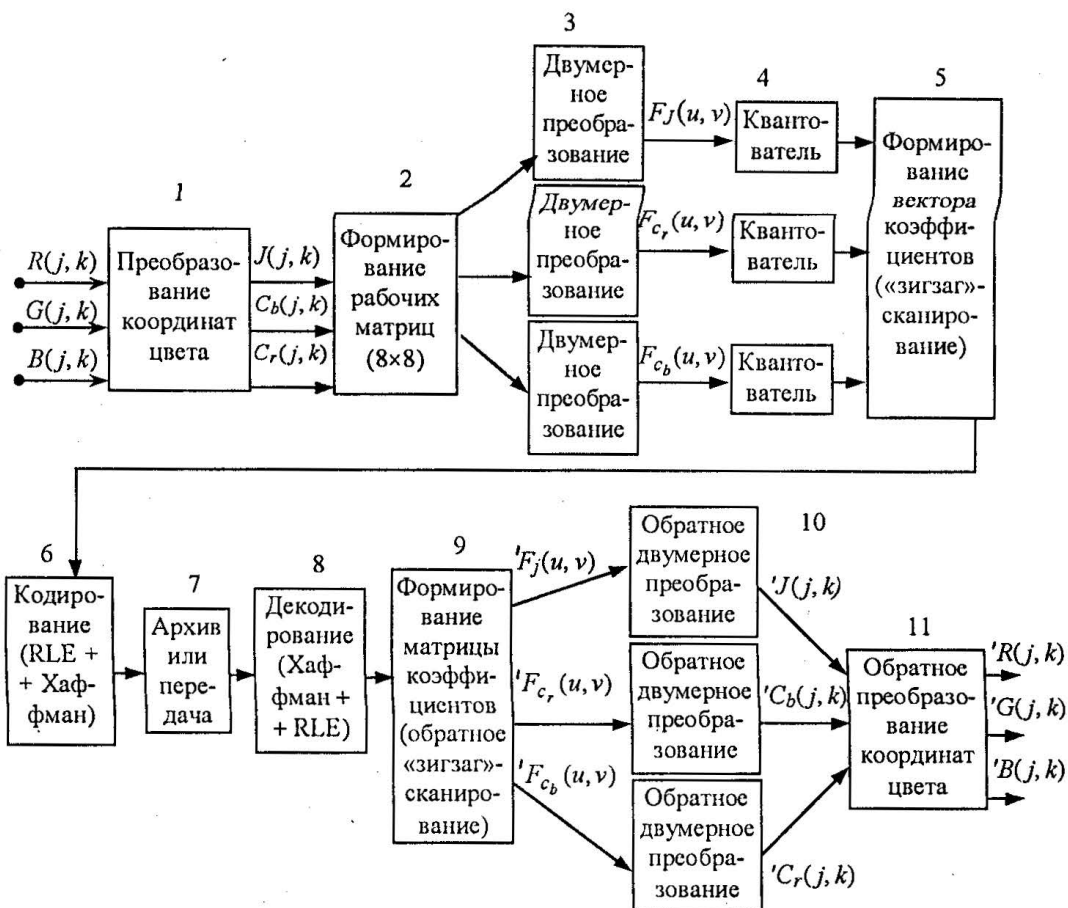


Рис. 1

В этой схеме цветное изображение представлено сигналами координат цвета источника $R(j, k)$, $G(j, k)$, $B(j, k)$, определяющими соответственно красную, зеленую и синюю составляющие для каждого элемента изображения с координатами (j, k) . На первом шаге система координат цвета источника преобразуется в другую трехмерную систему координат $J(j, k)$, $C_r(j, k)$, $C_b(j, k)$. В ней $J(j, k)$ — яркостная составляющая, C_r и C_b — компоненты, отвечающие за цвет (хроматический красный и хроматический синий). Преобразование координат определяется следующим образом:

$$\begin{matrix} C_b \\ C_r \end{matrix} \begin{vmatrix} J(j, k) \\ C_b(j, k) \\ C_r(j, k) \end{vmatrix} = \begin{vmatrix} 0,299 & 0,587 & 0,114 \\ 0,5 & -0,4187 & -0,0813 \\ 0,1687 & -0,3313 & 0,5 \end{vmatrix} \times \begin{vmatrix} R(j, k) \\ G(j, k) \\ B(j, k) \end{vmatrix} + \begin{vmatrix} 0 \\ 128 \\ 128 \end{vmatrix}. \quad (2)$$

Сигналы J , C_r и C_b удобнее кодировать, чем координаты цвета R , G , B , по той причине, что эти сигналы в значительной степени декоррелированы и основная часть их энергии приходится на компоненту J . В результате появляется возможность применить более эффективный квантователь. За счет того, что человеческий глаз менее чувствителен к цвету, чем к яркости, появляется возможность хранить массивы для компонент C_r и C_b с большими потерями и соответственно большими коэффициентами сжатия.

На втором шаге алгоритма формируются рабочие матрицы. С этой целью исходное изображение разбивается на матрицы 8×8 и из каждой формируется по три рабочие матрицы с глубиной цвета 8 бит отдельно для каждой компоненты. При увеличении степени сжатия изображение делится по компоненте J , как и в первом

случае, а для компонент C_r и C_b матрицы набираются через строчку и через столбец, т.е. из исходной матрицы размером 16×16 получается только одна рабочая матрица для ортогонального Фурье-преобразования. При этом, как нетрудно заметить, мы теряем 75 % полученной информации о цветовых составляющих изображения и с учетом яркостной составляющей получаем общее сжатие в два раза. На результирующем RGB -изображении, как показала практика, это сказывается несущественно.

