ЛЕКЦИЯ 9
Стандарт JPEG (сжатие с потерей качества)

Ниже приведено краткое описание стандарта JPEG для черно-белых изображений. Для цветных изображений один тот же процесс кодирования используется для яркостной и двух цветоразностных компонент в формате YUV (4:1:1).

Изображение делится на непересекающиеся блоки, каждый размером 8×8 точек. Если размеры изображения не кратны 8, то точки последней строки или соответственно последнего столбца повторяются. Каждый блок подвергается двумерному ДКП. Коэффициенты ДКП квантуются и кодируются неравномерным кодом. Кодирование неравномерным кодом выполняется в две ступени. На первой ступени выполняется дифференциальное кодирование для коэффициентов постоянного тока и кодирование длин серий для остальных коэффициентов. На второй ступени производится кодирование кодом Хаффмена.

Пусть DC_i и DC_{i-1} обозначают коэффициенты постоянного тока в блоках i и $i-1$. Из-за высокой корреляции значений коэффициентов постоянного тока в соседних блоках, JPEG использует для них дифференциальное кодирование. Для черно-белых изображений(или компонент Y,U или V цветного изображения) точка изображения соответствует 8 битам и разность $DC_i - DC_{i-1}$ принимает значение в диапазоне -2047 до 2047. Этот диапазон подразделяется на 12 категорий, где $i-$ я категория включает все разности, которые могут быть представлены в двоичной форме последовательностью длины i бит. Это первые 12 категорий из таблицы 1. После просмотра таблицы каждый коэффициент постоянного тока может быть описан парой (*категория, амплитуда).* Если значение $DC_i - DC_{i-1} \geq 0$, то амплитуда это просто двоичное представление этой разности с числом бит, задаваемым категорией. Если же $DC_i - DC_{i-1} < 0$, то амплитуда кодируется обратным кодом абсолютного значения разности. Из указанной пары : *(категория, амплитуда)* только категория кодируется кодом Хаффмена.

Для 8-битовых изображений остальные коэффициенты, называемые коэффициентами переменного тока, могут принимать значения в диапазоне -1023 до 1023. Как и для коэффициентов постоянного тока этот диапазон делится на 10 категорий и каждый коэффициент может быть описан парой: *(категория, амплитуда).* После квантования большинство коэффициентов переменного тока оказывается равными 0, так что кодировать необходимо только небольшое число ненулевых коэффициентов. Коэффициенты переменного тока обрабатываются в зигзагообразном порядке. На рис. 4 показано, каким образом производится считывание элементов матрицы размера 8×8. Этот порядок позволяет обеспечить более эффективное кодирование длин серий.

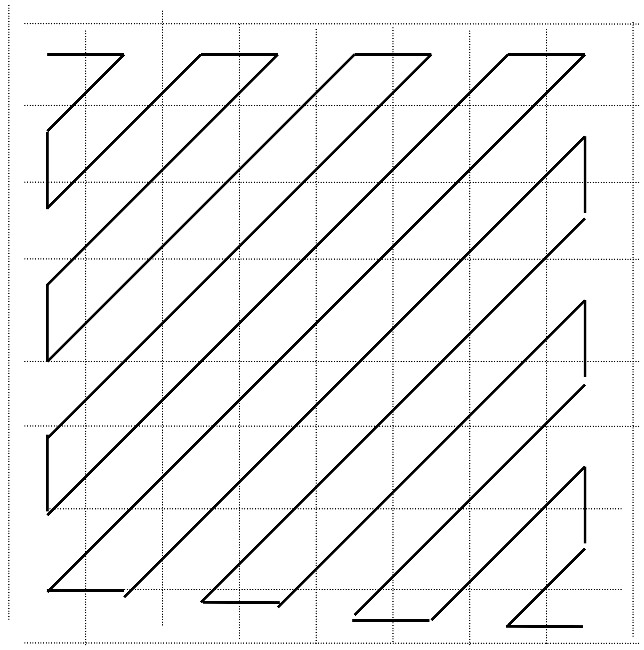

Рис.4. Зигзагообразный порядок считывания
коэффициентов преобразования

Кодер длин серий выдает ненулевое значение следующего за данным ненулевого коэффициента и длину серии нулей, т.е. число нулевых коэффициентов, предшествующих этому ненулевому коэффициенту. Таким образом, каждый ненулевой коэффициент может быть записан парой: *(длина серии/категория, амплитуда).* Значение длина серии/категория кодируется кодом Хаффмена, а значение амплитуды вычисляется как и в случае кодирования разностей коэффициентов постоянного тока и добавляется к коду. Например, пусть коэффициенту предшествует 6 нулей и он принимает значение -18. Из таблицы 1 следует, что -18 принадлежит к категории 5. Обратный код 18 равен 01101. Таким образом, коэффициент представляется парой (6/5,01101). Пара (6/5) кодируется кодом Хаффмена и 5-битовое значение для -18 добавляется к коду. Если слово кода Хаффмена для (6/5) есть 1101, то кодовое слово для -18 - это 110101101.

В кодировании коэффициентов переменного тока имеются два специальных случая.

