

РОССИЙСКАЯ АКАДЕМИЯ НАУК
СИБИРСКОЕ ОТДЕЛЕНИЕ
А В Т О М Е Т Р И Я

2002, том 38, № 4

УДК 621.372

Ю. С. Радченко
(Воронеж)

**МЕТОД СЖАТИЯ И ВОССТАНОВЛЕНИЯ ИЗОБРАЖЕНИЙ
НА ОСНОВЕ БЫСТРЫХ ЧЕБЫШЕВСКИХ ПРЕОБРАЗОВАНИЙ**

Предложен метод сжатия изображения, основанный на разложении по полиномам Чебышева и применении квадратурных формул Гаусса – Чебышева (алгоритм GDCT). Показано, что алгоритм GDCT является обобщением дискретного косинусного преобразования, используемого в алгоритмах JPEG и MPEG. Алгоритм имеет четыре ступени сжатия: прореживание исходного изображения, усечение спектра, обнаружение изменяющихся фрагментов, определение вектора движения фрагментов. На примерах обработки искусственных текстур и натуральных изображений продемонстрированы возможности алгоритма GDCT.

Введение. Передача по каналам цифровой связи потоков сообщений, телевизионных и компьютерных изображений предполагает применение процедуры сжатия сигналов. В настоящее время наибольший интерес вызывают алгоритмы, основанные на разложении сигналов по некоторому ортогональному базису [1, 2]. Однако ввиду огромной информационной емкости изображений остро стоит вопрос о создании процедур разложения сигналов, минимизирующих вычислительные затраты. Такие алгоритмы называются «быстрыми». Среди многообразия различных базисов можно выделить следующие преобразования, допускающие быструю реализацию: различные варианты дискретного преобразования Фурье; разложение по базису разрывных кусочно-постоянных функций (Уолша, Хаара и др.); вейвлет-преобразование. В последнее время выполнен ряд работ, в которых предлагается применять разложения по базису классических ортогональных полиномов Эрмита, Якоби (Лежандра, Чебышева) [3, 4]. Такие преобразования позволяют добиться, как показывают расчеты, максимального сжатия и допускают реализацию в виде «быстрых» алгоритмов преобразования. Кроме того, они обладают весьма компактными спектрами, которые чувствительны к сдвигу анализируемого фрагмента. В данной работе приводятся результаты исследования одного из вариантов полиномиальных преобразований – разложение и синтез по базису полиномов Чебышева 1-го и 2-го рода. Вычисление спектральных коэффициентов производится с помощью квадратурной формулы Гаусса – Чебышева, обладающей наивысшей степенью алгебраической точности при заданном числе отсчетов. Показано, что при этом получают преобразования, которые являются обобщением дискретного косинусного и синусного преобразований.

Разложение сигналов по базису ортогональных полиномов. Пусть в подобласти $\{x, y\} \in \Omega$ наблюдается поле $s(x, y, \tau)$, представляющее собой фрагмент $u(x, y)I_\Omega(x, y)$ пространственного сигнала. Здесь $I_\Omega(x, y)$ – индикаторная функция подобласти, $\tau = (\tau_x, \tau_y)$ – параметры сдвига фрагмента в данном кадре. Базисные функции $\varphi_{mk}(x, y)$, используемые для дискретного представления сигнала $s(x, y, \tau)$, определяются аналитическими свойствами самого сигнала и геометрической формой подобласти Ω . Однако реализация процедуры сжатия особенно для полиномиальных базисов существенно упрощается, если имеет место факторизация функций $\varphi_{mk}(x, y) = \varphi_m(x)\varphi_k(y)$. Здесь $\varphi_m(x), \varphi_k(y)$ – одномерные функции, основанные на ортогональных полиномах. Тогда для полезного сигнала $s(x, y, \tau)$ имеет место пара преобразований:

$$s(x, y, \tau) = \sum_m \sum_k C_{mk}(\tau) \varphi_m(x) \varphi_k(y), \quad (1)$$

$$C_{mk}(\tau) = \iint_{\Omega} s(x, y, \tau) \varphi_m(x) \varphi_k(y) dx dy.$$

Для разложения сигналов и изображений можно применять следующие классические ортогональные многочлены: Эрмита, Лежандра, полиномы Чебышева 1-го и 2-го рода. Если обозначить через a_x, a_y характерные размеры подобласти Ω , $z_1 = x/a_x, z_2 = y/a_y$, то (1) можно переписать с использованием ортогональных полиномов в виде