На выбор размерности рабочих матриц кодирования оказывают влияние два взаимно противоречивых требования. С одной стороны, выгодно увеличивать размеры одновременно кодируемых групп элементов изображения для лучшего использования связей между ними, с другой стороны — при этом быстро возрастают вычислительные затраты. Эксперименты показали, что при увеличении размеров более 16×16 для всех типов преобразований при фиксированном коэффициенте сокращения цифрового потока уменьшение среднего квадрата ошибки становится незначительным, а трудности реализации резко увеличиваются. Поэтому практически наиболее удобный компромисс — выбор матриц размерности 8×8 .

Эффективность схемы сжатия с использованием JPEG-технологий в целом зависит от того, насколько хорошо выполняет свою роль каждое из преобразований, производимое в процессе технологической цепочки кодирования. Так, трехмерное преобразование координат цвета перераспределяет энергию между тремя компонентами колориметрического описания изображения с целью ее концентрации. Затем двумерное преобразование перераспределяет энергию каждой компоненты в отдельности, концентрируя ее в узкой области спектра.

В результате прямого унитарного преобразования матрицы изображения $F(n_1, n_2)$ размера $N_1 \times N_2$ образуется матрица преобразованного изображения того же размера, элементы которой равны:

$$F(m_1, m_2) = \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} (F(n_1, n_2) A(n_1, n_2; m_1, m_2)), \quad (3)$$

где $A(n_1, n_2; m_1, m_2)$ — ядро прямого преобразования.

Так как преобразования Фурье разделимы, то результат воздействия двумерного унитарного преобразования можно находить в два этапа: сначала выполняется одномерное преобразование по всем строкам матрицы изображения, а затем — по всем столбцам полученной матрицы.

Ядро одномерного косинусного преобразования для пункта 3 (рис. 1) определяется следующим образом:

$$L_x(0) = \frac{1}{\sqrt{N}} \sum X(m); L_x(R) = \sqrt{\frac{2}{N}} \sum_{m=0}^{N-1} \left(X(m) \cos \frac{(2m+1)k\pi}{2N} \right). \quad (4)$$

Здесь $X(m)$ — значение отсчетов в строках соответствующих рабочих матриц, полученных на втором шаге JPEG-технологий, $L_x(k)$ — коэффициенты дискретного косинусного преобразования, m и k изменяются соответственно от 0 до $N-1$ и от 1 до $N-1$ с шагом 1. Следует заметить, что множество базисных векторов

$$\left\{ \frac{1}{\sqrt{N}}, \sqrt{\frac{2}{N}} \cos \frac{(2m+1)k\pi}{2N} \right\} \quad (5)$$

фактически образует класс дискретных многочленов Чебышева [7].

В случае использования преобразований Хаара для получения коэффициентов необходимо вычислить предварительно обобщенные промежуточные суммы Хаара

$$X_i^n = \sum_{k=2^{i-1}}^{2^i} X_k^{n-1}, \quad n=1, 2, \dots, \log(N-1), \quad i=1, 2, \dots, \frac{N}{2^n}, \quad (6)$$

а затем получить и сами коэффициенты

$$C_{mj} = \frac{1}{N} 2^{(m-1)/2} (X_k^{\log(N-1)-m} - X_{k+1}^{\log(N-1)-m}), \quad (7)$$

где в верхнем индексе $m=1, 2, \dots, \log N; j=1, 2, \dots, 2^{m-1}$, а для выражения, стоящего в скобках, — $m=m-1, k=2j-1$.

Свободный член определяется выражением

$$C_{01} = \frac{1}{N} (X_k^{\log(N-1)} + X_{k+1}^{\log(N-1)}), \quad (8)$$

причем значения X^0 в (8) — исходные данные матриц яркости и цветности.

При этом как в случае дискретного косинусного преобразования, так и в случае преобразования Хаара получаем матрицы, в которых коэффициенты в левом верхнем углу соответствуют низкочастотной составляющей изображения, а в правом нижнем — высокочастотной.

Процесс квантования в схеме на рис. 1 представляет собой простое деление матриц, полученных после преобразования, на матрицу квантования поэлементно. Для каждой компоненты J, C_r, C_b в общем случае задается своя матрица квантования (МК) $q(u, v)$:

$$Jq(u, v) = H \left[\frac{J(u, v)}{q(u, v)} \right], \quad (9)$$

где $H[\bullet]$ означает целую часть числа.

На этом шаге осуществляется управление степенью сжатия и происходят самые большие потери. Задавая МК с большими коэффициентами, получаем больше нулей и, следовательно, большую степень сжатия, что равносильно методу зонального отбора значащих коэффициентов.

В стандарт JPEG включены рекомендованные МК, которые построены опытным путем. Матрицы для большего или меньшего коэффициентов сжатия получают путем умножения МК на некоторое число (γ). При этом потери в низких частотах могут быть настолько велики, что изображение распадается на квадраты 8×8 . Потери в высоких частотах могут проявиться в так называемом «эффекте Гиббса», когда вокруг контуров с резким переходом цвета образуется своеобразный ореол.

