ИССЛЕДОВАНИЕ АЛГОРИТМОВ СЖАТИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ПОЛИНОМИАЛЬНЫХ ПРЕОБРАЗОВАНИЙ (АЛГОРИТМ GDCT)

Радченко Ю.С., Савинков А.Ю.

Воронежский государственный университет
394693, Воронеж, пл. Университетская, 1, кафедра радиофизики

Системы связи нового поколения ориентированы на передачу информационных потоков различного вида, то есть являются мультимедийными телекоммуникационными системами.[1,2] В связи с этим передача телевизионных и компьютерных изображений, видеоконференции и видеосотовая связь предполагают применение процедур сжатия изображений. В последнее время выполнен ряд работ, в которых предлагается применять разложения по базису классических ортогональных полиномов Эрмита, Якоби (Лежандра, Чебышева) [3,4]. Эти преобразования 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 I и II рода. Вычисление спектральных коэффициентов производится с помощью квадратурной формулы Гаусса – Чебышева, обладающей наивысшей степенью алгебраической точности при заданном числе отсчетов. Показано, что при этом получаются преобразования, которые являются обобщением дискретного косинусного и синусного преобразований.

1. 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, $\boldsymbol{\tau}=(\tau_x,\ \tau_y)$ параметр сдвига фрагмента в данном кадре. Базисные функции $\varphi_{mk}(x,y)$, используемые для дискретного представления сигнала $s(x,y,\boldsymbol{\tau})$, 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 $s(x,y,\boldsymbol{\tau})$ имеет место пара преобразований

$$s(x,y,\tau)=\sum_m\sum_k C_{mk}(\tau)\varphi_m(x)\varphi_k(y),\quad C_{mk}(\tau)=\int\int_\Omega s(x,y,\tau)\varphi_m(x)\varphi_k(y)dxdy. \qquad (1)$$

Для разложения сигналов и изображений можно применять следующие классические ортогональные многочлены: а) Эрмита; б) Лежандра; в) полиномы Чебышева I и II рода. Если обозначить 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

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

$$\begin{aligned}C_{mk}&=(d_md_k)^{-1}\iint_{\Omega^*}s(a_xz_1,a_yz_2)\rho(z_1)p_m(z_1)\rho(z_2)p_k(z_2)dz_1dz_2=\\&=(d_md_k)^{-1}\int\rho(z_1)p_m(z_1)dz_1\int s(a_xz_1,a_yz_2)\rho(z_2)p_k(z_2)dz_2.\end{aligned} \qquad (2)$$

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 на двумерный случай.

2. 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 $\int f(z)\rho(z)dz=\sum_{n=1}^{N}\lambda_nf(z_n)$. 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 ортогонального с весом $\rho(z)$ полинома $p_m(z)$. Для полиномов Чебышева получается формула Меллера (Гаусса-Чебышева)

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

Здесь $z_n=\cos(\pi(2n+1)/2N)$ –нули полинома Чебышева I рода $T_N(z)$; $\lambda_n=\pi/N$ – то есть все весовые коэффициенты одинаковы, норма полинома Чебышева $d_m=\pi/2$, если $m\neq0$, и $d_m=\pi$, если $m=0$. Учитывая, что $T_m(z_n)=\cos(m\cdot\arccos(z_n))=\cos(\pi m(n+0.5)/N)$, приходим к выражению для прямого и обратного преобразований

@2001, электронная версия подготовлена ЗАО АВТЭКС Санкт-Петербург, http://www.autex.spb.ru

$$C_m = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} s(z_n) cos(\pi m \frac{n+0.5}{N}), \quad C_0 = \sqrt{\frac{1}{N}} \sum_{n=0}^{N-!} s(z_n)$$
$$S_M(z) = g_m \sqrt{\frac{2}{N}} \sum_{m=0}^{M} C_m T_m(z) = g_m \sqrt{\frac{2}{N}} \sum_{m=0}^{M} C_m cos(m \cdot arccos(z)) \quad (4)$$