$$s(x, y) = \sum_{m, k} C_{mk} p_m(x/a_x) p_k(y/a_y);$$

$$C_{mk} = (d_m d_k)^{-1} \iint_{\Omega} s(a_x z_1, a_y z_2) \rho(z_1) p_m(z_1) \rho(z_2) p_k(z_2) dz_1 dz_2 = (2)$$

$$= (d_m d_k)^{-1} \int \rho(z_1) p_m(z_1) dz_1 \int s(a_x z_1, a_y z_2) \rho(z_2) p_k(z_2) dz_2.$$

Здесь d_m – норма ортогонального с весом $\rho(z)$ полинома $p_m(z)$. Поскольку для разложимых базисных функций $\varphi_{mk}(x, y) = \varphi_m(x)\varphi_k(y)$ вычисление спектральных коэффициентов производится последовательным интегрированием по координатам (x, y) , то для упрощения анализа рассмотрим сначала одномерные преобразования, а затем обобщим их для двумерного случая.

Алгоритм преобразований Чебышева. Пусть имеется процесс $f(z)$, квадратично интегрируемый с весом $\rho(z)$. Тогда можно записать квадратурную формулу гауссовского типа наивысшей алгебраической степени точности порядка $2N - 1$ как

$$\int f(z) \rho(z) dz = \sum_{n=1}^N \lambda_n f(z_n).$$

Здесь z_n – нули полиномов $p_N(z)$, ортогональных с весом $\rho(z)$; λ_n – числа Кристоффеля. Узлы и веса $\{z_n\}, \{\lambda_n\}$ однозначно определяются видом орто-

гонального с весом $\rho(z)$ полинома $p_m(z)$. Для полиномов Чебышева получается формула Меллера (Гаусса –Чебышева):

$$\frac{1}{d_m} \int_{-1}^1 \frac{s(z)T_m(z)}{\sqrt{1-z^2}} dz = \frac{\pi}{d_m N} \sum_{n=0}^{N-1} s(z_n)T_m(z_n). \quad (3)$$

Здесь $z_n = \cos(\pi(2n+1)/2N)$ – нули полинома Чебышева 1-го рода $T_N(z)$; $\lambda_n = \pi/N$, т. е. все весовые коэффициенты одинаковы; норма полинома Чебышева $d_m = \pi/2$, если $m \neq 0$, и $d_m = \pi$, если $m=0$. Учитывая, что

$$T_m(z_n) = \cos(m \arccos(z_n)) = \cos(\pi m(n+0,5)/N),$$

приходим к выражению для прямого и обратного преобразований:

$$C_m = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} s(z_n) \cos\left(\pi m \frac{n+0,5}{N}\right), \quad C_0 = \sqrt{\frac{1}{N}} \sum_{n=0}^{N-1} s(z_n); \quad (4)$$

$$S_M(z) = g_m \sqrt{\frac{2}{N}} \sum_{m=0}^M C_m T_m(z) = g_m \sqrt{\frac{2}{N}} \sum_{m=0}^M C_m \cos(m \arccos(z)).$$

Здесь $g_m = \begin{cases} \sqrt{0,5}, & m=0, \\ 1, & m>0. \end{cases}$ В формуле (4) использована симметричная нормировка базисных функций, удобная для практической реализации преобразований в матричном виде. Выражение (4) похоже на так называемое четное дискретное косинусное преобразование (ДКП) [2], но имеет три важных отличия.

1. Точки отсчета $z_n = \cos(\pi(n+0,5)/N)$ сигнала $s(z)$ берутся неравномерно.
2. Синтез сигнала $S_M(z)$ выполняется в произвольной точке $z \in [-1, 1]$, а не в дискретном наборе точек отсчета, как в ДКП.
3. Точность формулы (4) при преобразовании достаточно гладких функций существенно выше, чем у ДКП. Поэтому число отсчетов можно взять значительно меньше.