Дальнейшие шаги алгоритма требуют перевода матрицы коэффициентов преобразования (8×8) в 64-элементный вектор при помощи «зигзаг»-сканирования, т.е. выбора элементов с индексами (0,0), (0,1), (1,0), (2,0) и т.д. (рис. 2). Это позволяет получить достаточно однородную и регулярную двоичную структуру.

a_{00}	a_{01}	a_{02}	a_{03}	a_{04}	a_{05}	a_{06}	a_{07}
a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}
a_{20}	a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}	a_{27}
a_{30}	a_{31}	a_{32}	a_{33}	a_{34}	a_{35}	a_{36}	a_{37}
a_{40}	a_{41}	a_{42}	a_{43}	a_{44}	a_{45}	a_{46}	a_{47}
a_{50}	a_{51}	a_{52}	a_{53}	a_{54}	a_{55}	a_{56}	a_{57}
a_{60}	a_{61}	a_{62}	a_{63}	a_{64}	a_{65}	a_{66}	a_{67}
a_{70}	a_{71}	a_{72}	a_{73}	a_{74}	a_{75}	a_{76}	a_{77}

Рис. 2

Массив двоичных значений кодируется с использованием известных алгоритмов сжатия без потерь — алгоритма группового кодирования RLE и алгоритма Хаффмана.

Процесс восстановления изображений в схеме JPEG-сжатия полностью симметричен.

Двумерный многомасштабный вейвлет-анализ изображений представляет собой дальнейшее обобщение свойств системы функций Хаара.

Вейвлет-базис. Метод вейвлет-анализа изображений (иначе — многомасштабного анализа) заключается в разложении сигнала $x(t, s)$ по специальному базису, который имеет такие принципиально важные свойства.

1. Базис порождается всего одной функцией $\Psi(t, s)$ и состоит из ее масштабных преобразований, кратных числу 2, и целочисленных сдвигов по координатам:

$$\{\Psi_{i,j,k}(t, s) = 2^{-k} \Psi(2^{1k} t - i, 2^{-k} s - j) : i, j, k \in R\} \quad (10)$$

(множитель 2^{-k} нужен для нормировки).

Функция $\Psi(t, s)$ называется порождающим вейвлетом, а построенный по ней базис (10) — вейвлет-базисом. (Иногда удобно использовать несколько вейвлетов, каждый из которых определяет свою часть вейвлет-базиса по формулам, аналогичным формуле (10).)

2. Вейвлет $\Psi(t, s)$ локализован в пространстве, т.е. быстро убывает при $t^2 + s^2 \rightarrow \infty$. Это свойство объясняет название: слово «вейвлет» переводится как «маленькая волна». На практике, как правило, используются вейвлеты с ограниченным носителем.

3. Вейвлет, по возможности, должен быть как можно более гладкой симметричной функцией и иметь как можно больше нулевых моментов

$$M_{m,n} = \int \int t^n s^m x(t, s) dt ds, \quad m, n \in R.$$

Свойство 1 вейвлет-базиса позволяет построить простую и эффективную в вычислительном смысле рекуррентную процедуру разложения сигнала по вейвлет-базису:

$$x(t, s) = \sum_{i,j,k} (c_{i,j,k} \Psi_{i,j,k}(t, s)). \quad (11)$$

Эта процедура называется быстрым вейвлет-преобразованием, или алгоритмом Малла.

Построение вейвлет-базиса на прямой. Алгоритм Малла. Основная математическая задача вейвлет-анализа состоит в конструировании вейвлет-базисов, сочетающих простоту порождающего вейвлета с его удовлетворительной гладкостью и как можно большим количеством нулевых моментов. В настоящей работе используется наиболее распространенный способ построения вейвлетов на плоскости с помощью вейвлетов на прямой. Кратко остановимся на последних из них. (Более подробная информация содержится, например, в [2, 3] или в тех же обозначениях, что и ниже, в работе [4], где рассматривались возможности применения вейвлет-анализа к сжатию звуковых сигналов.)

Наибольшая трудность построения одномерного вейвлета заключается в отыскании конечной вещественной последовательности $\{h_i\}$, преобразование Фурье которой $h(\omega)$ удовлетворяет некоторому сложному функциональному уравнению. Например, для ортогонального вейвлет-базиса это уравнение выглядит следующим образом:

$$|h(\omega)|^2 + |h(\omega + \pi)|^2 = 2. \quad (12)$$

Далее с помощью найденной последовательности строится дуальная вейвлету скейлинг-функция $\varphi(t)$, а затем и сам вейвлет. Скейлинг-функцию получаем из уравнения

$$\varphi(t/2) = \sqrt{2} \sum_i (h_i \varphi(t-i)), \quad (13)$$

после чего сам вейвлет определяем соотношением

$$\psi(t) = \sqrt{2} \sum_i (g_i \varphi(t-i)), \quad (14)$$

где $\{g_i\}$ — последовательность, сопряженная к последовательности $\{h_i\}$, т.е. такая, что $g_i = (-1)^i h_{1-i}$.

Сопряженные последовательности $\{h_i\}$ и $\{g_i\}$ называют парой квадратурных зеркальных фильтров. Сразу же отметим, что в вычислительных формулах ни скейлинг-функция $\varphi(t)$, ни дуальный ей вейвлет $\psi(t)$ не участвуют, используются только квадратурные зеркальные фильтры.

Пусть $\{x_i\}$ — заданный дискретный сигнал. Фактически в ряд по вейвлет-базису разлагается функция

$$x(t) = \sum_i (x_i \varphi(t-i)),$$

представляющая собой очень удобную экстраполяцию дискретного сигнала $\{x_i\}$ до непрерывного.

