



# ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Научно-технический журнал  
№ 2/2006

Издается с 1999 года  
Выходит четыре раза в год

**ГЛАВНЫЙ РЕДАКТОР**  
Ю.Б. ЗУБАРЕВ

**ЗАМЕСТИТЕЛИ ГЛАВНОГО РЕДАКТОРА:**  
В.В. ВИТЯЗЕВ, В.П. ДВОРКОВИЧ

## РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

Б.А. Бабаян, Ю.И. Борисов,  
С.А. Грибачев, Г.В. Зайцев, Р.В. Зубарев,  
А.П. Кирпичников, М.И. Кривошеев,  
Н.А. Кузнецов, М.С. Куприянов, А.А. Ланнэ,  
В.Г. Мистюков, С.Л. Мишенков,  
А.А.Петровский, Ю.Н. Прохоров,  
А.Н. Соловьев, Ю.Г. Сосулин,  
В.В. Шахгильдян, Ю.С. Шинаков

## Адрес редакции:

107241 Москва, Щелковское ш., 23А  
Тел.: (495) 290-9088  
Факс: (495) 290-9085  
E-mail: editor@dspa.ru  
http://www.dspa.ru

## Для писем:

129090 Москва, а/я 48.

## Издатель:

КБ волоконно-оптических приборов  
Ответственный редактор: С.А. Задворнов  
Компьютерная верстка: В.В. Андряхина  
Дизайн: И.Е. Артюхина

Подписной индекс по каталогу  
ОАО «Роспечать» – **82185**

Подписано в печать 29.08.06  
Формат 60x90/8.

Гарнитура «Arial». Печать ризографическая.  
Бумага офсетная. Печ.л. 7. Тираж 700 экз.

Заказ № 4049. Отпечатано в  
ООО НПЦ «Информационные технологии»  
Рязань, ул. Островского, д. 21/1  
тел.: (4912) 98-69-84

Издание зарегистрировано в Министерстве  
Российской Федерации по делам печати,  
телерадиовещания и средств  
массовых коммуникаций.  
Свидетельство о регистрации ПИ № 77-1488  
от 14.01.2000

## УЧРЕДИТЕЛИ:

**ЗАО «Инструментальные системы»**  
**ФГУП «НИИ радио»**

## В НОМЕРЕ:

- Дворкович В.П., Дворкович А.В.*  
**Расчет банков фильтров дискретного вейвлет-преобразования и анализ их характеристик..... 2**
- Умняшкин С.В., Коплович Д.М., Черкасов И.В.*  
**Об использовании контекстного векторного квантования в области дискретных вейвлет-преобразований для компрессии изображений ..... 11**
- Авдеев О.В., Чобану М.К.*  
**Сжатие изображений с помощью частичной сортировки вейвлет-коэффициентов ..... 15**
- Кобелев В.Ю., Приоров А.Л.*  
**Применение неразделимых оптимизированных вейвлет-фильтров в задачах сжатия изображений ... 21**
- Арляпов С.А., Приоров А.Л., Хрящев В.В.*  
**Модифицированный критерий оценки качества восстановленных изображений ..... 27**
- Касаткин А.С., Шанцын Е.А., Мусихин В.В.*  
**Фрактальное кодирование цветных изображений и полутонных видеопоследовательностей ..... 34**
- Аксенов О.Ю.*  
**Обнаружение объектов на изображениях при изменяющихся условиях наблюдения..... 40**
- Алпатов Б.А., Бабаян П.В.*  
**Методы обработки и анализа изображений в бортовых системах обнаружения и сопровождения объектов ..... 45**
- Аринин В. А.*  
**Функциональный метод трассировки границ раздела двух сред, имеющих цилиндрическую геометрию ..... 52**

Подписной индекс по каталогу  
ОАО «Роспечать» – **82185**

## РАСЧЕТ БАНКОВ ФИЛЬТРОВ ДИСКРЕТНОГО ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ И АНАЛИЗ ИХ ХАРАКТЕРИСТИК

Дворкович В.П., Дворкович А.В.

Несмотря на то, что новое научное направление, связанное с разработкой непрерывных и дискретных вейвлет-преобразований, имеет незначительную временную историю [1-4], эти преобразования нашли широкое применение, в частности, при решении проблем обработки и сжатия статических и динамических изображений [5-8].

При дискретном вейвлет-преобразовании цифровых нестационарных сигналов чаще всего используют так называемые двухполосные фильтры, которые делят сигнал на два уровня – низкочастотную и высокочастотную составляющие. Характеристики этих фильтров обозначены на рис. 1 системными функциями  $H(z)$  и  $G(z)$  соответственно. Далее сигналы прореживаются в два раза и через блок памяти (или кодирования и передачи по каналам связи) поступают на схему синтеза. При восстановлении исходного сигнала между прореженными отсчетами добавляются нулевые отсчеты, и полученные таким образом дискретные отсчеты подаются на восстанавливающие фильтры, характеристики которых обозначены через  $Kh(z)$  и  $Kg(z)$  соответственно. Если отсчеты сигналов при кодировании дополнительно не квантуются, то после суммирования сигналов, возникающих на выходах фильтров  $Kh(z)$  и  $Kg(z)$ , исходный сигнал восстанавливается.

Приводятся расчеты параметров дискретных вейвлет-преобразований путем решения систем нелинейных уравнений и анализируются формы частотных характеристик соответствующих фильтров.

При восстановлении исходного сигнала между прореженными отсчетами добавляются нулевые отсчеты, и полученные таким образом дискретные отсчеты подаются на восстанавливающие фильтры, характеристики которых обозначены через  $Kh(z)$  и  $Kg(z)$  соответственно. Если отсчеты сигналов при кодировании дополнительно не квантуются, то после суммирования сигналов, возникающих на выходах фильтров  $Kh(z)$  и  $Kg(z)$ , исходный сигнал восстанавливается.

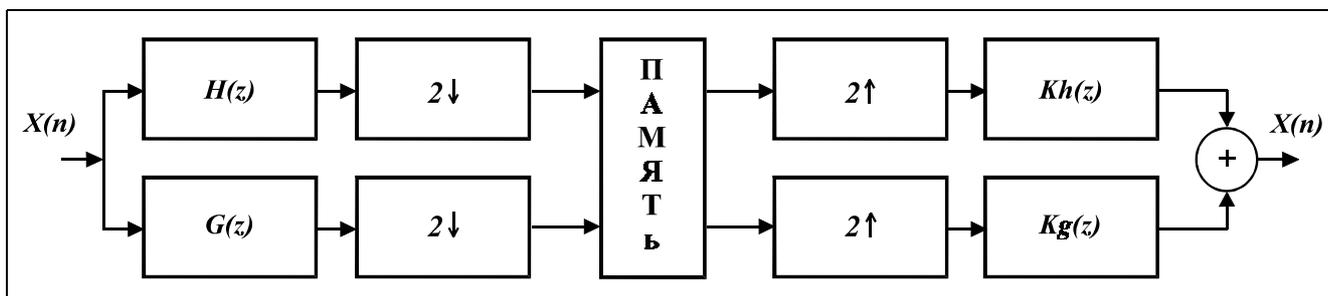


Рис. 1. Структурная схема двухполосного банка фильтров.

Пусть  $Z$ -преобразования двухполосных фильтров определяются соотношениями:

$$H(z) = \sum_{n=-N_1}^{N_2} h_n \cdot Z^{-n}, \quad G(z) = \sum_{n=-M_1}^{M_2} g_m \cdot Z^{-m}, \quad (1)$$

где  $Z = \exp(i\pi x)$ ,  $x = f \cdot T / 2$ ,  $0 \leq x \leq 1$ ,  $T$  – интервал дискретизации.

Для того, чтобы фильтр с системной функцией  $H(z)$  обеспечивал выделение низкочастотной составляющей сигнала (ФНЧ), необходимо выполнение условий:

$$H(0) = \sum_{n=-N_1}^{N_2} h_n = Const \text{ и } H(1) = \sum_{n=-N_1}^{N_2} (-1)^n \cdot h_n = 0. \quad (2)$$

Фильтр с системной функцией  $G(z)$  выделяет высокочастотную составляющую (ФВЧ) при

$$G(0) = \sum_{n=-M_1}^{M_2} g_m = 0 \text{ и } G(1) = \sum_{n=-M_1}^{M_2} (-1)^m \cdot g_m = Const. \quad (3)$$

Если на вход банка фильтров поступает один дискретный отсчет сигнала, изображенный на рис. 2а, то на выходах ФНЧ и ФВЧ формируются сигналы, показанные на рис. 2б и рис. 2в, 2г соответственно

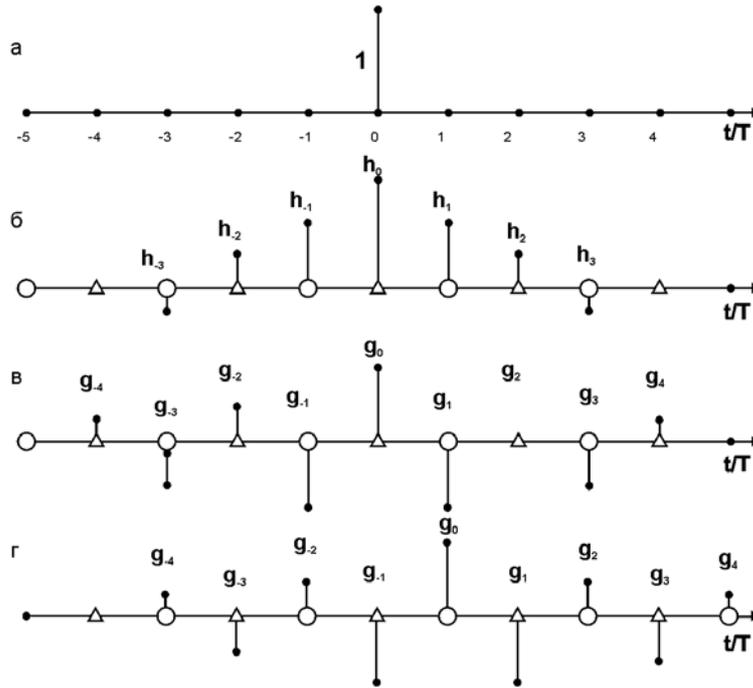


Рис. 2. Реакция низкочастотного (б) и высокочастотного (в, г) фильтров на дискретный отсчет входного сигнала (а).

В зависимости от того, каким образом осуществляется прореживание низкочастотного сигнала, после него выделяется сигнал, отмеченный на рис. 2б либо фигурой  $\Delta$ , либо фигурой  $\bigcirc$ :

$$H_{\Delta}(z) = \sum_{n=[-N_1/2]}^{[N_2/2]} h_{2n} \cdot Z^{-2n} \text{ или}$$

$$H_{\bigcirc}(z) = \sum_{n=[(-N_1+1)/2]}^{[(N_2+1)/2]} h_{2n-1} \cdot Z^{-(2n-1)}. \quad (4)$$

Если прореживание высокочастотного сигнала осуществляется синфазно, как показано на рис. 2в, то выделяются сигналы

$$G_{\Delta}(z) = \sum_{n=[-M_1/2]}^{[M_2/2]} g_{2n} \cdot Z^{-2n} \text{ или}$$

$$G_{\bigcirc}(z) = \sum_{n=[(-M_1+1)/2]}^{[(M_2+1)/2]} g_{2n-1} \cdot Z^{-(2n-1)}. \quad (5)$$

В этом случае входной сигнал будет восстановлен при расчете характеристик восстанавливающих фильтров  $Kh(z)$  и  $Kg(z)$  в соответствии с системой:

$$\begin{cases} H_{\Delta}(z) \cdot Kh(z) + G_{\Delta}(z) \cdot Kg(z) = 1, \\ H_{\bigcirc}(z) \cdot Kh(z) + G_{\bigcirc}(z) \cdot Kg(z) = 1. \end{cases} \quad (6)$$

В случае же, если прореживание высокочастотного сигнала смещено относительно прореживания низкочастотного сигнала на интервал дискретизации, как показано на рис. 2г, то восстановление сигнала будет осуществлено при выполнении условий:

$$\begin{cases} H_{\Delta}(z) \cdot Kh(z) + G_{\bigcirc}(z) \cdot Kg(z) = 1, \\ H_{\bigcirc}(z) \cdot Kh(z) + G_{\Delta}(z) \cdot Kg(z) = 1. \end{cases} \quad (7)$$

Вычитая и суммируя уравнения в соотношениях (6) и (7), можно получить стандартные соотношения [5]:

$$\begin{cases} \bar{H}(z) \cdot Kh(z) \pm \bar{G}(z) \cdot Kg(z) = 0, \\ H(z) \cdot Kh(z) + G(z) \cdot Kg(z) = 2, \end{cases} \quad (8)$$

где  $\bar{H}(z) = \sum_{n=-N_1}^{N_2} (-1)^n \cdot h_n \cdot Z^{-n}$ ,

$$\bar{G}(z) = \sum_{n=-M_1}^{M_2} (-1)^m \cdot g_m \cdot Z^{-m}.$$

Знак «+» в первом уравнении (8) соответствует соотношениям (6), а знак «-» – соотношениям (7).

Используя полученные соотношения, рассчитаем ряд банков КИХ-фильтров с нечетным и четным числом отсчетов цифровой решетки.

**А. КИХ-фильтры с нечетным числом отсчетов цифровой решетки и линейной (нулевой) ФЧХ**

Частотные характеристики таких фильтров, учитывая, что  $h_{-n} = h_n$  и  $g_{-n} = g_n$ , можно представить в виде:

$$H(x) = h_0 + 2 \sum_{n=1}^N h_n \cdot \cos nx, \quad G(x) = g_0 + 2 \sum_{n=1}^M g_m \cdot \cos mx, \quad 0 \leq x \leq 1,$$

$$\bar{H}(x) = h_0 + 2 \sum_{n=1}^N (-1)^n \cdot h_n \cdot \cos nx, \quad \bar{G}(x) = g_0 + 2 \sum_{n=1}^M (-1)^m \cdot g_m \cdot \cos mx, \quad (9)$$

$$H(0) = Const, \quad H(1) = 0, \quad G(1) = Const, \quad G(0) = 0.$$

Можно показать, что в данном случае восстановление сигнала возможно также с помощью КИХ-фильтров, если используется система уравнений:

$$\begin{cases} \bar{H}(x) \cdot Kh(x) - \bar{G}(x) \cdot Kg(x) = 0, \\ H(x) \cdot Kh(x) + G(x) \cdot Kg(x) = 2, \end{cases} \quad (10)$$

и число  $N + M = 2K + 1$  – нечетное.

Детерминант системы уравнений (10) равен:

$$\begin{aligned} \det(x) &= [h_0 + 2 \sum_{n=1}^N h_n \cdot \cos \pi n x] \cdot [g_0 + 2 \sum_{m=1}^M (-1)^m \cdot g_m \cdot \cos \pi m x] + \\ & [h_0 + 2 \sum_{n=1}^N (-1)^n \cdot h_n \cdot \cos \pi n x] \cdot [g_0 + 2 \sum_{m=1}^M g_m \cdot \cos \pi m x] = \\ & 2 \cdot \{h_0 \cdot g_0 + h_0 \sum_{k=1}^M [1 + (-1)^k] \cdot g_k \cdot \cos \pi k x + g_0 \sum_{k=1}^N [1 + (-1)^k] \cdot h_k \cdot \cos \pi k x + \\ & \sum_{n=1}^N \sum_{k=n-1}^{n-M} (-1)^n \cdot h_n \cdot g_{n-k} \cdot [1 + (-1)^k] \cdot \cos \pi k x + \sum_{n=1}^N \sum_{k=n+1}^{n+M} (-1)^n \cdot h_n \cdot g_{k-n} \cdot [1 + (-1)^k] \cdot \cos \pi k x\} = \\ & A_0 + \sum_{\forall k} A_k \cdot \cos \pi k x \end{aligned}$$

Из этого соотношения следует:

- коэффициенты  $A_k$  при всех нечетных значениях  $k$  равны нулю;

$$- A_0 = 2 \cdot (h_0 \cdot g_0 + 2 \sum_{n=1}^{\min\{N, M\}} (-1)^n \cdot h_n \cdot g_n); \quad (11)$$

- если  $A_0 = 2$  и все значения  $A_{2k}$  при  $k \neq 0$  приравнять нулю, то восстанавливающие фильтры являются КИХ-фильтрами и  $Kh(x) = \overline{G}(x)$ ;  $Kg(x) = \overline{H}(x)$ ; (12)

- можно показать, что если

$$\left\{ \begin{aligned} & A_{2k} = 0, \quad k \neq 0; \\ & H(0) = h_0 + 2 \sum_{n=1}^N h_n = \sqrt{2}; \quad H(1) = h_0 + 2 \sum_{n=1}^N (-1)^n h_n = 0; \\ & G(0) = g_0 + 2 \sum_{m=1}^M g_m = 0; \quad G(1) = g_0 + 2 \sum_{m=1}^M (-1)^m g_m = \sqrt{2}, \end{aligned} \right. \quad (13)$$

то  $h_0 \cdot g_0 + 2 \sum_{n=1}^{\min\{N, M\}} (-1)^n \cdot h_n \cdot g_n = 1$  и  $A_0 = 2$ .

Учитывая, что для вычисления неизвестных  $h_n$ ,  $0 \leq n \leq N$ , и  $g_m$ ,  $0 \leq m \leq M$ , требуется использовать  $N + M + 2$  уравнения, дополнительно можно приравнять нулю четные моменты функций  $H(x)$  и  $G(x)$ , т.е.

$$H^{(2r)}(x) \Big|_{x=0}, \quad r=1, 2, 3, \dots \quad \text{и} \quad G^{(2p)}(x) \Big|_{x=0}, \quad p=1, 2, 3, \dots$$

Нечетные производные в точке  $x = 0$  равны нулю по определению функций  $H(x)$  и  $G(x)$ . Равенство нулю четных производных функции  $H(x)$  в точке  $x = 0$  приводит к расширению спектра низкочастотной составляющей сигнала, равенство же нулю четных производных функции  $G(x)$  при  $x = 0$  обеспечивает уменьшение высокочастотной части спектра выделяемой составляющей сигнала. Зачастую для оптимизации формы спектра высокочастотной составляющей целесообразно подбирать значения некоторых производных  $G(x)$  в точке  $x = 0$ .

В Таблице 1 приведены параметры нескольких фильтров, рассчитанных по приведенной методике. Обозначение фильтров, приведенное в первом столбце, определяет число дискретных отсчетов фильтра при его реакции на одиночный входной сигнал:  $(2N + 1) / (2M + 1)$ . В ряде случаев система уравнений имеет несколько решений, наилучшие из которых (по мнению авторов) зафиксированы в Таблице 1.

Фильтры 5/3 и 9/7 по своим характеристикам идентичны соответствующим фильтрам, приведенным в стандарте JPEG2000 [2].

В качестве примера на рис. 3, для сравнения, приведены характеристики низкочастотного и высокочастотного фильтров для банков 5/3 и 13/11. Как следует из сравнения этих характеристик, высокочастотные составляющие сигнала на выходе банка 13/11 практически равны нулю в полосе частот до  $x = 0,25$ .

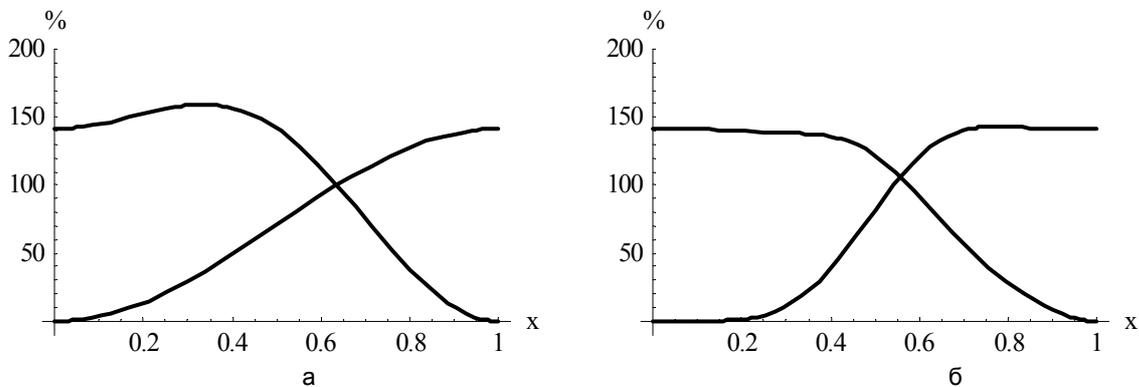


Рис. 3. АЧХ фильтров 5/3 (а) и 13/11 (б).

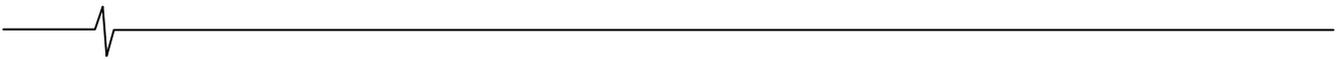
Таблица 1

Тип банка	$h_0$ $g_0$	$h_1$ $g_1$	$h_2$ $g_2$	$h_3$ $g_3$	$h_4$ $g_4$	$h_5$ $g_5$	$h_6$ $g_6$	$h_7$ $g_7$	$h_8$ $g_8$
3/5	$\frac{1}{\sqrt{2}}$	$\frac{1}{2\sqrt{2}}$							
5/3	$\frac{3}{2\sqrt{2}}$	$-\frac{1}{2\sqrt{2}}$	$-\frac{1}{4\sqrt{2}}$						
	$\frac{3}{2\sqrt{2}}$	$\frac{1}{2\sqrt{2}}$	$-\frac{1}{4\sqrt{2}}$						
5/7	$\frac{1}{\sqrt{2}}$	$-\frac{1}{2\sqrt{2}}$							
	0.91924	0.35355	-0.10607						
7/5	0.79549	-0.36681	-0.04419	0.01326					
	1.16168	0.33935	-0.22728	0.01421					
7/9	0.66291	-0.35355	0.02210						
	1.26091	0.25930	-0.27690	0.09426					
7/11	0.65488	-0.39145	0.03901	0.03790	-0.01290				
	0.79888	0.37480	-0.04589	-0.02125					
9/7	0.85202	-0.41007	-0.09961	0.06581	0.03146	-0.00929	-0.00430		
	1.01647	0.35355	-0.17678	0.0	0.02210				
9/11	0.70711	-0.39775	0.0	0.04419					
	1.19613	0.29344	-0.25647	0.06012	0.01196				
9/15	0.65833	-0.40098	0.03252	0.04905	-0.00813	-0.00162			
	1.0453	0.30266	-0.16208	0.05089	-0.00702				
11/9	0.72798	-0.39802	-0.00855	0.04776	-0.00277	-0.00317	0.00089	-0.00012	
	1.0843	0.35890	-0.23400	0.00097	0.04540	-0.00632			
11/13	0.67017	-0.39775	0.02462	0.04419	-0.00616				
	1.08054	0.37587	-0.24648	-0.01323	0.05976	-0.00908			
11/17	0.66295	-0.39614	0.02923	0.04178	-0.00703	0.00081	-0.00012		
	1.09347	0.29619	-0.19617	0.06208	0.003000	-0.00472			
13/11	0.70220	-0.40955	0.00676	0.06203	-0.00660	-0.00612	0.00243	0.00009	-0.00014
	0.93001	0.35053	-0.12210	-0.00042	0.01638	0.00344	0.00574		
13/15	-0.74971	-0.42142	-0.02840	0.07970	0.00710	-0.01184			
	1.19422	0.23649	-0.21254	0.11253	-0.03020	0.00453	-0.00081		
	0.69160	-0.40702	0.01424	0.05747	-0.00882	-0.00359	0.00234	-0.00042	

Ниже для пояснения приведена система уравнений, использованная при расчете банка фильтров 13/11. Первые пять уравнений обеспечивают выполнение условия  $\det(x) = \text{Const}$ . Эти уравнения с последующими четырьмя уравнениями обеспечивают выполнение условий (13). Следующее урав-

нение определяет равенство нулю второго момента низкочастотного фильтра, а последние три уравнения накладывают условия на второй, четвертый и шестой моменты высокочастотного фильтра, при этом второй и четвертый моменты приравнены нулю.

$$\begin{cases}
 -g_5 \cdot h_5 + g_4 \cdot h_6 = 0 \\
 -g_5 \cdot h_3 + g_4 \cdot h_4 - g_3 \cdot h_5 + g_2 \cdot h_6 = 0 \\
 -g_5 \cdot h_1 + g_4 \cdot h_2 - g_3 \cdot h_3 + g_2 \cdot h_4 - g_1 \cdot h_5 + g_0 \cdot h_6 = 0 \\
 g_4 \cdot h_0 - g_3 \cdot h_1 - g_5 \cdot h_1 + g_2 \cdot h_2 - g_1 \cdot h_3 + g_0 \cdot h_4 - g_1 \cdot h_5 + g_2 \cdot h_6 = 0 \\
 g_2 \cdot h_0 - g_1 \cdot h_1 - g_3 \cdot h_1 + g_0 \cdot h_2 - g_1 \cdot h_3 + g_4 \cdot h_2 - g_5 \cdot h_3 + g_2 \cdot h_4 - g_3 \cdot h_5 + g_4 \cdot h_6 = 0 \\
 h_0 + 2 \cdot h_1 + 2 \cdot h_2 + 2 \cdot h_3 + 2 \cdot h_4 + 2 \cdot h_5 + 2 \cdot h_6 = \sqrt{2} \\
 h_0 - 2 \cdot h_1 + 2 \cdot h_2 - 2 \cdot h_3 + 2 \cdot h_4 - 2 \cdot h_5 + 2 \cdot h_6 = 0 \\
 g_0 + 2 \cdot g_1 + 2 \cdot g_2 + 2 \cdot g_3 + 2 \cdot g_4 + 2 \cdot g_5 = 0 \\
 g_0 - 2 \cdot g_1 + 2 \cdot g_2 - 2 \cdot g_3 + 2 \cdot g_4 - 2 \cdot g_5 = \sqrt{2} \\
 h_1 + 4 \cdot h_2 + 9 \cdot h_3 + 16 \cdot h_4 + 25 \cdot h_5 + 36 \cdot h_6 = 0 \\
 g_1 + 4 \cdot g_2 + 9 \cdot g_3 + 16 \cdot g_4 + 25 \cdot g_5 = 0 \\
 g_1 + 4 \cdot 4 \cdot g_2 + 9 \cdot 9 \cdot g_3 + 16 \cdot 16 \cdot g_4 + 25 \cdot 25 \cdot g_5 = 0 \\
 g_1 + 4 \cdot 4 \cdot 4 \cdot g_2 + 9 \cdot 9 \cdot 9 \cdot g_3 + 16 \cdot 16 \cdot 16 \cdot g_4 + 25 \cdot 25 \cdot 25 \cdot g_5 = -100
 \end{cases}$$



На рис. 4 приведены сквозные характеристики банка фильтров 13/11 в точках до суммирования составляющих (см. рис. 1), соответствующие первому (а) и второму (б) уравнениям системы (7). В данном случае:

$$H_{\Delta}(z) \cdot Kh(z) = H_{\Delta}(x) \cdot \bar{G}(x); \quad G_{\otimes}(z) \cdot Kg(z) = G_{\otimes}(x) \cdot \bar{H}(x);$$

$$H_{\otimes}(z) \cdot Kh(z) = H_{\otimes}(x) \cdot \bar{G}(x); \quad G_{\Delta}(z) \cdot Kg(z) = G_{\Delta}(x) \cdot \bar{H}(x).$$

Заметим, что в первом случае характеристика высокочастотной составляющей проходит через нуль при  $x = 0,5$ , во втором же случае через нуль в этой точке проходит характеристика низкочастотной составляющей.

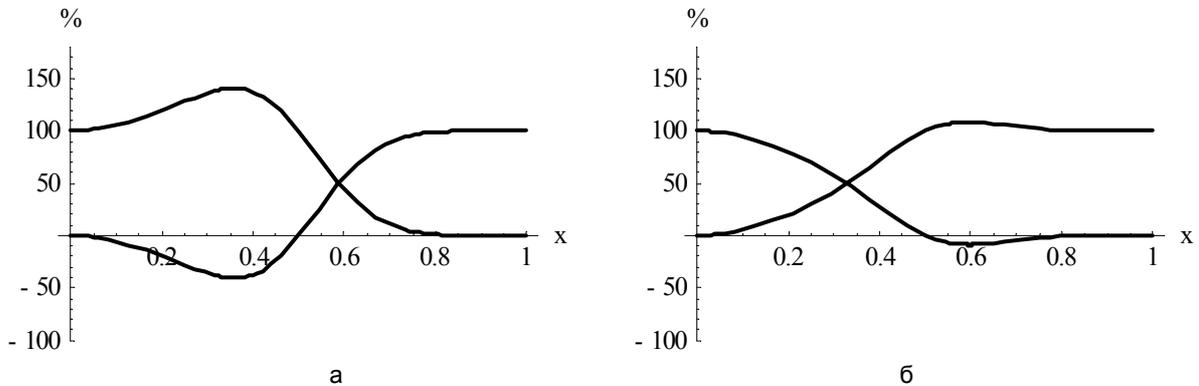


Рис. 4. Сквозные характеристики банка фильтров 13/11.

### Б. Ортогональные КИХ-фильтры с четным числом отсчетов цифровой решетки

При четном числе отсчетов цифровой решетки низкочастотного и высокочастотного фильтров банка для

того, чтобы число отсчетов справа и слева от оси ординат было одинаково, необходимо ее установить по середине между центральными отсчетами, как показано на рис. 5.

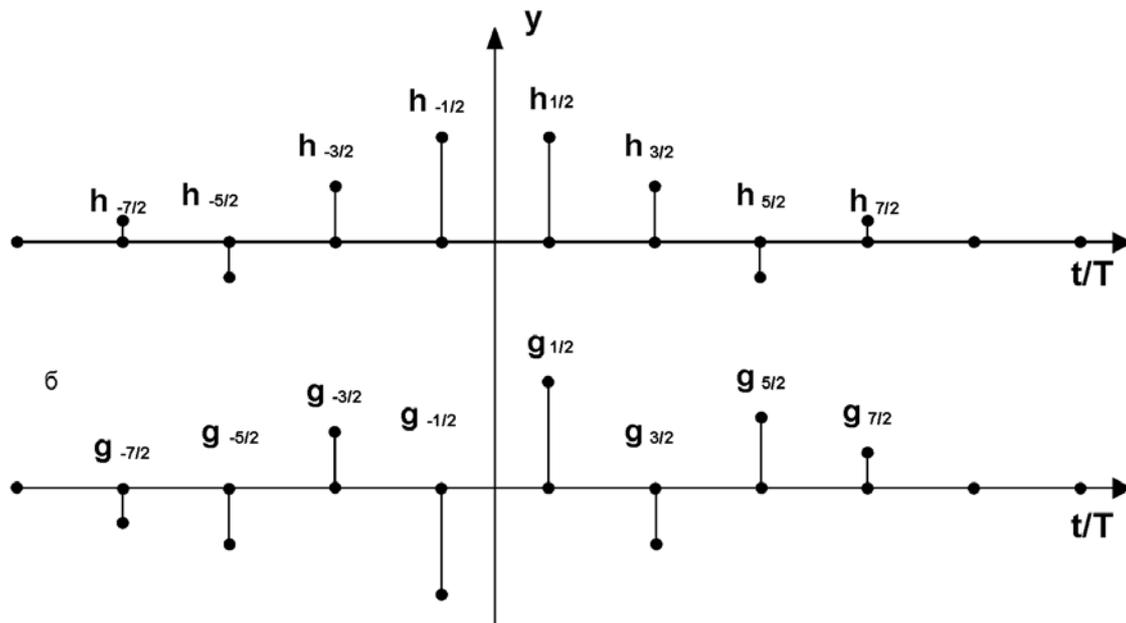


Рис. 5. Структура низкочастотного (а) и высокочастотного (б) КИХ-фильтров при четном количестве отсчетов цифровой решетки.