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) весьма 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 $z_j=\cos(\pi(j+0.5)/L)$, $j=0,\ldots L-1$, то обратное преобразование принимает вид

$$S_M(z_j) = g_m \sqrt{\frac{2}{N}} \sum_{m=0}^{M} C_m cos(\pi m \frac{j+0.5}{L}). \quad (5)$$

Преобразование (5) выглядит аналогично ДКП, однако спектральные коэффициенты отличаются от спектральных коэффициентов обычного ДКП. В том случае, если нас интересует равномерная сетка отсчетов $z_j=2j/(L-1) -1 +\delta$, где δ - некоторый сдвиг, $j=0,\ldots L-1$, то

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

3. Блочные преобразования (GDCT)

Перейдем теперь к двумерному варианту преобразований сигнала s(x,y). Пусть в пределах блока из $N1\times N1$ точек, берутся отсчеты сигнала $N\times N$ штук по закону

$$x_n = ROUND(0.5 \cdot (N1-1) \cdot (1+cos(\pi(n+0.5)/N)))$$
$$y_k = ROUND(0.5 \cdot (N1-1) \cdot (1+cos(\pi(k+0.5)/N))) \quad (7)$$

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

$$\mathbf{C} = \|C_{m,l}\| = \hat{\varphi}\mathbf{S}\hat{\varphi}^T \qquad \mathbf{S} = \Psi^{\mathbf{T}} C \Psi \quad (8)$$

Матрицы прямого и обратного преобразования GDCT имеют вид

$$\hat{\varphi} = \|\varphi_m(n)\|_{NM} = \sqrt{\frac{2}{N}} \begin{bmatrix} \sqrt{0.5} \\ cos(\pi m \frac{(n+0.5)}{N} \end{bmatrix}$$
$$\Psi = \|cos(m \cdot arccos(z_j))\|_{LM} = \left\| cos(m \cdot arccos(\frac{2j}{L-1} - 1 + \delta)) \right\| \quad (9)$$

В частном случае можно полагать $\delta=0$. Дискретизация аргументов {x,y} по правилу (7) 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 четырем ближайшим отсчетам.

4. 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 ступени I 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 II ступени процедуры. При восстановлении изображения формируется блок размером $L\times L$ элементов. При восстановлении размер блока $L\times L$ может быть меньше, равен или больше размеров исходного блока $N1\times N1$.

Для экспериментов был выбран набор искусственных изображений и натурных 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=int(0.299R+ 0.587G +0.114), U=int(R-Y=0.701R-0.587G-0.114B), V= B-Y=int(-0.299R-0.587G+0.886B). 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 **C** для реализации второй ступени сжатия изображения. 4) Для упрощения сопоставления исходного и восстановленного изображений проводилось восстановление блока размером N1×N1, то есть было выбрано N1=L. Затем производилась визуальная оценка качества изображения. На фигурах 1-6 показаны результаты преобразований объекта типа «портрет» с помощью алгоритмов GDCT и JPEG. Исследования показали, что в зависимости от типа изображений алгоритм GDCT в статическом режиме позволяет увеличить степень сжатия до 3 раз по сравнению с алгоритмом JPEG.

Литература

1. А.В. Дворкович, В.П. Дворкович, Б.Н. Мохин и др. Единые принципы сжатия цветных динамических изображений различного разрешения. Цифровая обработка сигналов, №1, 1999, с. 27-35.

2. Цифровая обработка телевизионных и компьютерных изображений./ Под ред. Ю.Б. Зубарева и В.П. Дворковича. М.: МЦНТИ, 1997, - 217с.

3. Ю.С. Радченко, М.Ю. Радченко Оптимальные быстрые алгоритмы представления изображений в базисе ортогональных полиномов. Труды I международной конференции «Цифровая обработка сигналов и ее применения» DSPA'98, 1998, Москва, т. III, с. 163-166.

