

И.Н. Айзенберг, Л.П. Ярославский

ДВУХЭТАПНЫЙ ТРАНСФОРМАЦИОННЫЙ МЕТОД КОДИРОВАНИЯ ИЗОБРАЖЕНИЙ

1. Введение

Практически все известные методы кодирования изображений, направленные на устранение избыточности их цифрового описания, построены по следующей классической схеме: линейное преобразование - изображения или его фрагментов, поэлементное квантование полученных коэффициентов преобразования (рис. 1).

В соответствии с основными представлениями теории [1] информации самым общим методом квантования, позволяющим достичь теоретического нижнего предела

объема описания - эpsilon энтропии, является блочное квантование. Однако практические трудности осуществления блочного квантования побуждают отыскивать такое преобразование, после которого в блочном квантовании не было бы надобности и оно могло бы быть заменено существенно более простым в практической реализации поэлементным квантованием. Очевидно, что для этого преобразование должно быть таким, чтобы отсчеты сигнала после преобразования стали бы статистически независимыми. Найти такое преобразование обычно не удается, ибо оно требует измерения многомерных распределений отсчетов сигнала, что практически невозможно. Если же ограничиться только классом линейных преобразований, то максимум чего можно добиться, так это декоррелированности отсчетов сигнала. Такое декоррелирующее преобразование называется преобразованием Карунена-Лоэва. В практике кодирования изображений используются различные ортогональные преобразования, в той или иной мере приближающиеся по эффективности декорреляции к преобразованию Карунена-Лоэва. Одним из наилучших среди них является так называемое косинусное преобразование [2].

Но декоррелированность еще не означает статистическую независимость. Сохраняющаяся статистическая избыточность коэффициентов преобразования для раз-



Рис. 1. Схема классического трансформационного кодирования изображений

личных видов базисов преобразований может проявляться по-разному. Общим решением проблемы устранения этой избыточности по-прежнему остается блочное квантование. Существует, однако, ряд базисов, достаточно близких по своим декоррелирующим свойствам к базису Карунена-Лозва и характеризующихся тем, что остающаяся статистическая зависимость коэффициентов преобразования проявляется очень просто: при естественном для них упорядочивании коэффициентов преобразования модуля этих коэффициентов обладают устойчивой тенденцией к быстрому убыванию с ростом их номера. Это известное и характерное свойство спектров изображений в базисах преобразования Фурье, Уолша, косинусного преобразования и других. Более того, из эксперимента известно, что достаточно значительное искажение квадрата модуля спектра изображения в таких базисах незначительно сказывается на качестве воспроизведения изображения [3]. Так мы приходим к идее двухэтапного трансформационного способа кодирования изображений, состоящей в том (рис. 2), что на первом этапе кодирования выполняется первое преобразование, затем вычисляются амплитуда (модуль) и знак (или фаза) коэффициентов преобразования, знаки (фаза) подвергаются поэлементному квантованию, а последовательность амплитуд коэффициентов поступает на второй этап кодирования, например, снова посредством того или иного преобразования и последующего поэлементного квантования.

Из сказанного ясно, что двухэтапный способ кодирования потенциально обещает более высокую степень сжатия, чем классические трансформационные способы, основанные на преобразовании Карунена-Лозва и его аппроксимациях. Достижимая степень сжатия для этого метода определяется количеством бит на элемент, необходимых для передачи знака (фазы) и амплитуды спектральных компонент изображения. Для передачи знака компонент требуется, очевидно, не более 1 бита на компоненту. Допустимость существенного искажения амплитуд компонент позволяет надеяться, что, используя этот метод кодирования, можно сжать объем описания изображений до величины порядка 1 бита на элемент изображения. В данной статье описаны результаты экспериментов по исследованию возможностей различных вариантов такого двухэтапного метода кодирования изображений. Эксперименты проводились с изображениями размером 256×256 элементов при 8 битах (256 уровней) на элемент изображения.

