

## КОМПРЕССИЯ ИЗОБРАЖЕНИЙ НА ПРИМЕРЕ СТАНДАРТА JPEG

Хунцария Дж. М., Хирьянов Ю. А., Хунцария Л. Дж.  
Грузинский Технический Университет

### *Резюме*

*Рассмотрен один из вариантов компактного представления неподвижных цветных изображений на основе стандарта JPEG, позволяющий эффективно использовать объемы соответствующих запоминающих устройств. Проанализированы отдельные этапы реализации процесса кодирования и декодирования на основе алгоритма JPEG: формирование яркостного  $Y$  и двух цветоразностных  $C_R$  и  $C_B$  составляющих из цветовых  $R$ ,  $G$  и  $B$  сигналов изображения; представление  $Y$ ,  $C_R$  и  $C_B$  составляющих в цифровой форме и их деление на отдельные блоки (фрагменты)  $8 \times 8$ ; поблочное (двумерное) преобразование каждого составляющего на основе дискретного косинусного преобразования (ДКП); квантование коэффициентов трансформанты ДКП; представление двумерного массива коэффициентов трансформанты в виде одномерного путём их зигзаг сканирования; кодирование с помощью кода Хаффмена полученной двоичной (цифровой) последовательности символов. Показано, что декодирование производится в обратном направлении. Приводится пример кодирования и декодирования как для фрагмента изображения, так и для шести тестовых цветных изображений. Представлены значения коэффициентов сжатия и результаты объективной и субъективной оценок сжатых тестовых изображений.*

**Ключевые слова:** кодирование, декодирование, дискретное косинусное преобразование, трансформанта, код Хаффмена.

Одной из наиболее актуальных и перспективных отраслей цифровой телекоммуникации считается обработка, передача и хранение (архивация) визуальной информации, что, в свою очередь, требует сжатие этой информации применением того или иного метода сжатия (компрессии).

Существуют два вида алгоритмов сжатия (компрессии) визуальной информации (изображений): сжатие без потерь информации (неразрушающая компрессия); сжатие с потерями (разрушающая компрессия). Первыми для архивации изображений стали применяться неразрушающие алгоритмы сжатия, которые из-за низкой степени сжатия в дальнейшем перестали удовлетворять требованиям, предъявляемым к системам архивации визуальной информации. Поэтому в настоящее время в основном применяются алгоритмы с потерями информации. При этом в каждом конкретном случае задача сжатия решается в результате компромисса между степенью сжатия и качеством сжатых изображений.

Алгоритм одного из наиболее эффективных методов сжатия неподвижных изображений излагается в стандарте JPEG (Joint Photographic Experts Group). Он принят Международной организацией стандартизации (ISO), определяет последовательность и параметры этапов кодирования и декодирования неподвижных цветных изображений и используется в основном для записи и хранения изображений с целью экономии объема запоминающих устройств [1].

Реализация алгоритма JPEG позволяет уменьшить объем только реальных полутоновых и цветных изображений примерно в 5...10 раз без заметного ухудшения их качества. Он не

предназначен для сжатия таких изображений, как рисунки и чертежи, имеющих два уровня яркости.

Алгоритм оперирует блоками изображения размерами  $8 \times 8$ , на которых яркость и цвет меняются достаточно плавно. Полутоновое (чёрно-белое) изображение разбивается на блоки  $8 \times 8$  пикселей, которые далее кодируются один за другим.