В этом случае положение каждого отсчета относительно оси ординат будет определяться интервалом, равным  $(2n-1)T/2$ ,  $n = \dots, -2, -1, 0, 1, 2, \dots$ . По этой причине обозначим отсчеты фильтра нижних частот через  $h_{(2n-1)/2}$ , а фильтра верхних частот – через  $g_{(2n-1)/2}$ . В случае, если

$h_{(2n-1)/2} = h_{-(2n-1)/2}$ , а  $g_{(2n-1)/2} = -g_{-(2n-1)/2}$  при  $n = 1, 2, 3, \dots$  комплексные частотные характери-

ки соответствующих фильтров можно записать в виде:

$$H(x) = 2 \sum_{n=1}^N h_{(2n-1)/2} \cdot \cos\left(\pi \frac{2n-1}{2} x\right);$$

$$G(x) = -2j \sum_{m=1}^M g_{(2m-1)/2} \cdot \sin\left(\pi \frac{2m-1}{2} x\right). \quad (14)$$

В данном случае в Z-представлении величины  $H_{\Delta}(z)$ ,  $H_{\otimes}(z)$ ,  $G_{\Delta}(z)$ ,  $G_{\otimes}(z)$  соответственно равны:

$$\begin{aligned}
 H_{\Delta}(z) &= \sum_{n=1}^{[N/2]} h_{(4n-1)/2} \cdot Z^{\frac{4n-1}{2}} + \sum_{n=1}^{[(N+1)/2]} h_{(4n-1)/2} \cdot Z^{-\frac{4n-1}{2}}; \\
 H_{\otimes}(z) &= \sum_{n=1}^{[(N+1)/2]} h_{(4n-3)/2} \cdot Z^{\frac{4n-3}{2}} + \sum_{n=1}^{[N/2]} h_{(4n-3)/2} \cdot Z^{-\frac{4n-3}{2}}; \\
 G_{\Delta}(z) &= -\sum_{m=1}^{[M/2]} g_{(4m-1)/2} \cdot Z^{\frac{4m-1}{2}} + \sum_{m=1}^{[(M+1)/2]} g_{(4m-1)/2} \cdot Z^{-\frac{4m-1}{2}}; \\
 G_{\otimes}(z) &= -\sum_{m=1}^{[(M+1)/2]} g_{(4m-3)/2} \cdot Z^{\frac{4m-3}{2}} + \sum_{m=1}^{[M/2]} g_{(4m-3)/2} \cdot Z^{-\frac{4m-3}{2}}.
 \end{aligned} \tag{15}$$

Подставив эти соотношения в систему уравнений (6), можно рассчитать параметры банка фильтров.

Эквивалентно использование соотношения:

$$\begin{cases} \bar{H}(x) \cdot Kh(x) + \bar{G}(x) \cdot Kg(x) = 0, \\ H(x) \cdot Kh(x) + G(x) \cdot Kg(x) = 2, \end{cases} \tag{16}$$

где  $\bar{H}(x) = 2j \sum_{n=1}^N (-1)^n \cdot h_{(2n-1)/2} \cdot \sin(\pi \frac{2n-1}{2} x)$ ;

а  $\bar{G}(x) = -2 \sum_{m=1}^M (-1)^m \cdot g_{(2m-1)/2} \cdot \cos(\pi \frac{2m-1}{2} x)$ .

Число  $N + M = 2K$  должно быть четным.

Детерминант системы уравнений (16) можно записать в виде:

$$\begin{aligned}
 \det(x) &= 2 \sum_{k=n-1}^N (-1)^n \cdot [(-1)^k - 1] \cdot h_{(2n-1)/2} \cdot \sum_{n=1}^{n-M} g_{(2(n-k)-1)/2} \cdot \cos(\pi kx) + \\
 & 2 \sum_{k=n}^{n+M-1} (-1)^n \cdot [(-1)^{k+1} + 1] \cdot h_{(2n-1)/2} \cdot \sum_{n=1}^{n+M-1} g_{(2(k-n)-1)/2} \cdot \cos(\pi kx) = A_0 + \sum_{\forall k} A_k \cdot \cos \pi kx.
 \end{aligned} \tag{17}$$

Исследуя это соотношение, аналогично изложенному выше, можно показать, что:

– коэффициенты  $A_k$  при всех нечетных значениях  $k$  равны нулю;

$$A_0 = 4 \sum_{n=1}^{\min\{N,M\}} (-1)^n \cdot h_{(2n-1)/2} \cdot g_{(2n-1)/2}; \tag{18}$$

– если  $A_0 = 2$  и все значения  $A_{2k}$  при  $k \neq 0$  приравнять нулю, то восстанавливающие фильтры являются КИХ-

фильтрами и  $Kh(x) = \bar{G}(x)$ ;  $Kg(x) = \bar{H}(x)$ ; (19)  
 – из определения функций  $H(x)$ ,  $\bar{H}(x)$ ,  $G(x)$ ,  $\bar{G}(x)$  следует, что  $H(1) = 0$ ,  $\bar{H}(0) = 0$ ,  $G(0) = 0$ ,  $\bar{G}(1) = 0$ ;  
 – можно доказать также, что если

$$\begin{cases} A_{2k} = 0, \quad k \neq 0; \\ H(0) = 2 \sum_{n=1}^N h_{(2n-1)/2} = \sqrt{2}; \\ G(1) = g_0 + 2 \sum_{m=1}^M (-1)^m g_{(2m-1)/2} = \sqrt{2}; \end{cases}$$

то  $2 \sum_{n=1}^{\min\{N,M\}} (-1)^n \cdot h_{(2n-1)/2} \cdot g_{(2n-1)/2} = 1$  и  $A_0 = 2$ .

В Таблице 2 приведены параметры нескольких ортогональных фильтров, рассчитанных по приведенной методике. Обозначение фильтров, приведенное в первом столбце, определяет число дискретных отсчетов фильтра при его реакции на одиночный входной сигнал:  $2N/2M$ . Существенно, что прореживание (децимация) сигналов на выходах низкочастотного и высокочастотного фильтров в данном случае должно осуществляться синфазно.

Заметим, что нечетные производные функции  $H(x)$  в точке  $x = 0$  равны нулю, а в точке  $x = 1$  равны нулю четные производные этой функции. Противоположный характер имеют производные функции  $G(x)$ .

На рис. 6 приведены амплитудно-частотные характеристики фильтров 2/2 (а, вейвлет Хаара) и 18/18 (б).

В отличие от фильтров, рассмотренных в разделе А, в данном случае характеристики цепей после децимации имеют близкие к равномерным частотные характеристики и нелинейные фазовые характеристики.

При этом АЧХ могут быть рассчитаны по формулам:

$$\begin{aligned}
 ModH_{\Delta}(x) &= ModH_{\otimes}(x) = \\
 & \sqrt{\sum_{n=1}^N \sum_{m=1}^N h_{(2n-1)/2} \cdot h_{(2m-1)/2} \cdot \cos \pi x \cdot [n - \frac{1}{2} - (-1)^{n+m} \cdot (m - \frac{1}{2})]} \tag{20}
 \end{aligned}$$

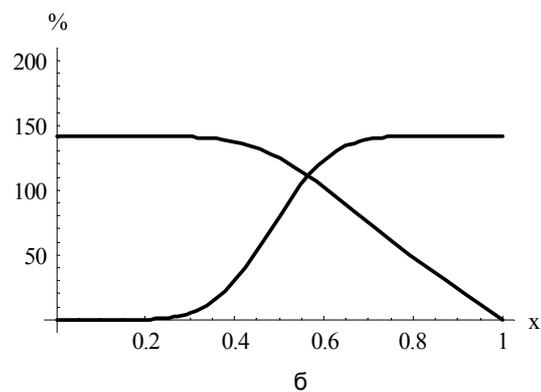
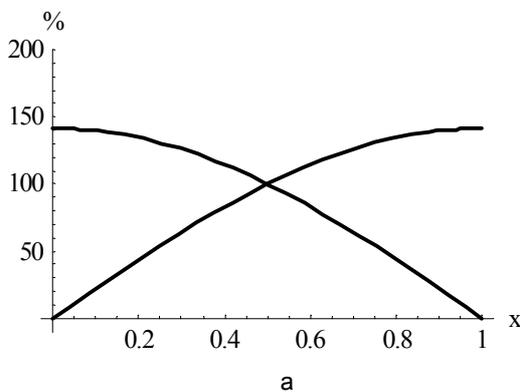


Рис. 6. АЧХ фильтров 2/2 (а) и 18/18 (б)

Таблица 2

Тип банка	$h_{1/2}$ $g_{1/2}$	$h_{3/2}$ $g_{3/2}$	$h_{5/2}$ $g_{5/2}$	$h_{7/2}$ $g_{7/2}$	$h_{9/2}$ $g_{9/2}$	$h_{11/2}$ $g_{11/2}$	$h_{13/2}$ $g_{13/2}$	$h_{15/2}$ $g_{15/2}$	$h_{17/2}$ $g_{17/2}$
2/2	$\frac{1}{\sqrt{2}}$								
2/6	$\frac{1}{\sqrt{2}}$								
4/4	$\frac{1}{\sqrt{2}}$	$-\frac{1}{8\sqrt{2}}$	$-\frac{1}{8\sqrt{2}}$						
	$\frac{3}{2\sqrt{2}}$	$\frac{1}{2\sqrt{2}}$							
	$\frac{3}{4\sqrt{2}}$	$-\frac{3}{4\sqrt{2}}$							
4/8	0.79549	-0.08839							
	0.64302	-0.13037	-0.05966	0.00663					
6/6	0.92685	-0.19178	-0.02796						
	0.57159	-0.15865	-0.02313						
6/10	0.79502	0.04443	-0.13234						
	0.61023	-0.13676	-0.06635	-0.00665	0.01982				
8/8	0.85243	-0.07903	-0.12946	0.06317					
	0.59770	-0.17570	-0.04455	0.02174					
8/12	0.87834	-0.12222	-0.13234	0.08333					
	0.59247	-0.19503	-0.046422	0.03240	-0.00097	0.00061			
10/10	0.95972	-0.20193	-0.153082	0.11281	-0.01041				
	0.56320	-0.20458	-0.02744	0.03043	-0.00281				
10/14	0.72294	0.06177	0.09762	0.01435	0.00566				
	0.66927	-0.12086	-0.08869	-0.00005	0.00523	-0.00062	-0.00024		
12/12	0.98458	-0.23183	-0.16533	0.13003	-0.00284	-0.00749			
	0.55896	-0.21308	-0.02582	0.04000	-0.00054	-0.00143			
12/16	0.83729	-0.06918	-0.16953	0.085012	0.02353	-0.01692			
	0.59260	-0.19976	-0.05335	0.03983	0.00345	-0.00343	0.00061	-0.00044	
14/14	0.88664	-0.11406	-0.17114	0.10839	0.01655	-0.02133	0.00206		
	0.58656	-0.20810	-0.05111	0.04621	0.00316	-0.00602	0.00058		
16/16	0.85855	-0.12399	-0.08598	0.05468	0.02137	-0.01716	-0.00278	0.00241	
	0.60752	-0.20251	-0.07290	0.05207	0.01350	-0.01105	-0.00133	0.00116	
18/18	0.78189	0.04881	-0.03671	0.00673	0.00548	-0.00085	-0.00073	0.00006	0.00005
	0.64549	-0.17213	-0.10328	0.03688	0.02869	-0.00671	-0.00567	0.00061	0.00054

Для функций  $Mod G_{\Delta}(x)$  и  $Mod G_{\otimes}(x)$  формула аналогична, но вместо произведения

$h_{(2n-1)/2} \cdot h_{(2m-1)/2}$   
необходимо подставить  $g_{(2n-1)/2} \cdot g_{(2m-1)/2}$ .

Для низкочастотного и высокочастотного фильтров банка 18/18  $Mod G_{\Delta}(x)$  и  $Mod G_{\otimes}(x)$  приведены на рис. 7.

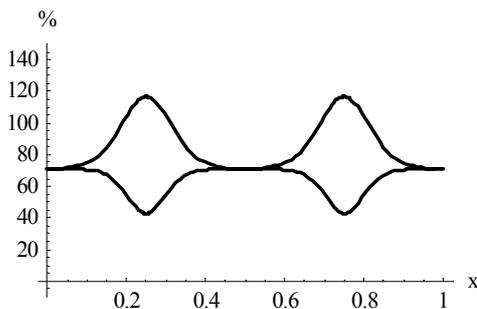


Рис. 7. АЧХ низкочастотного (сверху) и высокочастотного (снизу) фильтров банка 18/18 после децимации.

Групповое время запаздывания (производная фазовой характеристики) этих фильтров после децимации имеет вид:

$$t_{gr}[H_{\Delta}(x)] = -t_{gr}[H_{\otimes}(x)] = \frac{\pi \cdot T}{Mod H_{\Delta}(x)}$$

$$\left[ \sum_{n=1}^N \sum_{m=1}^N \frac{2n-1}{2} \cdot (-1)^{n+1} h_{(2n-1)/2} \cdot h_{(2m-1)/2} \cdot \cos \pi x \cdot \left[ n - \frac{1}{2} - (-1)^{n+m} \cdot \left( m - \frac{1}{2} \right) \right] \right] \quad (21)$$

Формулы расчета  $t_{gr}[G_{\Delta}(x)] = -t_{gr}[G_{\otimes}(x)]$  могут быть получены при замене

$h_{(2n-1)/2} \cdot h_{(2m-1)/2}$  на  $g_{(2n-1)/2} \cdot g_{(2m-1)/2}$

и  $Mod H_{\Delta}(x)$  на  $Mod G_{\Delta}(x)$ .

На рис. 8а и рис. 8б приведены относительные характеристики группового времени запаздывания фильтров 18/18 после децимации, соответствующие первому и второму уравнениям системы (6).

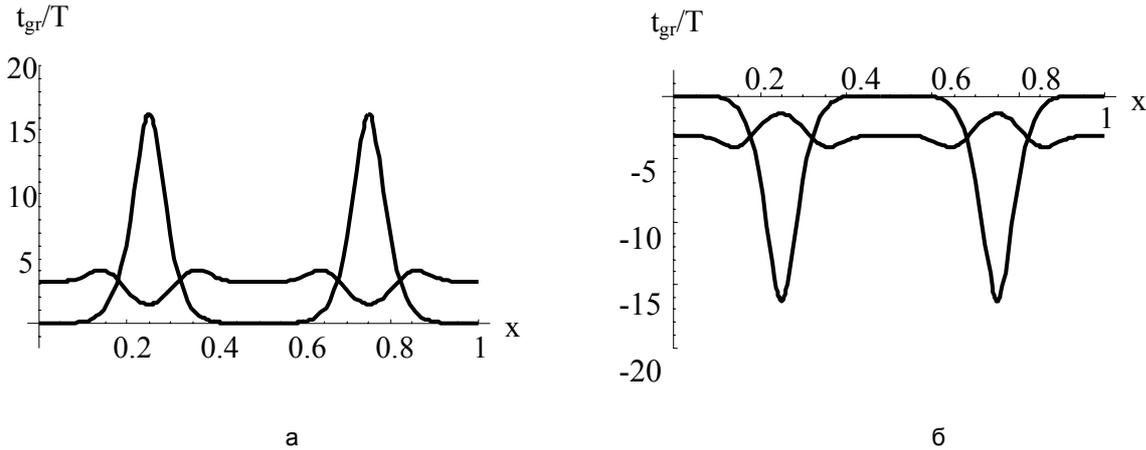


Рис. 8. Относительные характеристики группового времени запаздывания фильтров 18/18 после децимации (менее изменяющаяся характеристика соответствует низкочастотному, а более изрезанная – высокочастотному фильтру).

Аналогичным же образом можно рассчитывать квадратурно-зеркальные фильтры, для которых частотные характеристики в Z-представлении имеют вид:

$$H(z) = \sum_{n=1}^N \left( h_{\frac{2n-1}{2}} \cdot Z^{\frac{2n-1}{2}} + h_{\frac{2n-1}{2}} \cdot Z^{-\frac{2n-1}{2}} \right),$$

$$G(z) = \sum_{n=1}^N (-1)^n \left( h_{\frac{2n-1}{2}} \cdot Z^{\frac{2n-1}{2}} - h_{\frac{2n-1}{2}} \cdot Z^{-\frac{2n-1}{2}} \right). \quad (22)$$

В качестве примера приведем характеристики квадратурно-зеркального фильтра 10/10.

Система уравнений в данном случае дает четыре равнозначных решения, приведенных в Таблице 3. На рис. 9 приведены АЧХ этих фильтров.

Таблица 3

$h_{-9/2}$	$h_{-7/2}$	$h_{-5/2}$	$h_{-3/2}$	$h_{-1/2}$	$h_{1/2}$	$h_{3/2}$	$h_{5/2}$	$h_{7/2}$	$h_{9/2}$
0.16010	0.60383	0.72431	0.13843	-0.24229	-0.03224	0.07757	-0.00624	-0.01258	0.00334
0.00334	-0.01258	-0.00624	0.07757	-0.03224	-0.24229	0.13843	0.72431	0.60383	0.16010
0.02733	0.0295	-0.03913	0.19940	0.72341	0.63398	0.01660	-0.17533	-0.02110	0.01954
0.01954	-0.02110	-0.17533	0.01660	0.63398	0.72341	0.19940	-0.03913	0.02952	0.02733

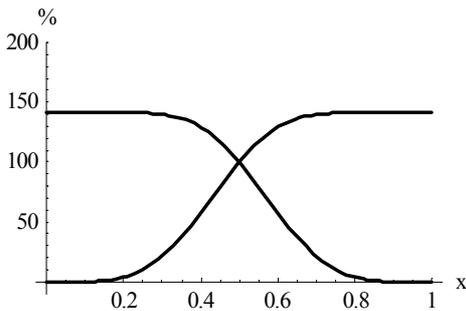


Рис. 9. АЧХ квадратурно-зеркальных фильтров банка 10/10.

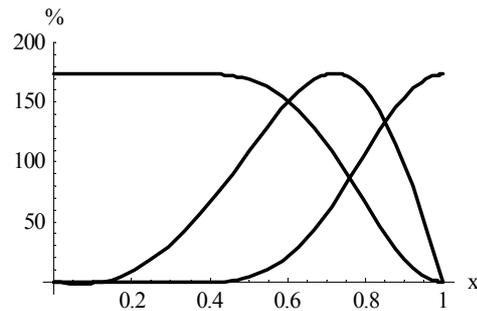


Рис. 10. Частотные характеристики трехполосного банка фильтров 9/9/9.

**Заключение**

Предложенные варианты расчета банка фильтров дискретного вейвлет-преобразования позволяют выбрать оптимальные варианты преобразований сигналов (в том числе и многомерных) по эффективности преобразований и требований к вычислительной мощности соответствующих цифровых систем.

Предложенный принцип расчета может быть использован также и для расчета фильтров многополосных вейвлет-преобразований.

В качестве примера приведем характеристики рассчитанного трехполосного банка фильтров 9/9/9,

содержащего фильтры нижних частот  $H(x)$ , средних частот  $B(x)$  и верхних частот  $G(x)$ . В данном случае выбраны следующие соотношения:

- для фильтра нижних частот  $h_n = h_{-n}$ ;
- для фильтра средних частот  $b_n = -b_{-n}, b_0 = 0$ ;
- для фильтра верхних частот  $g_n = g_{-n}$ .

Частотные характеристики этих фильтров приведены на рис.10.

В таблицах 4 и 5 приведены соответственно значения дискретных компонентов этих фильтров и восстанавливающих цепей  $Kh(x), Kb(x)$  и  $Kg(x)$ .

Таблица 4

$n$	0	1	2	3	4
$h_n$	1.25844	0.378886	-0.189443	0.0541266	-0.00676582
$b_n$	0.0	0.654084	-0.351267	0.0328772	-0.0125453
$g_n$	0.473608	-0.378886	0.189443	-0.0541266	0.00676582

Таблица 5

$n$	0	1	2	3	4	5	6	7
$\bar{h}_n$	0.577350	0.458638	0.204140	0.0	-0.0583726	-0.028569	0.0	0.001515
$\bar{b}_n$	0.0	-1.28250	-0.729767	0.0	0.038708	0.0	0.0	0.0
$\bar{g}_n$	0.577350	-1.69498	-1.15859	0.0	0.013910	-0.028569	0.0	0.001515

### Литература

1. S. Mallat A theory for multiresolution signal decomposition: the wavelet representation. IEEE Pattern Anal. and Machine Intel, 1989, vol. 11, N 7, pp. 674-693
2. И. Добеши Десять лекций по вейвлетам. Пер. с англ. Е. В. Мищенко под ред. А. П. Петухова. М.: РХД, 2001
3. C. Sidney Burrus, Ramesh A. Gopinath, and Haitao Guo Introduction to Wavelets and Wavelet Transforms. A Primer. Prentice-Hall, Upper Saddle River, New Jersey 07458, 1998
4. В. И. Воробьев, В. Г. Грибунин Теория и практика вейвлет-преобразований. С-Пб.: ВУС, 1999
5. В. П. Дьяконов Вейвлеты. От теории к практике. М.: Солон-Р, 2002
6. ITU-T Recommendation T.800 / Information Technology -- JPEG 2000 Image Coding System: Core Coding System, 2004
7. ISO/IEC 14496-2 – Information Technology -- Coding of Audio-Visual Objects -- Part 2: Visual, 2004 (MPEG-4 Video)
8. Р. Гонсалес, Р. Вудс Цифровая обработка изображений. Пер. с англ. под ред. П. А. Чочиа. М.: Техносфера, 2005

### Уважаемые коллеги !

Приглашаем вас принять участие в формировании тематических выпусков журнала «Цифровая обработка сигналов» в 2006\2007 гг.

Научно-технический журнал «Цифровая обработка сигналов» выпускается на регулярной основе с 2000 года. Цель его издания – содействие развитию и широкому применению методов цифровой обработки сигналов (ЦОС) и современных информационных технологий реального времени в разработках отечественных ученых и специалистов. Журнал «Цифровая обработка сигналов» внесен в дополнительный Перечень периодических изданий, в которых рекомендуется публикация основных результатов диссертаций на соискание ученой степени доктора технических наук.

Просим вас выразить свое отношение к данному предложению по сотрудничеству с редакцией журнала. Желательно заранее сообщить ориентировочное название статей и их предполагаемый объем (максимальный объем – до 20 страниц для обзорных статей и до 12 страниц для оригинальных работ).

**При подготовке ваших предложений просим учесть, что предпочтение отдается статьям обзорного характера и работам, посвященным решению проблем эффективного проектирования и технической реализации систем ЦОС с применением сигнальных процессоров и ПЛИС, а также различным приложениям технологий обработки сигналов и изображений в реальном времени.**

По всем вопросам, связанным с требованиями и порядком предоставления материалов статей, обращайтесь на сайт журнала: [www.dsra.ru](http://www.dsra.ru).

Предварительные заявки и материалы работ на рецензирование и публикацию принимаются по E-mail: [tor@rgta.rayzan.ru](mailto:tor@rgta.rayzan.ru).

Ждем ваших предложений.

Зам. главного редактора, профессор

В.В. Витязев

УДК 621.397.2

## ОБ ИСПОЛЬЗОВАНИИ КОНТЕКСТНОГО ВЕКТОРНОГО КВАНТОВАНИЯ В ОБЛАСТИ ДИСКРЕТНЫХ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЙ ДЛЯ КОМПРЕССИИ ИЗОБРАЖЕНИЙ

Умняшкин С.В., Коплович Д.М., Черкасов И.В.

### Введение

По применению функциональных базисов вейвлетов для компрессии цифровых изображений предложено немало алгоритмов и методов сжатия, действует основанный на дискретном вейвлет-преобразовании (ДВП) международный стандарт, который определяет формат сжатого представления изображений – JPEG-2000.

Естественно, что использование векторного квантования (ВК) [1-3] для обработки изображений в области ДВП не могло быть обойдено стороной, и этой теме также посвящено немало публикаций (см., например, [4-6]).

При ВК задается некоторое правило  $Q$ , по которому каждому вектору  $x$  исходного пространства  $R^N$  ставится в соответствие некоторый вектор  $c$  из того же пространства. Число векторов  $c$  конечно. Множество из всех  $m$  возможных значений кодовых векторов  $C = \{c_1, c_2, \dots, c_m\}$  является кодовой книгой. Таким образом, передаче и хранению вместо исходного вектора  $x$  подвергается индекс  $j$  вектора  $c_j = Q(x)$ .

При том, что потенциально ВК является весьма мощным инструментом для эффективного представления данных, его использование на практике сопряжено с проблемами построения оптимальных кодовых книг и значительной вычислительной сложностью методов сжатия данных, усложняющих реализацию методов на основе ВК [1-3]. Вычислительная сложность в основном сопряжена с проблемой поиска наилучших кодовых векторов в кодовой книге.

Для снижения вычислительной сложности ВК до приемлемого на практике уровня квантуемые векторы должны иметь невысокую размерность, а соответствующие кодовые книги – по возможности меньший объем. Рассмотрим возможность использования ВК с небольшим объемом кодовой книги в области ДВП.

Пусть  $X$  – матрица, элементы которой являются значениями яркости отсчетов (пикселей) дискретного изображения, а  $Y$  – матрица, полученная из  $X$  в результате двумерного ДВП. В наших экспериментах мы использовали 5-уровневое биортогональное ДВП 9/7 [7], выполненное над полутоновым изображением размерности  $512 \times 512$  пикселей, см. рис. 1.

Прежде всего, необходимо определиться с тем, как формировать векторы для ВК из матрицы  $Y$ . Общее правило состоит в том, что компоненты вейвлет-спектра, выбранные для совместной обработки посредством ВК, должны быть максимально коррелированы [8]. Статистические связи между компонентами ДВП-спектра носят двойственный характер [9]: это зависимо-

*Анализируются возможности использования контекстного векторного квантования при кодировании изображений с использованием дискретных вейвлет-преобразований. Предлагается использовать достаточно простой и вычислительно эффективный алгоритм кодирования.*

сти между соседними коэффициентами внутри саббэнда (субполосы) и между «родителями» и «потомками» – коэффициентами из разных саббэндов (см. рис. 2). В наших исследованиях мы разбивали каждый из саббэндов 7-15 (нумерацию см. на рис. 1) на 1024 квадратных блока (для ВК), а саббэнды 0-6 кодировались скалярно. То есть блоки для ВК формировались по принципу объединения соседних коэффициентов. Саббэнды с номерами 7-9 разбивались, таким образом, на блоки размерности  $2 \times 2$ , а саббэнды 13-15 – на блоки  $8 \times 8$ .

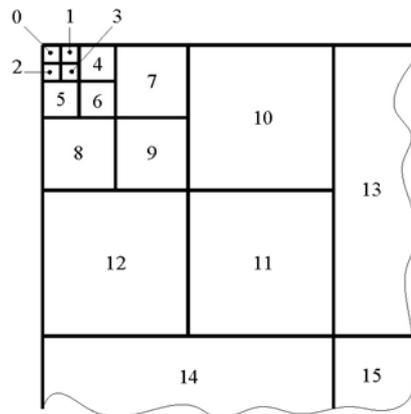


Рис. 1. Нумерация саббэндов пятиуровневого ДВП

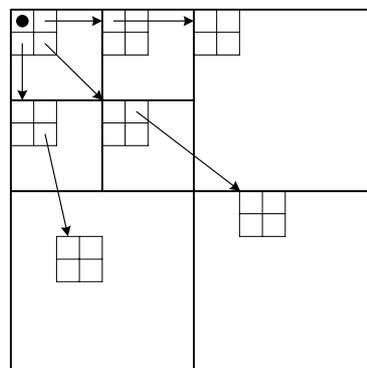


Рис. 2. Структура связей «родитель-потомки» между базисными функциями вейвлет-спектра. Выполнено три шага преобразования

### Базовый алгоритм вейвлет-компрессии

Каждому саббэнду с номерами 7–15 соответствовала своя начальная кодовая книга, полученная в результате обработки тестовых изображений с использованием алгоритма построения кодовых книг, описанного в [8], который представляет собой модификацию алгоритма ECVQ [10]. Саббэнды с номерами 0-6 квантовались скалярно, а саббэнды с номерами 7–15 – с использованием адаптивного ВК, алгоритм которого приведен в [8]. Детальное описание используемой при этом процедуры скалярного квантования и покомпонентного кодирования приведено в работе [11], откуда целиком была заимствована схема скалярной обработки коэффициентов. Отметим, что в [11] для всех саббэндов используется равномерное квантование с одним и тем же значением кванта  $q$ , после чего используется контекстное многомодельное арифметическое кодирование.

При построении начальных кодовых книг по тестовым последовательностям для каждого саббэнда задавался максимально возможный объем (количество векторов) кодовой книги, равный  $N_{нач}=32$ . При адаптации кодовой книги в процессе кодирования саббэнда допускалось ее «разрастание» до  $N_{макс}=64$  векторов, после чего добавление нового вектора в кодовую книгу осуществлялось одновременно с удалением самых «старых» и редко встречающихся векторов. Отметим, что увеличение значений  $N_{нач}$  и  $N_{макс}$  приводит к некоторому повышению качества восстановленного изображения, однако влечет за собой повышение битовых затрат и замедление обработки.

### Модифицированный алгоритм вейвлет-компрессии

Для ускорения времени обработки изображения необходимо по возможности максимально уменьшить объем кодовых книг. Кроме того, фотографические изображения обычно обладают нестационарной статистикой, и для повышения качества обработки, в смысле достижимого баланса между битовыми затратами и уровнем внесенной при обработке ошибки, целесообразно каким-то образом «адаптировать» векторный квантователь под локально изменяющиеся статистические характеристики обрабатываемых блоков. Например, вместо одной кодовой книги могут использоваться несколько книг меньшего объема, соответствующих разным классам векторов [12, 13]. Использование данного подхода в области ДВП описано в работах [14, 15]. Следуя идеям работ [12-15], мы постарались найти способ выделения из кодовой книги некоторого подмножества, *контекстной кодовой книги*, которая включала бы в себя небольшое число наиболее ожидаемых кодовых векторов из общей кодовой книги. Основу алгоритма составляют два метода кодирования: контекстное скалярное и *контекстное векторное*, которые применяются к каждому блоку саббэнда, и при кодировании саббэнда для каждого блока из них выбирается наилучший в смысле минимума RD-функции.

