

ОЦЕНКА ВЕКТОРА СДВИГА ФРАГМЕНТОВ ИЗОБРАЖЕНИЙ В АСИНХРОННОМ СПЕКТРАЛЬНОМ БАЗИСЕ

Ю. С. Радченко, С. В. Миляев

Воронежский государственный университет

В работе рассмотрены алгоритмы оценки вектора сдвига фрагментов динамических изображений в спектральной области. Алгоритмы оценки основаны на неинвариантности обобщённых спектров изображений к сдвигу сигнала. В качестве базисных функций выбраны полиномы Чебышева и функции Эрмита. Исследована структура экстремального функционала при анализе различного числа спектральных мод. Предложена процедура нахождения экстремума в виде «спектрального дискриминатора».

КЛЮЧЕВЫЕ СЛОВА: обработка изображений, спектральный анализ сигналов, вектор смещения изображения, алгоритмы оценки сдвига.

ВВЕДЕНИЕ

Современные телекоммуникационные системы ориентированы на передачу мультимедийной информации, наиболее емкой частью которой являются динамические видеоизображения. Без существенного информационного сжатия подобных сигналов практически невозможна передача цифрового видео по каналам связи в реальном времени. Стандарты MPEG4, H264 осуществляют процедуру компрессии видео за счёт использования межкадровой корреляции и предсказания движения фрагментов динамических изображений. Однако существующие подходы к синтезу алгоритмов оценки вектора сдвига фрагментов, основанные на обработке видеопоследовательности в пространственной области, приводят к весьма трудоёмким вычислительным процедурам.

В работах [1, 2] показано, что дискретные преобразования в базисе ортогональных полиномов обладают свойством неинвариантности обобщённых спектров к сдвигу сигналов. Это делает возможным реализацию спектральных алгоритмов оценки вектора сдвига и оценки изменений типа «линейный сдвиг» в изображениях. На основе анализа синхронных обобщённых спектров предложены в [3] алгоритмы оценки вектора сдвига в виде дискриминаторов. Реализация таких алгоритмов требует наличия банка образцов спектров при разных сдвигах.

Данная работа посвящена синтезу и анализу алгоритмов оценки вектора сдвига фрагментов изображений при использовании «асинхронного спектрального базиса». Под асинхронным спектром понимается обобщённый спектр

блока, взятого в произвольном месте изображения. Результаты исследований приведены для случаев дискретных преобразований в базисах полиномов Чебышева и функций Эрмита. Выполнен анализ логарифма функционала отношения правдоподобия при различном числе мод реального изображения.

МОДЕЛЬ СИГНАЛА

Пусть в моменты t и $t + \Delta t$ наблюдается поле $u^{(1)}(r; t)$ и $u^{(2)}(r; t + \Delta t)$, $r = (x, y)$. Некоторые части динамического изображения испытывают преобразование типа «линейный сдвиг». Обозначим через Ω_0 подобласть, занятую анализируемым фрагментом $s(r; t) = u^{(1)}(r; t)I_{\Omega_0}(r)$ в момент t , а через Ω область поиска, занятую тем же фрагментом $s(r - \tau_0; t + \Delta t)$ в момент $t + \Delta t$. Здесь $I_{\Omega_0}(r)$ — индикаторная функция подобласти, $\tau_0 = (\tau_{0x}, \tau_{0y})$ — неизвестный вектор сдвига. В области поиска поле можно представить в виде

$$\begin{aligned} \xi(r; t + \Delta t) &= u^{(2)}(r; t + \Delta t)I_{\Omega}(r) = \\ &= s(r - \tau_0; t) + \eta(r; t + \Delta t). \end{aligned} \quad (1)$$

В зависимости от постановки задачи $\eta(r; t + \Delta t)$ может представлять из себя: а) белый пространственный шум, б) сумму белого пространственного шума и неподвижного фона. В дальнейшем дискретный аргумент t будет в выражениях опущен. Для полезного сигнала $s(r) \equiv s(x, y)$ имеет место пара преобразований

$$\begin{aligned} s(x, y) &= \sum_m \sum_k C_{mk} \varphi_m(x) \varphi_k(y), \\ C_{mk} &= \iint_{\Omega_0} s(x, y) \varphi_m(x) \varphi_k(y) dx dy. \end{aligned} \quad (2)$$