Цветное изображение может быть представлено как в формате его цветовых составляющих  $R$ ,  $G$  и  $B$ , так и в формате яркостного  $Y$  и двух цветоразностных составляющих  $C_B$  и  $C_R$ . В первом случае для каждого пиксела задаются значения трёх основных цветов (красного, зелёного и синего) и каждый блок  $8 \times 8$  пикселей представляются тремя блоками  $8 \times 8$  чисел. Кодирование данных каждого из трёх цветов выполняется также, как для полутонового монохромного (чёрно-белого) изображения. Следует отметить, что предпочтительным является второй вариант представления цветного изображения, когда для каждого пиксела задаются значения яркости и цветоразностных сигналов. Его предпочтительность заключается в том, что в этом случае возможно уменьшение числа блоков цветоразностных составляющих путём уменьшения числа их отсчетов (пикселей) как по горизонтали, так и по вертикали в два раза, что соответствует формату  $4:2:0$ . Следовательно, на каждые четыре блока пикселей яркостного сигнала будет приходиться по одному блоку пикселей цветоразностных сигналов  $C_B$  и  $C_R$ . Поэтому по сравнению с форматом  $RGB$  полное число кодируемых блоков уменьшится в два раза, но заметного ухудшения качества изображения при этом не произойдёт, так как человеческое зрение не воспринимает искажения цвета мелких деталей изображения. Таким образом, в стандарте JPEG цветоразностные составляющие кодируются с большей погрешностью, чем яркостная составляющая.

При разложении блоков (фрагментов) изображения на основе дискретного косинусного преобразования значимыми оказываются лишь небольшое количество коэффициентов преобразованного пространства (трансформанты), сохранение которых требует небольшой объем по сравнению с исходным изображением.

Последовательность этапов алгоритма JPEG, количество которых для каждого из трех составляющих сигнала цветного изображения равно 7, можно пояснить структурной схемой, представленной на рис.1.

Этап 1. Как отмечалось выше, для цветных изображений наилучшее сжатие может быть достигнуто не при едином сжатии  $R$  (красного),  $G$  (зелёного) и  $B$  (синего) составляющих сигнала цветного изображения, а при раздельном кодировании его яркостного (чёрно-белого)  $Y$  и двух цветоразностных ( $C_R$  и  $C_B$ ) составляющих, которые, как известно, формируются из пространства  $R$ ,  $G$  и  $B$  (см. рис.1) следующим образом:

$$\left. \begin{aligned} Y &= 0,2989R + 0,5866G + 0,1145B, \\ C_R &= 0,5R - 0,4184G - 0,0816B, \\ C_B &= -0,1688R - 0,3312G + 0,5B. \end{aligned} \right\} \quad (1)$$

Этап 2: Каждое из составляющих цветового сигнала, полученные на первом этапе, преобразуется в цифровой сигнал путём восьмиразрядного кодирования, в результате чего суммарное число битов на один пиксел изображения равно 24. Затем цифровые составные изображения разбиваются на блоки (фрагменты)  $8 \times 8$  так, что на каждые четыре блока составляющего  $Y$  будет приходиться по одному блоку составляющих  $C_R$  и  $C_B$ , в результате чего суммарное число пикселей окажется в два раза меньше по сравнению с форматом  $RGB$  (см. рис.1). Действительно, четыре блока (по два блока по горизонтали и по вертикали) составляющего  $Y$  содержит  $4 \times 8 \times 8 = 256$  пикселей в то время, когда соответствующие им по одному блоку составляющих  $C_R$  и  $C_B$  будут содержать по  $2 \times 8 \times 8 = 128$  пикселей (в сумме 384 пиксела) в то время, когда в формате  $RGB$  суммарное число пикселей равно  $3 \times 4 \times 8 \times 8 = 768$ , т.е. в два раза больше по сравнению с форматом  $Y C_R C_B$ . Формированные таким образом

цифровые сигналы  $Y$ ,  $C_R$  и  $C_B$  в дальнейшем обрабатываются в соответствующих кодерах JPEG.

Возможны два варианта последовательности кодирования блоков цветного изображения. По первому варианту сначала кодируются все блоки пикселей сигнала  $Y$ , затем - все блоки пикселей сигнала  $C_R$ , а затем - все блоки сигнала  $C_B$ . По второму варианту, например, в случае формата 4:2:0, сначала кодируются четыре блока  $Y$ , образующие матрицу 2x2, затем соответствующий им один блок  $C_R$ , затем - один блок  $C_B$ , затем следующие четыре блока  $Y$  и т.д.