Таким образом, сначала для каждого саббэнда с номерами 7-15 (см. рис. 1) формируются исходные кодовые книги  $\{C_i\}$ , ( $i=7, \dots, 15$ ) по тестовым изображениям

при помощи алгоритма, описанного в [8]. Затем саббэнды последовательно обрабатываются в порядке, определенном на рис. 1, при котором коэффициенты-«родители» обрабатываются перед «потомками». Коэффициенты саббэнда 0 подвергаются двумерной дифференциальной импульсно-кодовой модуляции и арифметическому кодированию; вейвлет-коэффициенты саббэндов 1-6 обрабатываются при помощи контекстного скалярного квантования (КСК), схема которого заимствована из [11]. Обработка каждого блока саббэндов с номерами 7-15 может включать уже не только КСК, но и контекстное ВК (КВК).

В качестве параметра, по которому формируется контекстная кодовая книга из общей кодовой книги, предлагается использовать «энергию» векторов, т.е. величину  $E = E(y) = \|y\|_E^2$  (квадрат евклидовой нормы вектора). При обработке очередного вектора  $y_k$  вычисляется средневзвешенная энергия уже обработанных векторов-блоков

$$E_{\Sigma}^k = \frac{\sum_{i=1}^3 e_i^k w_i^k + \sum_{i=1}^9 E_i^k W_i^k}{\sum_{i=1}^3 w_i^k + \sum_{i=1}^9 W_i^k},$$

где  $e_i^k$  – энергия соседнего с  $y_k$  вектора,  $w_i^k$  – его вес;

аналогично,  $E_i^k$  и  $W_i^k$  – энергия и вес вектора из «родительского» саббэнда. На рис. 3 показано расположение векторов-блоков и их веса для вычисления величины  $E_{\Sigma}^k$ . Контекстная кодовая книга

$$\tilde{C} = \left\{ c \in C \mid \|c\|_E^2 \in [E_{\Sigma}^k - \Delta; E_{\Sigma}^k + \Delta] \right\}$$

формируется в процессе обработки из векторов общей кодовой книги  $C$ , где параметр  $\Delta$  носит эмпирический характер.

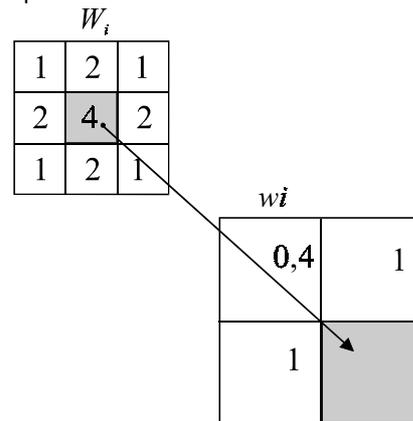


Рис. 3. Вычисление величины  $E_{\Sigma}^k$  для  $k$ -го вектор-блока (помечен темно-серым цветом). Родительский блок помечен светло-серым цветом

Опишем алгоритм кодирования саббэндов с номерами  $i=7 \dots 15$ , каждый из которых содержит по 1024 блока, просматриваемых слева направо, сверху вниз. Для каждого блока  $y_k^i$  ( $k=0 \dots 1023$ ) вычисляется функция Лагранжа, связанная с КСК:

$$J_k^{SQ} = \sum_{y_m \in y_k^i} (y_m - \hat{y}_m)^2 - \lambda (\sum_{y_m \in y_k^i} \log_2 v(\hat{y}_m) + \log_2 v_{SQ})$$

где  $\hat{y}_m$  – проквантованное значение компоненты

$y_m \in y_k^i, v(\hat{y}_m)$  – оцениваемая по накопленной статистике частота появления значения  $\hat{y}_m$  в контекстно выбранной модели скалярного кодирования (правило выбора – см. [11]), а  $v_{SQ}$  – частота использования скалярного метода кодирования блоков в данном саббэнде. Кроме того, вычисляется функция Лагранжа, связанная с КСК,

$$J_k^{VQ} = \min_{c_j^i \in C_i} \left( \|y_k^i - c_j^i\|_E^2 - \lambda \left( \log_2 \frac{v(c_j^i)}{\sum_{c_j^i \in C_i} v(c_j^i)} + \log_2 v_{VQ} \right) \right),$$

где  $v(c_n^i)$  – оцениваемая по накопленной статистике частота появления кодового вектора  $c_n^i$  в модели кодирования индексов  $i$ -ой кодовой книги, а  $v_{VQ}$  – частота использования векторного метода кодирования блоков в данном саббэнде,  $v_{VQ} + v_{SQ} = 1$ .

Параметр  $\lambda$  представляет собой величину, определяющую баланс между величиной сжатия данных и уровнем внесенной при кодировании ошибки, и должен выбираться таким же, что и при первоначальном построении кодовых книг с использованием алгоритма [8].

Если  $J_k^{SQ} < J_k^{VQ}$ , то производится скалярное кодирование  $k$ -го блока, и вектор из скалярно проквантованных компонент добавляется в кодовую книгу  $C_i$  (если достигнут максимальный объем кодовой книги, то перед

этим из нее удаляется кодовый вектор, который встречался реже остальных).

В результате применения описанной процедуры для обработки ряда широко распространенных тестовых полутоновых изображений были получены результаты, приведенные на графиках рис. 4, где ошибка восстановления изображения оценивается по величине СКО (MSE) через отношение пикового значения сигнала к шуму (ошибке),

$$PSNR = 10 \lg \frac{255^2}{MSE} \text{ [дБ]},$$

а величина сжатия данных выражена через битовые затраты в пересчете на один пиксел изображения. В экспериментах использовалось биортогональное преобразование 9/7 [7]. Для сравнения на приведенных графиках рис. 4 отражены также характеристики, которые дает реализация стандарта JPEG 2000 в различных программах обработки изображений (при определении битовых затрат длина заголовка файлов, создаваемых программами, отнималась от общего размера файла).

В силу того, что формируемые контекстно кодовые книги имеют небольшой объем, предложенный алгоритм является простым и вычислительно эффективным. Полученные характеристики предлагаемого алгоритма компрессии изображения, несмотря на его простоту, показывают результаты, аналогичные методу JPEG 2000. Дальнейшее направление исследований мы видим в создании метода компрессии изображений, объединяющего описанный подход с «подрезанием ветвей» вейвлет-коэффициентов [11] и более эффективно использующего контекстное скалярное квантование [16].

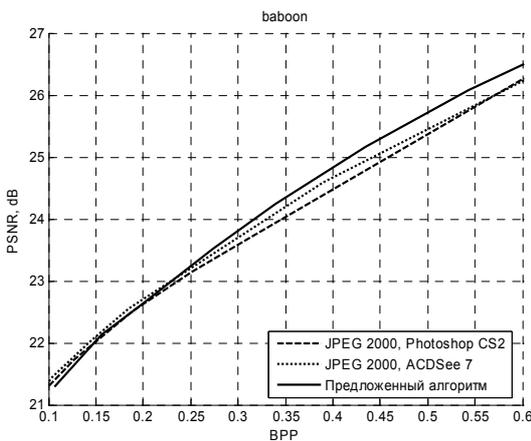


Рис. 4а. Результаты обработки тестового изображения Baboon

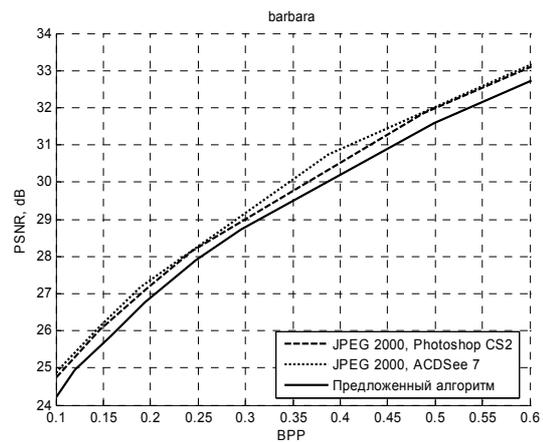


Рис. 4б. Результаты обработки тестового изображения Barbara

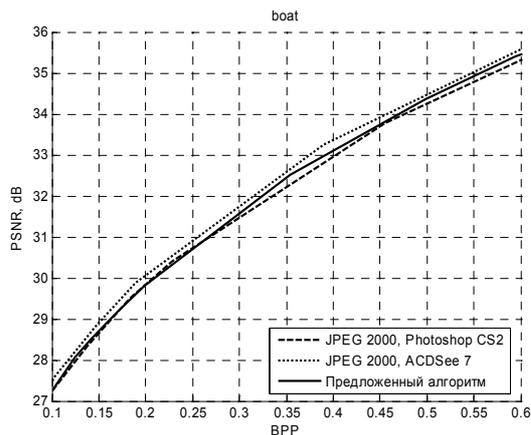


Рис. 4в. Результаты обработки тестового изображения Boat

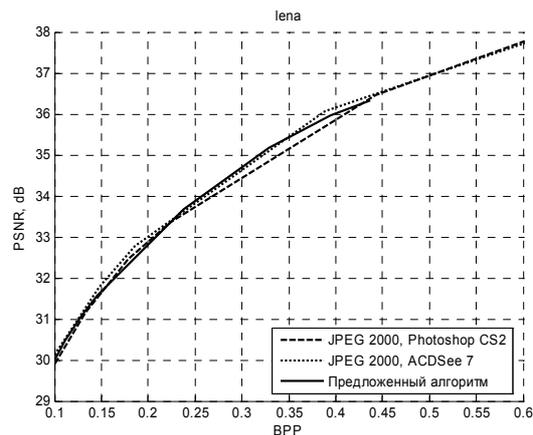


Рис. 4д. Результаты обработки тестового изображения Lena

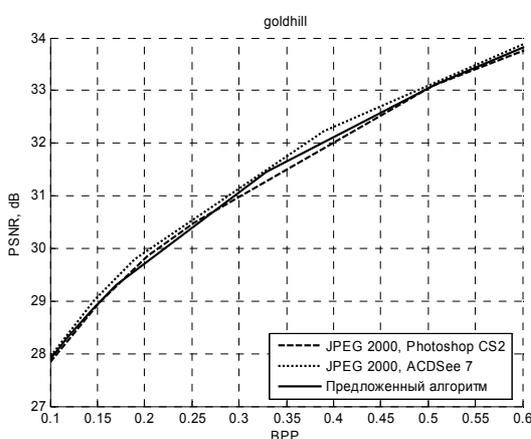


Рис. 4е. Результаты обработки тестового изображения Goldhill

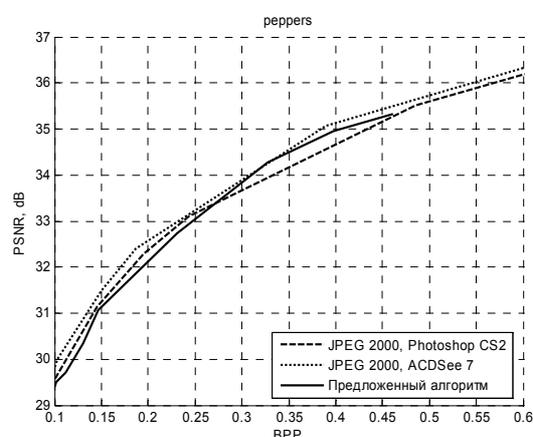


Рис. 4е. Результаты обработки тестового изображения Peppers

### Литература

1. Gray R. M., Neuhoff D. Quantization // IEEE Trans. Inf. Theory. – Oct.1998. – Vol. 44. – No. 6. – P. 2325-2383
2. Gray R. M. Vector Quantization // IEEE ASSP Magazine, April 1984, pp. 4-29.
3. Linde Y., Buzo A., Gray R. M. An Algorithm for Vector Quantizer Design. // IEEE Trans. Comm., vol. COM-28, No.1, January 1980, pp.84-95.
4. Antonini M., Barlaud M., Mathieu P., Daubechies I.. Image coding using wavelet transform. IEEE Trans. Image Processing, vol. 1, pp. 205-220, April 1992.
5. Cosman P., Gray R.M., Vetterli M. Vector quantization of image subbands: A survey // IEEE Trans. Image Proc. – Vol. 5, February 1996. – pp. 202-225.
7. Gu Q., Budge S. E. Rate-distortion adaptive vector quantization for wavelet image coding // Proc. ICASSP'00, June 2000, pp. 1903 - 1906.
8. Cohen A., Daubechies I., Feauveau J.-C.. Biorthogonal bases of compactly supported wavelets. Comm. Pure and Appl. Math., vol. 45, pp. 485-560, 1992.
9. Умняшкин С.В., Коплович Д.М. Метод компрессии изображений на основе векторного квантования коэффициентов в области дискретных преобразований // Известия вузов. Электроника. – №4-5. – 2005. – С. 149-155.
10. Chrysafis C., Ortega A. Efficient Context-Based Entropy Coding for Lossy Wavelet Image Compression // Proc. Data Compression Conference. – Snowbird (Utah), 1997. – P. 241-250
11. Chou P. A., Lookabaugh T., Gray R. M.. Entropy-constrained vector quantization. IEEE Transactions on ASSP, vol.37, No.1, January 1989, pp.31-42.
12. Умняшкин С. В. Вейвлет-компрессия цифровых изображений с прогнозированием статистических моделей //Известия вузов. Электроника. - №5. - 2001. - С.86-94.
13. Ramamurthi B., Gersho A. Classified vector quantization of images // IEEE Trans. Comm. – Vol. COMM-34, No. 11, 1986, pp. 1105–1115.
14. Ngan K. N., Koh H. C. Predictive classified vector quantization //IEEE Trans. Image Proc., vol. 1, pp. 269 - 280, July 1992.
15. Jafarkhani H., Farvardin N. Adaptive image coding using spectral classification // IEEE Trans. Image Proc., vol. 7, pp. 605 - 610, April 1998.
16. Joshi R. L., Jafarkhani H., Kasner J. H., Fischer T. R., Farvardin N., Marcellin M. W., Bamberger R. H. Comparison of different methods of classification in subband coding of images // IEEE Trans. Image Proc., vol. 6, pp. 1473 - 1486, November 1997.
17. Yoo Y., Ortega A., Yu B. Image subband coding using context-based classification and adaptive quantization // IEEE Trans. Image Proc., vol. 8, pp. 1702 - 1715, December 1999.

УДК 621.397.2

## СЖАТИЕ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ ЧАСТИЧНОЙ СОРТИРОВКИ ВЕЙВЛЕТ-КОЭФФИЦИЕНТОВ

*Авдеев О.В., Чобану М.К.*

### Введение

Цифровая обработка изображений с потерями позволяет достичь сжатия в десятки и даже сотни раз при удовлетворительном качестве их воспроизведения.

Целью сжатия изображений во многих случаях является не только компактное хранение данных, но и их последующая передача по каналам связи с максимальной скоростью. При этом в системах, где пропускная способность канала ограничена, а скорость передачи крайне важна, полезным оказывается свойство *шкалируемости качества* для рассматриваемого семейства алгоритмов. Это означает, что сначала надо передавать копию изображения с меньшим качеством, но с сохранением наиболее важной информации об изображении, а затем постепенное «уточнение» менее важных деталей – так называемая возможность постепенной передачи.

Во многих медицинских приложениях сжатие должно обеспечивать возможность анализа больших массивов изображений (наборов трехмерных данных, последовательностей изображений, каталогов) в интерактивном режиме, для исследования зависимых от контекста деталей изображения, а также количественного анализа результатов измерений. Поэтому в данной области существуют жесткие требования к качеству обрабатываемых изображений: недопустимо отбрасывать любые детали при обработке медицинских данных. Например, детали изображений, полученные с помощью проективной радиографии могут отображать повреждения внутренних органов, причем различить их можно только по еле заметным изменениям яркости изображения.

Концепция постепенной передачи изображений здесь также крайне важна, так как позволяет быстро определить по полученным ранее фрагментам с малым разрешением важные для постановки диагноза области изображения.

### Принцип вложенного кодирования

Семейство алгоритмов, построенных на сортировке вейвлет – коэффициентов по степени их вклада в качество результирующего изображения и последующей их постепенной передаче излагается в [1]. Эти алгоритмы ре-

*Рассматривается алгоритм сжатия двумерных сигналов, работающий с вейвлет-образом исходного изображения и основанный на принципе частичной сортировки вейвлет-коэффициентов, легко адаптируемый под широкий круг задач. Рассматриваются преимущества и ограничения выбранного метода в задачах сжатия изображений с потерями.*

ализуют вейвлет-декомпозицию исходного изображения в древовидную структуру, состоящую из частотных субполос. Такая структура может быть представлена также в виде двумерной матрицы, в которой субполосы расположены (декомпозиция Мала).

Затем эта матрица сжимается при помощи иерархического алгоритма, основанного на принципе частичной сортировки вейвлет-коэффициентов по степени их вклада в качество восстановленного изображения. Наиболее важные коэффициенты передаются первыми, менее важные – во вторую очередь, либо не передаются вообще. Качеством результата можно легко управлять, варьируя количество передаваемых бит. Выходной битовый поток затем сжимается арифметическим кодером. Общая структура системы показана на рис. 1.

Алгоритм вложенных нуль-деревьев EZW [1] показал очень хорошие результаты, однако его высокие требования к объему памяти и производительности ЭВМ заставили исследователей искать более простые схемы. Важным свойством этого алгоритма и других, основанных на тех же принципах, алгоритмов является возможность последовательной передачи коэффициентов таким образом, что наиболее важные (с точки зрения результата декодирования) коэффициенты передаются первыми. Это дает возможность гибко управлять соотношением качества результата в зависимости от количества переданных бит.

Такой алгоритм называется вложенным кодированием, так как при кодировании изображения сообщение фактически содержит в первых битах копию всего изображения, но с меньшим качеством (см. рис. 2). Через  $R_1$ ,  $R_2$  и  $R_3$  обозначено количество бит, требуемых для передачи последовательно улучшаемых копий сигнала.

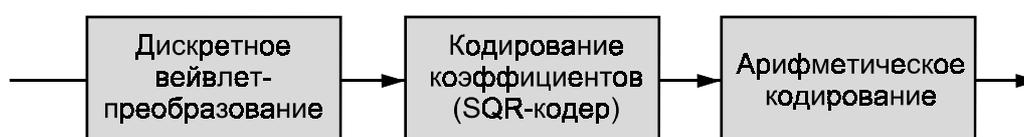


Рис. 1. Общая структура алгоритма

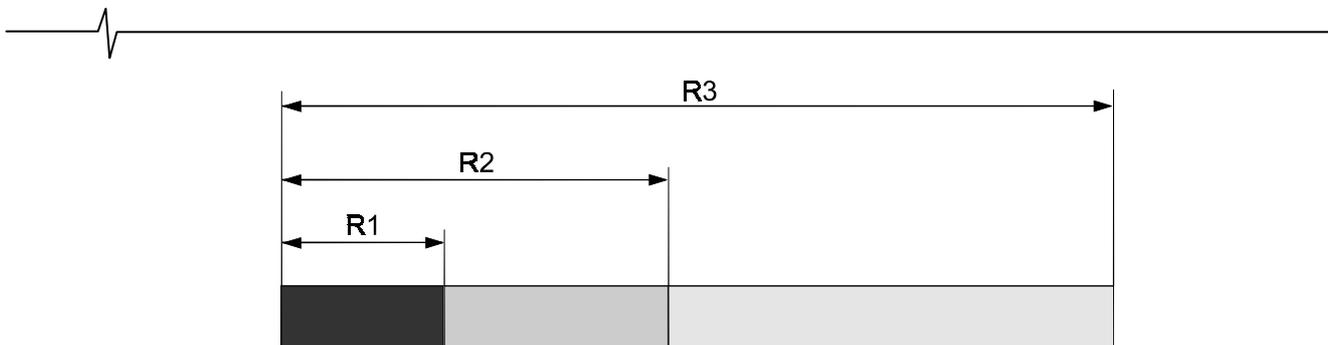


Рис. 2. Вложенное кодирование

### Метод пространственно упорядоченных иерархических деревьев

Наибольшую известность приобрел алгоритм «разбиения множеств в древовидных структурах» (SPIHT) [2]. Полученные результаты не уступали EZW, однако структура алгоритма была существенно проще, а скорость обработки выше. Несколько вариантов этого алгоритма были разработаны и оптимизированы в МЭИ ([3]).

Алгоритм заключается в том, что производится иерархическое разбиение вейвлет-образа на области по границам субполос и осуществляется частичная сортировка коэффициентов: прежде всего выбираются коэффициенты, «значимые» относительно набора пороговых значений.

При сортировке также используется тот факт, что внутри субполосы соседние коэффициенты сильно коррелированы, а также подобие субполос более высоких и более низких уровней (например, 1LH и 2LH на рис.3 и рис.4.б). В результате из сигнала размером  $N \times N$  (см. рис. 4.а) получаем его вейвлет-образ того же размера (см. рис. 4.б).

Алгоритм получил широкую известность и стал де-факто одним из стандартных алгоритмов сжатия изо-

бражений на основе вейвлетов. Однако сложность его реализации на архитектурах с ограниченными ресурсами заставила его создателей продолжить поиск, результатом которого стал алгоритм SPECK. После незначительных изменений он был включен в стандарт JPEG 2000 под названием SBHP (Иерархическое разбиение блоков субполос).

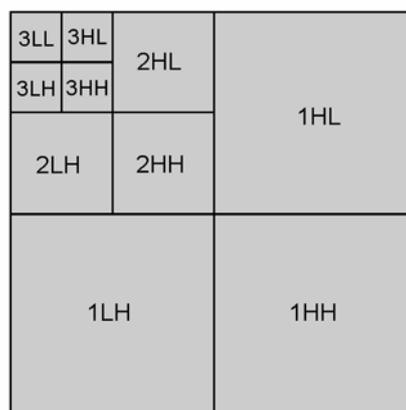
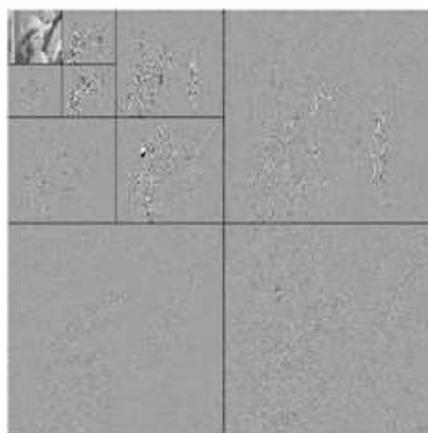


Рис. 3. Обозначение подполос вейвлет-образа сигнала



(а) Исходное изображение



(б) вейвлет - коэффициенты

Рис. 4. Пример вейвлет – разложения

Почти одновременно со SPECK, независимо от него был разработан аналогичный ему алгоритм SQP. Несмотря на отличие в подходах к проблеме, результаты работы обоих алгоритмов очень схожи. Поскольку SQP имеет более простую структуру, он лучше масштабируется и является более перспективным с точки зрения дальнейшего улучшения производительности.

Алгоритм вложенного кодирования вейвлет – коэффициентов EZW [1] использует простую обобщенную модель для

описания распределения коэффициентов в вейвлет-образе. Эта модель основывается на гипотезе «нуль-дерева», которая предполагает, что, если вейвлет-коэффициент  $w$ , соответствующий определенному значению шкалы масштаба, является «незначимым» по отношению к данному пороговому значению  $T$ , то есть  $|w| < T$ , то все коэффициенты, расположенные подобным образом и имеющие такие же относительные координаты на более низких уровнях разложения, также незначимы по отношению к  $T$ .

Квантование последовательным приближением (SAQ) применяется в работе [1] для определения значимости вейвлет-коэффициентов по отношению к множеству пороговых значений  $T_i$ . Координаты значимых и незначимых коэффициентов соответственно отражены в бинарных матрицах для каждого значения  $T_i$ . Как доказано в работе [1], вне зависимости от того насколько эффективно вейвлет-преобразование, блок квантования (или энтропийный кодер) и кодирование матриц значимости вносят значимый вклад в эффективность всего кодирования. Метод кодирования матриц значимости предложен в работе [1] в виде так называемого нуль-дерева, которое позволяет эффективно предсказывать положение незначимой информации на всех значениях масштаба. С помощью этого приема кодирование матриц значимости производится путем группировки незначимых коэффициентов в древовидных структурах и кодирования их с помощью специального алфавита.

Более сложный алгоритм, - «разбиение множеств в иерархических деревьях» (SPIHT) [2], используется для кодирования матриц значимости. При этом используется принцип частичной сортировки коэффициентов по модулю, разбиение множеств коэффициентов с помощью деревьев (то есть сортировка деревьев на основе их значимости по отношению к каждому пороговому значению), упорядоченная передача битовых плоскостей младших битов модуля коэффициентов, а также свойство самоподобия коэффициентов на разных масштабах, как и EZW. Важным отличием SPIHT от EZW является способ разбиения деревьев коэффициентов и то, как передается информация о значимости.

Рассматриваемый здесь способ применяет структуру квадродеревьев для кодирования положения «значимых» коэффициентов в вейвлет-изображении. Изображение делится на несколько двумерных однородных (с точки зрения степени корреляции соседних отсчетов) блоков переменного размера. Кодирование черно-белых изображений, основанное на сходных принципах разбиения и линейной модели функции яркости внутри блоков, было представлено в работе [4]. Более совершенный алгоритм опубликован в работе [5], где пороговые значения выбирались с помощью решения задачи оптимизации. Алгоритмы для построения оптимальной структуры деревьев с помощью оптимизации, где в качестве критерия использовалось среднеквадратичное отклонение и абсолютная разность изображений, были рассмотрены в [6].

**Алгоритм SQP**

В отличие от работ [4], [5], [6], в рассматриваемом алгоритме используется декомпозиция с помощью квадродеревьев в пространстве вейвлет – образа. Для определения значимости вейвлет – коэффициентов по отношению к множеству монотонно убывающих пороговых значений применяется последовательное приближение с помощью шагов квантования. Обозначим через  $(p,q)$  координаты некоторого коэффициента  $w(p,q)$  в вейвлет – образе, где  $p$  и  $q$  обозначают соответственно строку и

столбец в массиве коэффициентов. Значимость  $w(p,q)$  относительно заданного порогового значения  $T$  хранится в бинарной матрице значимости в элементе с координатами  $(p,q)$ . Значимым коэффициентам соответствует значение «1» в матрице значимости, незначимым – «0».

Пусть  $W_k$  - квадратная область шириной  $v$ , содержащая вейвлет-коэффициенты, расположенные в некоторой подполосе  $W_k \in \{W X_k, W Y_k, W X Y_k, 1 \leq k \leq J\}$ . Предположим, что ширина  $v$  этой области является степенью 2, т.е. удовлетворяет условию  $v = 2^\alpha, \alpha \in \mathbb{N}$ . Также обозначим через  $i$  и  $j$  индексы, соответствующие левому верхнему углу  $W_k$ , как показано на рис. 5.

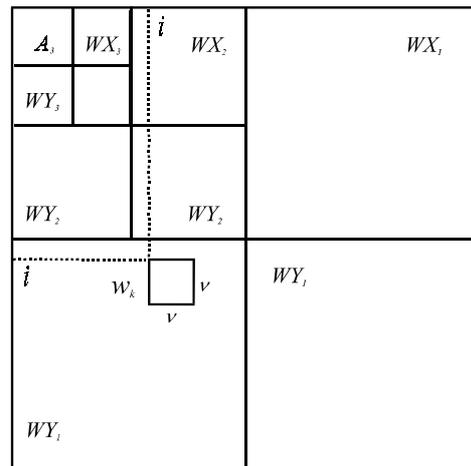


Рис. 5. Обозначение блоков

Определим матрицу размером  $v \times v$  содержащую вейвлет-коэффициенты  $w(p,q) \in W_k$ , расположенные в области  $W_k$ :

$$SQ(i,j,v) = w(p,q), i \leq p < i+v-1, j \leq q < j+v-1. \tag{1}$$

Значимость относительно порогового значения  $T$  коэффициентов, принадлежащих  $SQ(i,j,v)$ , определяется функцией  $f$ , которая отображает любое значение  $w(p,q)$  из  $SQ(i,j,v)$  в множество  $\{0,1\}$ :

$$\forall w(p,q) \in SQ(i,j,v), f(w(p,q)) = \begin{cases} 0, & \text{если } |w(p,q)| < T \\ 1, & \text{если } |w(p,q)| \geq T \end{cases} \tag{2}$$

Бинарная матрица, соответствующая  $SQ(i,j,v)$  и содержащая информацию о значимости коэффициентов относительно  $T$ , определена как

$$SQ_b(i,j,v) = f(w(p,q)), \tag{3}$$

$$i \leq p < i+v-1, j \leq q < j+v-1.$$

Зададим правило, разбивающее матрицу  $SQ_b(i,j,v)$  на четыре соседних минора:

$$P : SQ_b(i,j,v) = \begin{pmatrix} SQ_b(i,j,v/2) & SQ_b(i,j+v/2,v/2) \\ SQ_b(i+v/2,j,v/2) & SQ_b(i+v/2,j+v/2,v/2) \end{pmatrix}. \tag{4}$$



С помощью правила  $P$  матрица  $SQ_b(l,j,v)$  разбивается на 4 минора, а затем то же правило рекурсивно применяется к тем минорам, которые содержат ненулевые элементы. Процесс разбиения останавливается, когда все ненулевые элементы  $SQ_b(l,j,v)$  локализованы, то есть когда все миноры, содержащие ненулевые элементы имеют вид  $SQ_b(m,n,1)$ . С помощью правила  $P$  координаты  $(m,n)$  всех ненулевых элементов матрицы значимости однозначно определены.

Правило разбиения  $P$  представляет матрицу  $SQ_b(l,j,v)$  в виде иерархической структуры матриц. Исходя из этого, можно определить древовидную структуру, узлы которой соответствуют каждой матрице, разбиваемой с помощью  $P$ , и четыре потомка каждого узла соответствуют минорам, получаемым с помощью  $P$ . Пример для матрицы  $SQ_b(l,j,4)$  размером  $4 \times 4$  (рис. 6.а) приведен на рис. 6.б. Из рисунка видно, как правило  $P$  применяется только к минорам, содержащим ненулевые элементы.

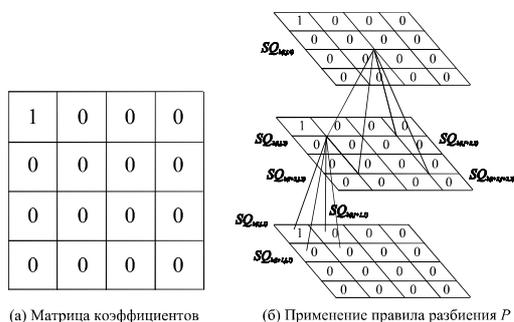


Рис. 6. Построение древовидной структуры символов

Кодирование координат значимых коэффициентов в матрице значимости фактически заключается в кодировании дерева матриц. Для этого, как кодер, так и декодер должны разделять матрицы, содержащие только нулевые элементы и матрицы, подлежащие разбиению. Каждая нулевая матрица обозначается символом «<незначимый>» (NSG), который показывает, что значимость соответствующих коэффициентов можно точно предсказать. С другой стороны, любая матрица, содержащая ненулевые элементы, обозначается символом «<значимый>» (SGN), что означает что хотя бы один коэффициент матрицы будет значимым.

