

Воронежский государственный университет, кафедра радиофизики

1. Разложение сигналов по базису ортогональных полиномов

Пусть в подобласти $\{x,y\} \in \Omega$ наблюдается поле $s(x,y,\tau)$, представляющее собой фрагмент изображения. Здесь $\tau=(\tau_x, \tau_y)$ параметры сдвига фрагмента в данном кадре. Процедура сжатия, существенно упрощается, если базисные функции, используемые для дискретного представления сигнала, имеют вид $\varphi_{mk}(x,y) = \varphi_m(x) \cdot \varphi_k(y)$, где $\varphi_m(x)$, $\varphi_k(y)$ – одномерные функции, основанные на ортогональных полиномах. Тогда для полезного сигнала $s(x,y,\tau)$ имеет место пара преобразований

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

$$C_{mk} = \alpha \iint_{\Omega^*} s(a_x z_1, a_y z_2) \rho(z_1) p_m(z_1) \rho(z_2) p_k(z_2) dz_1 dz_2 =$$

$$= \alpha \int \rho(z_1) p_m(z_1) dz_1 \int s(a_x z_1, a_y z_2) \rho(z_2) p_k(z_2) dz_2. \tag{1}$$

В (1) $\alpha=1/(d_m d_k)$, d_m – норма ортогонального с весом $\rho(z)$ полинома $p_m(z)$, a_x, a_y – характерные размеры подобласти Ω , $z_1=x/a_x$, $z_2=y/a_y$. Для последовательного вычисления интегралов в (1)

целесообразно использовать квадратурную формулу гауссовского типа $\int s(z) \rho(z) dz = \sum_{n=1}^N \lambda_n s(z_n)$.

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 полинома $p_m(z)$. Для полиномов Чебышева выражение для прямого и обратного преобразований имеет вид

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

где $g_m = 1$ при $m>0$ и $g_m = 0.5$ при $m=0$. $S_M(z)$ - сигнал, восстановленный по M спектральным коэффициентам. Согласно (2), точки отсчета $z_n = \cos(\pi(n+0.5)/N)$ сигнала $s(z)$ берутся неравномерно. Синтез сигнала $S_M(z)$ выполняется в произвольной точке $z \in [-1,1]$, а не в дискретном наборе точек отсчета, как в ДКП. Точность формулы (2) при преобразовании достаточно гладких функций существенно выше, чем у ДКП. Поэтому число отсчетов можно взять значительно меньше. В том случае, если используется сетка отсчетов $z_j = 2j/(L-1) - 1 + \delta$, где δ - некоторый сдвиг, $j=0, \dots, L-1$, а $L \neq N$ то восстановленное изображение подвергается геометрическому масштабированию [2]. Алгоритм преобразования (2), назовем обобщенным дискретным косинусным преобразованием (GDCT) [1,2].

Если использовать разложение по полиномам Эрмита, (функции параболического цилиндра) $\varphi_m(x) = \exp(-x^2/2a^2) H_m(x/a) / \sqrt{\sqrt{\pi} 2^m m! a}$, то получаем

$$C_m = \frac{1}{\sqrt{a \sqrt{\pi} 2^m m!}} \int_{-\infty}^{\infty} s(x) \exp(-\frac{x^2}{2a^2}) H_m(\frac{x}{a}) dx =$$

$$\sqrt{\frac{2a}{\sqrt{\pi} 2^m m!}} \sum_{k=1}^N \lambda_k s(z_k a \sqrt{2}) H_m(z_k \sqrt{2}) \tag{3}$$

Здесь z_k нули полинома $H_N(z)$, а $\lambda_k = \sqrt{\pi} \cdot 2^{N+1} N! / [H'_N(z_k)]^2$. Разложение (3) целесообразно использовать для обнаружения движения фрагментов изображения и оценки сдвига τ .

В двумерном преобразовании GDCT в пределах блока из $N_1 \times N_1$ точек берутся $N \times N$ отсчетов сигнала по закону

$$x_n = \text{ROUND}(0.5 \cdot (N1 - 1) \cdot (1 + \cos(\pi(n + 0.5) / N))),$$

$$y_k = \text{ROUND}(0.5 \cdot (N1 - 1) \cdot (1 + \cos(\pi(k + 0.5) / N)))$$

При полиномиальных преобразованиях изображения неравномерность отсчетов достигается пропуском части пикселей. 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 $L \times M$.