3. Ю.С. Радченко, А.Ю. Кожин, М.Ю. Радченко Обнаружение и оценка параметра сдвига сжатых с помощью ортогональных полиномов сигналов. Радиотехника, №6, 1999, с.17-19.

RESEARCH OF IMAGE SUPPRESSION ALGORITHMS IN BASIS OF POLYNOMIAL CONVERSIONS (ALGORITHM GDCT)

Radchenko Y.S., Savinkov A.Y.

Physics Department, Voronezh State University,
1 Universitetskaya Sq., Voronezh 394693, Russia

In this paper we introduce the results of the research of one of the variants of polynomial conversions – expansion and synthesis in basis of Chebyshev polynomials of I and II types. It is shown that the formed conversions are the generalization of the sampled cosine and sine conversions.

Suppose that we have s(x,y,τ) field as a fragment of the image in subdomain $\{x,y\}\in\Omega$. In order to expand signals and images we use the following classical polynomials: а) Hermite; б) Legendre; в) Chebyshev polynomials of I and II types. If a_x,a_y – typical sizes of the Ω subdomain, $z_1=x/a_x$, $z_2=y/a_y$, then (1) can be rewritten using the orthogonal polynomial as

$$\begin{aligned} s(x,y) &= \sum_{m,k} C_{mk}\, p_m(x/a_x)\, p_k(y/a_y) \\ C_{mk} &= (d_m d_k)^{-1} \iint_{\Omega^*} s(a_x z_1, a_y z_2)\rho(z_1)p_m(z_1)\rho(z_2)p_k(z_2)\,dz_1 dz_2 = \\ &= (d_m d_k)^{-1} \int \rho(z_1)p_m(z_1)dz_1 \int s(a_x z_1, a_y z_2)\rho(z_2)p_k(z_2)dz_2. \end{aligned} \tag{1}$$

Here d_m – norm of $p_m(z)$ orthogonal polynomial with ρ(z) weight.

Suppose that have quadrature-integrable process f(z) with ρ(z) weight. Then the quadrature formula of Gaussian kind with highest algebraic accuracy about 2N-1 can be written as $\int f(z)\rho(z)dz = \sum_{n=1}^{N} \lambda_n f(z_n)$. Here z_n – zeros of p_N (z) orthogonal polynomials with ρ(z) weight; λ_n – Kristoffel numbers. Nodes and weights $\{z_n\}$, $\{\lambda_n\}$ are uniquely determined by $p_m(z)$ orthogonal polynomial with ρ(z) weight. We have Meller formula (Gaussian-Chebyshev) for Chebyshev polynomials

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

Here $z_n=\cos(\pi(2n+1)/2N)$ –zeros of Chebyshev polynomial of I type $T_N(z)$; $\lambda_n=\pi/N$, norm of Chebyshev polynomial $d_m=\pi/2$, if $m\neq0$, and $d_m=\pi$, if m=0. Taking into account that $T_m(z_n)=\cos(m\cdot\arccos(z_n))=$ $=\cos(\pi m(n+0.5)/N)$, we arrive at expression for the direct and reverse conversions

$$C_m = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} s(z_n) cos(\pi m \frac{n+0.5}{N}), \quad C_0 = \sqrt{\frac{1}{N}} \sum_{n=0}^{N-!} s(z_n)$$
$$S_M(z) = g_m \sqrt{\frac{2}{N}} \sum_{m=0}^{M} C_m T_m(z) = g_m \sqrt{\frac{2}{N}} \sum_{m=0}^{M} C_m cos(m \cdot arccos(z)) \tag{3}$$

Here g_m=1 for m>0, $g_m = \sqrt{0.5}$ for m=0. The expression (4) is similar to the so-called even sampled cosine conversion (SCC). However it has three main differences. 1) $z_n = cos(\pi(n+0.5)/N)$ reference points of s(z) signal are taken irregularly. 2) Synthesis of the $S_M(z)$ signal is performed rather at the arbitrary point $z \in [-1,1]$, than in the sampled set of reference points, as in SCC. 3) The formula accuracy (4) with the conversion of the sufficiently smooth functions is substantially higher than in SCC. Therefore the number of samples can be less.