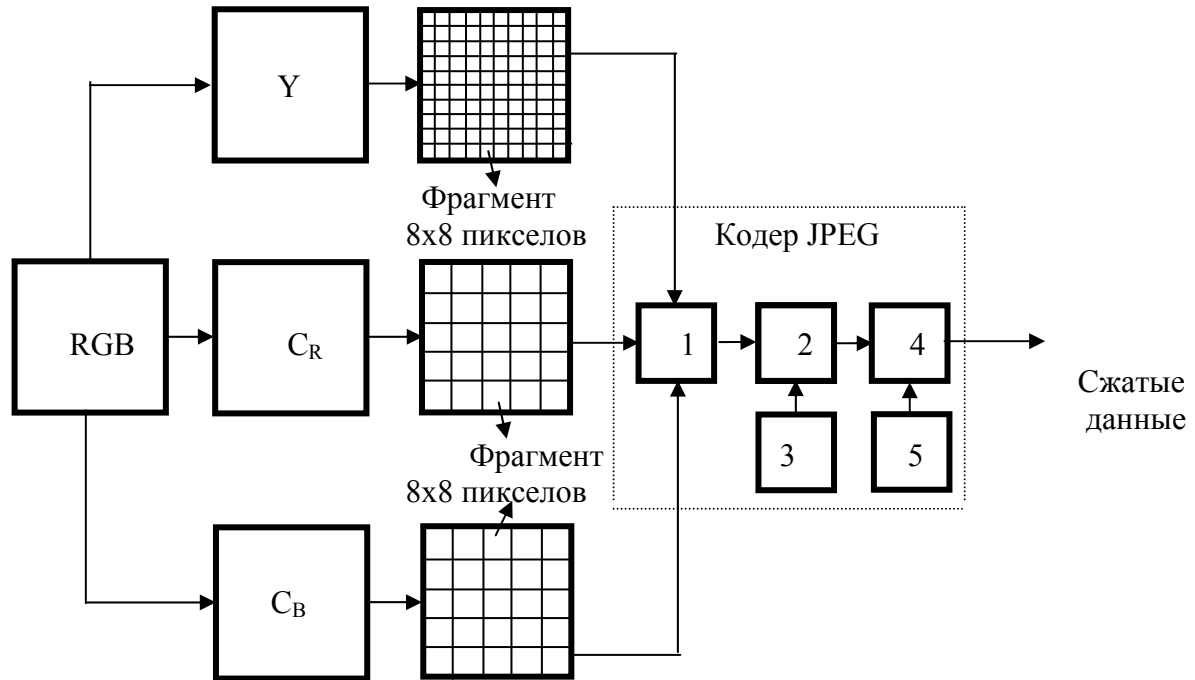


Рис.1. Структурная схема кодирования цветных изображений по стандарту JPEG  
 Обозначения операций на рис 1: 1-Дискретное косинусное преобразование; 2-Квантователь;  
 3-Таблица квантования Q; 4-Кодер Хаффмена; 5-Таблица кодов

Этап 3: В стандарте JPEG применяют прямое двумерное дискретное косинусное преобразование (ДКП) исходных данных, которые имеют вид матриц 8x8 пикселей сигналов  $Y$ ,  $C_R$  и  $C_B$ . При этом пиксели (элементы) выражены восьмиразрядными целыми положительными двоичными числами. Коэффициенты прямого двумерного ДКП (коэффициенты трансформанты) для блоков 8x8 вычисляются по следующим формулам:

$$F(u,v) = (C_u C_v \sum_{x=0}^7 \sum_{y=0}^7 f(x,y) (\cos((2x+1)u\pi)/16) (\cos((2y+1)v\pi)/16)) / 4, \quad (2)$$

где:  $x,y$ -координаты пикселей изображения в блоке  $f(x,y)$ ;  $f(x,y)$ -значения яркостного или цветоразностных сигналов;  $u,v$ - координаты коэффициентов ДКП в спектральной области (в трансформанте)  $F(u,v)$ ;  $F(u,v)$ -значения коэффициентов преобразования.