Здесь $\varphi_k(x)$, $\varphi_m(y)$ — ортогональные базисные функции. Если обозначить как a_x, a_y — характерные размеры подобласти Ω_0 , и ввести нормированные переменные $z_1 = x/a_x$, $z_2 = y/a_y$, то (2) можно переписать с использованием ортогональных полиномов в виде

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

$$C_{mk} =$$

$$= \alpha \iint_{\Omega_0} s(a_x z_1, a_y z_2) \rho(z_1) p_m(z_1) \rho(z_2) p_k(z_2) dz_1 dz_2 =$$

$$= \alpha \int \rho(z_1) p_m(z_1) dz_1 \int s(a_x z_1, a_y z_2) \rho(z_2) p_k(z_2) dz_2. \quad (3)$$

Здесь $\alpha = 1/(d_k d_m)$, d_m — норма ортогонального с весом $\rho(z)$ полинома $p_m(z)$. Поскольку для разложимых базисных функций $\varphi_{km}(x, y) = \varphi_k(x) \varphi_m(y)$ вычисление спектральных коэффициентов производится последовательным интегрированием по координатам x и y , то отметим особенности одномерных преобразований, а затем обобщим их на двумерный случай.

Для вычисления интегралов в (2), (3) целесообразно использовать квадратурную формулу гауссовского типа $\int s(z) \rho(z) dz = \sum_{n=1}^N \lambda_n s(z_n)$, обладающую наивысшей алгебраической степенью точности порядка $2N - 1$. 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 $p_N(z)$.

Для полиномов Чебышева I рода получается формула Меллера (Гаусса—Чебышева)

$$\frac{1}{d_m} \int_{-1}^1 \frac{s(z) T_m(z)}{\sqrt{1-z^2}} dz = \frac{\pi}{d_m N} \sum_{n=0}^{N-1} s(z_n) T_m(z_n). \quad (4)$$

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 $m \neq 0$, 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

$$C_m = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} s(z_n) \cos\left(\pi m \frac{n+0.5}{N}\right),$$

$$C_0 = \sqrt{\frac{1}{N}} \sum_{n=0}^{N-1} s(z_n),$$

$$S_M(z) = g_m \sqrt{\frac{2}{N}} \sum_{m=0}^M C_m \cos(m \cdot \arccos(z)). \quad (5)$$

Здесь $g_m = 1$ при $m > 0$ и $g_m = 0.5$ при $m = 0$. $S_M(z)$ — сигнал, восстановленный по M спектральным коэффициентам. Алгоритм преобразования (5) называется обобщенным дискретным косинусным преобразованием (GDCT) [1, 2].

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

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

$$= \sqrt{\frac{2a}{\sqrt{\pi} 2^m m!}} \sum_{k=1}^N \lambda_n s(z_n a \sqrt{2}) H_m(z_n \sqrt{2}). \quad (6)$$

В случае формулы Гаусса—Эрмита z_n нули полинома $H_N(z)$, а $\lambda_n = \sqrt{\pi} \cdot 2^{N+1} N! / [H'_N(z_n)]^2$.

ЭКСТРЕМАЛЬНЫЙ ФУНКЦИОНАЛ

Для нахождения неизвестного вектора сдвига может быть сформирован один из экстремальных функционалов, например, вида

$$\Lambda(\tau, \tau_0) = \int_{\Omega_0(\tau)} [\xi(r; \tau_0) - s(r - \tau)]^2 dr =$$

$$= \int_{\Omega_0} [\xi(r - \tau; \tau_0) - s(r)]^2 dr. \quad (7)$$

При переходе к спектральному описанию экстремальный функционал (7) может быть представлен в виде

$$\Lambda_M(\tau, \tau_0) = \sum_{k,m=0}^M [X_{km}(\tau; \tau_0) - C_{km}]^2, \quad (8)$$

где $X_{mk}(\tau; \tau_0) = \int_{\Omega_0} \xi(x, y; \tau; \tau_0) \varphi_m(x) \varphi_k(y) dx dy$.

Отличие экстремального функционала в спектральном представлении (8) от функционала (7) состоит в конечности суммы. Поэтому открывается возможность варьировать число слагаемых M и ускорять вычислительный процесс по оценке вектора сдвига.

В том случае, когда поле $\eta(r)$ — белый пространственный шум, экстремальные функционалы $\Lambda(\tau, \tau_0)$, $\Lambda_M(\tau, \tau_0)$ представляют собой логарифмы функционала правдоподобия неизвестного параметра τ , и полученные оценки являются оптимальными.