Оказывается, выполнение равенства (12) (при некоторых несобременительных вспомогательных условиях) гарантирует, что функцию $x(t)$ можно единственным образом представить в виде суммы двух сигналов:

$$x(t) = x_{-1}(t) + y_0(t), \quad (15)$$

$$x_{-1}(t) = \frac{1}{\sqrt{2}} \sum_i (b_i \varphi(t/2 - i)), \quad y_0(t) = \frac{1}{\sqrt{2}} \sum_i (c_i \psi(t - 1)).$$

Можно сказать, что первый из этих сигналов имеет масштаб 2, т.е. масштаб, вдвое больший, чем у исходного сигнала $x(t)$. Второй сигнал $y_0(t)$ представляет собой ту часть разложения сигнала $x(t)$ по вейвлет-базису, которая имеет масштаб 1. Таким образом, сигнал $x(t)$ представлен в виде разложения $y_0(t)$ по вейвлетам, имеющим масштаб этого сигнала, и остаточного члена $x_{-1}(t)$, масштаб которого вдвое больше исходного. Алгоритм Малла заключается в многократном повторении этой процедуры, применяемой каждый раз к остаточному члену предыдущего шага. При этом каждый раз от остаточного члена отщепляются члены разложения по вейвлет-базису того же масштаба, что и остаточный член, а масштаб нового остаточного члена вдвое больше предыдущего.

Формулы для отыскания коэффициентов в представлении (15) очень просты — они осуществляют замену базиса, определяемую формулами вида (13), (14) в соответствующем масштабе. Если $\{a_i\}$ — коэффициенты разложения остаточного члена, а $\{b_i\}$ и $\{c_i\}$ — коэффициенты соответственно нового остаточного члена и разложения по вейвлет-базису в текущем масштабе, то

$$b_i = \sum_s (h_s a_{2i+s}), \quad c_i = \sum_s (g_s a_{2i+s}). \quad (16)$$

Аналогично, если известны коэффициенты $\{b_i\}$ и $\{c_i\}$, то коэффициенты $\{a_i\}$ можно восстановить по формуле

$$a_i = \sum_s (h_s b_{i-s}) + \sum_s (g_s c_{i-s}). \quad (17)$$

Формула разложения (16) и формула синтеза (17) — основные. Они позволяют, разложив сигнал по вейвлет-базису, сохранить или передать его в таком виде, а затем снова восстановить.

На практике обычно используются биортогональные вейвлет-базисы. Отличие от ортогонального случая заключается в том, что для разложения и восстановления сигнала используются две разные пары квадратурных зеркальных фильтров: пара $\{\tilde{h}, \tilde{g}\}$ при разложении и пара $\{h, g\}$ при восстановлении. Иначе говоря, в формуле (16) над h и g нужно поставить тильду, а формулу (17) оставить без изменения. Основное уравнение, аналогичное уравнению (12), в этом случае имеет вид

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

Для этого уравнения имеется семейство решений, зависящих от двух параметров:

$$\tilde{h}_{n,m}(\omega) = \sqrt{2} \left(\frac{1+e^{i\omega}}{2} \right)^n; \quad h_{n,m}(\omega) = \sqrt{2} \left(\frac{1+e^{i\omega}}{2} \right)^2 P_m \left(\sin^2 \frac{\omega}{2} \right) e^{-im\omega}, \quad (18)$$

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

Фильтры, отвечающие этим решениям, симметричны, а построенные по ним разлагающие вейвлеты имеют достаточно нулевых моментов, количество которых увеличивается с ростом параметра n . Гладкость восстанавливающих вейвлетов повышается при увеличении параметра m , что сопровождается усложнением поведения вейвлета и соответственно ростом длины фильтра.

Построение вейвлет-базиса на плоскости. Рассмотрим наиболее простой и распространенный способ построения вейвлет базиса на плоскости. В качестве двумерной скейлинг-функции возьмем произведение одномерных скейлинг-функций $\Phi(t, s) = \varphi(t)\varphi(s)$, так что исходный дискретный сигнал $\{x_{i,j}\}$ заменяется непрерывной функцией $x(t, s) = \sum_{i,j} (x_{i,j}\varphi(t-i)\varphi(t-j))$.

Скейлинг-функция $\Phi(t, s)$ порождает три вейвлета

$$\Psi_{LH}(t, s) = \varphi(t)\psi(s), \quad \Psi_{HL}(t, s) = \psi(t)\varphi(s), \quad \Psi_{HH}(t, s) = \psi(t)\psi(s). \quad (19)$$

Это означает, что на каждом шаге алгоритма Малла сигнал разлагается на четыре составляющие: три — это ряды по целочисленным сдвигам вейвлетов (19) в масштабе, определяемом номером шага, а четвертый — остаточный член, представленный в виде ряда по целочисленным сдвигам скейлинг-функции с вдвое большим масштабом (рис. 3).

Коэффициенты этих четырех рядов находим с помощью формул (16). Нужно применить квадратурные сопряженные фильтры $\{h_i\}$ и $\{g_i\}$ сначала к строкам остаточного члена, а затем к получившимся столбцам. Порядок применения фильтров определяет вейвлет, по которому проводится разложение. Так, например, применяя к строкам исходного сигнала фильтр $\{h_i\}$, а к столбцам — $\{g_i\}$, получаем коэффициенты разложения по сдвигам вейвлета $\Psi_{LH}(t, s)$.