Если при восстановлении взять неравномерную сетку отсчетов по закону $z_j = \cos(\pi(j+0,5)/L)$, $j=0, \dots, L-1$, то обратное преобразование принимает вид

$$S_M(z_j) = g_m \sqrt{\frac{2}{N}} \sum_{m=0}^M C_m \cos\left(\pi m \frac{j+0,5}{L}\right). \quad (5)$$

Преобразование (5) выглядит аналогично обычному ДКП, однако их спектральные коэффициенты отличаются. В случае если нас интересует равномерная сетка отсчетов $z_j = 2j/(L-1) - 1 + \delta$, $j=0, \dots, L-1$, где δ – некоторый сдвиг, то

$$S_M(z_j) = g_m \sqrt{\frac{2}{N}} \sum_{m=0}^M C_m \cos(m \arccos(2j/(L-1) - 1 + \delta)). \quad (6)$$

Алгоритм преобразования (4), (6) назовем обобщенным дискретным косинусным преобразованием (GDCT). В формулах (5), (6) число точек L в блоке восстановленного изображения может быть произвольным. В таком случае возникает эффект масштабирования восстановленного изображения по сравнению с размерами исходного блока с изображением.

Если использовать разложение по полиномам Чебышева 2-го рода, то получим

$$C_m = \frac{2}{\pi} \int_{-1}^1 \sqrt{1-z^2} s(z) U_m(z) dz = \frac{2}{N+1} \sum_{k=0}^{N-1} s(z_k) \sin\left(\pi \frac{(k+1)(m+1)}{N+1}\right) \sin\left(\pi \frac{k+1}{N+1}\right); \quad (7)$$

$$S_M(z) = \sum_{m=0}^M C_m U_m(z).$$

Здесь учтено, что $z_k = \cos(\pi(k+1)/(N+1))$ и

$$U_m(z_k) = \frac{\sin((M+1)\arccos(z_k))}{\sqrt{1-z_k^2}} - \frac{\sin(\pi(k+1)(m+1)/(N+1))}{\sin(\pi(k+1)/(N+1))}. \quad (8)$$

Алгоритм преобразований (7) сходен с дискретным синусным преобразованием.

Для использования квадратурных формул типа Гаусса (3) при переходе от интегральных преобразований к дискретным требуется определить число узлов N , обеспечивающее вычисление необходимого количества спектральных коэффициентов C_m с заданной точностью. Как указывалось выше, формулы типа Гаусса являются точными для всех полиномов $I_{2N-1}(z)$ степени $2N-1$, а порядок ошибки формулы (3) $\epsilon \sim |s^{(2N)}(z)|$. Однако это только качественная оценка. Кроме того, ряд моделей полезных сигналов имеет особые точки, в которых функция не дифференцируема. Поэтому необходимы расчеты для типовых моделей сигналов, с помощью которых можно оценить нужное число отсчетов N . Вопросы точности аппроксимации сигналов и выбора числа узлов квадратурной формулы исследовались на примере следующих тестовых моделей: непрерывных дифференцируемых и недифференцируемых на множестве точек, финитных, разрывных (импульсных).

Например, если в качестве модельного сигнала взять $s(z) = \sqrt{1-z^2}$, то точные значения спектральных коэффициентов следующие:

$$C_0^T = \frac{i}{\pi} \int_{-1}^1 \frac{s(z)}{\sqrt{1-z^2}} dz = \frac{2}{\pi};$$

$$C_m^T = \frac{2}{\pi} \int_{-1}^1 \frac{s(z) T_m(z)}{\sqrt{1-z^2}} dz = \frac{1}{\pi} \left[\frac{1}{m+1} T_{m+1}(z) - \frac{1}{m-1} T_{m-1}(z) \right]_{-1}^1 = \quad (9)$$

$$= \begin{cases} \frac{-4}{\pi(m^2 - 1)} & \text{при четном } m > 0, \\ 0 & \text{при нечетном } m. \end{cases}$$

При вычислении по формуле (4) соответствующие коэффициенты имеют значения:

$$C_0 = \frac{1}{N} \sum_{k=0}^{N-1} \sin\left(\pi \frac{k+0,5}{N}\right); \quad (10)$$

$$C_m = \frac{2}{N} \sum_{k=0}^{N-1} \sin\left(\pi \frac{k+0,5}{N}\right) \cos\left(\frac{\pi m(k+0,5)}{N}\right).$$

В таблице приведены значения C_m^T и C_m при различном числе квадратурных точек N ($N = 4, 8, 12, 16$). Как видно из таблицы, достаточно точно вычисляются коэффициенты с номером $m \leq 3N/4$.