Для нахождения оценки параметра τ необходимо исследовать структуру экстремальных функционалов (7) и (8).

На рис. 1 приведен кадр видеопоследовательности «певица». На данном рисунке выде-



Рис. 1

лены типичные блоки динамического изображения — изменяющиеся и малоподвижные, которые были подвергнуты обобщенному спектральному анализу в базисе полиномов Чебышева. Для данных блоков были экспериментально построены экстремальный функционал в пространственной и ЛФОП спектральной области.

На рис. 2 приведен экстремальный функционал (7) фрагмента «глаз», построенный по изображению в пространственной области.

Рисунки 3а, 3б соответствуют функционалу (8), сформированному с учетом ограниченного числа мод $M=3, 4$ для фрагмента «глаз». Для удобства отображения произведена нормировка.

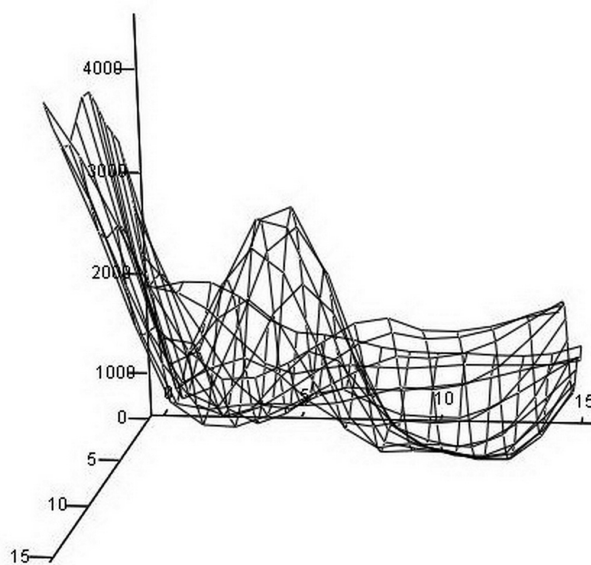


Рис. 2

Алгоритмы оценки. Решение экстремальной задачи по нахождению абсолютного минимума функционалов $\Lambda(\tau, \tau_0)$, $\Lambda_M(\tau, \tau_0)$ может быть выполнено методами безусловной оптимизации [4]. Популярным алгоритмом для целей поиска экстремума является алгоритм Нелдера—Мида с деформируемым симплексом [4]. При этом базовый вариант алгоритма нуждается в доработке, поскольку для цифровых изображений экстремальный функционал задается на сетке параметра τ . Широко применяемым на практике является подход, основанный на сеточных шаблонах. Анализируется функционал на шаблоне,