Дерево матриц может быть представлено в виде структуры соответствующих символов. Например, на рис. 7 приведена трехмерная структура символов для рис. 6.а.

Для получения выходного потока символов (списка значимости) необходима некоторая процедура для обхода трехмерного дерева символов и сохранения их в поток. При описании алгоритма обхода дерева введем понятия «прямых потомков» узла дерева - это четыре узла следующего уровня (сам узел для них является «родительским»), и для любого узла - «прямого потомка» (другие 3 «прямых потомка» являются «соседними»).

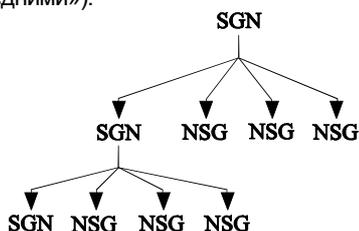


Рис. 7. Пример древовидной структуры символов

Процедура обхода выглядит следующим образом.

1. Инициализация: начинаем с корня дерева (самый высокий уровень,  $i=Q$ ):

1.1 если узел помечен символом «NSG», то добавляем его в список значимости и прекращаем процедуру обхода, поскольку у узла нет потомков;

1.2 если узел помечен символом «SGN», то добавляем его в список значимости и переходим к шагу 2.1.

2. Рекурсивный блок:

2.1 выбираем с уровня  $(i-1)$  четыре узла-потомка, соответствующих текущему узлу уровня  $i$ , и уменьшаем  $i$  на единицу:  $i \leftarrow i-1$ ;

2.2 затем начинаем обход узлов, выбранных в шагах 2.1 или 2.4г (в растровом порядке);

2.3 для каждого обходимого узла с уровня  $i$ , который не был добавлен в список значимости:

а. добавляем соответствующий символ в список значимости;

б. если символ - «SGN», то:

• если  $i \neq 1$  (текущий уровень - не самый низкий), переходим к шагу 2.1;

• иначе продолжаем;

в. иначе (текущий символ «NSG») продолжаем;

2.4 если все 4 коэффициента с уровня  $i$  были добавлены в список значимости, то:

а. переходим к родительскому узлу;

б.  $i \leftarrow i+1$ ;

в. если  $i=Q$ , прекращаем обход;

г. иначе выбираем набор узлов, который включает текущий и его соседей и переходим к 2.2.

На рис. 8 приведен порядок помещения символов в список значимости для дерева на рис. 7.

Из рисунка видно, что сначала обходится левый верхний минор, так как он, как правило, содержит больше всего информации об изображении (LL подполоса); последним кодируется правый нижний - подполоса HH.

С помощью этого алгоритма области в матрице значимости, содержащие единицы, постепенно локализируются путем группировки нулевых элементов в квадратных матрицах переменного размера. Если значимые коэффициенты сосредоточены в определенных областях матрицы значимости, то большие массивы незначимых коэффициентов могут быть закодированы небольшим количеством бит. В такой ситуации данный алгоритм кодирования коэффициентов должен показывать хорошие результаты, поскольку он способен изолировать важные ненулевые детали изображения, исключая большие незначимые области. С другой стороны, если распределение нулей и единиц в матрице значимости равномерное, кодирование координат ненулевых элементов становится крайне неэффективным, поскольку для этого требуется большое количество шагов разбиения. Однако, на практике ненулевые значения в матрице значимости, как правило, распределены вдоль резких переходов цветов или яркости изображения, а нулевые обычно соответствуют однородным областям.

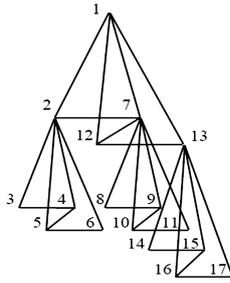


Рисунок 8. Порядок обхода узлов дерева

Для определения значимости коэффициентов относительно набора пороговых значений ( $T_i$ ),  $0 < i < N+1$  используется принцип квантования последовательным приближением. Пороговые значения выбираются таким образом, что  $T_{i+1} = T_i / 2$ , для каждого  $i$ ,  $1 \leq i < N$ . Максимальное значение

$$T_1 = 2^{\lceil \log_2 W_{\max} \rceil},$$

где  $W_{\max}$  – максимальное значение вейвлет - коэффициентов в изображении, взятое по модулю, и  $[x]$  – целая часть  $x$ . Минимальное пороговое значение  $T_N = 1$ , где

$$N = \lceil \log_2 W_{\max} \rceil + 1.$$

Описанный выше алгоритм кодирования матриц значимости используется для кодирования координат значимых коэффициентов  $SQ(0,0,L)$ . Входными данными являются вейвлет – образ ( $SQ(0,0,L)$ ) и текущее пороговое значение  $T_i$ .

Важным моментом является монотонное убывание значений  $T$  от  $T_1$  до  $T_N$ . Если коэффициент оказывается значимым относительно  $T_i$ , то он будет значимым и относительно всех  $T_j$ ,  $j > i$ . Это утверждение верно и для матриц: если в процессе кодирования в какой-то момент матрице присвоен символ SGN (поскольку она содержит значимый коэффициент), то на дальнейших шагах ей будет присваиваться тот же символ. Другими словами, если символ «SGN» был присвоен матрице, то этот символ будет добавлен в список значимости для нее только однажды.

Рассмотрим некоторый коэффициент, незначимый по отношению к множеству значений  $T_j$ ,  $1 \leq j < i$ , и значимый относительно  $T_i$ . Как только его координаты кодируются на шаге  $i$ , к процедуре кодирования добавляется этап уточнения значения коэффициента. Далее, на каждом шаге кодирования  $k$ ,  $k > i$ , значение, соответствующее битовой плоскости  $k$  двоичного представления коэффициента, добавляется в список уточнения. Это эквивалентно последовательному уточнению значения модуля коэффициента на стороне кодера на каждом шаге кодирования  $k$ ,  $k > i$ .

При кодировании используется еще один список, который содержит знаки коэффициентов. Добавление значений в него происходит после кодирования положения элемента в списке значимости. Заметим, что при данном подходе сначала кодируются коэффициенты с максимальным абсолютным значением. Таким образом достигается эффект последовательного вложенного кодирования, которое позволяет управлять точностью

передачи коэффициентов, передавая первые  $N$  бит из потока. Кодирование начинается с обработки коэффициентов самых верхних уровней пирамиды вейвлет – разложения и заканчивается самыми низкими уровнями.

Очевидно, декодирование потока производится аналогично процессу кодирования: процесс повторяется, для построения деревьев матриц используется информация из полученного битового потока. Для правильного декодирования принимающая сторона должна повторить в точности процесс кодирования – для этого ей нужно знать результаты тестирования блоков на значимость, которые она читает из переданного битового потока.

Необходимо заметить, что данный алгоритм может быть легко использован не только для сжатия двумерных сигналов. Например, в трехмерном случае вместо матрицы коэффициентов рекурсивно разбивается трехмерный массив на 8 подмассивов, и т.д.

Эффективность кодирования может быть повышена путем кодирования списков с помощью некоторого энтропийного кодера, для этой цели использовался адаптивный арифметический кодер.

### Результаты моделирования

На рисунках 10, 11 приведены графики зависимости качества результата сжатия (по критерию пикового отношения сигнал/шум - PSNR) от количества переданных бит, а также аналогичные показатели для других распространенных алгоритмов сжатия с потерями (JPEG и JPEG2000). Вейвлет – декомпозиция производилась с помощью банка фильтров 9/7. Надо заметить, что сложность данного алгоритма существенно меньше, чем у остальных рассмотренных.

Хотя алгоритм несколько проигрывает более сложному и медленному алгоритму JPEG2000, он значительно лучше JPEG (и сравним с ним по скорости), особенно для высоких степеней сжатия.

Для тестирования использовались четыре изображения размером  $512 \times 512$  в градациях серого (из набора тестовых изображений Waterloo Grayset). Сами изображения приведены на рис. 9.

В табл. 1 представлены результаты режимов работы кодера изображения «Mandrill» с арифметическим кодированием и без него. Несложно заметить, что выигрыш от арифметического кодирования (случай с АК) оказался незначительным, в числовом отношении он не превышает 0.1 – 0.12 дБ по показателю PSNR. Это вызвано, прежде всего, сложностью точного предсказания появления тех или иных символов в битовом потоке (и, соответственно, построения оптимальной статистической модели для адаптивного кодера). На практике эти показатели можно несколько улучшить, но для этого требуется усложнить функцию, моделирующую статистическую модель распределения символов в потоке, что повлечет за собой снижение производительности кодера. Например, в работе [7] использование более эффективного арифметического кодера привело к увеличению показателя PSNR на 0.15–0.25 дБ, которое также трудно назвать значительным.

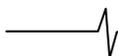


Таблица 1. Выигрыш от арифметического кодирования.

Бит/пиксел	1	0.5	0.25	0.125
Результат с АК (PSNR), дБ	27,372	24,427	22,711	21,284
Результат без АК (PSNR), дБ	27,314	24,37	22,604	21,254



Рис. 9. Тестовые изображения

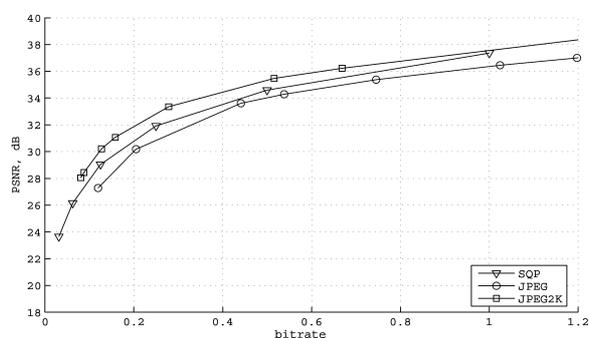


Рис. 10. Зависимость PSNR от количества переданных бит для тестового изображения «Peppers»

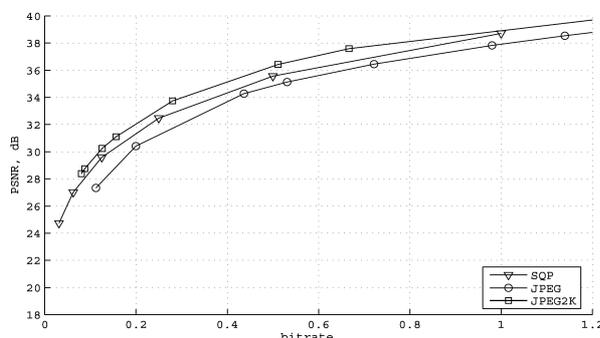


Рис. 11. Зависимость PSNR от количества переданных бит для тестового изображения «Lena»

Данный метод обладает набором важных свойств: он позволяет сжимать широкий спектр изображений, полученных из различных источников (цифровая фотография, рентгенография, и др.), гибко варьируя степень сжатия в зависимости от требуемого качества результата (для сжатия с потерями), а также эффективно сжимать изображения без потерь (в случае, когда передается весь выходной поток кодера). Кодер обладает свойством вложенного кодирования, которое позволяет максимально быстро передать копию изображения с низким качеством, но с сохранением основных черт, и затем итеративно передавать остальные детали.

Представленный алгоритм немного проигрывает по степени сжатия новейшим стандартам, таким как JPEG2000. По сути, он является несколько упрощенным их вариантом. Однако, важным качеством является скорость кодирования, хорошая адаптируемость под архитектуры современных процессоров (в том числе тех, которые используются в различных встраиваемых системах) за счет минимизации случайного доступа к областям памяти. В системах с кэш-памятью это может давать очень большой выигрыш в производительности. Данный алгоритм также легко адаптировать под задачи сжатия многомерных сигналов, которые в настоящее время становятся всё более актуальными, например, в медицинских приложениях. Также довольно очевидны пути адаптации настоящего метода под задачи сжатия очень больших объемов данных, поскольку именно про-

стая структура алгоритма позволяет варьировать его параметры под широкий круг задач.

#### Список литературы

1. J. M. Shapiro, «Embedded image coding using zerotrees of wavelet coefficients», IEEE Trans. Signal Processing 41, 3445–3462 (1993).
2. A. Said, W. Pearlman, «A new fast and efficient image codec based on set partitioning in hierarchical trees», IEEE Trans. on Circuits and Systems Video Technology, 6, 243–250 (1996).
3. Черников А., Чобану М. Современный метод сжатия изображений на базе вейвлет-преобразования и иерархического алгоритма кодирования. «Цифровая обработка сигналов», 2005, 2(16), сс. 40-49.
4. P. Strobach, «Quadtree-structured recursive plane decomposition coding of images», IEEE Trans. Signal Processing, 39, 1380–1397 (1991).
5. E. Shusterman and M. Feder, «Image compression via improved quadtree decomposition algorithms», IEEE Trans. Image Processing, 3, 207–215 (1994).
6. G. J. Sullivan and R. L. Baker, «Efficient quadtree coding of images and video», IEEE Trans. Image Processing, 3, 327–331 (1994).
7. A. Munteanu, J. Cornelis, G. Van der Auwera, P. Cristea «A Wavelet Based Lossless Compression Scheme with Progressive Transmission Capability» 1999.

УДК 621.397.2

## ПРИМЕНЕНИЕ НЕРАЗДЕЛИМЫХ ОПТИМИЗИРОВАННЫХ ВЕЙВЛЕТ-ФИЛЬТРОВ В ЗАДАЧАХ СЖАТИЯ ИЗОБРАЖЕНИЙ

Кобелев В.Ю., Приоров А.Л.

### Введение

Методы обработки изображений с использованием вейвлетов получают все большее распространение [1-5]. Не в последнюю очередь это связано с разработкой эффективных алгоритмов вейвлет-разложения [4, 5] и представления вейвлет-фильтров с оптимизированными частотно-временными свойствами [3, 6, 7].

Способом обработки изображений, получившим наибольшее распространение, является последовательное разложение изображений по строкам и столбцам. Применительно к двумерным вейвлет-фильтрам такой вид обработки соответствует использованию двумерных разделимых фильтров, импульсная характеристика которых есть тензорное произведение импульсных характеристик соответствующих одномерных фильтров [8, 9].

Более сложный вариант синтеза двумерных вейвлет-функций – выбор неразделимого базиса. Неразделимый вейвлет-фильтр позволяет более гибко подойти к задаче синтеза вейвлет-фильтра с заданной частотной избирательностью. Классический способ представления неразделимого вейвлет-базиса – построение двумерной масштабирующей функции  $\varphi(x, y)$  и соответствующих ей трех ортогональных вейвлет-функций  $\psi_1(x, y)$ ,  $\psi_2(x, y)$ ,  $\psi_3(x, y)$  [8].

Такое представление неразделимого вейвлет-базиса не является единственным, в литературе представлены также алгоритмы формирования двумерных неразделимых вейвлет-базисов с использованием масштабирующей функции  $\varphi(x, y)$  и одной или двух вейвлет-функций (например, методы двумерного лифтинга [8], “шахматная” схема разложения [5]).

### Расчет вейвлет-фильтров с оптимизированной частотной избирательностью

Пусть  $F^0$  – исходное изображение размером  $N_1 \times N_2$  пикселей,  $F^1$  – сжатое в 4 раза изображение,  $C'_{lm}$  – импульсная характеристика двумерного вейвлет-фильтра, тогда связь между этими компонентами представим в виде [8]:

$$F^1_{jk} = \sum_{l,m} F^0_{2j+1, 2k+m} C'_{lm} \quad (1)$$

Приводится расчет оптимизированного двумерного неразделимого вейвлет-базиса, соответствующего классической схеме разложения на четыре частотные плоскости. Представленный алгоритм синтеза оптимального двумерного вейвлет-фильтра осуществляет оптимизацию его частотной избирательности в соответствии с амплитудным спектром изображения с целью минимизации среднеквадратичной ошибки восстановленного изображения, причем при восстановлении используются только низкочастотные коэффициенты разложения. Рассмотрены также проблемы применения двумерных вейвлет-фильтров с оптимизированной частотной избирательностью для многоуровневого вейвлет-разложения изображения.

Соотношение (1) – свертка с последующей децимацией. Через операцию фильтрации оно может быть представлено следующим образом:

$$F^1(j\omega_1, j\omega_2) = F^0(j\omega_1, j\omega_2) H'(j\omega_1, j\omega_2) * K(j\omega_1, j\omega_2), \quad (2)$$

где  $H'(j\omega_1, j\omega_2)$  – Фурье-образ последовательности  $C'_{lm}$ ,

$$C'_{lm} = C_{N_1-l, N_2-m}, \text{ а } K(j\omega_1, j\omega_2) -$$

Фурье-образ последовательности  $K_{lm}$ , состоящей из нулей и единиц ( $K_{lm} = 1$ , при  $l, m$  – четных одновременно,  $K_{lm} = 0$  – при прочих значениях  $l, m$ ). Следует отметить, что в выражении (2) размеры двумерных последовательностей  $F^1$  и  $F^0$  одинаковы, поскольку после процедуры децимации недостающие отчеты заменены нулями. Опуская математические выкладки, запишем спектр сжатого сигнала

$$\begin{aligned} F^1(j\omega_1, j\omega_2) = & \frac{1}{2} F^0(j\omega_1, j\omega_2) H'(j\omega_1, j\omega_2) + \\ & \frac{1}{2} F^0(j(\omega_1 + \pi), j\omega_2) H'(j(\omega_1 + \pi), j\omega_2) + \\ & \frac{1}{2} F^0(j\omega_1, j(\omega_2 + \pi)) H'(j\omega_1, j(\omega_2 + \pi)) + \\ & \frac{1}{2} F^0(j(\omega_1 + \pi), j(\omega_2 + \pi)) H'(j(\omega_1 + \pi), j(\omega_2 + \pi)) \end{aligned} \quad (3)$$

Восстановленное изображение определится посредством свертки [8]

$$F^{0*}_{jk} = \sum_{l,m} F^1_{lm} C_{j-2l, k-2m} \quad (4)$$

В спектральной области выражение (4) представимо в виде

$$F^{0*}(j\omega_1, j\omega_2) = F(j\omega_1, j\omega_2) H(j\omega_1, j\omega_2), \quad (5)$$

где  $H(j\omega_1, j\omega_2) = \sum_{n_1, n_2} C_{n_1, n_2} e^{-j\omega_1 n_1} e^{-j\omega_2 n_2}$ .



Применив операцию фильтрации (5) к выражению (3), получим:

$$\begin{aligned}
 2F^{0*}(j\omega_1, j\omega_2) &= 2H(j\omega_1, j\omega_2)F^1(j\omega_1, j\omega_2) = \\
 &= F^0(j\omega_1, j\omega_2)H'(j\omega_1, j\omega_2)H(j\omega_1, j\omega_2) + \\
 &+ F^0(j(\omega_1 + \pi), j\omega_2)H'(j(\omega_1 + \pi), j\omega_2)H(j\omega_1, j\omega_2) + \\
 &+ F^0(j\omega_1, j(\omega_2 + \pi))H'(j\omega_1, j(\omega_2 + \pi))H(j\omega_1, j\omega_2) + \\
 &+ F^0(j(\omega_1 + \pi), j(\omega_2 + \pi))H'(j(\omega_1 + \pi), j(\omega_2 + \pi))H(j\omega_1, j\omega_2).
 \end{aligned} \quad (6)$$

При сопоставлении (6) со спектром исходного сигнала видно, что в восстановленном сигнале присутствуют компоненты, обусловленные эффектом наложения (последние три слагаемых), отрицательно влияющие на восстановленный сигнал. Частотные характеристики  $H'(j\omega_1, j\omega_2)$  и  $H(j\omega_1, j\omega_2)$

связаны соотношением

$$H'(j\omega_1, j\omega_2) = e^{-j\omega_1(N_1-1)} e^{-j\omega_2(N_2-1)} H(-j\omega_1, -j\omega_2). \quad (7)$$

С целью сокращения на фазовый множитель, сделаем предположение, что восстановленное изображение отличается от оригинала только фазовой задержкой

$$F^{0*}(j\omega_1, j\omega_2) = e^{-j\omega_1(N_1-1)} e^{-j\omega_2(N_2-1)} F^0(j\omega_1, j\omega_2). \quad (8)$$

После подстановки (7), (8) в (6) и некоторых преобразований получим:

$$\begin{aligned}
 2F^0(j\omega_1, j\omega_2) &= F^0(j\omega_1, j\omega_2)H(-j\omega_1, -j\omega_2)H(j\omega_1, j\omega_2) + \\
 &+ F^0(j(\omega_1 + \pi), j\omega_2)H(-j(\omega_1 + \pi), -j\omega_2)H(j\omega_1, j\omega_2) + \\
 &+ F^0(j\omega_1, j(\omega_2 + \pi))H(-j\omega_1, -j(\omega_2 + \pi))H(j\omega_1, j\omega_2) + \\
 &+ F^0(j(\omega_1 + \pi), j(\omega_2 + \pi))H(-j(\omega_1 + \pi), -j(\omega_2 + \pi))H(j\omega_1, j\omega_2).
 \end{aligned} \quad (9)$$

В связи с тем, что рассматриваемые вейвлет-функции являются ортогональными, дополнительно на частотные свойства соответствующих им вейвлет-фильтров накладывается ограничение (условие квадратурно-зеркальности) [8]:

$$\begin{aligned}
 |H(j\omega_1, j\omega_2)|^2 &+ |H(j\omega_1 + \pi, j\omega_2)|^2 + \\
 |H(j\omega_1, j\omega_2 + \pi)|^2 &+ |H(j\omega_1 + \pi, j\omega_2 + \pi)|^2 = 1
 \end{aligned} \quad (10)$$



а)

Уравнение (9) с учетом (10) может быть преобразовано к виду

$$\begin{cases}
 F^0(j\omega_1, j\omega_2)H(j(\omega_1 + \pi), j\omega_2) - F^0(j(\omega_1 + \pi), j\omega_2)H(j\omega_1, j\omega_2) = 0; \\
 F^0(j\omega_1, j\omega_2)H(j\omega_1, j(\omega_2 + \pi)) - F^0(j\omega_1, j(\omega_2 + \pi))H(j\omega_1, j\omega_2) = 0; \\
 F^0(j\omega_1, j\omega_2)H(j(\omega_1 + \pi), j(\omega_2 + \pi)) - F^0(j(\omega_1 + \pi), j(\omega_2 + \pi))H(j\omega_1, j\omega_2) = 0.
 \end{cases} \quad (11)$$

Рассмотрим модуль выражений (11)

$$\begin{cases}
 F^0(\omega_1, \omega_2)H(\omega_1 + \pi, \omega_2) = F^0(\omega_1 + \pi, \omega_2)H(\omega_1, \omega_2); \\
 F^0(\omega_1, \omega_2)H(\omega_1, \omega_2 + \pi) = F^0(\omega_1, \omega_2 + \pi)H(\omega_1, \omega_2); \\
 F^0(\omega_1, \omega_2)H(\omega_1 + \pi, \omega_2 + \pi) = F^0(\omega_1 + \pi, \omega_2 + \pi)H(\omega_1, \omega_2).
 \end{cases} \quad (12)$$

Возведем в квадрат правые и левые части системы (12), сложим почленно и применим свойство квадратурно-зеркальности (10). В результате получим конечное выражение

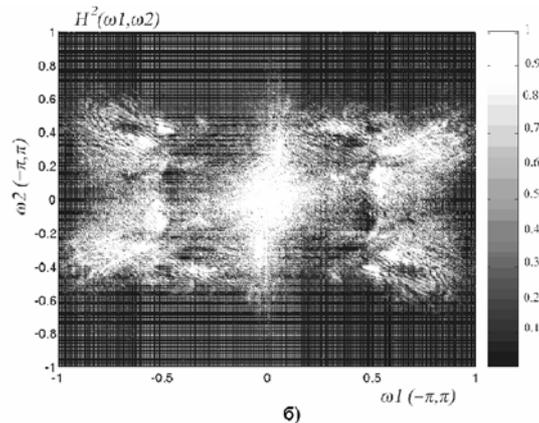
$$H^2(\omega_1, \omega_2) = \frac{F^2(\omega_1, \omega_2)}{F^2(\omega_1 + \pi, \omega_2) + F^2(\omega_1, \omega_2 + \pi) + F^2(\omega_1 + \pi, \omega_2 + \pi) + F^2(\omega_1, \omega_2)}. \quad (13)$$

Выражение (13) позволяет найти амплитудно-частотную характеристику (АЧХ) двумерного вейвлет-фильтра, обеспечивающего восстановление изображений, используя только низкочастотные компоненты вейвлет-разложения. Расчет АЧХ фильтра согласно (13) осуществляется только на основании значений амплитудного спектра исходного изображения. На рис. 1 представлены примеры тестовых изображений и расчет квадратов АЧХ соответствующих вейвлет-фильтров  $H^2(\omega_1, \omega_2)$ .

Рассмотрим подробнее процедуру децимации. В нашем случае процедура децимации осуществляется, используя поэлементное умножение значений изображения на оператор  $K_{Im}$ . Знаменатель выражения (13) представляет собой квадрат амплитудного спектра изображения после операции децимации

$$\begin{aligned}
 F_1^2(\omega_1, \omega_2) &= F^2(\omega_1 + \pi, \omega_2) + F^2(\omega_1, \omega_2 + \pi) + \\
 &+ F^2(\omega_1 + \pi, \omega_2 + \pi) + F^2(\omega_1, \omega_2)
 \end{aligned} \quad (14)$$

где  $F_1(\omega_1, \omega_2)$  – амплитудный спектр изображения после децимации,  $F(\omega_1, \omega_2)$  – амплитудный спектр исходного изображения.



б)

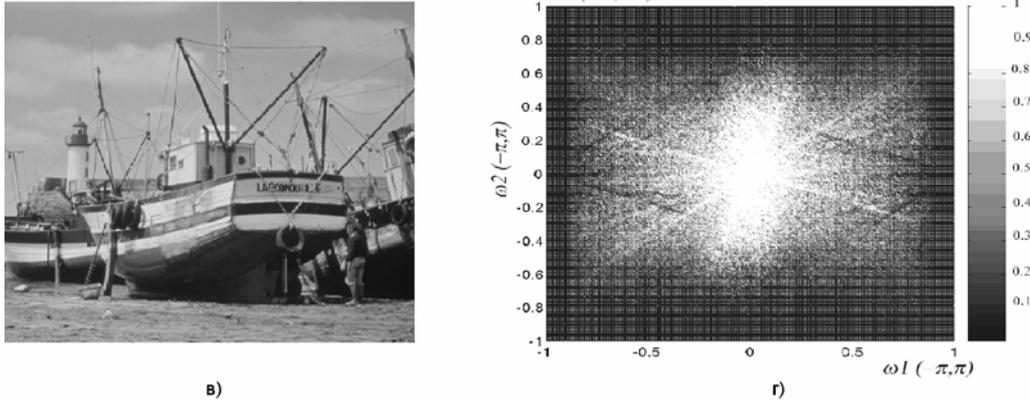


Рис. 1. Тестовые изображения (а, в) и квадраты АЧХ соответствующих “согласованных” вейвлет-фильтров (б, г)

Основное отличие  $F_1(\omega_1, \omega_2)$  от  $F(\omega_1, \omega_2)$  заключается в наличии ярко выраженного эффекта наложения спектра вследствие процедуры децимации. В данном случае мощность ошибки наложения максимальна (т.к. никакие промежуточные фильтры, снижающие воздействие эффектов наложения, принципиально не применялись). Таким образом, выражение (14) описывает изображение, максимально искаженное процедурой сжатия. Фильтр  $H(j\omega_1, j\omega_2)$ , АЧХ которого удовлетворяет выражению (13), получен из соотношения исходного неискаженного изображения и изображения, искаженного сжатием. Он обеспечивает максимально возможное отношение сигнал/шум для восстановленного изображения и попадает под определение согласованного фильтра [3]. Поэтому можно считать, что двумерный вейвлет-фильтр с АЧХ, определяемой выражением (13), представляет согласованный двумерный вейвлет-фильтр. Дополнительно стоит отметить, что полученный вейвлет-фильтр  $H(j\omega_1, j\omega_2)$  является также квадратурно-зеркальным. Из рис. 1 видно, что квадрат АЧХ согласованного вейвлет-фильтра имеет сильно изрезанную структуру, и поэтому его порядок сопоставим с размером исследуемого изображения ( $N \times N$  пикселей).

Следующей задачей является расчет двумерного вейвлет-фильтра с параметрами, близкими к параметрам согласованного вейвлет-фильтра, но имеющего сравнительно низкий порядок. Под термином *двумерный оптимизированный вейвлет-фильтр* представим вейвлет-фильтр  $\tilde{H}(j\omega_1, j\omega_2)$  порядка  $M \times M < N \times N$ , удовлетворяющий условию

$$\lim_{M \rightarrow N} \iint |H^2(\omega_1, \omega_2) - \tilde{H}^2(\omega_1, \omega_2)| d\omega_1 d\omega_2 \rightarrow 0. \quad (15)$$

Таким образом, оптимизированный вейвлет-фильтр понимается как фильтр с оптимизированной частотной избирательностью.

Расчет такого фильтра заданного порядка включает следующие этапы:

- аппроксимацию квадрата АЧХ тригонометрическим полиномом меньшего порядка;
- расчет соответствующей фазовой характеристики фильтра (решение фазовой задачи выполнено с использованием модифицированного алгоритма Герхберга [10]).

На рис. 2 представлен амплитудный спектр оптимизированного вейвлет-фильтра, порядка  $60 \times 60$ , рассчитанный для тестового изображения “Барбара”.

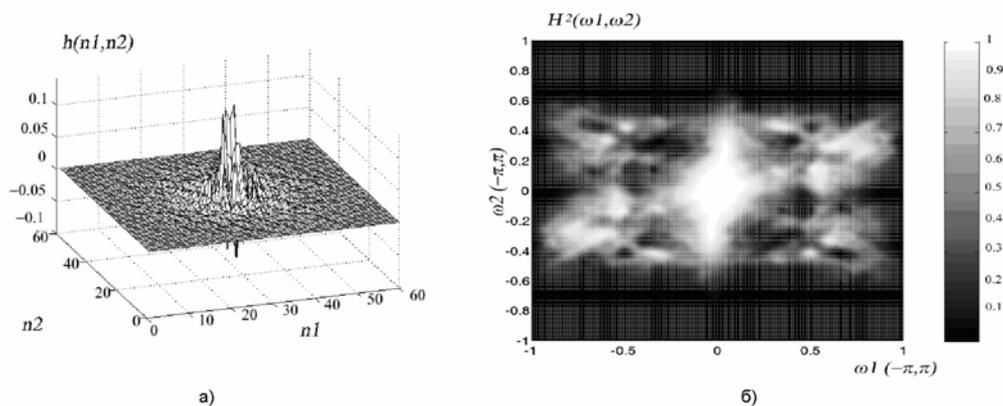


Рис. 2. Пример вейвлет-фильтра с оптимизированной частотной избирательностью: а) - импульсная характеристика фильтра порядка  $60 \times 60$ , б) – квадрат его АЧХ

### Многоуровневое вейвлет-разложение изображений с использованием оптимизированных вейвлет-фильтров

Оценка эффективности применения оптимизированных вейвлет-фильтров к задаче сжатия изображений выполнена на базе большого набора тестовых изображений с использованием как одноуровневой, так и многоуровневой схем обработки. Экспериментально получено, что приемлемый порядок вейвлет-фильтра для обработки изображения, разрешением  $512 \times 512$  пикселей, составляет  $30 \times 30$ . Представленный порядок при достаточно точном описании особенностей амплитудного спектра изображения позволяет ограничиться сравнительно небольшим числом сохраняемых отчетов импульсной характеристики фильтра. В работе за оценку качества восстановленного изображения принято нормированное среднеквадратичное отклонение восстановленного изображения от исходного тестового изображения:

$$\varepsilon_{НСКО} = -10 \lg \frac{\sum_{n=0}^{N-1} \sum_{m=0}^{N-1} (F_{n,m}^0 - F_{n,m}^{0*})^2}{\sum_{n=0}^{N-1} \sum_{m=0}^{N-1} (F_{n,m}^0)^2}, \text{ [дБ]}$$

где  $F_{n,m}^{0*}$  - восстановленное изображение,  $F_{n,m}^0$  - оригинальное изображение.