Аналогично с помощью формулы (17) можно восстановить исходный сигнал, повторяя все шаги разложения в обратном порядке.

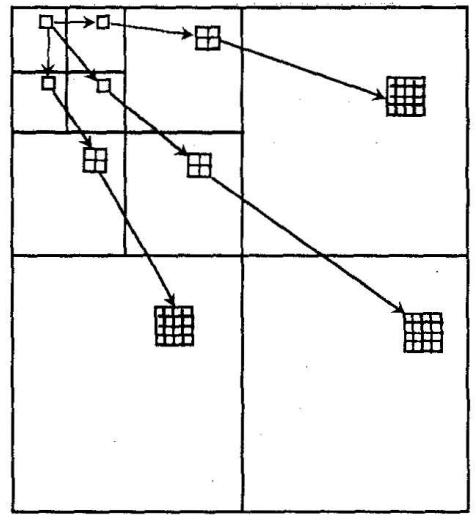


Рис. 3

Практические результаты. Вейвлет-преобразование исследуемых изображений проведено с использованием двух пар фильтров h и \tilde{h} , полученных из выражений (18). Пример фильтров, применяемых для конкретных вычислений, показан на рис. 4 и 5. Как видно из этих рисунков, каждое семейство фильтров симметрично (фильтры h (а) и \tilde{h} (б) при $N=3, M=9$ (рис. 4) и при $N=2, M=2$ (рис. 5)).

В силу особенностей разлагающих фильтров \tilde{h} , вейвлет-преобразования, которые были реализованы при помощи пары фильтров, представленных на рис. 4, будем называть «длинными» (long), а с помощью пары фильтров, представленных на рис. 5, — «короткими» (short).

После получения вейвлет-коэффициентов при заданном уровне преобразования D целесообразно перейти от матричной формы представления данных к векторной. В силу того, что после каждого шага преобразования качественный состав коэффициентов меняется, необходимо учитывать расположение коэффициентов, полученных на предшествующих шагах преобразования. Для последовательного расположения коэффициентов, полученных на различных этапах преобразования, целесообразно использовать методику, представленную на рис. 6 (а — спектр всего изображения, б — детализация одного из масштабов).