Были исследованы ошибки восстановления сигналов с использованием метрик

$$d_p(S_M, S) = \left\{ \int |S_M(x) - S(x)|^p dx \right\}^{1/p}$$

в пространстве L^p измеримых и интегрируемых функций для значений $p = 1, 2, \infty$. Расчеты для различных классов сигналов показали достаточность 6–8 точек для использования чебышевских преобразований. При этом размеры фрагмента изображения, к которому применяется полиномиальное преобразование, могут быть в 1,5–3 раза больше размеров блоков, применяемых в алгоритмах JPEG, MPEG.

Преобразование на основе ортогональных полиномов может быть реализовано в виде «быстрых» процедур. В основе таких процедур лежит минимизация числа отсчетов сигнала, симметрия узлов и весов квадратурных формул гауссовского типа и рекуррентный способ вычисления ортогональных

m	C_m^T	$C_m (N=4)$	$C_m (N=8)$	$C_m (N=12)$	$C_m (N=16)$
0	0,673	0,653	0,641	0,638	0,638
2	-0,424	-0,383	-0,416	-0,421	-0,422
4	-0,085	0	-0,075	-0,081	-0,083
6	-0,036	0,383	-0,023	-0,032	-0,034
8	-0,02	-1,307	0	-0,015	-0,018
10	-0,013	0,383	0,023	-0,0065	-0,01
12	-0,0089	-	0,075	0	-0,0055

полиномов. В [5] был описан векторно-матричный алгоритм полиномиальных преобразований с минимальным числом операций, требующий незначительных расходов памяти. Все операции выполняются с действительными числами. Установление связи формул Гаусса – Чебышева с дискретным косинусным и синусным преобразованиями позволяет воспользоваться существующими способами быстрого преобразования Фурье.

Пространственные (блочные) преобразования GDCT. Перейдем теперь к двумерному варианту преобразований сигнала $s(x, y)$. Пусть в пределах блока из $N_1 \times N_1$ точек выбирается матрица из $N \times N$ отсчетов по закону

$$x_n = \text{ROUND}(0,5(N_1 - 1)(1 + \cos(\pi(n + 0,5)/N))), \quad (11)$$

$$y_k = \text{ROUND}(0,5(N_1 - 1)(1 + \cos(\pi(k + 0,5)/N))).$$

Отсчеты сигнала образуют матрицу $S = \|s_{nk}\| = \|s(x_n, y_k)\|$. Затем эта матрица преобразуется в матрицу спектральных коэффициентов C размером $M \times M$. При обратном преобразовании может использоваться прямоугольная матрица размером $L \times M$, т. е. восстановленный блок имеет размеры $L \times L$. Прямое и обратное преобразования Чебышева (GDCT) в матричном виде определяются операциями

$$C = \|C_{m,i}\| = \hat{\Phi} S \hat{\Phi}^T, \quad S = \Psi^T C \Psi. \quad (12)$$

Матрицы прямого и обратного преобразований GDCT имеют вид

$$\hat{\Phi} = \|\Phi_m(n)\|_{NM} = \sqrt{\frac{2}{N}} \begin{bmatrix} \sqrt{0,5} \\ \cos\left(\pi m \frac{(n + 0,5)}{N}\right) \end{bmatrix}; \quad (13)$$

$$\Psi = \|\cos(\text{arccos}(z_j))\|_{LM} = \left\| \cos\left(\text{arccos}\left(\frac{2j}{L-1} - 1 + \delta\right)\right) \right\|.$$

В частном случае можно полагать $\delta = 0$. Дискретизация аргументов $\{x, y\}$ по правилу (11) приводит к появлению погрешности квадратурной формулы и неточному вычислению спектральных коэффициентов. Для уточнения расчетов можно произвести интерполяцию значения поля $s(x, y)$ в точках расположения нулей полиномов Чебышева по ближайшим отсчетам. Например, для этого можно использовать четырехточечную интерполяционную формулу Бесселя:

$$s(x_n, y_k) = 0,25(s_{00} + s_{01} + s_{10} + s_{11}) + 0,5(u - 0,5)(s_{10} - s_{00} + s_{11} - s_{01}) + \\ + 0,5(v - 0,5)(s_{01} - s_{00} + s_{11} - s_{10}) + (u - 0,5)(v - 0,5)(s_{11} - s_{10} - s_{01} + s_{00}), \quad (14)$$