Возможности предложенного алгоритма были исследованы экспериментально. В блоке изображений размером $N1 \times N1$ элементов выбирается матрица $N \times N$ отсчетов ($N < N1$). В зависимости от типа изображения выбирались следующие значения параметров: $N1/N = 16/8, 12/8, 12/6, 8/6$. Затем с помощью преобразования GDCT вычисляется матрица $M \times M$ из спектральных коэффициентов, которые в дальнейшем преобразуются или при помощи матрицы квантования, или путем регулируемой низкочастотной фильтрации. При восстановлении изображения формируется блок размером $L \times L$ элементов. Отношение размеров восстановленного и исходного блоков выбиралось $L/N1 = 1-2-4-6$ и реализовывалась, таким образом, функция геометрического масштабирования восстановленного изображения. Сжатие изображения в алгоритме GDCT в 1.8-2.5 раза больше, чем в JPEG.

2. Обнаружение изменений фрагментов

Свойство неинвариантности полиномиальных преобразований к сдвигу сигнала можно использовать для построения обнаружителя изменений. Структура алгоритма сжатия должна включать следующие этапы: 1) разбиение изображений на блоки, 2) преобразование фрагмента изображения в дискретную форму в базисе ортогональных полиномов, 3) проверка гипотезы о наличии изменений в блоке (назовем это полезным сигналом), 4) оценка параметров сдвига сигнала. Наилучшие решения поставленных задач целесообразно искать на основе статистической теории решений. Будем считать, что в кадре полезный сигнал наблюдается на фоне аддитивного белого гауссовского шума со спектральной плотностью мощности $N_0/2$. Рассмотрим следующие гипотезы

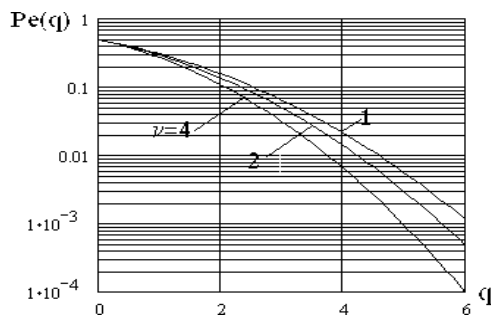
$$H_1: \xi(r) = s_1(r) + \eta(r), \quad H_2: \xi(r) = s_2(r, \tau) + \eta(r). \quad (4)$$

Здесь $r = (x, y)$, $s_1(r) = s(r)$ – сигнал в предыдущем кадре, $s_2(r, \tau) = s(r - \tau)$ – сигнал в текущем кадре, $r \in \Omega$. Правило проверки гипотез (4) при анализе конечного числа мод имеет вид. Подставляя в (4) векторное разложение наблюдаемого поля $\xi(x, y)$ и сигнала $s(x, y, \tau)$ по базисным функциям $\phi_m(x), \phi_k(y)$ получаем структуру логарифма ФОР в виде

$$\max \Lambda(\tau) = \frac{2}{N_0} \sum_{k,m} \left(2X_{km} (C_{km}^{(1)} - C_{km}^{(2)}(\tau)) \right) > \Lambda_0, \quad (5)$$

где $X_{km} = \int \xi(x, y) \phi_m(x) \phi_k(y) dx$, $\Lambda_0 = h_0 + (E_{1v} - E_{2v}) / N_0$ – порог принятия решения, h_0 – выбирается в соответствии с одним из критериев оптимальности, E_{jv} – энергия сигнала $S_{jv}(r)$, $j=1,2$, содержащего v мод. Распределение статистики $\max \Lambda(\tau)$ при $N_0 \rightarrow 0$ в первом приближении можно считать гауссовским [3]. В таком случае вероятность ошибки различения

$$P_e = 1 - \Phi \left(\sqrt{\frac{E_{1v} + E_{2v} - 2\sqrt{E_{1v}E_{2v}}R_v}{2N_0}} \right) = 1 - \Phi \left(\sqrt{\frac{q_1^2 \varepsilon_{1v} + q_2^2 \varepsilon_{2v} - 2q_1q_2 \sqrt{\varepsilon_{1v}\varepsilon_{2v}} R_v}{2}} \right). \quad (6)$$