1. После некоторого ненулевого коэффициента все остальные символы равны 0. В этом случае передается специальный символ (0/0), который означает конец блока (EOB).
2. Встретилась комбинация длина серии/категория, отсутствующая в таблице кода Хаффмена. В этом случае передается специальное кодовое слово, называемое escape-код, за которым равномерным кодом передается значение длины серии, далее следует равномерный код ненулевого значения.

Для завершения описания стандарта необходимы таблицы кодов и матрицы квантования. Эти данные мы не приводим, поскольку они заняли бы слишком много места и не существенны для понимания сути алгоритмов. Кроме того, стандарт допускает использование произвольных матриц квантования и неравномерных кодов для кодирования битовых категорий коэффициентов постоянного и переменного тока. В этом случае матрицы и кодовые таблицы должны быть переданы в составе битового потока в оговоренном стандартом формате.

Таблица 1.

Битовые категории целых чисел.

Категория	Числа
0	0
1	-1,1
2	-3,-2,2,3
3	-7,...,-4,4,...7
4	-15,...,-8,8,...15
5	-31,...-16,16,...31
6	-63,...-32,32,...63
7	-127,...-64,64,...127
8	-255,...-128,128,...255
9	-511,...-256,256,...511
10	-1023,..-.512,512,...1023
11	-2047,...-1024,1024,...2047
12	-4095,...-2048,2048,...4095
13	-8191,...-4096,4096,...8191
14	-16383,...-8192,8192,...16383
15	-32767,...-16384,16384,...32767
16	32768

Банки фильтров. Преобразование на основе банков фильтров

Известно, что последовательность $y(nT)$ на выходе ЦФ может быть выражена через входную последовательность $x(nT)$ и импульсную характеристику $h(nT)$ ЦФ с помощью уравнения свертки, т.е.

$$y(nT) = \sum_{k=0}^{n} h(kT)x(nT - kT) \qquad (7.1)$$

В матричном виде (7.1) может быть представлено следующим образом

$$\mathbf{y} = T\mathbf{x},$$

где

$$T = \begin{pmatrix} h(0) & 0 & 0 & 0 & ... \\ h(1) & h(0) & 0 & 0 & ... \\ h(2) & h(1) & h(0) & 0 & ... \\ h(3) & h(2) & h(1) & h(0) & ... \\ ... & h(3) & h(2) & h(1) & ... \end{pmatrix}$$

Другими словами, операция фильтрации может быть представлена как умножение на матрицу с постоянными диагоналями. В этом смысле операция фильтрации представляет собой некоторое линейное преобразование.

Рассмотрим простейший нерекурсивный ЦФ, задаваемый следующей формулой

$$y(nT) = 1/2x(nT) + 1/2x(nT - T). \qquad (7.2)$$

Коэффициенты этого фильтра $h(0) = 1/2$, $h(1) = 1/2$. Очевидно, что этот фильтр представляет собой КИХ фильтр и длина его импульсной характеристики равна 2. ЦФ, задаваемый уравнением (7.2), определяет так называемое скользящее среднее, так как в каждый момент времени выход ЦФ представляет собой среднее арифметическое текущего отсчета входного сигнала и его отсчета, взятого в предыдущий момент времени.

Найдем частотную характеристику ЦФ, задаваемого формулой (7.2). Вспоминая, что частотная характеристика ЦФ $H(e^{j\omega})$ есть преобразование Фурье от его импульсной характеристики, получаем

$$H(e^{j\omega T}) = \sum_n h(nT)e^{-j\omega nT} = 1/2 + 1/2e^{-j\omega T}. \qquad (7.3)$$

Вынося $e^{-j\omega T/2}$ из (7.3), получаем

$$H(e^{j\omega T}) = (1/2e^{-j\omega T/2} + 1/2e^{j\omega T/2})e^{-j\omega T/2}.$$

Таким образом, АЧХ ЦФ $\left|H(e^{j\omega T})\right| = \cos\left(\frac{\omega T}{2}\right)$, а ФЧХ $\varphi(\omega) = -\frac{\omega T}{2}$. Очевидно, что рассматриваемый фильтр представляет собой фильтр нижних частот. Это пример фильтра с линейной ФЧХ. Последнее свойство играет очень важную роль в процессе фильтрации. Оно означает, что фильтрация не вносит никаких фазовых искажений в выходной сигнал $y(nT)$, а приводит лишь к задержке выходного сигнала $y(nT)$ относительно входного $x(nT)$.

Теперь рассмотрим пример высокочастотного ЦФ. Это фильтр определяет скользящую разность отсчетов входной последовательности, т.е. задается формулой

$$y(nT) = 1/2x(nT) - 1/2x(nT - T). \qquad (7.4)$$

Его частотная характеристика определяется следующим образом

$$H(e^{j\omega T}) = \sum_n h(nT)e^{-j\omega nT} = 1/2 - 1/2e^{-j\omega T}.$$

Вынося $e^{-j\omega T/2}$,получаем, что АЧХ определяется по формуле

$$\left|H(e^{j\omega T})\right| = \sin\left(\frac{\omega T}{2}\right), \text{а ФЧХ } \varphi(\omega) = \frac{\pi}{2} - \frac{\omega T}{2}.$$

Нетрудно видеть, что ЦФ действительно представляет собой фильтр верхних частот. Он имеет линейную фазу, однако, ФЧХ имеет разрыв в точке $\omega = 0$.