где $u = x_n - x_{00}$; $v = y_k - y_{00}$; s_{ij} ($i, j = 0, \dots, 1$) — значения сигнала в ближайших четырех пикселах, окружающих точку x_n, y_k . Применение интерполяции улучшает передачу мелкомасштабной текстуры изображений при реализации процедуры сжатия-восстановления сигналов.

Результаты исследования алгоритма GDCT. Итак, в блоке изображений размером $N1 \times N1$ элементов выбирается матрица $N \times N$ отсчетов ($N < N1$). Так происходит сжатие на первой ступени процедуры. Затем с помощью преобразования GDCT вычисляется матрица $M \times M$ из наиболее значимых спектральных коэффициентов, которые в дальнейшем прореживаются или при помощи матрицы квантования, или путем регулируемой низкочастотной фильтрации. Так происходит сжатие изображения на второй ступени процедуры. При восстановлении изображения формируется блок размером $L \times L$ элементов. Этот размер может быть меньше, равен или больше размеров исходного блока $N1 \times N1$.

Для экспериментов был выбран набор искусственных изображений и натуральных фотографий. Цветные изображения были представлены в 24-битовом формате BMP. С этими изображениями осуществлялись следующие операции.

1. Перекодировка 8-битовых R -, G -, B -компонент цвета в Y, U, V по правилу:

$$Y = \text{int}(0,299R + 0,587G + 0,114B),$$

$$U = \text{int}(0,500R - 0,419G - 0,081B) + 128,$$

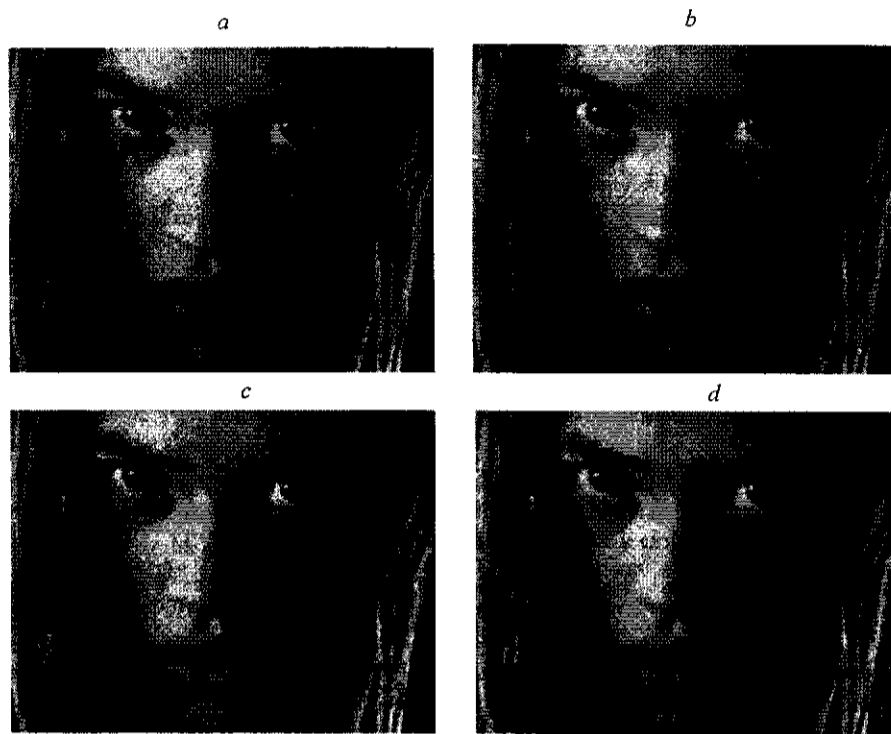
$$V = \text{int}(-0,169R - 0,331G + 0,500B) + 128.$$

2. Субдискретизация матриц Y, U, V по одному из законов: $4 : 4 : 4, 4 : 2 : 2, 4 : 2 : 0$.