2. Выбор первого преобразования и схемы кодирования амплитудной компоненты спектра

Для обеспечения кодирования амплитудной компоненты спектра необходимо, чтобы последовательность амплитуд преобразования была возможно ближе к моно-

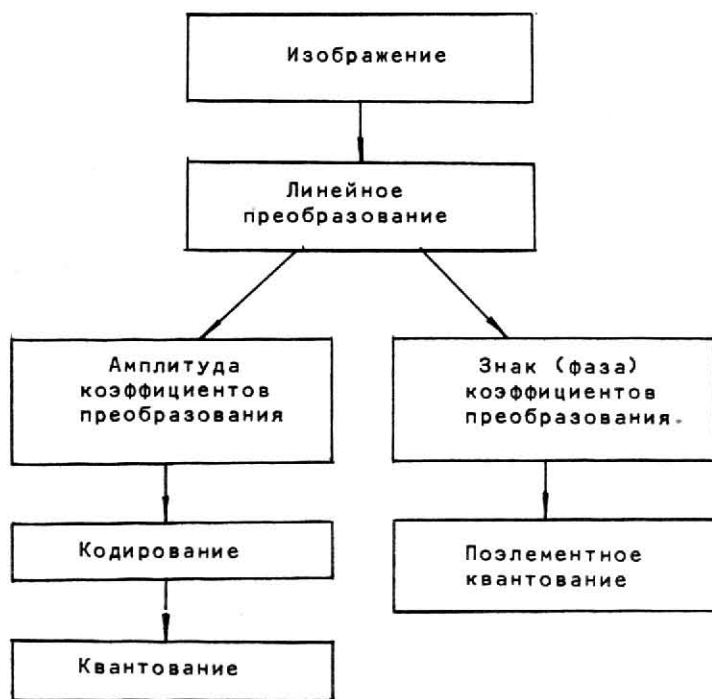


Рис. 2. Схема двухэтапного трансформационного метода кодирования изображений

тонно убывающей и в то же время убывала максимально быстро. Нами исследовались в качестве базисов разложения для первого преобразования базисы преобразования Уолша и косинусного преобразования. При этом косинусное преобразование рассматривалось как дающее потенциально наиболее быстро убывающий спектр благодаря тому, что при его использовании минимизируются краевые эффекты преобразования последовательностей конечной длины [4].

Для кодирования последовательности отсчетов амплитуды спектра косинусного преобразования изображений была выбрана схема, предусматривающая нелинейное предискажение значений отсчетов и последующую нелинейную коррекцию при восстановлении. Нелинейное предискажение вводится для уменьшения влияния ошибок аппроксимации амплитудного спектра на качество реконструкции (декодирования) спектра. Вид предискажающей функции определяется критерием точности реконструкции спектра. Пусть $a(k, l)$ - отсчеты амплитуды спектра; f - функция преобразования; $b(k, l)$ - преобразованные значения, подвергающиеся аппроксимации:

$$b = f(a), \quad (1)$$

$\Delta a(k, l)$ - ошибка аппроксимации, последовательности $a(k, l)$, $\Delta b(k, l)$ - ошибка аппроксимации преобразованных отсчетов $b(k, l)$. Очевидно, при малых $\Delta a(k, l)$, $\Delta b(k, l)$

$$\Delta b(k, l) = \frac{df}{da} \Delta a(k, l). \quad (2)$$

Мы исходим из порогового критерия:

$$\frac{\Delta a(k, l)}{a(k, l) + a_0} \leq \epsilon, \quad (3)$$

где a_0 - некоторая константа. Он означает, что при малых значениях ($a \ll a_0$) ограничивается абсолютная ошибка аппроксимации и при больших ($a \gg a_0$) - относительная ошибка. Статистический критерий ограничивает среднее по всем "a" значение такой же ошибки.

Найдем оптимальный вид предискажающей функции $f(a)$ для критерия (3). Очевидно, что из (3) следует необходимость поддержания ошибок на уровне порога:

$$\Delta a(k, l) = \epsilon [a(k, l) + a_0].$$

Пусть ошибка аппроксимации $b(k, l)$ не зависит от номера отсчета (k, l). Тогда из (2) и (3) следует, что

$$\frac{df}{da} = \frac{\Delta b}{\Delta a} = \frac{\Delta b}{\epsilon [a + a_0]}. \quad (4)$$

Откуда получаем:

$$\frac{b - b_{\min}}{b_{\max} - b_{\min}} = \frac{\ln \frac{a + a_0}{a_0}}{\ln \frac{a_{\max} + a_0}{a_0}}, \quad (5)$$

где $a_{\max}$ - максимальное значение a , которое передается без ограничения; $b_{\min}$,

$b_{\max}$ - минимальное и максимальное значения b . Без ограничения общности можно выбрать $b_{\min} = 0$, при этом величина $b_{\max}$ становится масштабным коэффициентом, который можно выбрать произвольно. Тогда основными параметрами преобразования становятся величины a_0 и $a_{\max}$. Величина a_0 выбиралась из условия, величина $a_{\max}$ выбиралась по фактическому максимальному значению отсчетов амплитуд спектра. При реконструкции спектра после декодирования выполнялось преобразование, обратное (5).