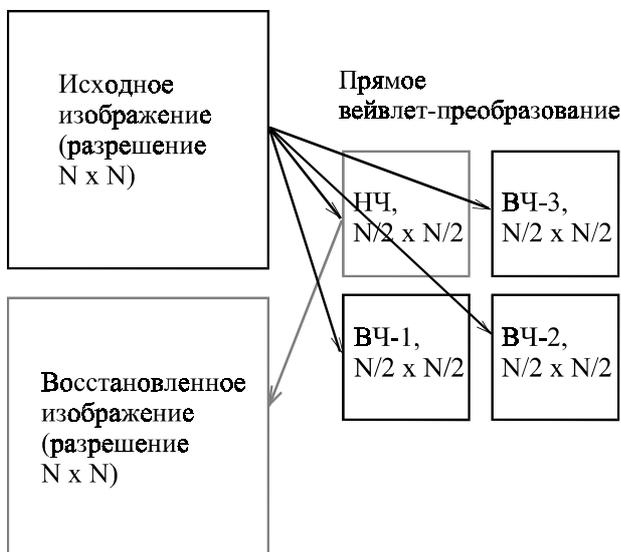


Рис. 3. Схема одноуровневого вейвлет-разложения изображения

Наиболее интересные результаты применения оптимизированных вейвлет-фильтров получены относительно многоуровневой схемы обработки. Относительно одноуровневой схемы (вариант реализации – рис. 3) приведем лишь обобщающие выводы:

- из всего набора стандартных тестовых изображений, наименьшая ошибка НСКО восстановленного изображения получена на тестовом изображении “Барбара” – 21 дБ (при использовании вейвлетов Добеши-8, отношение сигнал/шум восстановленного изображения составило 19,3 дБ). В первую очередь это объясняется наличием ярко выраженных высокочастотных особенностей (“полосы на одежде и скатерти”), для выделения которых происходит оптимизация АЧХ вейвлет-фильтра;
- при обработке прочих стандартных тестовых изображений применение вейвлет-фильтров с оптимизированной АЧХ улучшает отношение С/Ш восстановленного изображения лишь на доли децибел. Это объясняется либо отсутствием ярко выраженных специфических особенностей на изображениях, либо недостаточностью порядка используемого вейвлет-фильтра;
- другая группа изображений, где представляемый алгоритм показал значительный выигрыш (порядка  $2 \div 4$  дБ), – это изображения, восстановленные с использованием высокочастотных коэффициентов вейвлет-разложения (в первом приближении - это контуры стандартных изображений).

На основе приведенных выводов реализация многоуровневого вейвлет-разложения изображений выполнена по следующей схеме (рис. 4). Используя стандартные вейвлет-фильтры (например, Добеши-8), выделяется высокочастотная (ВЧ) часть обрабатываемого изображения на первом уровне разложения. Для обработки ВЧ-части рассчитывается оптимизированный вейвлет-фильтр, позволяющий с минимальными искажениями сжать ВЧ-часть изображения в 4 раза. Далее осуществляется переход на следующий уровень обработки, и указанная процедура повторяется. На рис. 4 представлена двухуровневая схема вейвлет-разложения. В данном случае исходное изображение представляется тремя наборами коэффициентов вейвлет-разложения (обведены кружком):

- низкочастотная часть, разрешением  $128 \times 128$  пикселей – блок А2;
- высокочастотная часть №1, разрешением  $256 \times 256$  пикселей – блок В1;
- высокочастотная часть №2, разрешением  $128 \times 128$  пикселей – блок В2.

Далее полученные коэффициенты вейвлет-разложения (А1, В1, В2) сжимаются с использованием стандартного кодека JPEG. В связи с наличием большого количества элементов с почти нулевыми значениями в высокочастотных блоках В1, В2, достигается их высокий коэффициент сжатия указанным кодеком. В табл. 1 представлены результаты многоуровневой вейвлет-обработки тестового изображения “Барбара”.

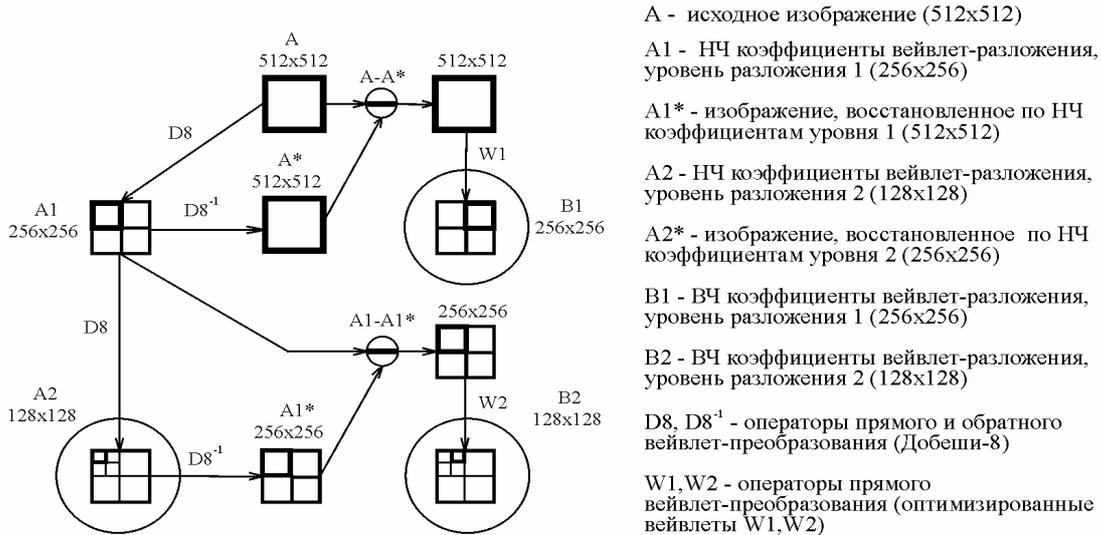


Рис. 4. Схема многоуровневого вейвлет-разложения с использованием вейвлет-фильтров с оптимизированной частотной избирательностью

На рис. 5 представлен результат восстановления тестового изображения “Барбара”. По сравнению со стандартными методами сжатия достигается улучшение отношения С/Ш на 0.9 дБ (используются стандартные

вейвлеты Добеши-8) и 1,7 дБ (используется JPEG). Однако, несмотря на небольшой итоговый выигрыш, воспроизведение высокочастотной части изображения выполнено с гораздо большей точностью.



Рис. 5. Тестовое изображение “Барбара”, восстановленное с использованием оптимизированных вейвлет-фильтров по многоуровневой схеме (брр = 0.48 бит/пиксел, НСКО = 18.23 дБ)

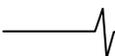
Таблица 1.

Результаты многоуровневой вейвлет-обработки изображения

№	Наименов.	Разрешение (пикс.)	Тип вейвлет фильтра*	Объем ** (Кб)	Сумм. объем (Кб)	Бит на пиксель
1	Блок А1	128 × 128	Доб-8.- 16 × 16	6,42	15,46	0.48
2	Блок В1	256 × 256	Опт.- 30 × 30	5,80		
3	Блок В2	128 × 128	Опт.- 22 × 22	2,84		
4	ИХ фильтра 1	30 × 30		0,20***		
5	ИХ фильтра 2	22 × 22		0,10***		

\* - указан тип и порядок вейвлет-фильтра, используемого для получения вейвлет-коэффициентов в соответствующих блоках;

\*\* - занимаемый коэффициентами объем памяти указан после применения кодера JPEG; \*\*\* - для кодирования импульсной характеристики (ИХ) оптимизированных вейвлет-фильтров использовался только метод неравномерного квантования (центральная область ИХ квантуется на 256 уровней, средняя – на 64 уровня, крайняя – на 8 уровней).



На рис. 6 представлены ВЧ-части изображения (контуры) уровня разложения 1 (а) и уровня разложения 2 (б). Соответственно на рис. 6 (в, г) пред-

ставлены квадраты АЧХ вейвлет-фильтров с оптимизированной избирательностью для представленных частей изображения.

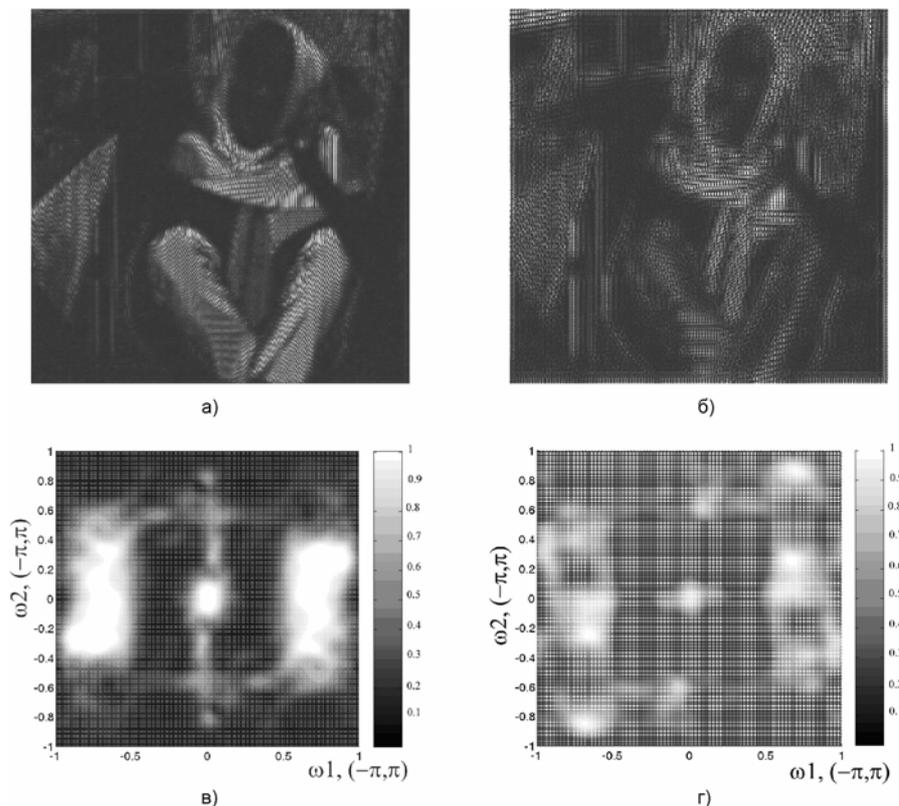


Рис. 6. Квадрат АЧХ оптимизированных вейвлет-фильтров для обработки ВЧ-частей тестового изображения: (а) – уровень разложения 1, (б) – уровень разложения 2; квадраты АЧХ соответствующих вейвлет-фильтров: (в) – порядок фильтра  $30 \times 30$ , (г) – порядок фильтра  $22 \times 22$

### Заключение

На основе полученных результатов можно сделать следующие выводы:

1. Используя закономерности преобразования спектров изображений при сжатии, получено выражение, позволяющее найти АЧХ двумерного вейвлет-фильтра, обеспечивающего качественное восстановление изображений, используя только низкочастотные компоненты вейвлет-разложения. Вейвлет-фильтр со свойством такого восстановления получил название “согласованного”.

2. Выполнен расчет оптимизированного двумерного вейвлет-фильтра на основе АЧХ “согласованного” вейвлет-фильтра, ключевым моментом которого является алгоритм понижения порядка “согласованного” фильтра до заданного пользователем. Приведен алгоритм расчета фазовой характеристики на базе модифицированного итерационного алгоритма Герхберга.

3. Разработанный алгоритм эффективен при обработке высокочастотных частей изображения в рамках многоуровневой схемы вейвлет-разложения, что позволяет добиться улучшения отношения С/Ш восстановленного изображения на  $0.9 \div 1.7$  дБ.

### Литература

1. A. Moislilovic, M.V. Popovic, D. Rackov. On the selection of an optimal wavelet basis for texture characterization // IEEE Trans. Image Process. 2000. V. 9, № 12. P. 2043-2050.

2. Starck J.-L., Murtagh F., Candes E., Donoho D. Gray and color image contrast enhancement by the curvelet transform // IEEE Trans. Image Process. 2003. V. 12, № 6. P. 706-717.

3. Новиков Л.В. Адаптивный вейвлет-анализ сигналов // Научное приборостроение. 1999. Т. 9, № 2. С. 1-13.

4. Sprljan N., Grgic S., Grgic M. Modified SPIHT algorithm for wavelet packet image coding // Real-time Imaging. 2005. № 11. P. 378-388.

5. Толкова Е.И. Wavelet-анализ изображений // Оптический журнал. 2001. Т. 68, № 3. С. 49-59.

7. Кобелев В.Ю., Корепанов И.В., Буралков Д.В. Параметризация и особенности представления вейвлет-фильтров на z-плоскости // Докл. 6-ой междунар. конф. “Цифровая обработка сигналов и ее применение” (DSPA’04). Москва, 2004. Т. 2, С. 122-124.

8. Кобелев В.Ю., Моисеев А.А., Волохов В.А., Смоляков А.В. Синтез двумерных неразделимых вейвлет-фильтров с перестраиваемыми коэффициентами // Докл. 8-ой междунар. конф. “Цифровая обработка сигналов и ее применение” (DSPA’06). Москва, 2006. Т. 2, С. 389-392.

9. Daubechies I. Ten lectures on wavelets. CBMS-NSF conference series in applied mathematics. SIAM Ed., 1992.

10. Чуи К. Введение в вейвлеты. Пер. с англ. - М.: Мир, 2001.

11. Прэтт У. Цифровая обработка изображений. - М.: Мир, 1982. Т. 1, 2.

УДК 621.397

## МОДИФИЦИРОВАННЫЙ КРИТЕРИЙ ОЦЕНКИ КАЧЕСТВА ВОССТАНОВЛЕННЫХ ИЗОБРАЖЕНИЙ

Арлянов С.А., Приоров А.Л., Хрящев В.В.

### Введение

Сложность и дороговизна субъективных экспертных оценок делают привлекательным автоматическое измерение качества с помощью определенных алгоритмов. Многие разработчики систем цифровой обработки изображений часто опираются на так называемые объективные (алгоритмические) меры качества [1-6], которые должны хорошо согласовываться с результатами субъективных оценок для широкого класса изображений, не имея при этом чересчур сложных вычислений. Кроме того, желательно, чтобы эти метрики имели простую аналитическую форму и их можно было бы применять в качестве критериев оптимальности при выборе параметров системы обработки изображений [7, 8].

На данный момент наиболее популярной объективной мерой служит пиковое отношение сигнал/шум (ПОСШ), определяемое по следующей формуле [9]:

$$\text{ПОСШ} = 20 \lg \frac{255}{\sqrt{\text{СКО}}}; \quad \text{СКО} = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2,$$

где  $x_i$  и  $y_i$  - значения пикселей исходного и восстановленного изображений соответственно, а СКО - среднеквадратичная ошибка. Величина ПОСШ обычно используется для сравнения различных алгоритмов обработки, например, для сравнения производительности различных методов сжатия или для изучения влияния разных параметров на производительность того или иного алгоритма [10].

Широкое применение СКО и ПОСШ в практических задачах обусловлено тем, что подсчет этих характеристик математически прост и не требует больших вычислительных затрат. Однако необходимо отметить и то, что величина ПОСШ не может в полной мере отразить воздействие на изображение различных видов помех [11], то есть при наличии в изображении разных видов шумов ее значение может оставаться одним и тем же, хотя качество изображения при этом может существенно изменяться. Следовательно, величина ПОСШ плохо согласуется с субъективными критериями качества, определяемыми рекомендацией ITU-R 500 [12], и не является универсальной. Все это говорит о том, что при построении систем цифровой обработки изображений необходимо пользоваться другими метриками оценки качества [13, 14].

В последние годы выполнено множество работ по созданию более приемлемых алгоритмических мер визуального качества обрабатываемых изображений, которые были бы близки к результатам субъективных оценок [15]. Предлагается использование универсального индекса качества (УИК), определяющего подобие восстановленного изображения некоторому оригиналу.

*Рассматривается применение нового метода оценки качества в задачах восстановления цифровых изображений. В отличие от широко применяемых алгоритмов на основе среднеквадратичной ошибки, предлагаемый подход учитывает искажения яркости и контраста, а также степень коррелированности между двумя изображениями. Результаты моделирования показывают хорошую коррелированность данного критерия с визуально воспринимаемым качеством изображений.*

Универсальность данного критерия заключается в том, что он отражает не только некоторую схожесть обработанного изображения по отношению к оригиналу, но и должным образом учитывает различные виды искажений. В отличие от алгоритмов, предложенных в работах [16, 17], данный метод не привязан к специфике изображения и искажениям, присутствующим в нем. Он основывается на статистическом анализе отдельных блоков входного сигнала и дальнейшем сравнении полученных результатов со значениями эталонного изображения.

### Определение УИК для одномерных сигналов

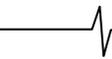
Рассмотрим две последовательности  $\{x_i\}$  и  $\{y_i\}$ , где  $i = 1, 2, \dots, N$ , соответствующие стационарному тестовому сигналу и его оригиналу. Тогда универсальный индекс качества будет определяться комбинацией статистических характеристик соответствующих последовательностей по следующей формуле [18]:

$$\text{УИК} = \frac{4 \sigma_{xy} \bar{x} \bar{y}}{[\sigma_x^2 + \sigma_y^2][(\bar{x})^2 + (\bar{y})^2]}, \quad (1)$$

где

$$\begin{aligned} \bar{x} &= \frac{1}{N} \sum_{i=1}^N x_i, \quad \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i, \\ \sigma_x^2 &= \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2, \quad \sigma_y^2 = \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2, \\ \sigma_{xy} &= \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) \end{aligned}$$

Таким образом, получается, что значение УИК изменяется в интервале  $[-1, 1]$ , при этом УИК = 1 соответствует наилучшему качеству сигнала. Это возможно только тогда, когда  $y_i = x_i$  ( $i = 1, 2, \dots, N$ ). Минимальное значение УИК = -1 достигается в случае, если  $y_i = 2\bar{x} - x_i$ , где  $i = 1, 2, \dots, N$ .



Определенный таким образом индекс качества учитывает в себе три искажающих фактора: степень коррелированности отчетов двух сигналов, изменение значений математического ожидания и среднеквадратичных отклонений сигнала относительно оригинала. Такое представление позволяет наиболее точно оценивать качество восстановленного сигнала и величину искажений, присутствующих в нем. В результате (1) может быть представлено в виде произведения трех множителей:

$$УИК = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \cdot \frac{2\bar{x}\bar{y}}{(\bar{x})^2 + (\bar{y})^2} \cdot \frac{2\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}. \quad (2)$$

Первый множитель - коэффициент взаимной корреляции сигналов  $x$  и  $y$ , определяющий степень их линейной подобности, принимает значения в интервале  $[-1, 1]$ , его максимальное значение достигается в случае, если  $y_i = ax_i + b$ , где  $a > 0$ ,  $b > 0$ . Второй множитель определяет соотношение между математическими ожиданиями входного и восстановленного сигналов и принимает значения в интервале  $[0, 1]$ , при этом максимум достигается в случае их равенства. Если сигналом является изображение, этот коэффициент отвечает за изменения его яркости. Значения среднеквадратичных отклонений  $\sigma_x$  и  $\sigma_y$  определяют динамический диапазон изменения сигналов  $x$  и  $y$  соответственно, вследствие чего они могут рассматриваться как характеристики контраста, если сигналом является изображение. Таким образом, третий множитель в выражении (2) соответствует искажению контрастности восстановленного сигнала относительно оригинала и принимает значения в интервале  $[0, 1]$ .

Предложенный индекс качества обладает хорошей согласованностью со значениями субъективных оценок и имеет абсолютное значение, равное единице.

### Применение УИК для оценки качества изображений

Изображение является двумерным нестационарным процессом и прямое использование формул (1,2) для вычисления УИК в данном случае нерационально. Для решения этой проблемы предлагается выделять в изображении локальные области размером  $A \times B$ , в пределах которых сигнал можно считать стационарным и вычислять статистические характеристики внутри этих областей, а затем с их использованием вычислять значение  $УИК_i$ , соответствующее выделенному блоку. Такой подход обеспечивает наиболее правильное и точное суждение о качестве восстановленного изображения.

В результате, процесс нахождения УИК для конкретного изображения будет сводиться к следующему: «скользящее окно» размером  $A \times B$ , начиная из верхнего левого угла, проходит пиксель за пикселем по всем строкам и столбцам этого изображения, при этом на каждом шаге вычисляется значение статистических характеристик и собственно сам  $УИК_i$  для текущего положения окна. Общий УИК, характеризующий качество изображения в целом, будет определяться как среднее арифметическое значений частных  $УИК_i$ :

$$УИК = \frac{1}{M} \sum_{i=1}^M УИК_i,$$

где  $M$  - число блоков, внутри которых вычислялся УИК. Такой подход к нахождению численных оценок качества особенно хорош для изображений, имеющих блочную структуру. Как известно, одним из таких типов изображений является формат JPEG, основанный на применении дискретного косинусного преобразования (ДКП) отдельных блоков размером  $8 \times 8$  пикселей. В результате этого для оценки качества, например JPEG-изображений, возможно использование УИК с размером «скользящего окна» идентичного размеру блоков ДКП.

Проиллюстрируем применение предложенного подхода к оценке качества на некоторых задачах цифровой обработки изображений.

### Оценка качества изображений при равенстве ПОСШ

Проведем сравнительный анализ значений УИК с равными значениями ПОСШ для ряда изображений, подверженных различным искажающим факторам. На рис. 1 представлены изображения с различными видами искажений, при этом величина ПОСШ для всех случаев приблизительно одинакова и составляет 23.4 дБ. Это говорит о том, что с точки зрения данной метрики изображения считаются визуально равноценными. Таким образом, получается, что, например, качество изображения с явно выраженной блочной структурой, обусловленной высоким коэффициентом сжатия JPEG-кодера (рис. 1е), является таким же, как для изображения с видоизмененной гистограммой (рис. 1б), хотя очевидно, что качество последнего изображения существенно выше. Если же судить о качестве изображений с точки зрения предложенного УИК, то видно, что изображение с блочной структурой имеет более низкое качество, чем изображение с трансформированной гистограммой. Для других изображений рис. 1 показания УИК также соответствуют субъективному восприятию.

### Оценка яркостных искажений

Как было отмечено ранее, в некоторых случаях обработки изображений возникают такие ситуации, когда показания ПОСШ и СКО совершенно не соответствуют субъективному восприятию, например, при видоизменении гистограмм, гамма-коррекции и других операциях изменения яркости. При линеаризации гистограммы изображения его качество визуально должно улучшаться благодаря повышению контраста и уменьшаться при его ослаблении. Однако на практике при численной оценке таких изображений с помощью стандартных метрик качества ПОСШ и СКО для некоторых изображений возникают противоречивые результаты, то есть изображение с низким контрастом оценивается как изображение с более высоким качеством. Типичный пример такой ситуации представлен на рис. 2. При оценке предложенных изображений с помощью УИК полученные результаты в большей степени соответствуют субъективной оценке и не противоречат суждению о превосходстве качества высококонтрастных изображений над изображениями с низким контрастом.



а) УИК = 1,0



б) УИК = 0,89; ПОСШ = 23,4 дБ



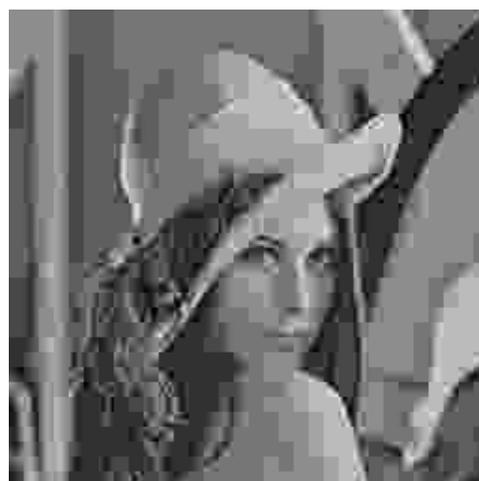
в) УИК = 0,78; ПОСШ = 23,4 дБ



г) УИК = 0,67; ПОСШ = 23,4 дБ

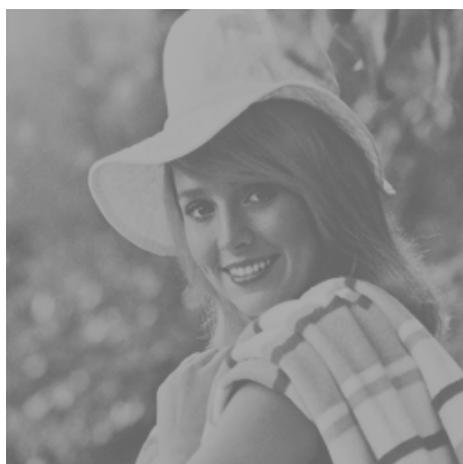
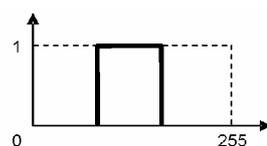
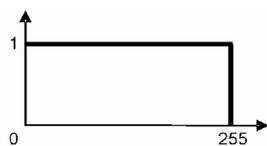


д) УИК = 0,43; ПОСШ = 23,4 дБ



е) УИК = 0,39; ПОСШ = 23,4 дБ

Рис. 1. Оценка качества изображений по значению УИК 7x7 при равенстве ПОСШ: а) оригинал; б) видоизменение гистограммы; в) обработка ФВЧ; г) влияние импульсного шума; д) влияние «белого шума»; е) JPEG-кодирование с высоким коэффициентом сжатия



а) УИК  $7 \times 7 = 0,80$ ; ПОСШ = 18,5 дБ

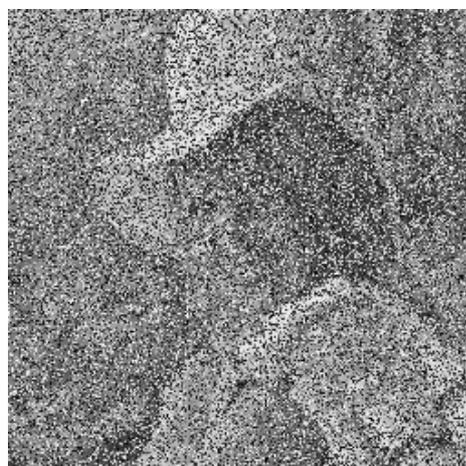
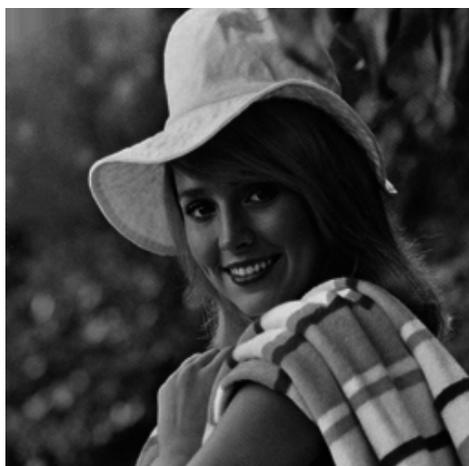
б) УИК  $7 \times 7 = 0,71$ ; ПОСШ = 19,9 дБ

Рис. 2. Оценка качества изображений с трансформацией гистограмм: а) линейаризация гистограммы; б) усечение гистограммы

Применение УИК особенно важно при гамма-коррекции изображений. Это связано с тем, что в большинстве таких случаев общая яркость изображения либо возрастает, либо уменьшается, в результате этого показателя СКО принимают достаточно большие значения (порядка 10000) и приводят к существенному уменьшению ПОСШ (менее 10 дБ), что свидетельствует об очень плохом качестве изображений. С точки зрения ПОСШ такие результаты соответствуют значительным искажениям (например, присутствию 60% импульсного шума с произвольными значениями выбросов) в изображении и соответственно указывают на отсутствие в нем какой-либо полезной информации. Однако на самом деле в случае гамма-коррекции только при достаточно малых значениях  $\gamma < 0.2$  происходит значимая потеря информации, содержащейся в изображении, в остальных случаях общее представление об изображении сохраняется, и говорить об отсутствии в нем полезной информации нельзя. Например, исходя

из показаний ПОСШ (рис. 3), можно сказать, что изображение, зашумленное импульсным шумом, обладает более высоким качеством, чем изображение после гамма-коррекции. Это говорит о том, что гамма-коррекция вносит искажения, измеримые с действием импульсного шума, что в свою очередь противоречит действительности, так как понятно, что качество первого изображения рис. 3а намного выше второго - рис. 3б. Показания УИК наиболее правильно характеризуют качество представленных изображений и соответствуют визуальному восприятию человека в большей степени.

На рис. 4. представлены графики зависимости качества изображения «LENA» от значения показателя  $\gamma$ . Из них следует, что в результате гамма-коррекции даже при небольших значениях  $\gamma \sim 0.7$ , при использовании метрики ПОСШ, качество изображения существенно ухудшается (ПОСШ < 19 дБ), хотя изображение при этом визуально практически не отличается от оригинала.