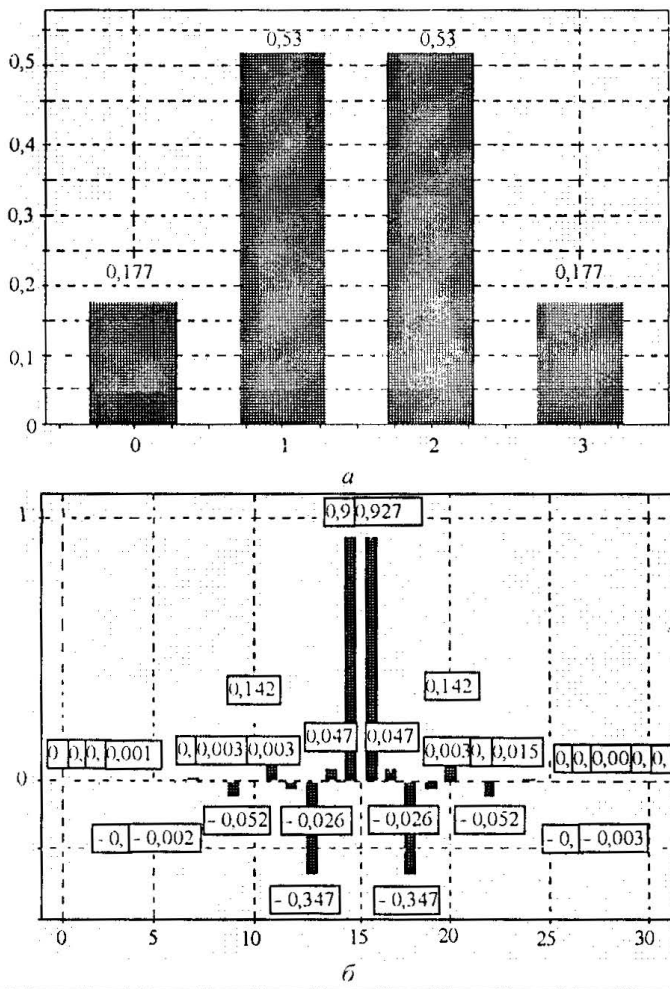


Рис. 4

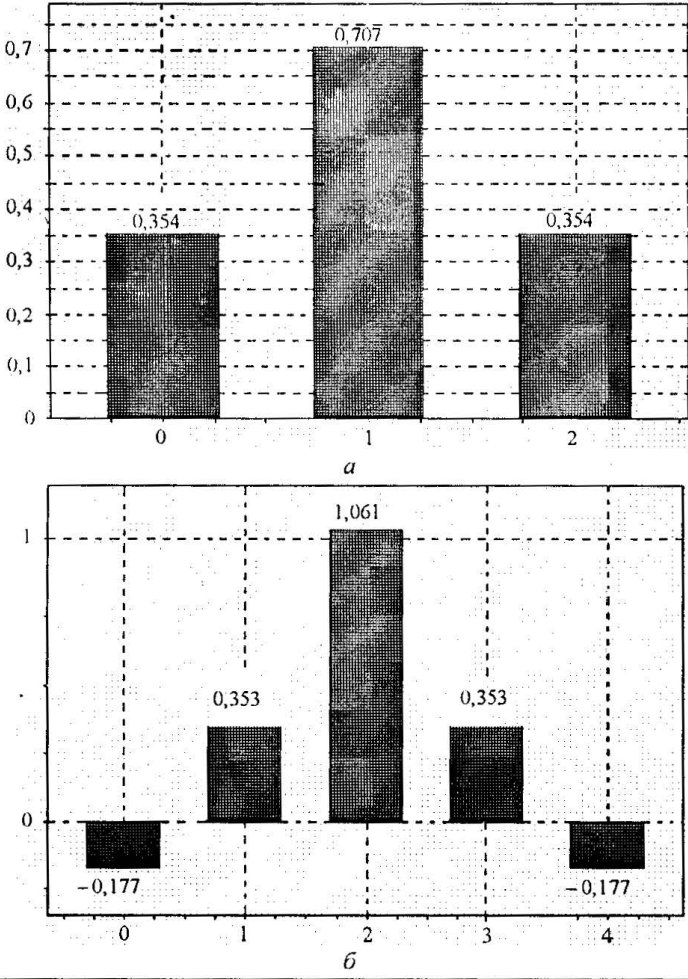
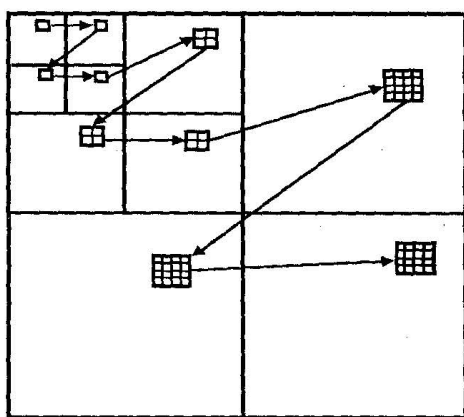


Рис. 5



a

0	1	4	5	16	17	20	21
2	3	6	7	18	19	22	23
8	9	12	13	24	25	28	29
10	11	14	15	26	27	30	31
32	33	36	37	48	49	52	53
34	35	38	39	50	51	54	55
40	41	44	45	56	57	60	61
42	43	46	47	58	59	62	63

б

Рис. 6

После считывания коэффициентов согласно схеме, приведенной на рис. 6, получаем последовательность коэффициентов, показанную на рис. 7.

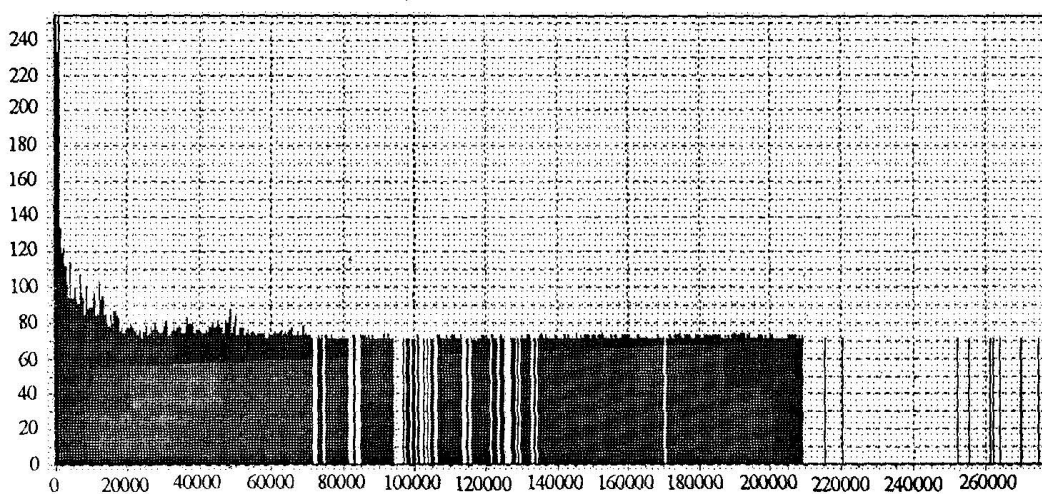


Рис. 7

Как видно из рис. 7, такой подход полностью оправдан с точки зрения получения более равномерного распределения коэффициентов. Это обстоятельство позволяет в дальнейшем эффективнее использовать методы компрессии без потерь.

Для оценки эффективности и показателей качества рассматриваемых преобразований из библиотеки стандартных изображений [16] выбраны два вида изображений. В качестве цветного изображения выбран файл *Peppers.bmp* с параметрами: размер — 512×512 точек, глубина цвета *RGB* — 8 бит, исходный размер файла — 768 кБ. Второй вид изображения представлял файл *Zelda.bmp* с параметрами: размер — 512×512 точек, глубина цвета — 8 бит в градации серого, исходный объем файла — 257 кБ.

Изменение коэффициента сжатия (K) — отношения количества двоичных знаков на входе технологической цепочки кодирования к количеству этих знаков, полученных на выходе этой цепочки, — оценивалось в зависимости от среднеквадратической ошибки (СКО) восстановленного изображения после преобразования, причем соответствующим интервалам значений СКО были субъективно присвоены значения «отлично», «хорошо», «удовлетворительно» и «плохо».