3. Кодирование двумерного спектра

Для аппроксимации модуля двумерного косинусного спектра после его преобразования по (5) использовались следующие функции:

$$f(i, j) = \frac{a(i^2 + j^2)}{b(i^3 + j^3)} - c \cdot i \cdot j, \quad (6)$$

$$f(i, j) = \frac{a}{b(i+j)} - c(i+j), \quad (7)$$

где i и j - номера отсчетов по осям абсцисс и ординат, соответственно $i, j = 1, 2, \dots, N$; $N = 256$. Экспериментально установлено, что параметры a, b и c , при помощи которых описывается аппроксимирующая функция, меняются в зависимости от конкретного класса изображений (например, крупно- или мелкодетальные). Функции (6) и (7) достаточно хорошо приближают двумерный косинусный спектр на интервалах $[N/8, N-1]$ по обеим координатам, несколько хуже - на интервалах $[N/16, N/8]$. На интервалах $[0, N/16]$, где последовательность модулей спектральных коэффициентов является наименее гладкой и имеет значительные скачки, поведение указанных функций довольно сильно расходится с поведением аппроксимируемой последовательности модулей спектра, т.е. в высокочастотной своей части спектр аппроксимируется хорошо, в области средних частот аппроксимация удовлетворительная, в области же низких частот - неудовлетворительная. Это иллюстрируется рис. 3, на котором представлен профиль модуля двух-

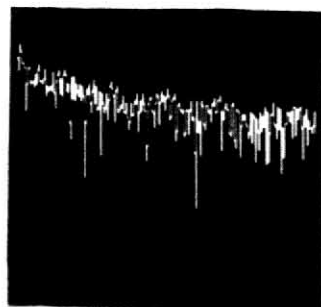


Рис. 3. График сечения двумерного косинусного спектра изображения типа "портрет крупным планом"

мерного спектра изображения "портрет крупным планом". Отсюда следует, что участок последовательности модулей, соответствующий низкочастотной части спектра (размером примерно $2^{N/16} \times 2^{N/16}$), должен быть закодирован каким-либо иным способом. Параметры в формулах (6) и (7) выбираются исходя из следующего. Параметры a и b определяют соответствие поведения аппроксимирующей функции поведению аппроксимируемой. Экспериментально установлено, что в (6) для $N=256$.

$$\begin{cases} 0.7 \leq a \leq 1.5 \\ 10 \leq b \leq 20 \end{cases}$$

и в (7)

$$\begin{cases} 5 \leq a \leq 10 \\ 0.6 \leq b \leq 1.5 \end{cases}$$

При этом для "а" область в районе максимума (для b - минимума) соответствует крупнодетальным изображениям, противоположные значения дают лучшие результаты на мелкодетальных изображениях. Коэффициент "с" не зависит от класса кодируемого сигнала и изменяется от низкочастотной области к высокочастотной следующим образом:

$$\begin{aligned} 0.015 \geq c \geq 0.0006 & \quad \text{для (6)} \\ 0.017 \geq c \geq 0.007 & \quad \text{для (7)} \end{aligned}$$

Функция (6) лучше приближает косинусный спектр на интервалах $[0, N/2]$ по обеим координатам, а (7) - на интервалах $[N/2, N]$.

На рис. 4 приведены тестовые изображения размером 256×256 элементов; на рис. 5 - результаты экспериментов по их кодированию. Эксперименты проводились следующим образом. Фазовая (знаковая) часть спектра сохранялась в первоначальном виде. Амплитуда подвергалась преоб-

разованию по (5), после чего аппроксимировалась функциями (6) - в низкочастотной части спектра и (7) - в высокочастотной. Модули спектральных коэффициентов, соответствующих наиболее низким частотам (или "заключенные" в интервале от $[0, 31]$ до $[0, 50]$ по обеим координатам) подвергались зональному квантованию. Эксперименты показали, что описанная реализация двухэтапного метода позволяет достичь сжатия до 1,3-1,5 бит/элемент.