а) УИК  $7 \times 7 = 0,62$ ; ПОСШ = 10,86 дБ

б) УИК  $7 \times 7 = 0,05$ ; ПОСШ = 11,50 дБ

Рис. 3. Изображения с малым значением ПОСШ: а) гамма-коррекция изображения  $\gamma < 0.40$ ; б) изображение с 60% импульсным шумом со случайными значениями импульсов

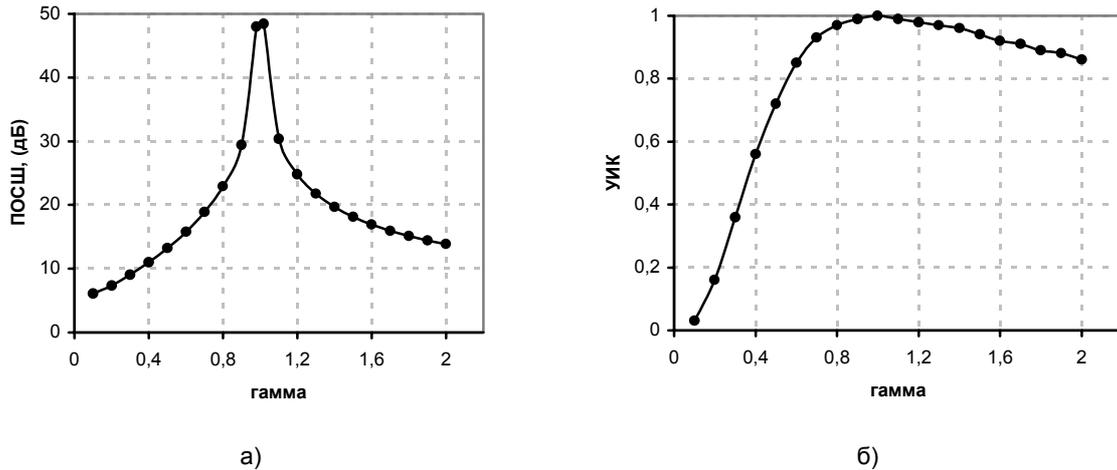


Рис. 4. Влияние показателя  $\gamma$  на качество изображения: а) по шкале ПОСШ; б) по шкале УИК

**Оценка качества изображений в случае медианной фильтрации**

Для подавления импульсных помех наиболее часто применяются различные типы медианных фильтров и их модификации, при этом для сравнения эффективности работы того или иного алгоритма обычно пользуются стандартными метриками качества типа ПОСШ. Оценим качество изображений, обработанных различными типами медианных фильтров, с помощью УИК и сравним полученные результаты с показаниями ПОСШ. В качестве тестируемых алгоритмов будем использовать следующие медианные фильтры: простой (МФ), адаптивный (АМФ), взвешенный (ВМФ) и адаптивный переключающийся (АПМФ) [19]. Важно отметить то, что качество изображения в дан-

ном случае будет зависеть от двух факторов: величины импульсной помехи в изображении и искажений, вносимых в результате применения того или иного алгоритма фильтрации. Результаты обработки изображений при разной степени зашумленности представлены на рис. 5. Если, например, считать, что изображения со значениями ПОСШ более 35 дБ являются изображениями с высоким качеством, то из графиков видно, что даже при наличии небольшого числа случайных выбросов (до 20%) значение ПОСШ говорит о высоком качестве только тех изображений, которые обработаны адаптивным переключающимся медианным фильтром. Можно сказать, что величина ПОСШ слишком чувствительна к искажениям, вносимым импульсными помехами (импульсным шумом).

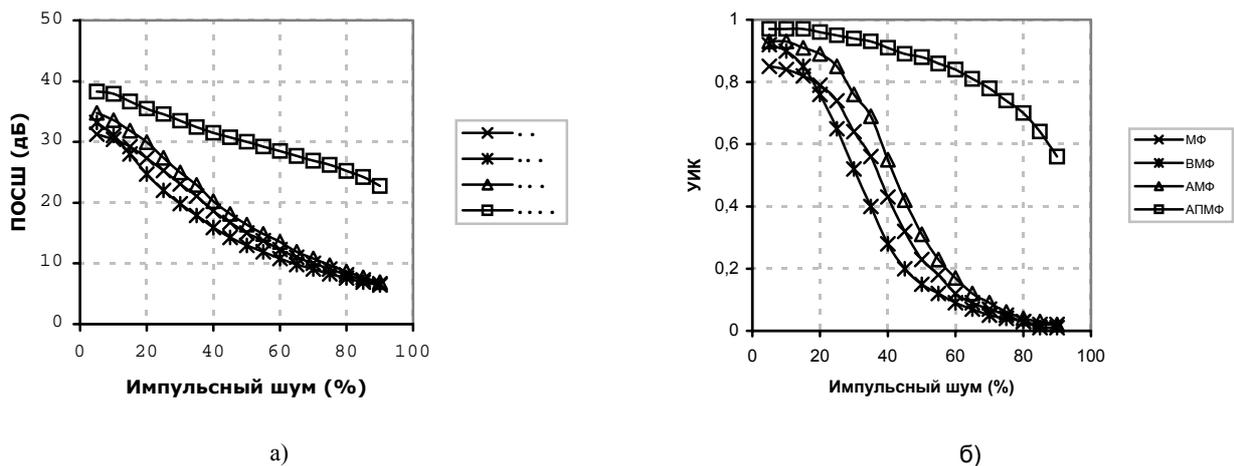


Рис. 5. Оценка изображений, обработанных медианными фильтрами: а) по шкале ПОСШ; б) по шкале УИК 7x7

При использовании метрики УИК резкий спад качества происходит только в диапазоне шума 20-60%, а в остальных случаях уменьшение качества происходит более медленно. Благодаря показаниям УИК (рис. 5) эффективность того или иного алгоритма медианной фильтрации просматривается наиболее отчетливо. Так например, разница в эффективности работы адаптивного медианного фильтра и взвешенного медианного фильтра для 40%-го импульсного шума составляет 0,3 по шка-

ле УИК в пользу первого алгоритма, а по шкале ПОСШ всего лишь 5 дБ (в процентном отношении это существенно меньше).

Для большей согласованности значений УИК с субъективной экспертной оценкой необходимо выбирать размер маски равный  $7 \times 7$  пикселей, так как при этом показания УИК наиболее точно соответствуют визуальному восприятию изображений (рис. 6). При использовании других размеров маски ( $3 \times 3$  и  $5 \times 5$ ) получается, например, что качество изобра-



жения с наличием 50%-го импульсного шума после обработки его адаптивным переключающимся фильтром оценивается как среднее, хотя изображение при этом слабо отличается от оригинала и, соответственно, должно обладать высоким качеством. В других случаях, например при использовании взвешенного медианного фильтра, получается, что показания УИК наоборот завышены и соответствуют изображению со средним качеством, хотя субъективно его качество оценивается как низкое (рис. 6г).

### Заключение

Предложен универсальный индекс качества для сравнительного анализа степени искажений в изображениях при различных типах шумовых воздействий и алгоритмах фильтрации. Установлено, что в ряде случаев использование УИК позволяет дифференцировать изображения, обладающие одинаковым качеством, с точки зрения стандартных метрик типа ПОСШ и СКО, при этом показания УИК достаточно хорошо согласуются с визуальным восприятием таких изобра-

жений. Это обусловлено тем, что алгоритм вычисления УИК имеет комбинированную структуру и учитывает различные виды искажений, в том числе искажения яркости и контрастности. Благодаря такому подходу, оценка качества изображений осуществляется наиболее точно и объективно.

С использованием предложенного индекса качества возможно построение новых алгоритмов обработки изображений и усовершенствование существующих. Возможно также построение адаптивных фильтров и использование показаний УИК в эвристических методах обработки изображений. Кроме того, он может быть использован как основа для построения новых способов сравнения различных алгоритмов обработки не только изображений, но и других цифровых сигналов, а также применяться с целью контроля качества изображений и видеоинформации.

Недостатком метода является то, что, как и в случае вычисления ПОСШ, УИК является эталонной мерой качества, то есть для его нахождения помимо тестового изображения необходимо иметь изображение-оригинал.



а) УИК 3x3 = 0,81; УИК 5x5 = 0,88;  
УИК 7x7 = 0,91;



б) УИК 3x3 = 0,63; УИК 5x5 = 0,60;  
УИК 7x7 = 0,55;



в) УИК 3x3 = 0,42; УИК 5x5 = 0,46;  
УИК 7x7 = 0,43;



г) УИК 3x3 = 0,41; УИК 5x5 = 0,33;  
УИК 7x7 = 0,28;

Рис. 6. Влияние выбора размера маски для вычисления УИК в случае медианной фильтрации: а) адаптивным переключающимся фильтром; б) адаптивным фильтром; в) классическим медианным фильтром; г) взвешенным медианным фильтром

## Список литературы

1. Eskicioglu A., Fisher P. Image quality measures and their performance // IEEE Trans. Communications. 1995. V. 43, № 12. P. 2959–2965.
2. Yu Z., Wu H., Winkler S., Chen T. Vision-model-based impairment metric to evaluate blocking artifact in digital video // Proc. of the IEEE. 2002. V. 90, № 1. P. 154–169.
3. Dijk A., Martens J., Watson A. Quality assessment for coded images using numerical category scaling // Proc. SPIE. 1995. V. 2451, № 3. P. 90-101.
4. Watson A. Visual detection of spatial contrast patterns: evaluation of five simple models // Optic Express. 2000. V. 6, № 1. P. 12-33.
5. Цифровая обработка телевизионных и компьютерных изображений. Под ред. Ю.Б. Зубарева и В.П. Дворковича. – М. 1997.
6. Ричардсон Я. Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения. – М.: Техносфера, 2005.
7. Сухарев А.Г., Тихонов А.В., Федоров В.В. Курс методов оптимизации: Учеб. пособ. – 2 изд. – М.: ФИЗМАТЛИТ, 2005.
8. Aubert G., Kornprobst P. Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations. – Springer Verlag, 2002.
9. Сэлмон Д. Сжатие данных, изображений и звука. – М.: Техносфера, 2004.
10. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005.
11. Wang Z., Bovik A. Modern Image Quality Assessment. Synthesis Lectures on Image, Video & Multimedia Processing. Morgan & Claypool, 2006.
12. Recommendation ITU-R BT.500-11, Methodology for the subjective assessment of the quality of television pictures. ITU-T. 2002.
13. Branden Lambrecht C., Verscheure O. Perceptual quality measure using spatio-temporal model of the Human Visual System. Digital Video Compression Algorithms and Technologies // Proc. SPIE. 2668. 1996.
14. Pappas T., Safranek R. Perceptual criteria for image quality evaluation // In Handbook of Image and Video Processing (A. Bovik, ed.). Academic Press, 2000.
15. Wade N., Swanston M. Visual Perception: An Introduction. 2nd ed. L.: Psychology Press, 2001.
16. Wang Z., Sheikh H., Bovik A. No-reference perceptual quality assessment of JPEG compressed images // Proc. IEEE International Conference on Image Processing. 2002. P. 477-480.
17. Marziliano P., Dufaux F., Winkler S., Ebrahimi T. Perceptual blur and ringing metrics: Application to JPEG2000 // Signal Processing: Image Communication. 2004. V. 19, № 2. P. 163–172.
18. Wang Z., Bovik A. A Universal Image Quality Index // IEEE Signal processing letters. 2002. V. 9, № 3. P. 81–84.
19. Приоров А. Л., Апальков И. В., Бухтояров С. С., Хрящев В.В. Применение переключающихся медианных фильтров для восстановления зашумленных изображений // Вопросы радиоэлектроники, серия общетехническая. 2006. Вып. 2. С. 137-147.

## Уважаемые авторы !

Редакция научно-технического журнала «Цифровая обработка сигналов» просит Вас соблюдать следующие требования к материалам, направляемым на публикацию:

### 1) Требования к текстовым материалам и сопроводительным документам:

- Текст – текстовый редактор Microsoft Word на базе версии WINDOWS' 95 или выше.
- Таблицы и рисунки должны быть пронумерованы. На все рисунки, таблицы и библиографические данные указываются ссылки в тексте статьи.
- Объем статьи до 12 стр. (шрифт 12). Для заказных обзорных работ объем может быть увеличен до 20 стр.
- Название статьи на русском и английском языках.
- Рукопись статьи сопровождается:
  - краткой аннотацией на русском и английском языках;
  - номером УДК;
  - сведениями об авторах (Ф.И.О., организация, телефоны, электронная почта).

### 2) Требования к иллюстрациям:

- Векторные (схемы, графики) – желательно использование графических редакторов Adobe Illustrator или Corel DRAW.
- Растровые (фотографии, рисунки) – М 1:1, разрешение не менее 300dpi, формат tiff.

Справки по телефонам: (495) 290-90-88, (4912) 96-10-95 или по электронной почте [info@dspa.ru](mailto:info@dspa.ru), [tor@rgta.ryazan.ru](mailto:tor@rgta.ryazan.ru)

## ФРАКТАЛЬНОЕ КОДИРОВАНИЕ ЦВЕТНЫХ ИЗОБРАЖЕНИЙ И ПОЛУТОНОВЫХ ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЕЙ

*Касаткин А.С., Шанцын Е.А., Мусихин В.В.*

### Введение

Постоянное возрастание объемов передаваемых данных требует усовершенствования прежних и разработку новых алгоритмов сжатия данных. Среди общего объема цифровых данных цифровые изображения занимают особое место. Связано это, во-первых, со статистическими характеристиками изображений, во-вторых, с психофизическим восприятием последних человеком.

В отличие от произвольного цифрового потока, представляемого в виде массива данных, поступающих на кодер, цифровое изображение правильнее представлять матрицей, обладающей пространственными корреляционными связями. Корреляционные характеристики, практически постоянные на всем интервале изображения, указывают на схожесть (предсказуемость) соседних элементов изображения. Чем выше корреляция пикселей изображения, тем меньше бит информации будет необходимо для их кодирования.

Однако использование одной статистической избыточности не позволяет достичь больших коэффициентов сжатия, например, кодирование файла bmp архиватором ZIP обычно дает сжатие в 1.5-3 раз. Поэтому все современные методы сжатия построены с учетом информационной избыточности изображения. Для качественного восприятия изображения достаточно его приближенной копии, задача кодера в этом случае состоит в выборе значащих составляющих изображения и эффективного их кодирования. Коэффициент сжатия в случае использования алгоритмов сжатия с частичной потерей данных при приемлемом восстановлении изображения достигает 20 раз и более.

Чем больше информации об изображении использует кодер, тем эффективнее происходит кодирование. Яркость или цветность в цифровом изображении передается посредством цепочки бит, поэтому наряду с пространственными корреляционными характеристиками в цифровом изображении присутствуют корреляционные связи между битами данных. Операции межбитовой обработки являются бинарными, т.е. могут быть аппаратно реализованы для ускорения процессов кодирования/ декодирования.

*Рассмотрен вопрос эффективности представления цифрового изображения в виде набора битовых сечений и приведены результаты фрактальных бинарных алгоритмов кодирования динамических полутонных и статических цветных цифровых изображений.*

### Битовые сечения цифрового изображения

Разрядность является одной из характеристик цифрового изображения и показывает, сколько бит данных приходится на кодирование каждого пикселя изображения. Чаще всего на практике из-за простоты используют равномерную импульсно-кодую модуляцию, для кодирования полутонного цифрового изображения достаточно 8 бит данных, цветного изображения – 24 бита, по восемь бит на каждый цвет.

Битовое сечение представляет собой сечение цифрового изображения по соответствующему разряду. Старшим сечением является сечение по восьмому биту, младшим – по первому. Разбиение на битовые сечения позволяет выделить бинарные составляющие цифрового изображения. Старшие битовые сечения в основном передают крупные и средние объекты (низкочастотные составляющие) цифрового сигнала, с уменьшением разряда битового сечения величина передаваемых объектов уменьшается (частота составляющих сигнала увеличивается). Установлено [1], что наибольший вклад в восприятие человеком изображения оказывают низкочастотные сигналы, формирующие основу изображения.

Равномерное квантование по сравнению с неравномерным увеличивает избыточность цифрового изображения, т.к. для кодирования информации используется цепочка бит фиксированной большей длины. Возрастание избыточности приводит к усилению корреляционных связей между битовыми сечениями. Связи в первую очередь возрастают между старшими битами, постепенно ослабевая с уменьшением номера разряда. В битовых сечениях, особенно в соседних, наблюдается множество похожих областей, полностью совпадающих, повернутых на определенный угол или инвертированных (рис.1,б,в).

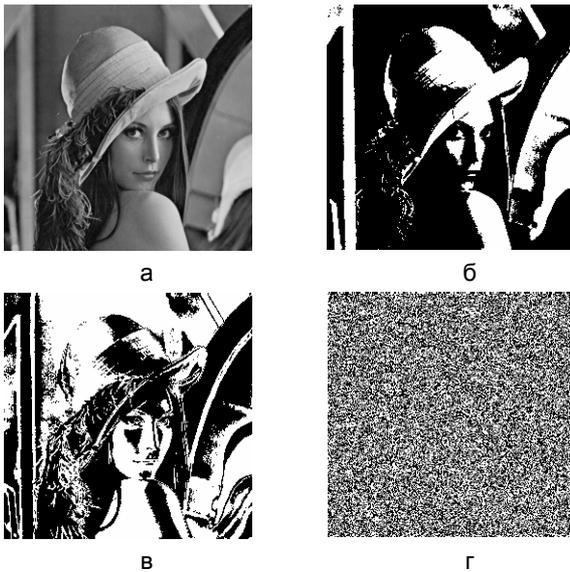


Рис.1

На рис.1 представлено полутоновое изображение «Лена» и его битовые сечения: 1,а – исходное изображение; 1,б – битовое сечение восьмого разряда; 1,в – битовое сечение седьмого разряда; 1,г – битовое сечение первого разряда.

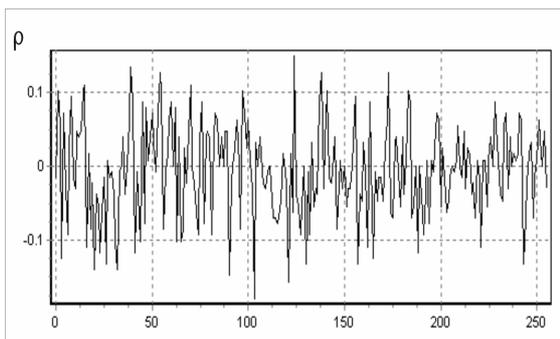
Было проведено исследование корреляционных характеристик изображения рис.1,а. Коэффициент корреляции в среднем между строками составлял 0,967, между столбцами – 0,934. Коэффициенты корреляции в сечениях старших и младших битов представлены в табл.1.

Таблица 1

Зависимость коэффициента корреляции

Коэффициент корреляции	Между строками	Между столбцами
8 бит	0.9102	0.8638
7 бит	0.8026	0.7240
2 бит	0.0051	0.0044
1 бит	0.0050	0.0000

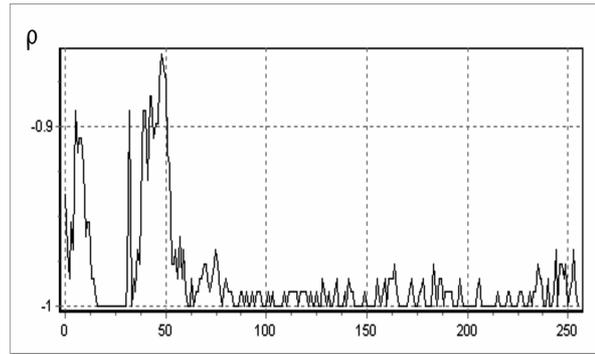
Корреляция между битовыми сечениями зависит от положения битовых сечений друг относительно друга и их разрядности. Наибольший коэффициент корреляции наблюдается между соседними старшими битовыми сечениями рис.2,в, практически все области инверсные с коэффициентом корреляции близким к -1.



а



б



в

Рис.2

На рис.2 представлены зависимость коэффициента корреляции между столбцами в битовом сечении цифрового изображения «Лена»: 2,а – зависимость коэффициента корреляции между столбцами 1 и 2 битовых сечений; 2,б – та же зависимость для 8 и 2 битовых сечений; 2,в – для 8 и 7 битовых сечений.

Уменьшение разрядности битовых сечений снижает коэффициент корреляции рис.2,а. Удаленность битовых сечений друг от друга также снижает совпадение бинарных областей рис.2,б.

Качество восстановленного из битовых сечений полутонового изображения зависит в первую очередь от точности восстановления старших битовых сечений, младшие (1-3 разряда) битовые сечения можно даже не восстанавливать, что заметно сократит вычислительные затраты и незначительно ухудшит качество восстановления.

Во время декодирования с целью повышения качества восстановленного изображения информацию о младших разрядах можно восстановить за счет аппроксимации яркостных составляющих изображения [4].

**Битовые сечения в динамических полутоновых и статических цветных цифровых изображениях**

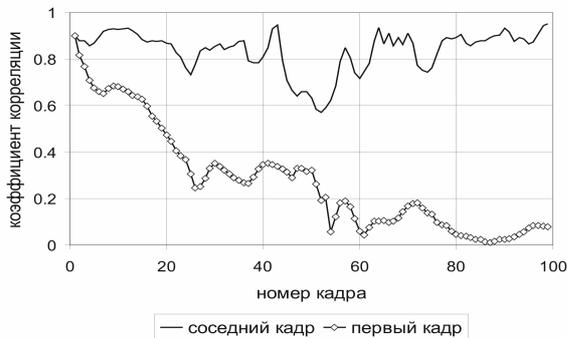
Корреляционные характеристики между битовыми сечениями не удастся описать простейшими линейными уравнениями, поэтому статистические связи довольно сложно прогнозировать и использовать при сжатии полутоновых изображений [3]. Более перспективным выглядит вариант использования корреляционных связей между битовыми сечениями в видеоизображениях и цветных изображениях.

### Корреляционные характеристики полутоновых цифровых видеоизображений

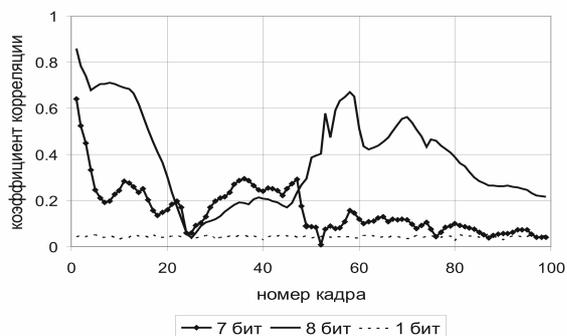
Видеоизображение представляет собой чередующийся с определенной скоростью набор статических изображений (кадров). Поэтому в видеоизображениях можно выделить пространственные корреляционные связи, ничем не отличающиеся от связей в статических изображениях рис.1, и временные корреляционные связи, определяющие большие коэффициенты компрессии при сжатии.



а



б



в

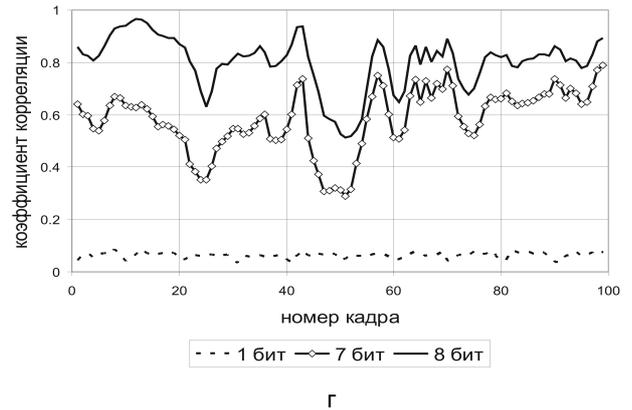


Рис.3

На рис.3 представлены корреляционные характеристики динамических изображений: 3,а – типовой кадр исследуемой видеопоследовательности; 3,б – зависимость коэффициента корреляции между кадрами полутонового видеоизображения; 3,в – зависимость коэффициента корреляции между битовыми сечениями (8, 7, 1 сечения) кадров видеоизображения при первом опорном кадре; 3,г – зависимость коэффициента корреляции между битовыми сечениями (8, 7, 1 сечения) кадров видеоизображения при соседнем опорном кадре.

На рис.3,а представлен типовой кадр исследуемой видеопоследовательности длительностью 100 кадров. Выбранная видеопоследовательность сочетала крупные объекты и довольно быструю съемку (динамику смены кадра). Исследовались две корреляционные зависимости: связь текущего кадра с первым (рис.3,б,в) и связь текущего кадра с соседним кадром (рис.3,б,г). Данные зависимости могут использоваться для выбора частоты следования опорных кадров в видеоизображении.

Корреляционные характеристики исследовались непосредственно в полутоновом изображении (рис.3,б) и в битовых сечениях кадров рис.3,в,г. С удалением от опорного кадра степень корреляции в старших битовых сечениях ослабевает несколько меньше, чем в целом полутоновом, основное совпадение наблюдается на начальных кадрах. Особенно сильные корреляционные связи просматриваются между соседними кадрами, в которых корреляция в старших сечениях повторяет зависимость коэффициента корреляции в полутоновом видеоизображении, незначительно уменьшаясь по абсолютной величине с уменьшением номера разряда сечения (рис.3,б, г).

### Корреляционные характеристики цветных цифровых изображений

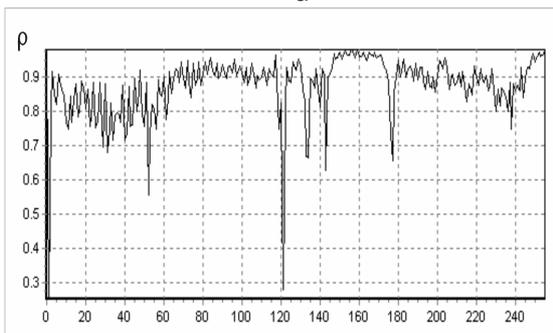
Цвет в цветном цифровом изображении передается комбинацией трех цветовых составляющих. Известно несколько цветовых координатных систем, наиболее известной и часто употребляемой в вычислительной

технике является система RGB, например, формат bmp файлов. В этом случае на каждую составляющую цвета выделяется одинаковое число бит – по восемь, т.е. каждую цветовую составляющую в отдельности можно считать полутоновым изображением. Корреляционные характеристики цветовой компоненты похожи на корреляционные характеристики полутоновых изображений.

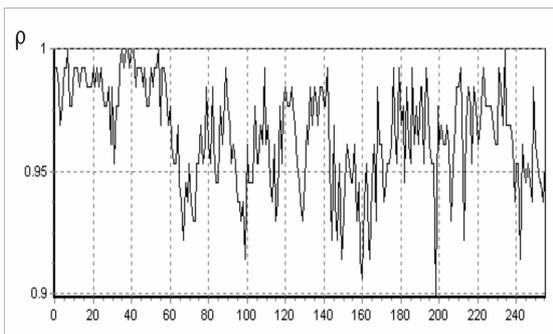
Можно работать и в других координатных системах, однако, например, перевод из системы RGB в XYZ требует последующей работы с числами с плавающей точкой, также после обработки необходимо будет для большинства графических редакторов снова переводить изображение в систему RGB.



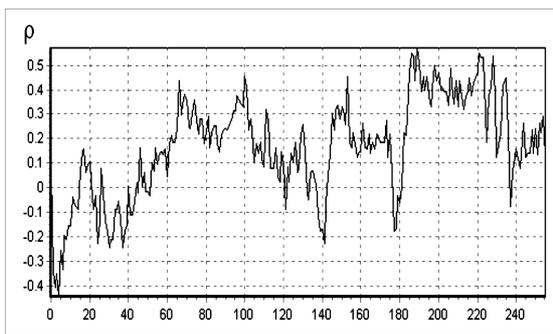
а



б



в



г

На рис.4 представлены корреляционные характеристики между столбцов цветовых составляющих цветного цифрового изображения: 4,а – исходное цветное изображение; 4,б – зависимость коэффициента корреляции между R и G цветовыми компонентами; 4,в – зависимость коэффициента корреляции между R и G цветовыми компонентами в 8-х битовых сечениях; 4,г – зависимость коэффициента корреляции между R и B цветовыми компонентами в 5-х битовых сечениях.

В JPEG кодерах цветное изображение разбивается на яркостную и два цветоразностных сигнала. Чем меньше информации содержится в цветоразностных составляющих тем лучше, так как они кодируются с большими потерями по сравнению с яркостным сигналом. Можно организовать кодер противоположным образом: выбирать координатную систему с максимальным совпадением цветовых составляющих – в этом случае кодируется не отличие, как в JPEG, а совпадение областей изображения. Выбор цветовой координатной системы определяется корреляционными связями между цветовыми составляющими. Коэффициенты корреляции цветовых составляющих по строкам и столбцам цветного изображения рис.4,а близки единице рис.4,б.

Корреляция между старшими битовыми сечениями цветовых составляющих практически совпадает с корреляцией между самими цветовыми составляющими рис.4,б,в. Статистические связи между другими битовыми сечениями цветовых составляющих сигнала зависят от изображения и могут соответствовать зависимостям как на рис.4,в, так и на рис.4,г.

### Фрактальное кодирование

Корреляционные связи, наблюдаемые между битовыми сечениями, можно использовать для сокращения кодируемых данных, например, с помощью простейшего адаптивного предсказания. Наиболее эффективным в данном случае является процесс кодирования, учитывающий совпадение пространственных областей, в качестве которого можно привести один из самых современных способов сжатия – фрактальный алгоритм сжатия.

Существующие аналоги [2] производят обработку полутонового или цветного изображения, работая с яркостными или цветовыми составляющими изображения. Разложение на битовые сечения позволяет работать с бинарными сигналами, значительно упрощая сравнение областей.

### Фрактальное кодирование динамических полутоновых изображений

Корреляционные зависимости между битовыми сечениями соседних кадров рис.3,в,г позволяют спроектировать довольно простой алгоритм фрактального кодирования: видеоизображение разбивается на группы, в основе которых лежит опорный кадр и последовательность кодируемых кадров, по аналогии с I и P кадрами в стандарте MPEG2; битовые сечения опорного кадра разбиваются на до-

Рис.4



менные области; сечения кодируемого кадра разбиваются на ранговые области; ранговые области битового сечения кодируются по доменным блокам соответствующего битового сечения (т.е. восьмой по восьмому, седьмой по седьмому и т.д.) опорного кадра с использованием выбранных преобразований; информация о преобразованиях сохраняется в результирующем файле.

В данной работе доменные и ранговые блоки (для простоты обработки) выбирались в виде квадратных областей, количество преобразований было минимальным: прямой перенос с масштабированием и инверсный перенос с масштабированием, т.к. более 70% блоков передается с помощью этих преобразований. Для уменьшения времени кодирования и увеличения суммарного коэффициента компрессии опорный кадр сжимался с помощью встроенного в CorelDraw11 плагина jpeg2000. Результирующий файл дополнительно сжимался архиватором ZIP.

Исходный алгоритм фрактального сжатия полутонового изображения [2] предусматривает при максимальном коэффициенте сжатия полный перебор подходящих доменных блоков, начиная с левого верхнего угла. Для ускорения времени кодирования, поиск подходящего домена начинался с координат местоположения рангового блока, информация записывалась в виде вектора смещения. Если в текущем окне обработки подходящего домена по заданному порогу соответствия не находилось, кодирование проводилось по стандартной схеме с возможным уменьшением размера рангового блока.

Кодируемая видеопоследовательность из 50 кадров разбивалась на группы по 15 кадров с одним опорным кадром, размер кадров составлял 256 на 256 пикселей. Параметры алгоритма: размер домена 16 на 16 пикселей, порог соответствия составлял 0,95.