На рис. 8 представлена зависимость коэффициента сжатия от СКО E изображения *Peppers* при использовании вейвлет-преобразований с двумя парами фильт-

ров (рис. 4 — long, рис. 5 — short), дискретного косинусного преобразования (ДКП) JPEG и преобразования Хаара в структуре JPEG-технологий. Ось СКО E условно разделена на четыре сектора, каждый из которых имеет соответствующую оценку качества изображения после его восстановления. Эти оценки позволяют дать более наглядную характеристику качества восстановленного изображения для конкретного диапазона значений СКО. Из рис. 8 видно, что на исследуемом отрезке значений вейвлет-преобразования имеют преимущество перед ДКП JPEG и преобразованием Хаара, реализованных в плоскости JPEG-технологий. Так, для фиксированного значения СКО E коэффициент сжатия K , полученный с использованием вейвлет-преобразований, превосходит аналогичный показатель для ДКП JPEG и преобразования Хаара в 1,3 раза при СКО, равной 18 %, и в 1,7 раза — при значениях СКО 9 %. Однако на интервале СКО E — «отлично», когда ошибка имеет малые значения, коэффициент сжатия для всех преобразований практически одинаков. Отметим также, что JPEG-технологии подразумевают переход от стандартных компонент цветности RGB к J (яркость), C_r (хроматическая красный) и C_b (хроматическая синий) и последующим прореживанием вдвое хроматических компонент C_r и C_b , что уже само по себе дает двукратное преимущество в сжатии перед энтропийным кодированием.

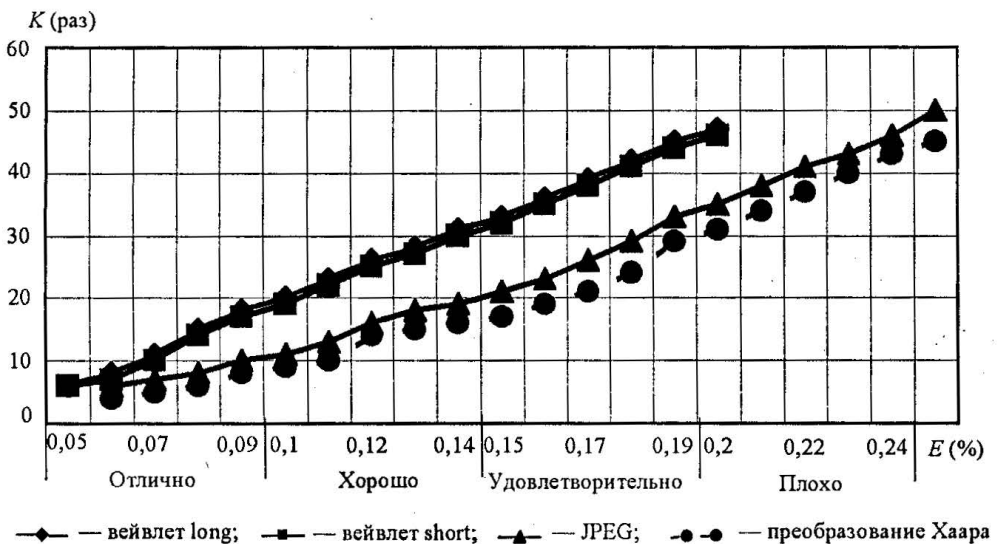


Рис. 8

Вейвлет-преобразования рассматривались только для RGB -компонент без предварительного прореживания, поэтому степень сжатия перед компрессией изображений без потери информации всегда равна 1. Из анализа графика зависимости коэффициента сжатия K от СКО E также следует, что вейвлет-преобразования, выполненные с применением «длинных» — long и «коротких» — short фильтров разложения, практически не отличаются друг от друга. Однако время вычисления вейвлет-коэффициентов с использованием «коротких» разлагающих фильтров значительно меньше. Это объясняется сокращением вычислений при выполнении операций свертки фильтра с исследуемых изображений (по строкам и столбцам).

На рис. 9 представлена зависимость изменения коэффициента сжатия K от СКО E изображения Zelda при использовании исследуемых преобразований. Ось СКО также разделена на диапазоны с оценкой качества восстановленных изображений. Как и на рис. 8, преимущество вейвлет-преобразований перед ДКП JPEG и преобразованием Хаара в терминах JPEG-технологий очевидно

и составляет 1,4–1,8 раза при том же качестве, что и для предыдущего изображения; в диапазоне с малыми ошибками все преобразования имеют приблизительно равные показатели. Следует отметить практически одинаковую эффективность обработки в базисе Хаара и ДКП JPEG, а анализ однородных областей в матрицах дисперсий яркостных составляющих преобразований Хаара говорит о том, что для этого преобразования должна быть найдена своя матрица квантования и соответствующий ей зигзаг сканирования коэффициентов в пункте 5 схемы кодирования на рис. 1. Диапазоны значений СКО, которые можно характеризовать одной оценкой качества изображения (отлично, хорошо, удовлетворительно, плохо), несколько отличаются от аналогичных диапазонов в изображении *Repers.bmp*. Этот факт еще раз свидетельствует о том, что СКО, полученная после восстановления одного класса изображений, не всегда точно характеризует качественные показатели, полученные после восстановления другого класса изображений.