В формуле (4)  $C_u(C_v)=1/2^{1/2}$ , при  $u(v)=0$  и  $C_u(C_v)=1$  при  $u(v)\neq 0$ .

В результате прямого ДКП образуется трансформанта в виде матрицы 8x8, элементы которой являются коэффициенты преобразования. В качестве примера приведём матрицы преобразуемого  $[f(x,y)]$  и преобразованного с помощью ДКП  $[F(u,v)]$  массивов, соответствующих яркостному составляющему  $Y$  изображения.

$$[f(x,y)] = \begin{bmatrix} 139 & 144 & 149 & 153 & 155 & 155 & 155 & 155 \\ 144 & 151 & 153 & 156 & 159 & 156 & 156 & 156 \\ 150 & 155 & 160 & 163 & 158 & 156 & 156 & 156 \\ 159 & 161 & 165 & 160 & 160 & 159 & 159 & 159 \\ 159 & 160 & 161 & 162 & 162 & 155 & 155 & 155 \\ 161 & 161 & 161 & 161 & 160 & 157 & 157 & 157 \\ 162 & 162 & 161 & 163 & 162 & 157 & 157 & 157 \\ 162 & 162 & 161 & 161 & 163 & 158 & 158 & 158 \end{bmatrix}$$

$$[F(u,v)] = \begin{bmatrix} 1260 & -1 & -12 & -6 & 2 & -2 & -2 & 2 \\ -22 & -17 & -6 & -3 & -3 & 0 & 1 & -1 \\ -11 & -10 & -1 & 2 & 1 & -1 & -1 & -1 \\ -7 & -2 & 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 1 & -1 & -1 & 1 & 2 \\ 2 & 0 & 1 & -1 & -1 & 2 & 2 & 0 \\ -1 & -1 & 0 & -1 & 0 & 2 & 1 & -1 \\ -3 & 1 & -3 & -1 & 2 & 1 & -1 & -1 \end{bmatrix}$$

Следует отметить, что коэффициенты в левом верхнем углу матрицы  $[F(u,v)]$  соответствуют низкочастотной составляющей изображения, а в правом нижнем - высокочастотной, причём плавное изменение яркости или цвета соответствует низкочастотной составляющей изображения, а резкие - высокочастотной.

Этап 4: Наибольший вклад при формировании большинства реальных изображений, как известно, вносят их низкочастотные составляющие, которые определяют формы и значения элементов основных объектов и фона. Высокочастотные составляющие создают резкие границы и контуры, а также мелкие детали изображения. Возможность уменьшения количества двоичных символов с помощью ДКП основана на указанных свойствах пространственно-частотного спектра реальных изображений и на ограниченной способности человеческого зрения воспринимать изменения и искажения мелкой структуры изображения. Поэтому в стандарте JPEG предусмотрено квантование коэффициентов трансформант ДКП с помощью таблицы квантования  $Q$ . Следовательно, на этом этапе производится деление коэффициентов матрицы  $[F(u,v)]$  на соответствующие элементы таблицы квантования  $Q$ , которую для яркостного  $Y$  и цветоразностных  $C_R$  и  $C_B$  составляющих изображения можно представить в виде матриц  $[Q_Y]$  и  $[Q_C]$  соответственно:

$$[Q_Y] = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

$$[Q_C] = \begin{bmatrix} 17 & 18 & 24 & 47 & 99 & 99 & 99 & 99 \\ 18 & 21 & 26 & 66 & 99 & 99 & 99 & 99 \\ 24 & 26 & 56 & 99 & 99 & 99 & 99 & 99 \\ 47 & 66 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \end{bmatrix}$$