Исследования алгоритма показали, что время кодирования, начиная с восьмого и до пятого битовых сечений, каждого последующего сечения увеличивается вдвое. Так как младшие битовые сечения содержат меньшую долю информации о восстанавливаемом кадре, то их можно исключать из кодирования. При смене кадра изображение получает некоторый сдвиг относительно предыдущего, который повторяется во всех битовых сечениях. Это приводит к появлению похожих областей в закодированных битовых сечениях. Таким образом, во время декодирования часть информации младших разрядов может быть восстановлена за счет старших. Результаты кодирования приведены в табл.2.

$$PSNR = 10 * \log_{10} \left( \frac{m^2}{\frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2} \right) \quad (1)$$

Таблица 2

Результаты кодирования полутоновой видеопоследовательности

Количество кодируемых сечений	PSNR, дБ	Коэффициент сжатия, раз
2	15.76 / 17.93	19.14
3	20.70 / 22.17	9.27
4	23.58 / 24.03	-

Во второй колонке табл. 2 слева приведено среднее пиковое отношение с/ш, рассчитанное по формуле (1) с m=255, при восстановлении только старших битовых сечений, справа – при восстановлении всего кадра по информации старших битовых сечений. Во второй колонке приведен коэффициент сжатия относительно последовательности bmp кадров. Различные импульсные помехи и блочные структуры при восстановлении кадров могут быть снижены за счет фильтрации и интерполяции.

Выбор количества кодируемых сечений зависит от пропускной способности канала связи и может управляться: при перегрузке канала связи и переполнении буфера данных устройство управления выдает сигнал на кодер об уменьшении числа кодируемых сечений; при восстановлении работы канала, количество кодируемых сечений увеличивается. При таком способе обмена информацией удаленный пользователь все время будет получать видеоизображение без заметных временных задержек.

### Фрактальное кодирование цветных изображений

Кодирование цифрового цветного изображения проводится по следующей схеме: по корреляционным характеристикам цветовых составляющих выбирается опорный цвет, имеющий наибольшие коэффициенты корреляции с двумя другими; битовые сечения опорного цвета разбиваются на доменные блоки, а кодируемых цветов – на ранговые блоки. Ранговые области битового сечения кодируются по доменным блокам соответствующего битового сечения (т.е. восьмой по восьмому, седьмой по седьмому и т.д.) опорного цвета с использованием выбранных преобразований; информация о преобразованиях сохраняется в результирующем файле.

Для простоты обработки области являлись квадратами, в качестве преобразований использовался параллельный перенос с масштабированием и инверсией. Опорный цвет кодировался с помощью встроенного в CorelDraw11 плагина jpeg2000. Файл результата дополнительно сжимался архиватором ZIP.

В качестве тестируемого было выбрано цветное изображение леса рис.4,а. Картинка содержит большое количество мелких деталей. Исследова-

ние результатов показало, что наиболее эффективным является использование параметров кодирования: размер домена 16 на 16 пикселей, порог соответствия 0,95. В табл.3 представлены результаты кодирования изображения рис.4,а. Мерой качества восстановления изображения была выбрана величина пикового отношения сигнал/шум, рассчитанная по формуле (1) с  $m=16777215 (2^{24}-1)$ .

Таблица 3  
Результаты кодирования цветного изображения

Количество сечений	PSNR, дБ/сжатие, раз	
	0,90	0,95
1	110.13 / 22.061	110.19 / 20.107
2	111.92 / 15.682	112.31 / 11.860
3	115.66 / 9.7815	116.49 / 7.2480
4	116.59 / 5.9114	117.78 / 4.425

Так как корреляционные характеристики между битовыми сечениями цветовых составляющих во многом похожи, это позволяет использовать закодированную информацию старших битов для восстановления младших разрядов цветного изображения, повышая качество восстановленного изображения (результаты для порога 0,95 представлены в табл.4).

Дополнительная обработка позволяет повысить величину пикового отношения с/ш на 1-4 дБ. Для сравнения в табл.4 приведены результаты кодирования встроенных в CorelDraw11 кодеров jpeg и jpeg2000. Снижение величины пикового отношения с/ш (при восстановлении четырех разрядов) объясняется частичным искажением и отсутствием информационных составляющих в младших разрядах опорного цвета, потерянных при jpeg кодировании.

Таблица 4  
Результаты кодирования цветного изображения

Количество сечений	PSNR, дБ	JPEG PSNR, дБ/сжатие, раз	JPEG2000 PSNR, дБ/сжатие, раз
1	116.17		117.42
2	116.23	121.45	
3	118.48	26.94	23.14
4	117.47		

Работа алгоритма была проверена на цветных изображениях [5], параметры алгоритма: размер домена 16 на 16, порог соответствия ранговых и доменных областей 0,95, кодируются три старших битовых сечения, используется восстановление в младших разрядах (табл.5).

Увеличение количества обрабатываемых битовых сечений до четырех незначительно (до 1,5 дБ) повысит величину пикового отношения с/ш, но существенно (примерно 1,5 раза) снизит коэффициент сжатия.

Таблица 5

Результаты кодирования цифровых цветных изображений

Количество сечений	PSNR, дБ	Сжатие, раз
Airplane	119.02	6.53
Baboon	117.73	3.86
House	119.23	7.44
Lenna	117.34	5.34

## Выводы

1. Рассмотрены корреляционные характеристики между битовыми сечениями цифровых изображений. Коэффициенты корреляции в старших битовых сечениях значительны, это можно использовать для кодирования цифровых изображений.

2. На основании корреляционных характеристик предложен алгоритм фрактального бинарного кодирования цифровых видеоизображений. При приемлемом качестве восстановления удается сжать видеопоследовательность в 10-20 раз (в зависимости от требуемого качества) относительно исходного размера bmp файлов. Преимуществом данного способа кодирования является бинарная обработка и возможность регулирования величины бинарного потока на выходе кодера.

3. Корреляционные характеристики между цветовыми составляющими цветного цифрового изображения могут достигать 90%, та же картина наблюдается между битовыми сечениями цветовых составляющих. Коэффициент сжатия существенно зависит от детальности изображения, в работе коэффициент сжатия составляет величину 4-7 раз при пиковом отношении сигнал/шум равном примерно 118 дБ. При передаче по сетям связи сжатие можно повысить в ущерб качеству, т.к. при восстановлении изображения глаз в первую очередь будет аппроксимировать яркостную составляющую.

## Литература:

1. У. Прэтт Цифровая обработка изображений в 2кн., М.: – «Мир», 1982.
2. Уэлстид С. Фракталы и вейвлеты для сжатия в действии. – М.: Триумф, 2003. – 320с.
3. Трубин И.С., Касаткин А.С. Фрактальное сжатие полутоновых изображений. Всероссийская научно-техническая конференция «Наука-производство-технология-экология». Сборник материалов. Том 2, ФАВТ, ФПМТ, Киров: Изд-во ВятГУ., 2004г., с.106-107.
4. Трубин И.С., Касаткин А.С. Методы борьбы с ложными контурами сжатого цифрового полутонового изображения. Вестник Вятского научного центра Верхнее-Волжского отделения Академии технологических наук РФ, Серия: Проблемы обработки информации, Выпуск 1(5)/2004, Вятский научный центр, Киров, 2004 г., с.50-55.
5. [http://en.wikipedia.org/wiki/Standard\\_test\\_image](http://en.wikipedia.org/wiki/Standard_test_image).

## ОБНАРУЖЕНИЕ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ ПРИ ИЗМЕНЯЮЩИХСЯ УСЛОВИЯХ НАБЛЮДЕНИЯ

Аксенов О.Ю.

Системы видеонаблюдения, в частности системы обнаружения объектов заданных классов на поверхности земли, обрабатывают последовательности изображений, формируемые телевизионными камерами, и решают задачу определения характеристик обнаруженных объектов. Это означает, что фон, на котором наблюдается объект, может иметь разнообразные свойства, и что способ обработки изображений должен обеспечивать работоспособность системы обнаружения в широком диапазоне условий наблюдения.

В большинстве случаев, имеющих практический интерес, особенно когда речь идет об изображениях естественных сцен, получаемых в «полевых условиях», результаты обнаружения зависят не только от используемых алгоритмов обработки, но и от условий наблюдения, а также от содержания наблюдаемой сцены. Поэтому, прежде всего, следует оговорить понятия «наблюдаемая сцена» и «условия наблюдения», остановиться на способах их формального описания [1].

**Наблюдаемая сцена** определяет характерный вид объекта и фона, для работы с которыми предназначена рассматриваемая система обработки изображений. На-



Рис. 1. Изображение наблюдаемой сцены

Наиболее полно яркостные характеристики наблюдаемой сцены определяет гистограмма распределения яркости. Для изображения рис.1, такая гистограмма показана на рис.2 темным цветом. Для более компактного описания яркостных характеристик наблюдаемой сцены можно использовать такой параметр, как модуляционный контраст, представляющий собой отношение ширины гистограммы к положению ее середины – средней яркости –  $K_m = 2 \cdot D / \bar{B}$  [2]. Здесь:  $D = B_{\max} - B_{\min}$  – ширина гистограммы распределения яркости,  $\bar{B} = (B_{\max} + B_{\min}) / 2$  – средняя яркость изображения,  $B_{\min}$ ,  $B_{\max}$  – минимальное и максимальное значения яркости на изображении. Величина  $K_m$  может изме-

*Рассматривается технология создания системы обнаружения объектов на изображениях естественных сцен. Применены два подхода к задаче обеспечения работоспособности системы в широком диапазоне условий наблюдения. Один из них связан с использованием адаптации при обработке изображений, а второй – с применением для обработки и тестирования используемых алгоритмов изображений, соответствующих различным условиям наблюдения. Применение обоих подходов демонстрируется на примере системы обнаружения объектов на изоб*

блюдаемая сцена характеризуется содержанием и яркостными параметрами. Содержание наблюдаемой сцены может быть задано словесно. Ниже приводится пример описания сцены, изображенной на рис.1. При описании наблюдаемой сцены необходимо определить следующее:

- характерный вид объекта (легковой автомобиль);
- относительный размер объекта (~10% линейного размера изображения);
- число объектов (1);
- минимальное расстояние между краями объектов, если их несколько (относительно линейного размера объекта);
- фон в окрестности объекта (неоднородный, отличный от цвета объекта)

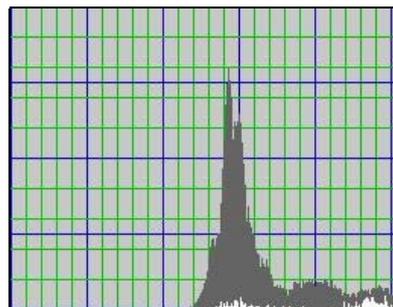


Рис. 2. Гистограмма распределения яркости: объект – светлый, фон – темный

няться в пределах 0...1. Однако, в данном приложении представляется более удобным использовать не значение контраста, а величины  $D$  и  $\bar{B}$  по отдельности.

На изображениях естественных сцен фон в окрестности объекта может заметно отличаться от фона, расположенного на значительном удалении от объекта. Поэтому в некоторых случаях представляется целесообразным для каждого объекта использовать понятие «локальный фон» – фон в ближайшей окрестности объекта. Положение этой окрестности иллюстрирует рис.3. На нем область локального фона имеет вид темной окантовки вокруг объекта (автомобиля). Эта область получена за счет последовательного утолще-

ния «окантовки равной ширины». Увеличение окантовки проводилось от значения «толщина в 1 пиксел» до тех пор, пока площадь окантовки приблизительно не сравнялась с площадью объекта. В данном случае при формате кадра 384×288 это 960...970 пикселов.

В дальнейшем параметры, соответствующие всему изображению, будут маркироваться индексом "f", объекту – индексом "o", фону – индексом "b", а локальному фону –

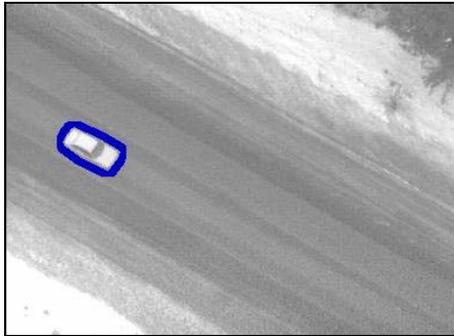


Рис. 3. Область локального фона (темная кайма вокруг автомобиля)

индексом "l". Так, например, всему изображению будут соответствовать параметры  $D_f$  и  $\bar{B}_f$ , а объекту -  $D_o$  и  $\bar{B}_o$ . В рассматриваемом случае, когда яркость отдельной точки изображения может принимать значения 0...255, параметры  $D$  и  $\bar{B}$  также лежат в указанном диапазоне. Поэтому вместо  $D$  и  $\bar{B}$  можно использовать нормированные безразмерные значения  $d = D/255$  и  $\bar{b} = \bar{B}/255$ , принимающие значения 0...1.

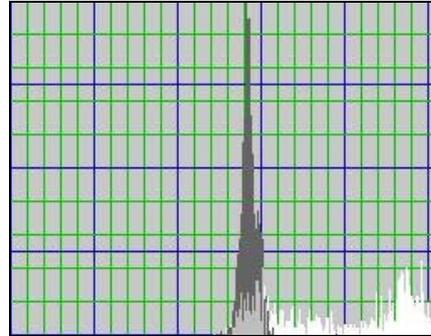


Рис. 4. Гистограммы яркости: объект – светлый, локальный фон – темный

Гистограммы распределения яркости объекта и фона (все изображение за исключением объекта) содержит рис.2. На нем гистограмма для объекта показана белым, для фона – темным. При этом ордината гистограммы объекта увеличена в 10 раз. Для локального фона соответствующую гистограмму отображает рис.4. В этом случае масштаб гистограмм для объекта и для локального фона одинаков.

Таким образом, наблюдаемую сцену предлагается задавать описанием (или типичным изображением) и набором параметров:

$$d_f \approx d_b, \bar{d}_o, d_l \leq d_b, \bar{b}_f \approx \bar{b}_b, \bar{b}_o, \bar{b}_l.$$

Можно считать, что этим задано некоторое «эталонное изображение», полученное реальной аппаратурой в

идеальных условиях наблюдения.

Возможные модификации наблюдаемой сцены и способы их параметризации в ходе обработки изображения содержит Таблица 1. Параметризация обеспечивается применением алгоритмов обработки, обеспечивающих инвариантность процесса обнаружения к ожидаемым изменениям наблюдаемой сцены – адаптацию системы обнаружения. Данная таблица не претендует на полноту описания всех возможных способов адаптации, а иллюстрирует приемы, использованные при создании конкретной системы видеонаблюдения, обнаружение в которой основано на распознавании объекта заданного класса по форме.

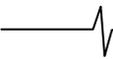
Таблица 1

Возможные модификации наблюдаемой сцены	Способы параметризации модификации наблюдаемой сцены при обработке изображения
Изменение условий наблюдения, влияющих на яркость и контрастность изображения	Нормализация яркости изображения на базе анализа гистограммы распределения яркости
Изменения цвета объекта	Замена полутонового изображения бинаризованным за счет выделения границ перепадов яркости при помощи локального оператора с
Изменение масштаба (размера) объекта на изображении	Выражение размеров элементов объекта в относительных единицах (в частях от наибольшего размера объекта)
Изменение положения объекта на изображении	Выделение для последующей обработки на изображении локальной области объекта
Изменение ориентации объекта на изображении	Приведение изображения объекта к «стандартной» ориентации за счет вращения локальной области изображения
Изменения облика, ракурса объекта (внутри класса)	Использование обучаемого распознавания формы объекта, представляемой в виде «проволочной модели»

*Примечания:* Увеличение диапазона яркости изображения (или локального участка) до максимально возможного значения. Каждая точка бинаризованного изображения может принимать одно из значений: «0» – в этой точке полутонового изображения нет перепада яркости или «1» – перепад яркости есть.

Чтобы отработать (и/или протестировать) систему видеонаблюдения и, в частности, процедуру нормализации яркости изображения (Таблица 1), требуются тестовые последовательности изображений, отличающиеся условиями наблюдений. Можно сказать, что от рассмотренного выше «идеального»

изображения нужно перейти к изображению, соответствующему условиям наблюдения, которые могут иметь место на практике. Наиболее типичными факторами, вызывающими ухудшение изображения, являются туман, осадки (дождь, снег), недостаток освещения.



Поскольку на практике не всегда есть возможность получить изображения, соответствующие требуемому диапазону указанных факторов, предлагается способ их имитации, позволяющий требуемым образом изменять эталонное изображение. В дальнейшем параметры, соответствующие измененным условиям наблюдения, будут маркироваться дополнительным индексом "r".

Для количественной оценки изменения средней яркости изображения и ширины гистограммы яркости удобно использовать параметры  $\Delta B_f = (B_f - B_{fr})$  и  $\Delta D_f = (D_f - D_{fr})$  соответственно. Или, переходя к безразмерным величинам,  $\Delta \bar{b}_f = (\bar{b}_f - \bar{b}_{fr})$  и  $\Delta d_f = (d_f - d_{fr})$ .

Обычно телевизионная камера, используемая для формирования изображений, имеет систему автоматической регулировки усиления, поддерживающую значение средней яркости изображения ( $\bar{B}_f$ ) неизменным. В этих условиях влияние «тумана» при достаточном освещении можно представить как уменьшение ширины гистограммы яркости всего изображения:  $\Delta \bar{b}_f = 0$ ,  $\Delta d_f > 0$ . При недостаточном освещении влияние тумана можно учесть путем уменьшения ширины гистограммы яркости и одновременного уменьшения средней яркости изображения:  $\Delta \bar{b}_f > 0$ ,  $\Delta d_f > 0$ .

**Недостаток освещения** в отсутствие тумана приводит сначала к снижению средней яркости изображения:  $\Delta \bar{b}_f > 0$ ,  $\Delta d_f = 0$ . После того, как минимальная яркость изображения достигнет нулевого значения, кроме снижения средней яркости изображения начнется уменьшение ширины гистограммы яркости:  $\Delta \bar{b}_f > 0$ ,  $\Delta d_f > 0$ .

**Осадки** можно представить в виде композиции тумана и появления на изображении совокупности капель – малоразмерных объектов, по форме близких к кругу, и имеющих яркость  $B_s$  (не рассматриваются такие специфические факторы, как блики, имеющие место при наблюдении, например, асфальтового покрытия при дожде или в свете фар транспортных средств). Для актуального случая, когда телевизионная камера, формирующая изображение, смотрит сверху вниз, можно принять, что диаметр отдельных капель имеет значение, равномерно распределенное в пределах

1...S точек изображения. Например, для изображения форматом 384×288 по результатам исследований размер капель принимается равным  $S = 5$ . При этом общее количество капель задается как вероятность ( $f_s$ ), с которой каждая из равномерно распределенных по площади изображения  $1/S^2$  точек является центром капли. В итоге совокупность капель задается параметрами

$$B_S = B_{\min} \dots B_{\max}, S \approx 5, f_s = 0 \dots 1.$$

Таким образом, для того, чтобы учесть влияние условий наблюдения ( $D_{fr}, \Delta \bar{B}_{fr}$ ) на исходное (эталонное) изображение ( $D_f, \bar{B}_r$ ) следует скорректировать яркость каждой его точки –  $b(x, y)$  в соответствии со следующим алгоритмом:

- оценить среднюю яркость ( $\bar{B}_f$ ) и ширину гистограммы яркости ( $D_f$ ) изображения  $b(x, y)$ ;
- скорректировать контрастность изображения  $b(x, y)$ , используя соотношение

$$b_1(x, y) = \left( b(x, y) \cdot \frac{D_{fr}}{D_f} \right);$$

- оценить  $\bar{B}_{f1}$  – среднюю яркость изображения  $b_1(x, y)$ ;

- скорректировать яркость изображения  $b_1(x, y)$  в соответствии с выражением

$$b_2(x, y) = b_1(x, y) - \Delta \bar{B}_{1f} + \bar{B}_f - \Delta \bar{B}_{fr}.$$

Изображение  $b_2(x, y)$ , сформированное из эталонного указанным способом, будет соответствовать заданным условиям наблюдения ( $D_{fr}, \Delta \bar{B}_{fr}$ ). Аналогично можно корректировать и другие указанные выше яркостные характеристики изображения.

Рис.5, 7 и 9 поясняют результаты, полученные рассмотренным способом из изображения, представленного на рис.1. Гистограммы распределения яркости объекта и локального фона для указанных изображений содержат рис.6, 8 и 10, соответственно. Результаты моделирования дождевых капель иллюстрируют рис.11 и 12. Яркостные характеристики рассмотренных изображений содержит Таблица 2.



Рис. 5. Ширина гистограммы уменьшена вдвое, средняя яркость сохранена

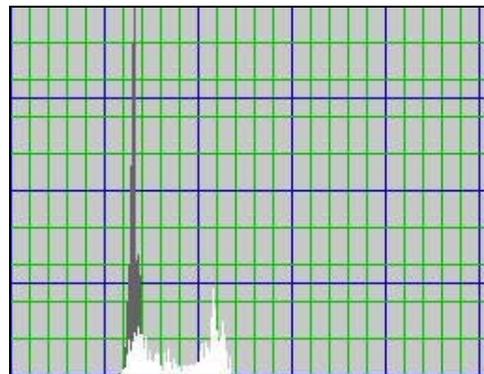


Рис. 6. Гистограммы яркости: объект – светлый, локальный фон – темный



Рис. 7. Ширина гистограммы сохранена, средняя яркость максимально уменьшена

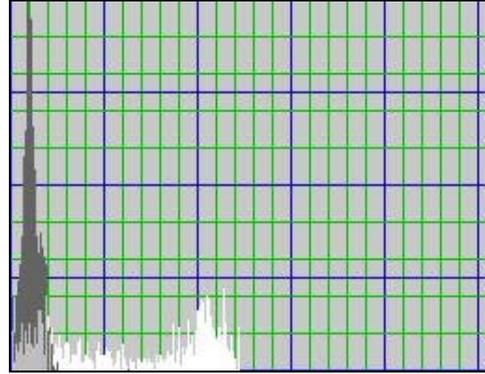


Рис. 8. Гистограммы яркости: объект (светлый), локальный фон (темный)



Рис. 9. Ширина гистограммы уменьшена вдвое, средняя яркость максимально уменьшена

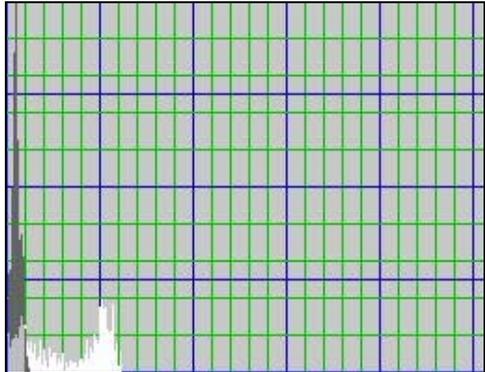


Рис. 10. Гистограммы яркости: объект – светлый, локальный фон – темный



Рис. 11. Осадки 10%



Рис. 12. Осадки 30%

Таблица 2

Количественные характеристики изображений

Изображение	$\bar{B}_{f(r)}$	$\bar{B}_{o(r)}$	$\bar{B}_{l(r)}$	$D_{f(r)}$	$D_{o(r)}$	$D_{l(r)}$	$\Delta\bar{B}_{f(r)}$	$\Delta D_{f(r)}$	$B_S$	$S$	$f_S$
Рис.1	168	196	143	120	120	40	0	0	–	–	0
Рис. 5	169	183	156	60	60	10	0	60	–	–	0
Рис. 7	36	64	11	120	120	20	132	0	–	–	0
Рис. 9	18	31	5	60	60	10	150	60	–	–	0
Рис. 11	168	196	143	120	120	40	0	0	$B$	5	0,1
Рис. 12	168	196	143	120	120	40	0	0	$B$	5	0,3

Примечание:  $B \approx |\bar{B} - B(x, y)|$



Предложенный подход к учету условий наблюдения позволяет проводить отработку и тестирование алгоритмов обработки изображения с целью обеспечения их работы в широком диапазоне условий наблюдения. При этом скорректированные указанным способом последовательности изображений подаются на вход системы обнаружения объектов.

Качество обнаружения оценивалось в виде отношения  $f_r / f$ , где  $f$  – число кадров, на которых произошло обнаружение для последовательностей исходных («эталонных») изображений ( $D_f, \bar{B}_f$ ), а  $f_r$  – число кадров, на которых происходило обнаружение для последовательностей изображений, скорректированных для заданных условий наблюдения ( $D_{f_r}, \bar{B}_{f_r}$ ). Относительная величина  $f_r / f$  характеризует изменения качества работы алгоритма обнаружения в заданных условиях наблюдения по сравнению с эталонными.

Результаты оценок величины  $f_r / f$  для случаев уменьшения ширины гистограммы яркости (усиление тумана), снижения средней яркости (уменьшение освещенности), увеличения интенсивности осадков содер-

жат Таблица 3, 4 и 5 соответственно. Эти таблицы иллюстрируют степень зависимости использованных алгоритмов обработки изображения от условий наблюдения, что позволяет использовать описанный подход для отработки алгоритмов обработки изображения.

Следует отметить, что Таблица 4 содержит значение  $f_r / f > 100\%$ . Это говорит о том, что изображение, принятое за эталонное, не является наилучшим с точки зрения обеспечения обнаружения объекта при использовании выбранного способа обработки. Следовательно, используемый подход позволяет не только оценить качество работы конкретной системы видеонаблюдения, но и определить наилучшие для нее условия наблюдения.

Таким образом, работоспособность системы видеонаблюдения в широком диапазоне условий наблюдения предлагается обеспечивать за счет использования при обработке изображений адаптивных алгоритмов, а отработку этих алгоритмов вести по изображениям, полученным на основании изложенного подхода к имитации условий наблюдения.

Таблица 3

Зависимость показателей обнаружения от ширины гистограммы яркости

$D_{f_r}$	$D_{f_r} / D_f$	$\bar{B}_{f_r}$	$f_r / f$
130	100%	168	100%
104	80% (1/1,25)	168	50%
86	66%(1/1,5)	168	25%
74	57%(1/1,75)	168	0%
65	50%(1/2)	168	0%

Таблица 4

Зависимость показателей обнаружения от средней яркости

$D_{f_r}$	$\bar{B}_{f_r}$	$\bar{B}_{f_r} / \bar{B}_f$	$f_r / f$
130	168	100%	100%
130	138	82%	187%
130	108	78%	100%
130	78	46%	25%

Таблица 5

Зависимость показателей обнаружения от интенсивности осадков

$f_s$	$B_s$	$S$	$f_r / f$
0	–	–	100%
10	$B$	5	87%
20	$B$	5	25%
30	$B$	5	12%

#### Литература

1. О.Ю.Аксенов. Описание наблюдаемой сцены. // VIII Международная конференция “Цифровая обработка сигналов и ее применение” (DSPА-2006). // Труды НТОРЭС им. А.С.

2. Ю.И.Фивенский, М.К.Наумкин. Аэрокосмическая фотография. М.,МГУ 1987 г. 140 с.

УДК 621.391.268

## МЕТОДЫ ОБРАБОТКИ И АНАЛИЗА ИЗОБРАЖЕНИЙ В БОРТОВЫХ СИСТЕМАХ ОБНАРУЖЕНИЯ И СОПРОВОЖДЕНИЯ ОБЪЕКТОВ

*Алпатов Б.А., Бабаян П.В.*

### Введение

Одним из направлений при создании систем обработки и анализа видеоинформации является разработка бортовых систем обнаружения и сопровождения объектов, которые предназначены для установки на автономных носителях, таких как самолёты, вертолёты, автомобили. Можно выделить две характерных черты этих бортовых систем технического зрения: необходимость работы в реальном масштабе времени и автономность режима функционирования этих систем при незначительных усилиях оператора.

Важной проблемой, связанной с разработкой бортовых систем обнаружения и сопровождения объектов, является проблема обнаружения и сопровождения объектов, находящихся в поле зрения датчика изображений. При этом информация о характеристиках объектов, которые требуется выделять, обычно очень скудна и, в лучшем случае, включает в себя лишь приблизительные размеры объектов.

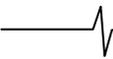
Перед тем как ввести классификацию методов решения задач обнаружения и сопровождения объектов, остановим внимание на иерархии уровней обработки информации, присутствующей в системах обнаружения и сопровождения объектов. Обычно рассматриваются следующие уровни обработки информации: восстановление изображений, выделение объектов, оценка параметров сегментов, слежение за объектами, оценка параметров, обнаружение, распознавание и сопровождение объектов. Под сегментом здесь понимается связанная совокупность точек бинарного изображения, полученного в результате той или иной обработки. Под объектами подразумеваются изображения реальных наблюдаемых объектов. Сегменты бинарного изображения обычно являются исходными данными для дальнейшего обнаружения объектов.

*Представлен обзор основных подходов к решению задач обработки и анализа изображений в современных системах обнаружения и сопровождения объектов. Предложен новый алгоритм выделения объектов в последовательности видеоизображений, учитывающий присутствие геометрических деформаций наблюдаемого изображения. Приведены результаты статистических исследований алгоритма.*

Пусть производится обработка и анализ последовательности односпектральных или мультиспектральных изображений. **Восстановление изображений** состоит в реконструкции изображения, которое было до этого искажено вследствие дефектов и шумов датчика изображения, неравномерности частотных характеристик канала передачи и т.д. **Выделение объектов** состоит в принятии решения о наличии объекта для каждой точки изображения. **Оценка параметров сегментов** состоит в определении таких характеристик сегментов, как координаты центра, размеры, яркость, форма. **Слежение** состоит в оценке траекторий движения объектов во времени, основываясь на анализе поведения сегментов во времени. **Оценка параметров объектов** состоит в определении таких характеристик объектов, как координаты центра, размеры, яркость, форма, скорость, ускорение. Оцененные характеристики могут быть использованы для принятия решения об **обнаружении** объектов заданного типа. Результаты оценки параметров можно использовать для решения задачи **сопровождения объектов**, не имеющей непосредственного отношения к анализу изображений. В результате сопровождения осуществляется корректировка направления оптической оси датчика изображений с целью поддержания направления на заданный объект. **Распознавание** объектов состоит в отнесении наблюдаемых изображений объектов к одному из заданных классов.



Рис. 1 – Иерархия уровней обработки информации в системах обнаружения и сопровождения объектов



Иерархия уровней обработки информации в системах обнаружения и сопровождения объектов представлена в виде пирамиды на рис. 1. В непосредственной близости от параллелепипедов, представляющих уровни иерархии, приведены краткие описания результатов анализа, получаемых на каждом уровне. Обычно в системе обнаружения и сопровождения объектов осуществляется переход от одного уровня иерархии к другому и, в конце концов, достигаются один или оба верхних уровня иерархии – сопровождение и распознавание объектов. В зависимости от задач, решаемых системой обнаружения и сопровождения, а также в зависимости от используемых для решения подходов, некоторые уровни иерархии в конкретной системе могут отсутствовать.

Выделим четыре основных класса методов анализа изображений, используемых в системах обнаружения и сопровождения объектов. Для каждого класса методов можно обозначить условия наблюдения, в которых применение данного класса методов наиболее эффективно. Кроме того, классы методов отличаются по количеству требуемой априорной информации об объекте и по вычислительной сложности реализации.