K (раз)

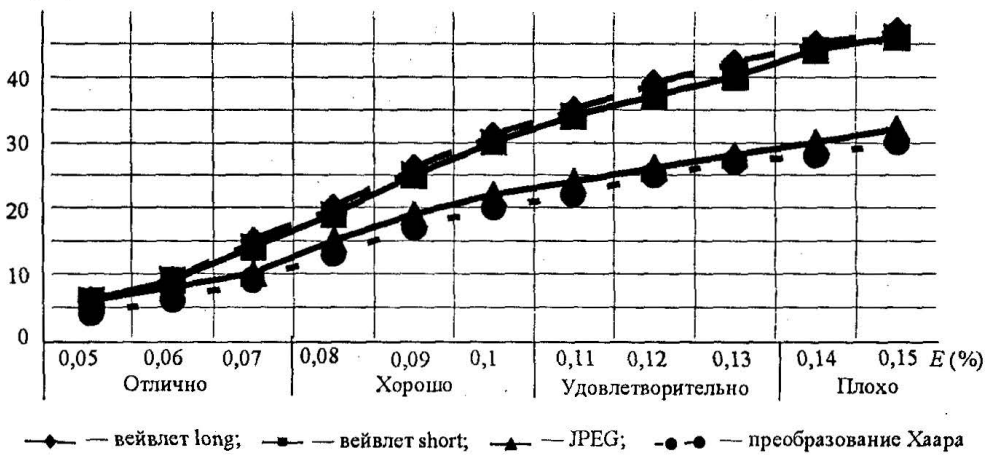


Рис. 9

Проведенные исследования и полученные результаты позволяют говорить о перспективности объединения статистических методов кодирования, методов на основе обобщенных ортогональных Фурье- и вейвлет-преобразований, а также структурных методов описания изображений, что позволит на практике решать задачи предельно сжатого представления объектов кодирования.

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

ФУР'Є- І ВЕЙВЛЕТ-АНАЛІЗ ЗОБРАЖЕНЬ У ПЛОЩИНІ JPEG-ТЕХНОЛОГІЙ

Наведено порівняльні результати комп'ютерного моделювання стиснення деяких класів зображень з використанням дискретного косинусного перетворення, перетворення Хаара і вейвлет-перетворень на базі JPEG-технологій, в основі яких покладено ідею заміни зображення як безпосереднього об'єкта кодування його двовимірним перетворенням Фур'є. При цьому зображення піддається унітарному математичному перетворенню й отримані в результаті коефіцієнти квантуються, кодуються і передаються каналом зв'язку або записуються в архів.

FOURIER AND WAVELET IMAGE ANALYSIS IN A PLANE OF JPEG-PROCESS ENGINEERINGS

Comparative outcomes of computer simulation on compression of some classes of maps with usage of discrete cosine conversion, conversion of Haar and wavelet-conversions are reduced on the basis of JPEG-process engineerings in which basis the idea of changeover of the map as immediate plant of coding, its bidimensional Fourier transform lays. Thus the map is exposed to unitary mathematical conversion and the factor obtained in outcome are quantized, encoded and transmitted on a data link or note in archive.

1. *Кравченко В.Ф., Рвачев В.А.* Wavelet-системы и их применение в обработке сигналов // *Зарубеж. радиоэлектроника.* — 1996. — № 4. — С. 3–20.
2. *Новиков М.Я., Стечкин С.Б.* Основы теории всплесков // *Успехи мат. наук.* — 1998. — 53, № 6. — С. 53–128.
3. *Воробьев В.И., Грибунин В.Г.* Теория и практика вейвлет-преобразования. — СПб. : Изд-во ВУС, 1999. — 208 с.
4. *Иванов В.Г., Любарский М.Г., Ломоносов Ю.В.* Применение вейвлет-анализа к сжатию звуковых сигналов // *Вісн. НТУ «ХПІ». 36. наук. праць : Системний аналіз, управління та інформаційні технології.* — 2003. — 1. — № 7. — С. 39–50.
5. *Andrews N.C., Pratt W.K.* Fourier transform coding of images // *Hawaii Intern. Conf. on Syst. Sci.*, Januar 1968. — P. 667-679.
6. *Пратт У., Кейн Р., Эндриус Н.* Кодирование изображений посредством преобразования Адамара. — ТИИЭР. — 1969. — 57, № 1. — С. 66–67.
7. *Ахмед Н., Рао К.Р.* Ортогональные преобразования при обработке цифровых сигналов. — М. : Связь, 1980. — 248 с.
8. *Претт У.* Цифровая обработка изображений. Кн. 2. — М. : Мир, 1982. — 480 с.
9. *Цифровое кодирование телевизионных изображений / И.И. Цуккерман, Б.М. Кац, Д.С. Лебедев, В.Г. Маковец, С.В. Сардыко, Е.З. Сорока, В.А. Хлебородов, Н.Н. Шостацкий.* // Под ред. И.И. Цуккермана. — М. : Радио и связь, 1981. — 240 с.
10. *Колесник В.Д., Полтырев Г.Ш.* Курс теории информации. — М. : Наука, 1982. — 416 с.
11. *Орищенко В.И., Санников В.Г., Свириденко В.А.* Сжатие данных в системах сбора и передачи информации / Под ред. В.А. Свириденко. — М. : Радио и связь, 1985. — 184 с.
12. *Свириденко В.А.* Анализ систем со сжатием данных. — М. : Связь, 1977. — 184 с.
13. *Ватанабе С.* Разложение Карунена–Лоэва и факториальный анализ // *Автоматический анализ сложных изображений.* — М. : Мир, 1969. — С. 254–275.
14. *Претт У.* Цифровая обработка изображений. Кн. 1. — М. : Мир, 1982. — 312 с.
15. *Иванов В.Г.* Формальное описание дискретных преобразований Хаара // *Проблемы управления и информатики.* — 2003. — № 5. — С. 68–74.
16. (<http://www.icsl.ucla.edu>).

Получено 17.11.2003