Классическая реализация трансформационных методов (2) предполагает отбрасывание части спектральных коэффициентов, вносящих наименьший вклад в общую энергию сигнала. Этот прием можно, в принципе, как дополнительный, применить и к описываемому двухэтапному методу. На рис. 6 представлены изображения, закодированные тем же способом, что и на рис. 5, однако из их спектров отброшена примерно половина коэффициентов, соответствующих самым высоким частотам. При этом проявляется естественный и неизбежный в таких случаях дефект - дефокусировка восстанавливаемого изображения.

При реализации излагаемого подхода могут встретиться случаи, когда плохо сходящийся спектр, даже после применения к нему процедуры (5), содержит значительное число скачков. Для преодоления такого "дефекта" предлагается следующий прием. Для класса изображений с "плохими" спектрами находится среднее значение модуля спектра $|S|$. Затем вычисляется разность $|S| - |S| = R$, где $|S|$ - модуль спектра кодируемого видеосигнала. $|R|$ будет, естественно, более гладкой функцией, чем $|S|$. Отсюда и вытекает возможность использовать $|R|$ в ка-



Рис. 4. Исходные тестовые изображения



Рис. 5. Двухэтапное трансформационное кодирование с двумерным преобразованием



Рис. 6. Двухэтапное трансформационное кодирование с двумерным преобразованием и отбрасыванием части коэффициентов

честве "нулевого" приближения к модулю спектра. Именно такой подход использован при кодировании изображений типа "портрет крупным планом" (см. рис. 5).

4. Кодирование одномерных спектров

По аналогии с двумерным случаем естественно предположить, что амплитуд-

ная компонента одномерных спектров строк (столбцов) изображений в ортогональных базисах также обладает значительной избыточностью и к ним также приложима идея двухэтапного метода кодирования. Кодирование одномерных спектров предпочтительнее кодирования двумерных в силу того, что кодирование одномерных спектров существенно проще технически. В то же время ясно, что оно

сопряжено с необходимостью точнее передавать аппроксимируемый спектр, чем в двумерном случае, ибо одномерные спектры отдельных строк изображения существенно разнятся друг от друга даже для соседних строк изображения (рис. 7).

Главная проблема реализации идеи двухэтапного трансформационного метода кодирования - выбор функции, аппроксимирующей амплитудную компоненту спектра. Если посмотреть на эту проблему с позиций статистической теории кодирования, то лучшей аппроксимацией множества амплитудных спектров строк изображения будет аппроксимация разложением по базису Карунена-Лозва. Нулевым приближением по этому базису будет приближение средними значениями модулей спектра строк. Вариант такого метода кодирования иллюстрируется на рис. 8. Здесь амплитудная компонента спектра строк изображения в пределах фрагмента из соседних строк заменялась усредненной амплитудной компонентой по всем строкам фрагмента, причем для повышения точности аппроксимации первые $N/10-N/8$ спектральных коэффициентов передавались точно, то есть брались из спектра данной строки, а не из усредненного спектра.

Несколько лучшие результаты по сжатию можно получить, если не сохранять механически модули коэффициентов, соответствующих низким частотам, а сохранять только те, которые в некотором смысле заметно отличаются от соответствующих коэффициентов усредненного спектра. Алгоритм обнаружения таких "особых" точек может быть следующим. Пусть $\epsilon > 0$ - некоторый порог. Точка k считается особой, если

$$|f(k) - \tilde{f}(k)| > \epsilon,$$

где $f(k)$ - модуль k -го спектрального коэффициента строки; $\tilde{f}(k)$ - усредненное значение того же коэффициента, причем

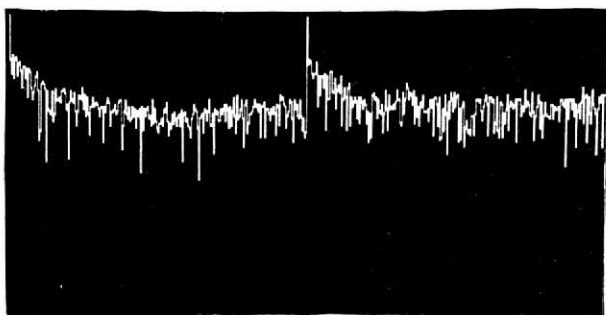


Рис. 7. Графики одномерных косинусных спектров соседних строк изображения типа "портрет крупным планом"

величина порога выбирается экспериментально и пропорционально стандартному отклонению модулей спектральных коэффициентов строки от усредненного спектра.