3. Спектральное преобразование матриц Y, U, V на основе GDCT. В зависимости от типа изображения на первой ступени сжатия проводилось прореживание в пропорции $N1/N = 16/6, 16/8, 8/6$. Затем проводилось преобразование спектральной матрицы S для реализации второй ступени сжатия изображения.

4. Для упрощения сопоставления исходного и восстановленного изображений проводилось восстановление блока размером $N1 \times N1$, т. е. было выбрано $N1 = L$. Затем производилась визуальная оценка качества изображения. На рисунке, $a-d$ показаны результаты преобразований объекта с помощью алгоритмов GDCT и JPEG. Степень сжатия за счет усечения спектров в алгоритмах JPEG и GDCT взята 5,6. Но в алгоритме GDCT проведено дополнительно сжатие первой ступени в 1,8 (b) и в 4 раза (c).

Поскольку алгоритмы сжатия и восстановления на основе чебышевских преобразований обладают свойством масштабирования восстановленных изображений по сравнению с оригиналами, было проведено исследование такой возможности GDCT. В зависимости от степени гладкости исходного изображения коэффициент пространственного увеличения ($L/N1$) может достигать до 20. Исследование преобразований типовых натуральных изображений с резкими границами и мелкомасштабной текстурой показало, что при коэффициенте увеличения ($L/N1$) > 4 начинает проявляться блочная структура изображений.



Результаты преобразований изображения: исходное (a), GDCT 8 → 6 (b), GDCT 12 → 6 (c), JPEG – 5,6 (d)

Заключение. Исследования показали, что изображения с плавными градиентами яркости и цвета могут быть подвергнуты сильному сжатию уже на первой ступени процедуры. Насыщенные мелкими деталями текстуры и натурные изображения испытывают искажения в GDCT, рекомендуемая степень сжатия первой ступени порядка 30–50 %. Скорость убывания спектров в GDCT такая же, как и в классическом дискретном косинусном преобразовании. Поэтому сжатие на второй ступени такого же порядка, как и в алгоритме JPEG. Так как положения узлов полиномов Чебышева образуют дробные числа, то исследовались две реализации GDCT – с округлением положения узлов до ближайшего пиксела и с межпиксельной линейной интерполяцией. Эксперименты показали, что интерполяция улучшает передачу тонких деталей и наклонных краев изображений. Выполнено сравнение в статическом режиме алгоритмов сжатия GDCT и классического JPEG для различных типов изображений. Показано, что для широкого класса изображений алгоритм GDCT в статическом режиме обеспечивает значительно более высокую степень сжатия по сравнению с классическим алгоритмом JPEG, сохраняя при этом одинаковый с ним уровень среднеквадратического отклонения восстановленного изображения от оригинала.

Автор выражает благодарность М. Ю. Радченко за помощь при разработке программного кода алгоритма GDCT на языке C++.

СПИСОК ЛИТЕРАТУРЫ

1. Дворкович А. В., Дворкович В. П., Мохин Б. Н. и др. Единые принципы сжатия цветных динамических изображений различного разрешения // Цифровая обработка сигналов. 1999. № 1. С. 27.
2. Цифровая обработка телевизионных и компьютерных изображений /Под ред. Ю. Б. Зубарсва, В. П. Дворковича. М.: МЦНТИ, 1997.
3. Радченко Ю. С., Радченко М. Ю. Оптимальные быстрые алгоритмы представления изображений в базисе ортогональных полиномов // Тр. I междунар. конф. «Цифровая обработка сигналов и ее применения» DSPA'98. М.: МЦНТИ, 1998. Т. III. С. 163.
4. Радченко Ю. С., Кожин А. Ю., Радченко М. Ю. Обнаружение и оценка параметра сдвига сжатых с помощью ортогональных полиномов сигналов // Радиотехника. 1999. № 6. С. 17.
5. Радченко М. Ю., Радченко Ю. С. Быстрые алгоритмы дискретизации изображений в базисе ортогональных полиномов // Межвуз. сб. науч. тр. «Синтез, передача и прием сигналов управления и связи». Воронеж: ВГТУ, 1998. Вып. 5. С. 33.

*Воронежский государственный университет,
E-mail: rad@sendmail.ru*

*Поступила в редакцию
13 июля 2001 г.*

Подписка на наш журнал – залог Вашего успеха!