Обозначим через H_0 ЧХ фильтра скользящего среднего и через H_1 - ЧХ фильтра скользящей разности, а матрицы соответствующих преобразований через T_0 и T_1, соответственно. Эти матрицы имеют вид

$$T_0 = \begin{pmatrix} 1/2 & 0 & 0 & 0 & \dots \\ 1/2 & 1/2 & 0 & 0 & \dots \\ 0 & 1/2 & 1/2 & 0 & \dots \\ 0 & 0 & 1/2 & 1/2 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \text{ и } T_1 = \begin{pmatrix} 1/2 & 0 & 0 & 0 & \dots \\ -1/2 & 1/2 & 0 & 0 & \dots \\ 0 & -1/2 & 1/2 & 0 & \dots \\ 0 & 0 & -1/2 & 1/2 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

Очевидно, что по отдельности оба рассмотренных ЦФ необратимы, то есть не существует $T_i^{-1}, i = 0,1$ (формально матрица обратного преобразования может быть записана, но она соответствует неустойчивому фильтру). Действительно не существует ЦФ, который мог бы восстановить входной сигнал $x(nT)$ из нулевой последовательности, т.е. обратимый

ЦФ должен иметь $H(\omega) \neq 0$ на всех частотах. Для рассмотренных фильтров $H_0(\pi) = 0$ и $H_1(0) = 0$.

Однако, если рассмотреть оба фильтра вместе, то они разделяют входной сигнал на полосы. Причем $T_1\mathbf{x}$ является дополнением для $T_0\mathbf{x}$. Говорят, что эти фильтры образуют *банк фильтров*. Можно показать, что существует обратный банк фильтров, т.е. преобразование, состоящее в разделении входной последовательности на полосы, является обратимым. Но теперь мы столкнулись с другой проблемой, а именно, длина выходной последовательности удваивается, так как каждая из компонент (низкочастотная и высокочастотная) имеют длину, равную длине входной последовательности. Эта проблема решается введением операции *децимации*. Пусть $\mathbf{y}_0 = T_0\mathbf{x}, \mathbf{y}_1 = T_1\mathbf{x}$, тогда операция децимации представляет собой выбрасывание всех нечетных компонент из векторов $\mathbf{y}_0$ и $\mathbf{y}_1$. Чтобы компенсировать выбрасывание всех нечетных компонент умножим оставшиеся компоненты на нормализующий множитель $\sqrt{2}$.

С учетом децимации матрицы преобразования T_0 и T_1 заменяются матрицами

$$L = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & \dots \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & \dots \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \dots \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix} \text{ и}$$

$$B = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & \dots \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & \dots \\ 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \dots \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

соответственно. Банк фильтров характеризуется матрицей вида

$$\begin{bmatrix} L \\ B \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & \dots & & & \\ & & 1 & 1 & \dots & \\ \dots & \dots & \dots & \dots & \dots & \\ -1 & 1 & \dots & & & .. \\ & & -1 & 1 & \dots & \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

Матрица обратная к матрице $\begin{bmatrix} L \\ B \end{bmatrix}$ существует и определяется следующим образом

$$\begin{bmatrix} L \\ B \end{bmatrix}^{-1} = \begin{bmatrix} L^T & B^T \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & & . & -1 & & . \\ 1 & & . & 1 & & . \\ & 1 & . & & -1 & . \\ & 1 & . & & 1 & . \\ & & . & & & . \\ & & . & & & . \end{pmatrix}.$$

Матрица $\begin{bmatrix} L^T & B^T \end{bmatrix}$ также задает банк фильтров, называемый банком восстанавливающих фильтров. Так как матрица обратного преобразования представляет собой транспонированную матрицу прямого преобразования, то рассмотренное преобразование является ортогональным. Очевидно, что

$$\begin{bmatrix} L^T & B^T \end{bmatrix} \begin{bmatrix} L \\ B \end{bmatrix} = L^T L + B^T B = I \, .$$

Таким образом, преобразование обеспечивает однозначное восстановление входной последовательности. Однако, прямое преобразование состояло из двух шагов, а именно фильтрации и децимации. Поэтому восстанавливающий банк фильтров также реализуется в виде двух операций: интерполяции и собственно фильтрации. Интерполяция состоит в том, что выброшенные нечетные компоненты векторов $\mathbf{y}_0$ и $\mathbf{y}_1$ заменяются 0. Затем выполняется фильтрация получившихся расширенных векторов $\tilde{\mathbf{y}}_0$ и $\tilde{\mathbf{y}}_1$ восстанавливающими фильтрами, задаваемыми, соответственно, следующими формулами

$$y(nT) = \frac{1}{\sqrt{2}}\left(x(nT) + x(nT - T)\right),$$

$$y(nT) = \frac{1}{\sqrt{2}}\left(x(nT - T) - x(nT)\right).$$

Сумма результатов фильтрации представляет собой задержанный входной сигнал.