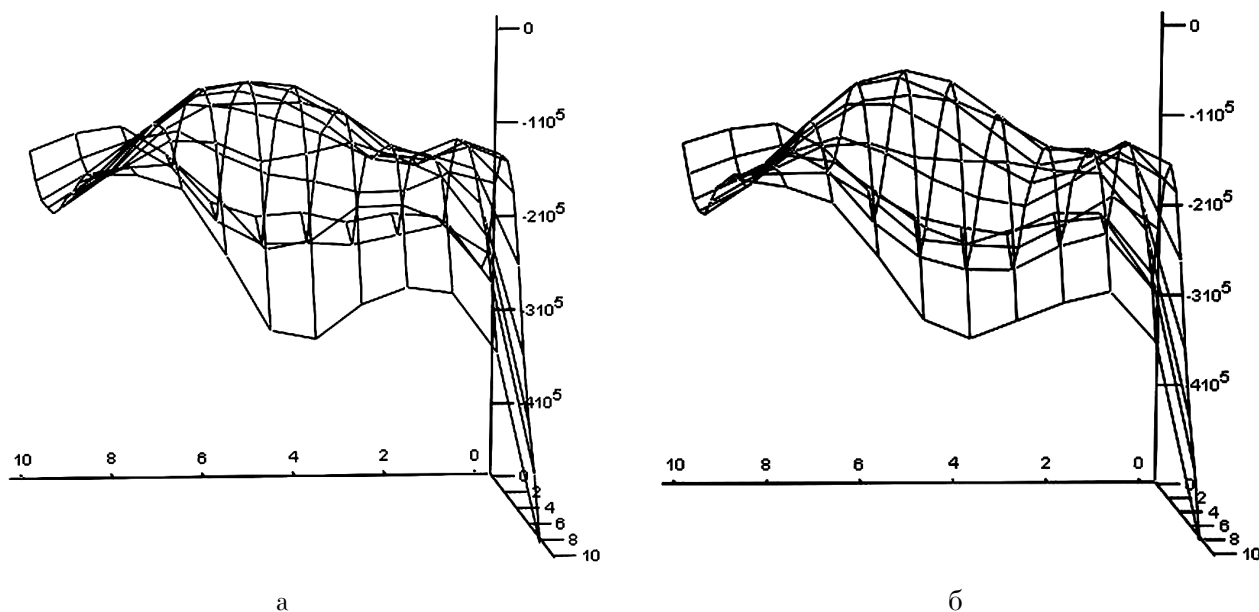


Рис. 3. ЛФОП: а — для 3 мод; б — для 4 мод

взятом с большим шагом. Затем шаг шаблона уменьшается, а его центр смещается в точку минимума функционала, и так далее. Наибольшее применение получил 3-шаговый алгоритм — TSS (Tree-Step Search) с 9-точечным шаблоном. Применяются и другие стратегии поиска экстремума (Cross search, orthogonal search, TSS, FSS, spiral search, hierarchical search) [5].

Применение спектрального описания открывает новые возможности в нахождении экстремума функционала $\Lambda_M(\tau, \tau_0)$, поскольку для ускорения вычислений можно оперировать с ограниченным числом спектральных мод на начальном этапе поиска экстремума, и увеличить их число на конечном этапе, когда необходима высокая точность оценки положения экстремума функционала.

Большой интерес представляет применение быстродействующих градиентных и ньютоновских методов. Положение минимума выражения (7) определяется из системы уравнений

$$\begin{aligned} \left. \left(\partial \Lambda(\vec{\tau}, \vec{\tau}_0) / \partial \tau_x \right) \right|_{\vec{\tau}=\vec{\tau}_m} &= 0, \\ \left. \left(\partial \Lambda(\vec{\tau}, \vec{\tau}_0) / \partial \tau_y \right) \right|_{\vec{\tau}=\vec{\tau}_m} &= 0. \end{aligned} \quad (10)$$

Дифференцируя логарифм функционала отношения правдоподобия (3) по параметру τ получаем систему уравнений правдоподобия

$$\begin{aligned} \sum_{m,k=0}^{M-1} (X_{mk}(\tau) - C_{mk}) \frac{\partial X_{mk}(\tau)}{\partial \tau_x} &= 0, \\ \sum_{m,k=0}^{M-1} (X_{mk}(\tau) - C_{mk}) \frac{\partial X_{mk}(\tau)}{\partial \tau_y} &= 0. \end{aligned} \quad (11)$$

Используя рекуррентные свойства ортогональных полиномов и тот факт, что производные ортогональных полиномов выражаются через аналогичные полиномы других порядков, выражение для производных $\partial X_{mk}(\tau) / \partial \tau_x$, $\partial X_{mk}(\tau) / \partial \tau_y$ можно переписать через спектральные коэффициенты других порядков. Так, для полиномов Эрмита выполняются соотношения

$$\begin{aligned} \frac{dH_k(z)}{dz} &= 2kH_{k-1}(z), \\ 2zH_k(z) &= H_{k+1}(z) + 2kH_{k-1}(z). \end{aligned}$$

Отсюда получаем следующую связь производных базисных функций

$$\begin{aligned} \frac{d\varphi_k(z)}{dz} &= \frac{1}{\sqrt{2}} \left(\sqrt{k}\varphi_{k-1}(z) - \sqrt{k+1}\varphi_{k+1}(z) \right), \quad k > 0, \\ \frac{d\varphi_0}{dz} &= -\varphi_1(z) / \sqrt{2}. \end{aligned}$$

Тогда

$$\begin{aligned} \partial X_{mk}(\tau) / \partial \tau_x &= \sqrt{m}X_{m-1,k}(\tau) - \sqrt{m+1}X_{m+1,k}(\tau), \\ \partial X_{mk}(\tau) / \partial \tau_y &= \sqrt{k}X_{m,k-1}(\tau) - \sqrt{k+1}X_{m,k+1}(\tau). \end{aligned} \quad (12)$$

Подставляя (12) в (11), получаем нелинейные уравнения правдоподобия для оцениваемых параметров.