If we take in conversion the irregular sample grid according to $z_j=cos(\pi(j+0.5)/N)$, then the reverse conversion assumes the form of $S_M(z_j) = g_m \sqrt{2/N} \sum_{m=0}^{M} C_m cos(\pi m(j+0.5)/L)$ (j=0,…L-1), that looks similarly to SCC. However the spectral coefficients differ from the spectral coefficients of common SCC. Provided the we are interested in the regular sample grid $z_j=2j/(L-1)-1+\delta$, where δ - some shift, j=0,…L-1, then

$$S_M(z_j) = g_m \sqrt{\frac{2}{N}} \sum_{m=0}^{M} C_m cos(m \cdot arccos(2j/(L-1)-1+\delta). \tag{4}$$

Let us pass on to the two-dimensional variant of the signal conversion s(x,y). Suppose that in the block the matrix $\mathbf{S} = \|s_{nk}\| = \|s(x_n, y_k)\|$ of samples of N×N size is generated from N1×N1 points according to

$$x_n = ROUND(0.5 \cdot (N1-1) \cdot (1 + cos(\pi(n+0.5)/N)))$$
$$y_k = ROUND(0.5 \cdot (N1-1) \cdot (1 + cos(\pi(k+0.5)/N))). \tag{5}$$

Then **S** is converted into the matrix of C spectral coefficients of M×M size. While the reverse conversion the rectangular matrix of L×M size is used. The recovered block is of L×L size. The direct and reverse Chebyshev conversion (GDCT) of the matrix kind are determined by the following operations

$$\mathbf{C} = \|C_{m,l}\| = \hat{\varphi}\mathbf{S}\hat{\varphi}^T \quad \mathbf{S} = \Psi^{\mathbf{T}} C \Psi \tag{6}$$

Matrixes of direct and reverse conversion GDCT are below

$$\hat{\varphi} = \|\varphi_m(n)\|_{NM} = \sqrt{\frac{2}{N}} \begin{bmatrix} \sqrt{0.5} \\ cos(\pi m \frac{(n+0.5)}{N} \end{bmatrix}$$
$$\Psi = \|cos(m \cdot arccos(z_j))\|_{LM} = \left\|cos(m \cdot arccos(\frac{2j}{L-1} - 1 + \delta))\right\| \tag{7}$$

Thus in the image block of N1×N1 size the matrix of N×N samples is selected (N<N1). In that way the image suppression is performed at the first procedure step. Then with the help of GDCT conversion the matrix M×M is calculated from the most significant spectral coefficients, which are then thinned out using the quantization matrix or the regulated low-frequency flirtation. In that way the image suppression is performed at the second procedure step. While image recovering the block of L×L size is generated. While recovering the block size (L×L) can be less, equal or more than the sizes of the initial block N1×N1.

The set of artificial images and natural photos were selected for the experiments. The colored images were of the 24-bit format BMP. With theses images 8-bit R, G, B color components are code-converted into Y,U,V and the Y,U,V matrixes subsampled according to one of the principles: 4:4:4, 4:2:2, 4:2:0. Depending on the image type at the first step of image suppression the thinning is performed in the ratio of N1/N=16/6, 16/8, 8/6. Then the recovering of **C** spectral matrix was carried out for the implementation of the second step of the image suppression. Further the arbitrary visual estimation of the image quality was performed. The research showed that depending on the image type, GDCT algorithm in the static mode allows us to increase the degree of suppression up to 3 times in comparison with JPEG algorithm.

References

1. Y.S. Radchenko, M.Y. Radchenko Optimal Fast Algorithms for Image Representation in Orthogonal Polynomial Basis. Proc. the I international conference «Digital Signal Processing and its applications» DSPA'98, 1998, Moscow, v, III-E, p. 104-107.