Здесь $R_v = \int S_{1v}(r) S_{2v}(r) dr / \sqrt{E_{1v} E_{2v}}$ – нормированный коэффициент корреляции двух сигналов, $\varepsilon_{jv} = E_{jv} / E_i$, $\Phi(u)$ – интеграл вероятности. На рисунке приведены зависимости $P_e(q)$, рассчитанные по формуле (6) при разном числе обрабатываемых мод $v=1, 2, 4$. Из расчетов следует, что для обнаружения изменений в блоке вполне достаточно проанализировать

3 Восстановление сжатых сигналов при наличии шумов.

Учет флуктуационных шумов, всегда реально существующих вместе с полезным сигналом, приводит к появлению оптимального числа мод, которые используются для восстановления изображения. Будем считать, что для восстановления используется K мод

разложения $\tilde{S}_k(x) = \sum_{m=0}^{K-1} \tilde{C}_m \phi_m(x)$. Для определения отличия $\tilde{S}_k(x)$ от $S_M(x)$ зададимся

евклидовой метрикой. Тогда

$$\langle \varepsilon^2 \rangle = \left\langle \int_{\Omega} (\tilde{S}_k(x) - S_M(x))^2 dx \right\rangle = \sum_{m=0}^{K-1} D(\tilde{C}_m) + \sum_{m=K}^M C_m^2 = \frac{2}{N_0} \left(K + \sum_{m=K}^M q_m^2 \right). \quad (7)$$

В выражении (7) $q_m^2 = 2C_m^2 / N_0$ - отношение сигнал/шум для моды с индексом m . Ошибка $\langle \varepsilon^2 \rangle$ состоит из двух компонент. Первая – флуктуационная ошибка, связанная с отклонением \tilde{C}_m от C_m , она возрастает с ростом K . Вторая компонента обусловлена энергией не восстановленных мод. Очевидно, что она убывает с ростом K . Если q_{\min}^2 - минимальное, а q_{\max}^2 - максимальное отношение сигнал/шум в остаточном модовом разложении сигнала, то можно записать очевидное неравенство

$$Mq_{\min}^2 + K(1 - q_{\min}^2) < \frac{\langle \varepsilon^2 \rangle N_0}{2} < Mq_{\max}^2 + K(1 - q_{\max}^2). \quad (8)$$

То есть нормированная ошибка $\mu(K) = \langle \varepsilon^2 \rangle N_0 / 2$, как функция переменной K , лежит между двумя мажорирующими функциями $f_1(K) = Mq_{\min}^2 + K(1 - q_{\min}^2)$, $f_2(K) = Mq_{\max}^2 + K(1 - q_{\max}^2)$. Возможны три случая поведения нормированной ошибки $\mu(K)$: а) $q_{\max} < 1$, б) $q_{\min} > 1$ в) $q_{\max} > 1$, $q_{\min} < 1$. В случае а) – функция $\mu(K)$ монотонно возрастающая. В случае б) - функция $\mu(K)$ монотонно уменьшается. Наконец, в случае в) поведение $\mu(K)$ немонотонное. Это легко увидеть, переписав (7) в виде $\mu(K) = \left(M + \sum_{m=K}^M (q_m^2 - 1) \right)$. Из этой формулы следует, что пока $q_m^2 > 1$ слагаемые под знаком суммы положительные, $\mu(K)$ убывает. Но как только q_m^2 станет < 1 , то ошибка $\mu(K)$ начнет возрастать. Следовательно оптимальное число восстанавливаемых мод K определяется соотношением $q_m^2 = 1$. Поэтому дальнейшее наращивание компонент сигнала $\tilde{S}_k(x)$ нецелесообразно.

Литература

1. Ю.С. Радченко Алгоритм сжатия изображений на основе полиномиальных преобразований. Цифровая обработка сигналов, №1, с.2-6.
2. Ю.С. Радченко Метод сжатия и восстановления изображений на основе быстрых чебышевских преобразований. Автометрия. 2002, №4, 32-40.
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.