$$\begin{aligned} \omega_x(\tau) &= \sum_{m,k=0}^{M-1} (X_{mk}(\tau) - C_{mk}) \times \\ &\times \left(\sqrt{m}X_{m-1,k}(\tau) - \sqrt{m+1}X_{m+1,k}(\tau) \right) = 0, \end{aligned} \quad (13)$$

$$\begin{aligned} \omega_y(\tau) &= \sum_{m,k=0}^{M-1} (X_{mk}(\tau) - C_{mk}) \times \\ &\times \left(\sqrt{k}X_{m,k-1}(\tau) - \sqrt{k+1}X_{m,k+1}(\tau) \right) = 0. \end{aligned} \quad (14)$$

Перегруппировывая слагаемые в суммах (13), (14), получаем более наглядную форму уравнения для оценки компонент вектора сдвига

$$\begin{aligned} \omega_x(\tau) &= \sum_{m,k=0}^{M-1} \sqrt{m+1} \times \\ &\times \left(X_{m+1,k}(\tau)C_{mk} - X_{m,k}(\tau)C_{m+1,k} \right) = 0, \end{aligned} \quad (15)$$

$$\begin{aligned} \omega_y(\tau) &= \sum_{m,k=0}^{M-1} \sqrt{k+1} \times \\ &\times \left(X_{m,k+1}(\tau)C_{mk} - X_{m,k}(\tau)C_{m,k+1} \right) = 0. \end{aligned} \quad (16)$$

Если C_{mk} рассматривать как спектральные коэффициенты в опорном кадре, а $X_{mk}(\tau)$ — в анализируемом кадре, то видим, что в алгоритме измерения сдвига τ вычисляются разности перекрестных произведений спектральных коэффициентов в двух кадрах. При отсутствии шумов при $\tau = \tau_0$ $\omega_x = 0$, $\omega_y = 0$. Функции

$$\begin{aligned} \omega_x(\tau) &= \partial \Lambda(\tau, \tau_0) / \partial \tau_x \equiv d_x(\tau, \tau_0), \\ \omega_y(\tau) &= \partial \Lambda(\tau, \tau_0) / \partial \tau_y \equiv d_y(\tau, \tau_0) \end{aligned}$$

называются дискриминационными характеристиками измерительного устройства. На рис. 4 приведен вид дискриминационной кривой в зависимости от числа анализируемых мод. Анализ дискриминационной характеристики показывает целесообразность использования для оценки параметров 3—5 мод разложения.

Решение систем уравнений (13), (14) или (15), (16) может быть получено с помощью метода Ньютона—Рафсона

$$\tau^{(i+1)} = \tau^{(i)} - H^{-1}(\tau^{(i)})W(\tau^{(i)}), \quad i = 0, 1, \dots, \quad (17)$$

где $H(\tau)$ — матрица Якоби, а $W(\tau)$ — вектор столбец функций

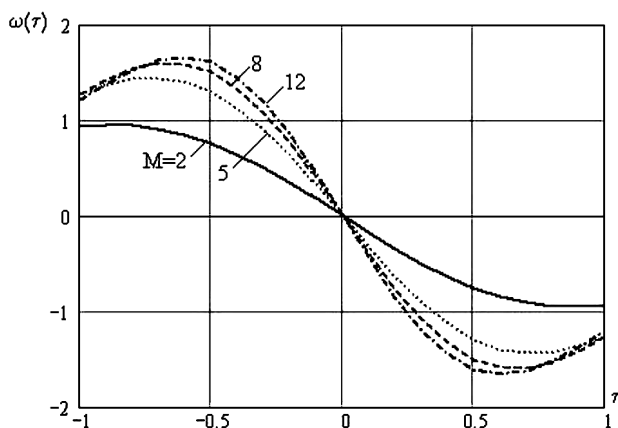


Рис. 4

$$H(\tau) = \begin{pmatrix} \partial \omega_x(\tau) / \partial \tau_x & \partial \omega_x(\tau) / \partial \tau_y \\ \partial \omega_y(\tau) / \partial \tau_x & \partial \omega_y(\tau) / \partial \tau_y \end{pmatrix},$$

$$W(\tau) = (\omega_x(\tau), \omega_y(\tau))^T.$$

Число итераций i для нахождения оценки максимального правдоподобия зависит от начальной расстройки $\tau_0 - \tau^{(0)}$ и формы фрагмента.