Из матриц квантования следует, что при квантовании низкочастотные составляющие квантуются более точно, чем высокочастотные. Результаты деления округляются до ближайшего целого числа. Так, например, поскольку матрица  $[F(u,v)]$  получена путём преобразования матрицы  $[f(x,y)]$ , соответствующей яркостной составляющей изображения, то её элементы (коэффициенты трансформанты) надо делить на элементы матрицы  $[Q_Y]$ , в результате чего получим квантованную матрицу коэффициентов  $[F'(u,v)]$ :

$$[F'(u,v)] = \begin{bmatrix} 79 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

С квантованием могут быть связаны специфические искажения на воспроизведённых изображениях, которые при потерях в низкочастотных составляющих проявляются в виде заметного разбиения изображения на блоки  $8 \times 8$ , а при потерях в высокочастотных составляющих - в виде так называемого "эффекта Гиббса", при котором вокруг контуров с резкими перепадами образуется своеобразный "нимб".

Этап 5: На этом этапе матрица  $[Q_Y]$  преобразуется в 64-элементный вектор при помощи зигзаг-сканирования, которое выполняется на основе таблицы 1.

Таблица 1. Зигзаг-сканирование квантованных коэффициентов ДКП

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

Как видно из таблицы 1, в результате зигзаг-сканирования в 64-элементном векторе коэффициенты трансформанты будут расположены в следующей последовательности:  $F(0,0)$ ;  $F(0,1)$ ;  $F(1,0)$ ;  $F(2,0)$ ;  $F(1,1)$ ;  $F(0,2)$ ;  $F(0,3)$ ;  $F(2,1)$ ; ...,  $F(7,7)$ . Например, после сканирования в 64-элементном векторе коэффициентов, соответствующем трансформанте  $[F'(u,v)]$ , коэффициенты окажутся расположены в следующей последовательности: 79; 0; -2; -1; -1; -1; 0; 0; -1; -1; 0; 0; 0; 0; 0; 0; 0; 0; ..., 0.

Этап 6: Из матрицы  $[F(u,v)]$  следует, что в результате квантования многие из коэффициентов ДКП становятся равными нулю, поэтому в векторе, полученном после сканирования, оказывается большое число нулевых коэффициентов. Каждый отличный от нуля коэффициент ДКП представляется в виде пары чисел, первое из которых показывает, сколько нулевых значений подряд прошло в последовательности перед данным ненулевым коэффициентом, а второе - значение самого квантованного коэффициента. В частности, для

рассмотренного выше вектора пары чисел будут иметь следующий вид: (0;79), (1;-2), (0;-1), (0;-1), (0;-1), (2;-1), (0;-1), (53;0). Поскольку для реальных изображений в результате квантования получается много нулевых и малых по абсолютной величине коэффициентов, поэтому такое кодирование, называемое кодированием с бегущей длиной (runlength coding), даёт значительный выигрыш, так как, во-первых, уменьшается общее количество чисел, представляющий кодируемый блок, а во вторых, уменьшается число двоичных символов для представления большинства чисел.

Этап 7: Полученная на шестом этапе кодирования последовательность двоичных чисел затем подвергается энтропийному кодированию, при котором чаще всего применяется метод Хаффмена, который заключается в построении такого кода с переменной длиной кодового слова, что чаще встречающимся символам ставятся в соответствие более короткие кодовые слова, а реже встречающимся символам - более длинные кодовые слова. Это позволяет дополнительно уменьшить количество двоичных символов без потери информации, имеющей после квантования коэффициентов ДКП. Заметим, что иногда вместо кодирования по Хаффмену может использоваться другой вид энтропийного кодирования - арифметическое кодирование.

Воспроизведение исходного изображения (декодирование) производится в обратном направлении, которое иллюстрируется на рис. 2.