Другой технически простой способ аппроксимации модулей спектров строк - использование метода предсказания модуля спектра строк по предыдущим строкам: амплитудная компонента первой строки передается полностью, для второй строки передаются только особые точки, определенные относительно первой строки, а значения амплитуды спектра в остальных точках полагаются равными значениям амплитуд в соответствующих точках первой строки и, вообще, если $f(i, j)$ - последовательность амплитуд спектра i -й строки, $\tilde{f}(i, j)$ - ее аппроксимация,

$$\text{то } \tilde{f}(1, j) = f(1, j), \quad j = \overline{1, N}$$

$$\tilde{f}(i, j) = \begin{cases} f(i, j), \\ \tilde{f}(i-1, j), \end{cases}$$

$$\text{если } |f(i, j) - \tilde{f}(i-1, j)| > \epsilon,$$

$$|f(i, j) - \tilde{f}(i-1, j)| \leq \epsilon, \\ i = \overline{2, N}.$$



Рис. 8. Двухэтапное трансформационное кодирование с одномерным преобразованием, вариант 1

Порог ϵ выбирается экспериментально так же, как в предыдущем случае. Эксперименты по такому методу кодирования иллюстрируются рис. 9. Здесь в качестве порога ϵ использовалась величина, пропорциональная стандартному отклонению наблюдаемого спектра от аппроксимации, причем коэффициент пропорциональности принимал следующие значения (j - номер отсчета): $2,5(0 \leq j < 70)$; $1,5(70 \leq j < 140)$; $1,0(140 \leq j < 200)$; $0,6(200 \leq j < 240)$; $0(240 \leq j \leq 256)$.

Техника передачи особых точек достаточно проста. Каждая особая точка полностью определяется своим номером в строке, номера же самих строк запоминать не нужно. В самом деле, пусть $\{k_j\}$ - последовательность, состоящая из номеров особых точек, взятых от 1-й строки до N -й. Ясно, что если номера $k_{i_s}, \dots, k_{i_{s+m}}$ соответствуют особым точкам одной строки, то $k_{i_{j+1}} > k_{i_j}$ при

$s < j < s+m$. Если же найдется i_j , для которого $k_{i_{j+1}} < k_{i_j}$, это означает, что номер $k_{i_{j+1}}$ соответствует особой точке

спектра следующей строки. Это следует из того, что особые точки сосредоточены в основном в низкочастотной области спектра и, как бы ни были похожи соседние строки, именно в этой области они проявляются обязательно, причем всегда в самом начале строки.

Вместо значений самих модулей особых точек спектра целесообразно кодировать их разности (этот прием можно применить и к последовательности координат). Незначительное искажение, вносимое при этом в спектр, не сказывается на качестве декодируемого изображения, цифровое же описание несколько упрощается.



Рис. 9. Двухэтапное трансформационное кодирование с одномерным преобразованием, вариант 2

5. Сравнение двумерного и одномерного двухэтапного трансформационного метода кодирования

Искажения изображений при использовании двухэтапного трансформационного метода кодирования проиллюстрируем на примере вариантов метода (см. рис. 5 и 9). На рис. 10 и 11 показаны усиленные для наглядности поэлементные разности для способов рис. 5 и 9 соответственно. Из сравнения этих рисунков можно сделать следующие выводы.

При использовании двумерного преобразования значительно сильнее проявляется дефокусировка изображения вследствие

отбрасывания части высокочастотных коэффициентов. Способ кодирования с одномерным преобразованием в этом смысле более устойчив, так как для достижения той же степени сжатия, что и для двумерного кодирования, он не нуждается в отбрасывании высокочастотных коэффициентов преобразования благодаря возможности простой аппроксимации высокочастотной части амплитудного спектра. Кроме того, благодаря автоматическому обнаружению "особых" точек в одномерном спектре удастся избежать сильного сглаживания формы спектра, которое имеет место в двумерном случае. Аналогичное обнаружение можно производить, конечно, и в двумерном случае, но тогда способ лишился бы своего главного достоинства - простоты.