ALGORITHMS OF CHANGABLE IMAGE SUPPRESSION, PROCESSING AND RECOVERY AT POLYNOMIAL CONVERSIONS

Radchenko Yu.

Voronezh State University, Radio Physic Department

In $\{x,y\} \in \Omega$ sub-domain $s(x,y,\tau)$ field being a fragment of an image is observed. Here $\tau=(\tau_x, \tau_y)$ parameters of the fragment shift in this short. Then for $s(x,y,\tau)$ signal there exists a pair of conversions in the basis of orthogonal polynomials, $\varphi_{mk}(x,y) = \varphi_m(x) \cdot \varphi_k(y)$

$$s(x,y) = \sum_{m,k} C_{mk} p_m(x/a_x) p_k(y/a_y), \quad C_{mk} = \alpha \int \rho(z_1) p_m(z_1) dz_1 \int s(a_x z_1, a_y z_2) \rho(z_2) p_k(z_2) dz_2. \quad (1)$$

In (1) $\alpha=1/(d_m d_k)$, d_m – norm of orthogonal with $\rho(z)$ weight of $p_m(z)$ polynomial, a_x, a_y – typical sizes of Ω sub-domain, $z_1=x/a_x$, $z_2=y/a_y$. For Chebyshev polynomials the conversion algorithm is referred to the generalized discrete cosine conversion (GDCT) [1,2]. In GDCT algorithm the reference points of $s(z)$ signal are taken non-uniformly. Non-uniformity of samples is reached by missing a part of pixels. Synthesis of the signal is performed rather at an arbitrary point, than in the discrete set of reference points as in DCT. The recovered image is easily subjected to the geometrical scaling [1,2].

Hermit polynomial decomposition (functions of parabolic cylinder) is preferred to use for detecting the movement of image fragments and estimating τ shift.

The capabilities of the proposed algorithm were researched experimentally. In the image block of $N_1 \times N_1$ elements size there selects a matrix of $N \times N$ samples ($N < N_1$). Then with the help of GDCT conversion $M \times M$ matrix is calculated from spectral coefficients. While recovering the image a block of $L \times L$ elements size is generated. The ratio of sizes of the recovered and initial blocks was selected as $L/N_1=1-2-4-6$. The image suppression in GDCT algorithm is in 1.8-2.5 times more than that in JPEG.

The feature of non-invariance of polynomial conversions to the signal shift can be utilized for building a changes detector. The structure of the suppression algorithm must include the following steps: 1) divide images into blocks, 2) convert the image fragment into the discrete form in the orthogonal polynomial basis; 3) check the hypothesis concerning the presence of changes in the block (let us call it a desired signal), 4) estimate parameters of the signal shift. Consider the following hypotheses:

$$H_1: \xi(r) = s_1(r) + \eta(r), \quad H_2: \xi(r) = s_2(r, \tau) + \eta(r).$$

Here $r=(x,y)$, $s_1(r) = s(r)$ – signal in the previous short, $s_2(r, \tau) = s(r - \tau)$ – signal in the current short, $r \in \Omega$, $\eta(r)$ – white noise. The rule of checking hypotheses at the analysis of the final number of modes is the following.

$$\max \Lambda(\tau) = \frac{2}{N_0} \sum_{k,m} \left(2X_{km} (C_{km}^{(1)} - C_{km}^{(2)}(\tau)) \right) \underset{<}{>} \Lambda_0, \quad (2)$$

where $X_{km} = \int \xi(x,y) \varphi_m(x) \varphi_k(y) dx$, Λ_0 – threshold of decision-making. From the calculations of the distinguishing errors' probability according to rule (2) follows that for detecting changes in the block it is quite enough to analyze the number of modes less than 4.

Consideration of fluctuation noises that actually always exist along with a desired signal leads to the appearing of optimal number of modes, which are used for the image recovery.

References

1. Yu.S. Radchenko Image suppression algorithm based on polynomial conversions. Digital signal processing, №1, c.2-6.
2. Yu.S. Radchenko Method of image suppression and recovery based on fast Chebyshev conversions. Autometria. 2002, №4, 32-40.