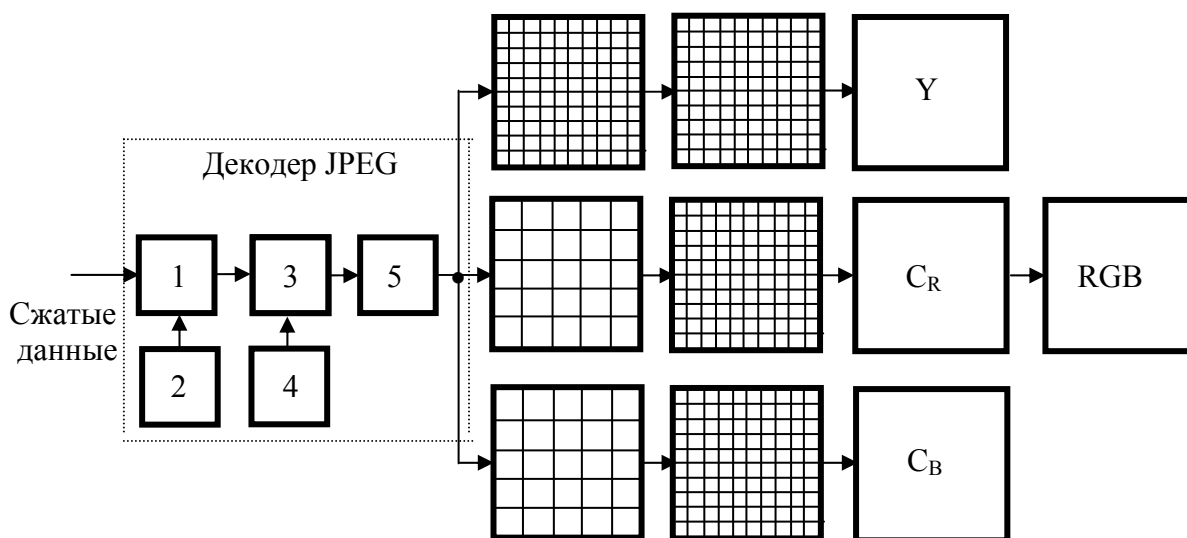


Рис. 2. Структурная схема декодирования цветных изображений по стандарту JPEG

Обозначения операций на рис 2:

- 1-Декодер Хаффмена; 2-Таблица кодов; 3-Деквантователь; 4-Таблица деквантования Q;  
5 - Обратное дискретное косинусное преобразование

Этот этап начинается обратным преобразованием слов кода Хаффмена, считываемых из файла сжатых данных, в последовательность чисел, по которым восстанавливаются значения квантованных коэффициентов ДКП. Следует отметить, что кодирование кодом Хаффмена и соответствующее декодирование не приводит к дополнительной потере информации. Затем производится деквантование коэффициентов трансформанты и обратное ДКП, в результате чего восстанавливаются пиксели  $Y$ ,  $C_R$  и  $C_B$  составляющие. Заметим, что восстановление пикселей цветоразностных  $C_R$  и  $C_B$  составляющих, пропущенных на втором этапе процесса кодирования, производится с помощью интерполяции (см. рис.2), после чего с учётом выражения 1 воспроизводятся цветовые составляющие R,G и B.

$$\left. \begin{aligned} R &= Y + 1,402 \times C_R, \\ G &= (0,837 \times Y - 0,598 \times C_R - 0,289 \times C_B) / 0,837, \\ B &= Y + 1,77 \times C_B. \end{aligned} \right\} \quad (3)$$

Для рассмотренного выше примера восстановленные в процессе декодирования матрицы коэффициентов трансформанты  $[F(u,v)]$ , пикселей исходного массива  $[f(x,y)]$ , полученного в результате обратного ДКП  $[F(u,v)]$  и погрешностей восстановления пикселей  $[\varepsilon(x,y)]$  представляются в виде следующих матриц:

$$[F(u,v)] = \begin{bmatrix} 1264 & 0 & -10 & 0 & 0 & 0 & 0 & 0 \\ -24 & -12 & 0 & 0 & 0 & 0 & 0 & 0 \\ -14 & -13 & 0 & 0 & 0 & 0 & 0 & 0 \\ -14 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[f(x,y)] = \begin{bmatrix} 142 & 144 & 147 & 150 & 152 & 153 & 154 & 154 \\ 149 & 150 & 153 & 155 & 156 & 157 & 156 & 156 \\ 157 & 158 & 159 & 161 & 161 & 160 & 159 & 158 \\ 162 & 162 & 163 & 163 & 162 & 160 & 158 & 157 \\ 162 & 162 & 162 & 162 & 161 & 158 & 156 & 155 \\ 160 & 161 & 161 & 161 & 160 & 158 & 156 & 154 \\ 160 & 160 & 161 & 162 & 161 & 160 & 158 & 157 \\ 160 & 161 & 163 & 164 & 164 & 163 & 161 & 160 \end{bmatrix}$$

$$[\varepsilon(x,y)] = \begin{bmatrix} 3 & 0 & -2 & -3 & -3 & -2 & -1 & -1 \\ 5 & -1 & 0 & -1 & -3 & 1 & 0 & 0 \\ 7 & 3 & -1 & -2 & 3 & 4 & 3 & 2 \\ 3 & 1 & -2 & 3 & 2 & 1 & -1 & -2 \\ 3 & 2 & 1 & 0 & -1 & 3 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 & -1 & -3 \\ -2 & -2 & 0 & -1 & -1 & 3 & 1 & 0 \\ -2 & -1 & 2 & 3 & 1 & 5 & 3 & 2 \end{bmatrix}$$

В рассмотренном случае среднеквадратическая ошибка восстановления  $\varepsilon^2 \approx 2,24$ .

На рис. 3-8 представлены исходные и восстановленные после эффективного кодирования (сжатия) на основе алгоритма JPEG тестовые изображения “Лена”, “Природа”, “Парк”, “Вишня”, “ТВС”, “Кросс”, а также изображения ошибок восстановления.



Рис.3. Исходное и сжатое изображения “Лена” и изображение ошибок



Рис.4. Исходное и сжатое изображения “Природа” и изображение ошибок



Рис.5. Исходное и сжатое изображения “Парк” и изображение ошибок



Рис.6. Исходное и сжатое изображения “Вишня” и изображение ошибок





Рис.7. Исходное и сжатое изображения “ТВС” и изображение ошибок



Рис.8. Исходное и сжатое изображения “Кросс” и изображение ошибок

В таблице 2 приводятся данные коэффициента (сжатия) компрессии и среднеквадратичной ошибки (СКО) восстановления для указанных сжатых изображений, с помощью которых проводится объективная оценка эффективности сжатия и качества кодированных изображений. Из таблицы 2 следует, что данные СКО для всех тестовых изображений находятся в пределах допустимых значений.

Таблица 2. Значения коэффициента компрессии и ошибки восстановления кодированных на основе стандарта JPEG тестовых изображений

Изображения	Коэффициент компрессии	Среднеквадратическая ошибка (СКО)
“Лена”	6,3	5,9
“Природа”	5,6	8,1
“Парк”	6,4	9,9
“Вишня”	5,1	7,2
“ТВС”	6,5	5,2
“Кросс”	4,9	9,6

В результате субъективной оценки сжатых изображений установлено, что качество восстановления в среднем оценивается 4,8 балом по пятибалльной шкале качества.

Отрицательной стороной алгоритма JPEG является то, что при больших значениях степени сжатия изображение распадается на отдельные блоки 8x8, проявляется так называемая эффект Гиббса - ореолы по границам резких переходов. Следует также отметить, что часто на восстановленных изображениях горизонтальные и вертикальные полосы абсолютно не видны и проявляются только в виде муарового узора.

**ЛИТЕРАТУРА**

1. *Ватолин Д.С.* Методы сжатия изображений. Интернет Университет информационных технологий. <http://www.intuit.ru>, 2007.
2. *Яшин В.В.* Анализ и обработка изображений// ISBN 5-89407-041-4.-М., Диалог-МГУ, 1999.
3. *Шульгин В.И.* Основы теории передачи информации// Учебное пособие. -Харьков, ХАИ, 2003.