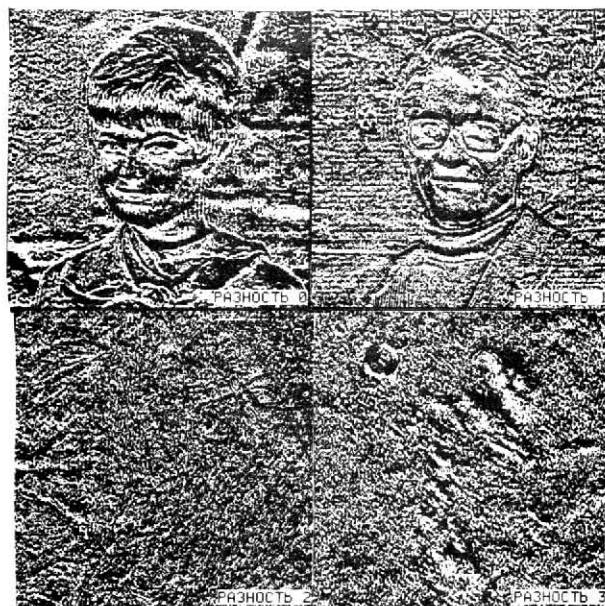


Рис. 10. Усиленная разность между оригиналами и изображениями, закодированными двухэтапным методом с двумерным преобразованием

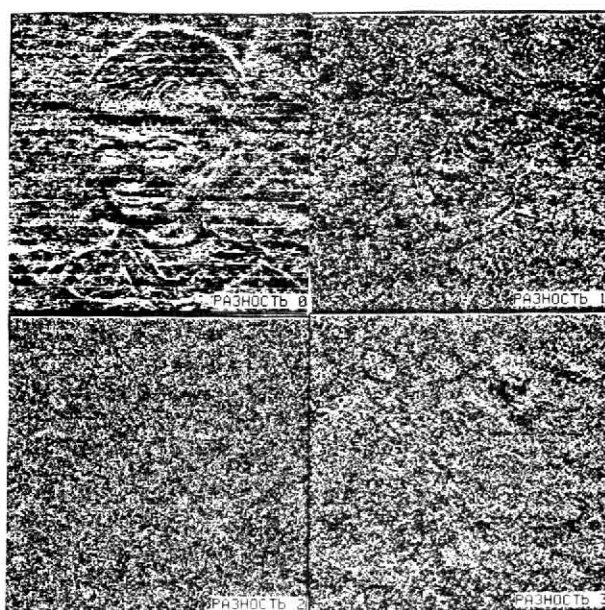


Рис. 11. Усиленная разность между оригиналами и изображениями, закодированными двухэтапным методом с одномерным преобразованием, вариант 2

Определенный интерес представляет также сравнение двухэтапного метода кодирования с известным трансформационным кодированием [2]. На рис. 12 иллюстрируется результат применения классического трансформационного метода к одному из тестовых изображений. Даже чисто визуальное сравнение говорит в

пользу двухэтапного метода. Эксперименты показывают, что при традиционном трансформационном кодировании сжатие до 1,8 бит/элемент сопряжено со значительно большими искажениями сигнала, чем сжатие до 1,4-1,5 бит/элемент для двухэтапного метода.



Рис. 12. Классическое трансформационное кодирование

Л и т е р а т у р а

1. Шеннон К. Математическая теория связи // Работы по теории информации и кибернетике. М., ИИЛ, 1963.
2. Ярославский Л.П. Введение в цифровую обработку изображений. М., Сов. радио, 1979.
3. Ахмед Н., Рао К.Р. Ортогональные преобразования при обработке цифровых сигналов. - М.: Связь, 1980.
4. Карнаухов В.Н., Ярославский Л.П. Коррекция четкости изображений при произвольной "гладкой" дефокусировке // Письма в ЖТФ, 1981. Т. 7. Вып. 15. с. 908-911.