S-образный вид дискриминационной характеристики $\omega(\tau)$, показывает, что алгоритм Ньютона—Рафсона недостаточно эффективен для нахождения оценки из-за узкого диапазона расстроек, большого числа необходимых итераций и даже опасности «зацикливания» итерационного процесса. Гораздо лучшими возможностями обладает его модификация с подвижным полюсом [6]. Одномерный вариант этого алгоритма имеет вид

$$\tau^{(i+1)} = \tau^{(i)} - \frac{\omega(\tau^{(i)})}{\omega'(\tau^{(i)}) - \frac{\omega(\tau^{(i)})}{g_i - \tau^{(i)}}};$$

$$g_i = \tau^{(i)} + \frac{\omega(\tau^{(i)}) - 2}{\omega'(\tau^{(i)})}, i = 0, 1.$$

Начальное значение расстройки $\tau^{(0)}$ может выбираться на основе предсказания по нескольким предыдущим кадрам. Так, при про-

гнозе по предыдущему кадру в качестве $\tau^{(0)}$ берется положение блока в этом кадре. При прогнозе второго порядка по двум предыдущим кадрам используется линейная экстраполяция движения фрагмента. В таблице приведены ошибки итерационного процесса (18) в зависимости от числа мод M и начальной расстройки $\Delta = \tau_0 - \tau^{(0)}$.

В данной таблице величина смещения или ошибки 0,1 соответствует 1 пикселю дискретного изображения. Итак, из таблицы следует, что за одну итерацию получается ошибка много меньше одного пикселя. Данные из таблицы также говорят, что брать число мод $M > 5$ нецелесообразно, поскольку диапазон начальных расстроек Δ уменьшается.

Уравнения, аналогичные (13) и (14), получаются и при использовании полиномов Чебышева I рода. Для таких полиномов

$$\partial X_{mk}(\tau) / \partial \tau_x = -2g_m \sum_{l=0}^{\lfloor \frac{v+1-m}{2} \rfloor} (m+2l+1) X_{m+2l+1,k}(\tau).$$

Здесь v — порядок высшей производной $s(x, y, \tau)$ по параметру τ .

ВЫВОДЫ

Таким образом, измерение вектора сдвига фрагмента изображения может быть реализовано в спектральной области при использовании «асинхронного базиса». Предложена реализация дискриминаторов сдвига в спектральной области, которая более предпочтительна из-за меньшего объема вычислений. Выполнен анализ поверхности экстремального функционала при обработке ограниченного числа спектральных мод и определена область существования одного экстремума. Предложена процедура нахождения экстремума в виде спектрального дискриминатора. Анализ спектрального дискриминатора показал, что достаточно обработки нескольких наиболее мощных спектральных мод для достижения предельной точности.

Таблица

	Δ	i			
		1	2	3	4
M = 5	0,5	$-4 \cdot 10^{-3}$	$-9,4 \cdot 10^{-4}$	$-2,1 \cdot 10^{-4}$	
	0,6	-0,047	-0,013	$-3,2 \cdot 10^{-3}$	$-7,3 \cdot 10^{-4}$
	0,7	-0,208	-0,095	-0,032	$-8,5 \cdot 10^{-3}$
M = 8	0,3	$6,5 \cdot 10^{-3}$	$1,5 \cdot 10^{-3}$	$3,6 \cdot 10^{-4}$	
	0,4	$-3,9 \cdot 10^{-4}$	$-9,3 \cdot 10^{-5}$		
	0,5	-0,043	-0,013	$-3,4 \cdot 10^{-3}$	$-8,3 \cdot 10^{-4}$

ЛИТЕРАТУРА

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

2. Радченко Ю. С. Метод сжатия и восстановления изображений на основе быстрых чебышевских преобразований / Ю. С. Радченко // Автометрия. — 2002. — № 4. — С. 32—40.

3. Радченко Ю. С. Синтез и анализ дискриминаторов сдвига динамических фрагментов изображе-

ний в пространственной и спектральной области / Ю. С. Радченко, Е. В. Овчинников // Изв. вузов. Радиоэлектроника. — 2006. — № 9. — С. 61—69.

4. Химмельблау Д. Прикладное нелинейное программирование / Д. Химмельблау. — М. : Мир, 1975. — 850 с.

5. Ричардсон Я. Видеокодирование. H.264 и MPEG-4 — стандарты нового поколения / пер. с англ. — М. : Техносфера, 2005. — 368 с.

6. Вержбицкий В. М. Основы численных методов / В. М. Вержбицкий. — М. : Высшая школа, 2002. — 840 с.