**1. Корреляционные методы** могут использоваться для измерения координат движущихся и неподвижных объектов, наблюдаемых на однородном и неоднородном фоне, при малых отношениях сигнал/шум. Практически это самый помехоустойчивый класс методов. Реализация корреляционных методов состоит в том, что известное заранее изображение объекта, которое обычно называют эталоном, перемещается по наблюдаемому изображению. Для каждого положения эталона вычисляется значение некоторого критерия. Положение эталона, при котором достигается максимальное значение критерия, рассматривается как измеренное положение объекта. Платой за помехоустойчивость корреляционного подхода является высокая вычислительная сложность, неустойчивость при геометрических изменениях, происходящих с объектом, и большой объём требуемой априорной информации о наблюдаемом объекте. Фактически, необходимо заранее знать изображение объекта наблюдения. Тем не менее, предложено немало способов преодоления этих недостатков [2,3].

**2. Методы статистической сегментации** предназначены для выделения движущихся и неподвижных объектов, наблюдаемых на относительно однородном фоне. Они основаны на использовании априорной информации о различии статистических свойств объекта и фона. Например, метод Байесовской сегментации [4] наиболее эффективен при выделении воздушных объектов на фоне неба. Он требует значительно меньше априорной информации, чем корреляционный. Не требуется знать изображение объекта, однако желательно начальное указание области изображения, в пределах которой следует искать объект. Вычислительная сложность метода Байесовской сегментации невысока, к другим его преимуществам относится устойчивость к геометрическим изменениям объекта. Отрицательной чертой подобных методов является низкая эффективность при наблюдении объектов на неоднородном фоне. Перспективным направлением дальнейшего развития методов статисти-

ческой сегментации считается применение многопризнаковой классификации с использованием дополнительных признаков или спектральных диапазонов видеонаблюдения [5].

**3. Методы выделения объектов с помощью пространственной фильтрации** [6] основаны на использовании различных операций линейной и нелинейной пространственной фильтрации изображений. Данный класс методов эффективен при выделении движущихся и неподвижных объектов на фоне ясного или облачного неба, в том числе и при малых отношениях сигнал/шум. Преимущество этого подхода заключается в невысоких требованиях к объёму априорной информации, описывающей объекты наблюдения. Начальное целеуказание не является обязательным, необходима лишь грубая подстройка алгоритма под размеры объектов. Недостатком методов пространственной фильтрации является то, что для их успешного использования требуется различие в свойствах пространственной структуры объекта и фона. Следует отметить, что вычислительная сложность методов пространственной фильтрации, как правило, меньше, чем у корреляционных методов, но больше, чем у статистических методов сегментации.

**4. Методы пространственно-временной фильтрации** изображений ориентированы на решение задачи выделения движущихся объектов, наблюдаемых на однородном или неоднородном фоне. Принцип работы подобных алгоритмов основан на выявлении таких изменений, происходящих с течением времени в наблюдаемой последовательности изображений. Теоретическое исследование основ этого класса методов было проведено, в частности, в работах [7,8]. Основным преимуществом рассматриваемого подхода является возможность организации полностью автоматического режима функционирования системы обнаружения и сопровождения объектов, причём сложность и неоднородность наблюдаемой сцены не является помехой для успешного решения возложенных на систему задач. Этот фактор, наряду с невысокой вычислительной сложностью реализации, делает указанный класс методов одним из наиболее перспективных. Главным недостатком известных методов, построенных на основе пространственно-временной фильтрации изображений, является невозможность их непосредственного применения при наличии геометрических искажений наблюдаемой последовательности изображений. Тем не менее, проблема использования пространственно-временных методов выделения объектов при наличии преобразований смещения была решена путём введения процедуры оценивания параметров геометрических преобразований и внесения усовершенствований в алгоритм выделения объектов [9]. Продолжаются исследования в направлении использования методов временной обработки изображений при более сложных типах геометрических преобразований. Один из подходов к решению этой задачи при евклидовых преобразованиях изображений рассматривается в данной статье.

Обзор основных классов методов показывает, что они не обеспечивают в полной мере решение всех задач, соответствующих уровням иерархии обработки информации, представленным на рис. 1. В частности, корреля-

ционный подход позволяет лишь выполнить оценку координат объектов, а все остальные подходы, по сути, являются методами выделения объектов. Для реализации процедур обработки информации, соответствующих другим уровням, разработаны специальные методы. Например, для восстановления изображений служат подходы, использующие пространственно-временную фильтрацию, методы нормализации гистограмм и другие [1]. Метод разметки и параметризации обеспечивает переход на уровень оценки параметров сегментов. С помощью методов траекторного анализа осуществляется слежение.

Оценка параметров объектов производится различными методами теории оптимальной фильтрации. Для реализации сопровождения объектов разработаны специальные методы управления приводами поворотного устройства, обеспечивающие высокую точность сопровождения даже при кратковременных пропаданиях объектов [10]. Распознавание объектов может производиться с использованием элементов теории решений или с помощью нейронных сетей. Структура типовой системы обнаружения и сопровождения объектов может иметь вид, представленный на рис. 2.

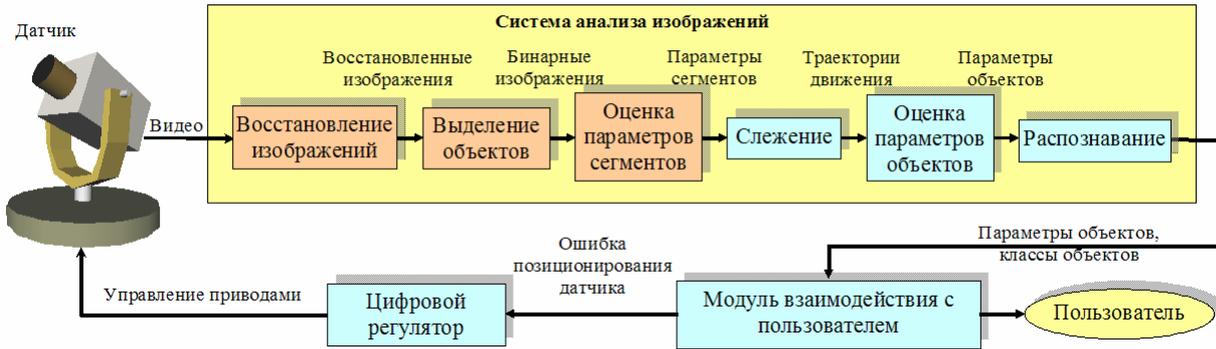


Рис. 2 – Структура системы обнаружения и сопровождения объектов

Реальная система обнаружения и сопровождения объектов часто объединяет в себе совокупность различных методик анализа визуальной информации, относящихся к различным описанным выше классам.

Значительный интерес представляет разработка методов выделения объектов на сложном фоне в условиях геометрических преобразований изображения.

**ПОСТАНОВКА ЗАДАЧИ.** Пусть  $g(x,y)$  – известное изображение фона,  $g_M(x,y)$  – изображение фона в системе координат, связанной с наблюдаемым изображением,  $(\alpha, \beta, \varphi)$  – неизвестный вектор параметров, описывающих евклидово преобразование наблюдаемого изображения относительно изображения фона  $g(x,y)$ .

Зафиксируем точку наблюдаемого изображения с координатами  $(x_0, y_0)$ . Определим для этой точки следующие величины:  $f$  – яркость фона, деформированного геометрическими преобразованиями,  $r$  – бинарный параметр, единичное значение которого соответствует наличию объекта в рассматриваемой точке,  $h$  – яркость объекта наблюдаемой сцены,  $l$  – наблюдаемое значение яркости,

$z_x \sim N(0, \sigma_z^2)$ ,  $z_y \sim N(0, \sigma_z^2)$  – некоррелированные гауссовы случайные величины с известной дисперсией, описывающие случайные деформации изображения, связанные с преломлением света в атмосфере,  $\xi$  – нормальный белый шум датчика:  $\xi \sim N(0, \sigma_\xi^2)$ . Тогда модель формирования яркости наблюдаемого изображения в точке  $(x_0, y_0)$  имеет следующий вид:

$$l = f(1 - r) + hr + \xi, \tag{1}$$

$$f = g_M(x_0 - z_x, y_0 - z_y), \tag{2}$$

$$g_M(x,y) = g \begin{pmatrix} x \cos \varphi + y \sin \varphi - \alpha, \\ -x \sin \varphi + y \cos \varphi - \beta \end{pmatrix} \tag{3}$$

Предполагается, что  $h \sim R(c_{\min}, c_{\max})$ , где  $c_{\min}, c_{\max}$  – минимально и максимально возможные яркости точек на изображении, причём

$$\sigma_\xi(x,y) \ll c_{\max} - c_{\min}.$$

Пусть, каким-либо методом [11], получены оценки параметров евклидова преобразования изображения  $(\hat{\alpha}, \hat{\beta}, \hat{\varphi})$  и известна плотность распределения ошибки оценивания. Предположим, что компоненты этого случайного вектора независимы и распределены нормально с известными дисперсиями:

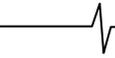
$$\begin{aligned} p(\alpha, \beta, \varphi) &= p(\alpha)p(\beta)p(\varphi), \\ p(\alpha) &\sim N(\hat{\alpha}, \sigma_{\alpha\beta}^2), \quad p(\beta) \sim N(\hat{\beta}, \sigma_{\alpha\beta}^2), \\ p(\varphi) &\sim N(\hat{\varphi}, \sigma_\varphi^2), \end{aligned} \tag{4}$$

где  $p(\alpha, \beta, \varphi)$  – совместная апостериорная плотность распределения параметров преобразования;  $\sigma_{\alpha\beta}^2$  – дисперсия оценки смещения,  $\sigma_\varphi^2$  – дисперсия оценки угла поворота. Дисперсии оценки смещения по  $x$  и  $y$  полагаем равными одной и той же величине  $\sigma_{\alpha\beta}^2$ .

Таким образом, известны величины:  $g(x,y)$ ,  $l$ ,  $\sigma_z^2$ ,  $\sigma_\xi^2$ ,  $(\hat{\alpha}, \hat{\beta}, \hat{\varphi})$ ,  $\sigma_{\alpha\beta}^2$ ,  $\sigma_\varphi^2$ . Необходимо в точке  $(x_0, y_0)$  принять решение о наличии объекта, то есть найти  $r$ , которая является оценкой для  $r$ . Требования к алгоритму задаются в форме Неймана-Пирсона:

$$\begin{aligned} P(\bar{r} = 1 / r = 0) &\leq p_-, \\ P(\bar{r} = 1 / r = 1) &= p_+ \rightarrow \max, \end{aligned} \tag{5}$$

где  $p_-$  – заданное значение вероятности ложного выделения,  $p_+$  – вероятность правильного выделения,



которая должна быть максимизирована путём оптимального выбора параметров алгоритма выделения объектов.

**РЕШЕНИЕ ЗАДАЧИ.** Оптимальное решающее правило для поставленной задачи определяется выражением:

$$u^*(l) = \begin{cases} 1, & \text{если } \Lambda(l) > \Lambda_0; \\ 0, & \text{в противном случае,} \end{cases} \quad (6)$$

где  $\Lambda(l)$  – отношение правдоподобия,

$$\Lambda(l) = \frac{p(l/r=1)}{p(l/r=0)}, \quad (7)$$

$\Lambda_0$  – пороговое отношение правдоподобия,  $p(l/r=1)$ ,  $p(l/r=0)$  – условные плотности распределения яркости наблюдаемого изображения при наличии и отсутствии объекта соответственно. При использовании критерия Неймана-Пирсона  $\Lambda_0$  вычисляется из условия:

$$\int_{\Lambda(l) > \Lambda_0} p(l/r=0) dl = p_- \quad (8)$$

Определим условную плотность  $p(l/r=1)$ . При  $r=1$  из (1) следует, что яркость наблюдаемого изображения  $l$  является композицией независимых распределений  $p(h)$  и  $p(\xi)$ :

$$p(l/r=1) = p(h) * p(\xi), \quad (9)$$

где  $*$  – символ свёртки. Поскольку  $\sigma_\xi(x, y) \ll c_{\max} - c_{\min}$ , то, пренебрегая влиянием шума  $\xi$  на плотность  $p(l/r=1)$ , будем считать, что

$$p(l/r=1) = p(h) = \begin{cases} \frac{1}{c_{\max} - c_{\min}}, & \text{если } c_{\max} \leq h \leq c_{\min}; \\ 0, & \text{иначе.} \end{cases} \quad (10)$$

Оценим условное распределение  $p(l/r=0)$ . Для этого вначале найдём распределение  $p(f)$  яркости деформированного фона. Из (2), (3) следует

$$f = g \left( \begin{matrix} (x_0 - z_x) \cos \varphi + (y_0 - z_y) \sin \varphi - \alpha, \\ -(x_0 - z_x) \sin \varphi + (y_0 - z_y) \cos \varphi - \beta \end{matrix} \right). \quad (11)$$

Обозначим аргументы функции  $g(x, y)$  через  $\mu_x$  и  $\mu_y$  и преобразуем их:

$$\begin{aligned} f &= g(\mu_x, \mu_y) \\ \mu_x &= x_0 \cos \varphi + y_0 \sin \varphi - (z_x \cos \varphi + z_y \sin \varphi) - \alpha \\ \mu_y &= -x_0 \sin \varphi + y_0 \cos \varphi - (-z_x \sin \varphi + z_y \cos \varphi) - \beta \end{aligned} \quad (12)$$

Рассмотрим величины, заключенные в скобки в (12), обозначив их  $\zeta_x$ ,  $\zeta_y$ . Поскольку система случайных величин  $\zeta_x$ ,  $\zeta_y$  получена в результате поворота системы величин  $z_x$ ,  $z_y$ , то  $\zeta_x$ ,  $\zeta_y$  являются независимыми гауссовыми случайными величинами с плотностями распределения, равными  $N(0, \sigma_z^2)$ .

Далее, выполнив в (12) следующую замену:  $\varphi = \hat{\varphi} + \Delta\varphi$ ,  $\alpha = \hat{\alpha} + \Delta\alpha$ ,  $\beta = \hat{\beta} + \Delta\beta$ , получим:

$$\begin{aligned} \mu_x &= x' \cos \Delta\varphi + y' \sin \Delta\varphi - \zeta_x - \alpha - \Delta\alpha \\ \mu_y &= -x' \sin \Delta\varphi + y' \cos \Delta\varphi - \zeta_y - \hat{\beta} - \Delta\beta \end{aligned} \quad (13)$$

где  $x' = x_0 \cos \hat{\varphi} + y_0 \sin \hat{\varphi}$ ,  $y' = -x_0 \sin \hat{\varphi} + y_0 \cos \hat{\varphi}$ .

Величины  $\Delta\alpha$ ,  $\Delta\beta$ ,  $\Delta\varphi$  являются отклонениями истинных значений параметров преобразований от их оценок. Они – нормальные, центрированные и независимые случайные величины с дисперсиями  $\sigma_{\alpha\beta}$ ,  $\sigma_{\alpha\beta}$ ,  $\sigma_\varphi$  соответственно. Принимая гипотезу о малости  $\Delta\varphi$ , заменим тригонометрические функции разложениями Тейлора, отбрасывая слагаемые второго и высших порядков:  $\cos \Delta\varphi \approx 1$ ,  $\sin \Delta\varphi \approx \Delta\varphi$ . Тогда (13) примет вид:

$$\begin{aligned} \mu_x &= x' - \hat{\alpha} + y' \Delta\varphi - \zeta_x - \Delta\alpha \\ \mu_y &= y' - \hat{\beta} - x' \Delta\varphi - \zeta_y - \Delta\beta \end{aligned} \quad (14)$$

Для нахождения распределения  $p(f)$ , введём в рассмотрение точку  $(\hat{x}, \hat{y})$ , заданную выражениями

$$\begin{aligned} \hat{x} &= x' - \hat{\alpha} = \hat{x} = x_0 \cos \hat{\varphi} + y_0 \sin \hat{\varphi} - \hat{\alpha} \\ \hat{y} &= y' - \hat{\beta} = -x_0 \sin \hat{\varphi} + y_0 \cos \hat{\varphi} - \hat{\beta} \end{aligned} \quad (15)$$

Эта точка представляет собой результат отображения точки  $(x_0, y_0)$  посредством евклидова преобразования, заданного оценками  $(\hat{\alpha}, \hat{\beta}, \hat{\varphi})$ . Применяя разложение Тейлора функции  $g(\mu_x, \mu_y)$  в окрестности точки  $(\hat{x}, \hat{y})$ , приводя подобные слагаемые, получим:

$$\begin{aligned} f &= g(\mu_x, \mu_y) \approx g(\hat{x}, \hat{y}) + \left( \frac{\partial g(\hat{x}, \hat{y})}{\partial x} y' - \frac{\partial g(\hat{x}, \hat{y})}{\partial y} x' \right) \Delta\varphi + \\ &+ \frac{\partial g(\hat{x}, \hat{y})}{\partial x} (-\zeta_x - \Delta\alpha) + \frac{\partial g(\hat{x}, \hat{y})}{\partial y} (-\zeta_y - \Delta\beta) \end{aligned} \quad (16)$$

В первом приближении  $f$  является гауссовой случайной величиной. Найдём её математическое ожидание и дисперсию:

$$E(f) = g(\hat{x}, \hat{y}) \quad (17)$$

$$\begin{aligned} D(f) &= \left( \frac{\partial g(\hat{x}, \hat{y})}{\partial x} y' - \frac{\partial g(\hat{x}, \hat{y})}{\partial y} x' \right)^2 \sigma_\varphi^2 + \\ &+ \left( \frac{\partial g(\hat{x}, \hat{y})}{\partial x} \right)^2 + \left( \frac{\partial g(\hat{x}, \hat{y})}{\partial y} \right)^2 (\sigma_{\alpha\beta}^2 + \sigma_z^2) \end{aligned} \quad (18)$$

С использованием понятия градиента изображения выражение (18) принимает вид

$$D(f) = \left[ \nabla g(\hat{x}, \hat{y}) \begin{pmatrix} y' \\ -x' \end{pmatrix} \right]^2 \sigma_\varphi^2 + |\nabla g(\hat{x}, \hat{y})|^2 (\sigma_{\alpha\beta}^2 + \sigma_z^2),$$

$$\text{где } \nabla g = \left( \frac{\partial g}{\partial x}, \frac{\partial g}{\partial y} \right). \quad (19)$$

Поскольку  $p(l/r=0) = p(f) * p(\xi)$ , то искомая плотность распределения  $p(l/r=0)$  близка к гауссовой со следующими математическим ожиданием и дисперсией:

$$E(l/r=0) = g(\hat{x}, \hat{y}), \quad (20)$$

$$\begin{aligned} D(l/r=0) &= \left[ \nabla g(x, y) \begin{pmatrix} y' \\ -x' \end{pmatrix} \right]^2 \sigma_\varphi^2 + \\ &+ |\nabla g(x, y)|^2 (\sigma_{\alpha\beta}^2 + \sigma_z^2) + \sigma_\xi^2 \end{aligned} \quad (21)$$

Подставляя найденную условную плотность вероятностей  $p(l/r=1)$  (7), используя (6) и (8), получим следующее правило принятия решения о наличии объекта в точке  $(x_0, y_0)$ :

$$u^*(l) = \begin{cases} 1, & \text{если } p(l/r=0) < \Lambda'; \\ 0, & \text{в противном случае,} \end{cases} \quad (22)$$

где порог  $\Lambda'$  определяется из соотношения:

$$\int_{p(l/r=0) < \Lambda'} p(l/r=0) dl = p_- \quad (23)$$

Таким образом, решение о наличии объекта принимается, если величина  $l$  не попадает в  $(1-p_-) \cdot 100\%$  доверительный интервал для распределения  $p(l/r=0)$ . В противном случае принимается решение о присутствии фона в точке  $(x_0, y_0)$ . Формально это правило можно задать следующим образом:

$$\hat{r} = \begin{cases} 1, & \text{если } |g(\hat{x}, \hat{y}) - l| > d(p_-) \sqrt{D(l/r=0)}; \\ 0, & \text{иначе.} \end{cases} \quad (24)$$

Здесь  $d(p_-)$  – полуширина  $(1-p_-) \cdot 100\%$  доверительного интервала для нормированной гауссовой случайной величины.

Таким образом, зная фоновое изображение  $g(x, y)$ , оценки величин  $(\alpha, \beta, \varphi)$ , дисперсии  $\sigma_z^2$ ,  $\sigma_{\alpha\beta}^2$ ,  $\sigma_\varphi^2$ ,  $\sigma_\xi^2$ , можно произвести оптимальное по критерию Неймана–Пирсона выделение объекта путём проверки попадания наблюдаемого значения яркости  $l$  в  $(1-p_-) \cdot 100\%$  доверительный интервал яркости, центрированный на величине  $g(\hat{x}, \hat{y})$ .



а) кадр № 50



б) кадр № 200

Рис. 3 – Кадры видеопоследовательности, использованной при исследованиях

Эталонные данные были сформированы вручную с помощью графического редактора для каждого из пяти кадров, по которым проводилась количественная оценка. Эталонные данные представляют собой бинарные изображения, единичные пиксели которых соответствуют наличию объекта, а нулевые – наличию фона в дан-

ной точке. Для упрощения эксперимента эталонные данные формировались не для кадров  $l(i, j, n)$ , приведённых на рис. 3, а для кадров, полученных после поворота этих изображений на угол  $(-\hat{\varphi})$ . Исходное и эталонное изображения для двухсотого кадра видеопоследовательности приведены на рис. 4.

Поясним качественный смысл выражений (21), (24). При деформациях изображения наибольшие ошибки измерения яркости присутствуют вблизи резких перепадов яркостей. Поэтому в таких областях ширину доверительного интервала необходимо увеличить. Напротив, в тех местах, где изображение однородно, деформации не вызывают больших ошибок, и даже небольшое изменение яркости свидетельствует о наличии объекта.

Заметим, что при неизвестном фоновом изображении  $g(x, y)$  и дисперсии аддитивного шума, для их оценки можно использовать подходы, предложенные в [12]. В целом предлагаемый подход является обобщением подхода [9] и сводится к нему при отсутствии вращения изображения.

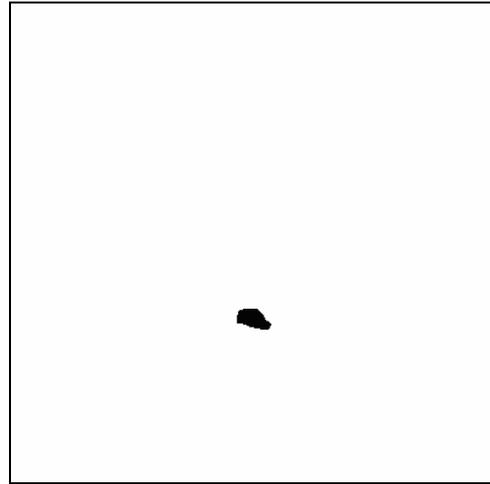
ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ. Для экспериментальных исследований был выбран видеосюжет продолжительностью 231 кадр (9,24 секунды), содержащий изображение движущегося автомобиля. Отношение сигнал/шум составляло около 12. При съёмке сюжета ось датчика поворачивалась оператором вслед за автомобилем и, кроме того, происходил крен датчика. Сюжет снят цифровой камерой DMK 21F04 с прогрессивной развёрткой. Размер каждого кадра видеосюжета составлял 640x480 точек.

Оценка параметров евклидова преобразования выполнялась по центральной области кадра размером 256x256 точек. Выделение объектов производилось в центральной части кадра размером 400x400 точек. Кадры видеосюжета были искажены вследствие турбулентности атмосферы.

Количественными параметрами, характеризующими качество выделения объектов были частота правильного выделения и частота ложного выделения. Для вычисления количественных параметров использовались кадры видеопоследовательности с номерами 50, 60, 100, 150, 200. Изображения кадров с номерами 50 и 200 представлены на рис. 3.



а) исходное изображение на кадре №200, подвергнутое повороту



б) эталонное изображение на кадре № 200

Рис. 4 – Исходное и эталонное изображения на двухсотом кадре видеосюжета

В процессе исследований были построены характеристики выделения для различных пар величин  $\sigma_\varphi$  и  $\sigma_s$ , где  $\sigma_s = \sqrt{\sigma_{\alpha\beta}^2 + \sigma_z^2}$  – суммарное СКО ошибок оценивания смещения. Для того, чтобы построить кривую характеристики выделения, варьировался параметр ширины доверительного интервала  $d$ , и при каждом его значении вычислялась частота правильного выделения  $\hat{p}_+$  и частота ложного выделения  $\hat{p}_-$ .

На рис. 5 приведены характеристики выделения, полученные при различных значениях величин  $\sigma_\varphi$  и  $\sigma_s$ . При вычислении частот истинного и ложного выделения производилось усреднение результатов по всем обрабатываемым кадрам. Крестами на кривых отмечены опорные точки, по которым строились графики. Числа около крестов обозначают значения ширины доверительных интервалов  $d$  в данных точках.

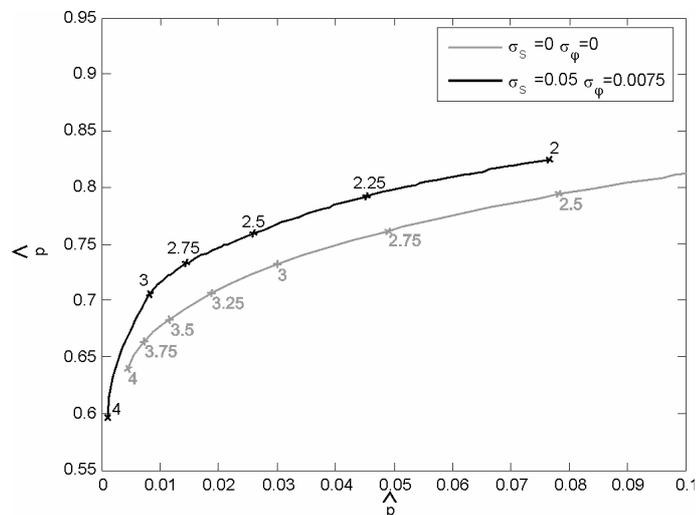


Рис. 5 – Характеристики выделения при различных значениях параметров

Анализ характеристик выделения показывает, что использование предлагаемого подхода позволяет повысить частоту правильного выделения на 4-6% и одновременно с этим снизить частоту ложного выделения более чем в 2 раза при одном и том же значении параметра  $d$ .

На рис. 6 приведены бинарные изображения, полученные в результате выделения объектов с различными параметрами алгоритма. Заметно значительно меньшее количество неверно выделенных точек на рис. 6 (б) по сравнению с рис. 6 (а). Частота правильного выделения на рисунках одна и та же и равна 0,7.

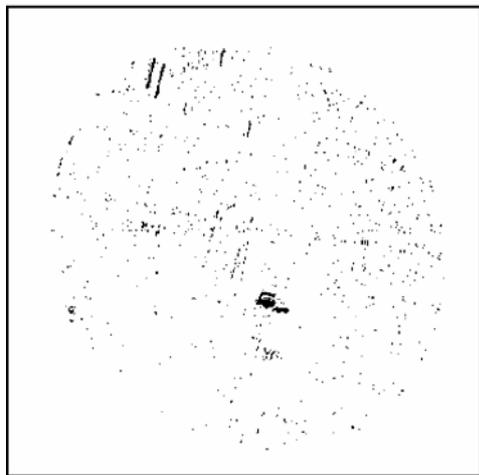
**РЕАЛИЗАЦИЯ.** Для проверки работоспособности рассмотренных подходов в реальных условиях наблюдения был создан макет системы обнаружения и сопровождения движущихся объектов. Разработанный макет состоит из видеокамеры, поворотного устройства, платы ввода изображений и двух компьютеров с процессорами Pentium 4 (3 ГГц). Первый компьютер решает задачи оценки параметров геометрических преобразований, а на второй возложены задачи выделения, обнаружения, анализа траекторий движения объектов, взаимодействия с пользователем и управления поворотным устройством. Обмен данными

между компьютерами осуществляется по сетям Ethernet или FireWire.

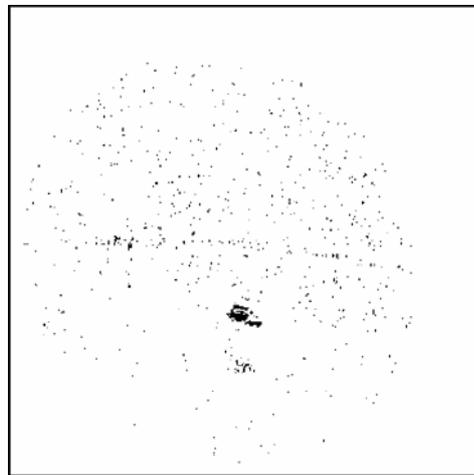
Для макета разработано специализированное программное обеспечение на языке C++, позволяющее работать с такими источниками видеoinформации как AVI-файлы, аналоговые видеокамеры (через фреймграббер

Matrox Meteor II), цифровые видеокамеры стандарта DCAM. При размере анализируемой части кадра 256x256 точек, частота обработки кадров составила 25 Гц.

Испытания макета системы обнаружения и сопровождения движущихся объектов подтвердили эффективность разработанных подходов.



а)  $d = 3,25$ ;  $\sigma_s = 0$ ;  $\sigma_\varphi = 0$



б)  $d = 3$ ;  $\sigma_s = 0,05$ ;  $\sigma_\varphi = 0,0075$

Рис. 6 – Результаты выделения объекта при различных параметрах алгоритма

Таким образом, решение задач обнаружения и сопровождения объектов требует организации обработки информации на нескольких иерархических уровнях, включающих восстановление изображений, выделение объектов, оценку параметров сегментов, слежение, оценку параметров, обнаружение, распознавание и сопровождение объектов.

Рассмотренный подход может быть использован при разработке систем обработки изображений реального времени, реализуемых на базе современных ПЛИС и сигнальных процессоров. В настоящее время разработанные алгоритмы используются при модернизации программно-алгоритмического обеспечения для системы обработки изображений реального времени «Охотник» [1].

#### Литература

1. Алпатов Б.А., Зеленюк Ю.И., Костяшкин Л.Н., Романов Ю.Н., Семенов В.П. Видеоинформационные технологии в бортовом радиоэлектронном оборудовании перспективных авиационных систем: концепция, алгоритм и решения // Радиотехника. – 2002. – №8. – С. 102-112.
2. Алпатов Б.А., Селяев А.А. Алгоритм оценки местоположения объекта на двумерном изображении // Изв. вузов. – Приборостроение. – 1988. – №5. – С. 3-5.
3. Алпатов Б.А., Хлудов С.Ю. Алгоритм последовательных испытаний в задаче совмещения двумерных изображений // Изв. вузов. Сер. Электромеханика. – 1988. – №7. – С. 87-91.
4. Применение методов сегментации изображений в автономных системах обнаружения, распознавания и сопровождения движущихся целей / Под ред. П. А. Бакута // За-

рубежная радиоэлектроника. – 1987. – №10. – С. 3-93.

5. Алпатов Б.А., Балашов О.Е. Исследование методов оценки фрактальной размерности и сегментация изображения // Цифровая обработка сигналов и ее применения. Тез. докл. 6-й междунар. конф. Том 2. – М.: 2004. – С. 97-99.
6. Алпатов Б.А., Муравьев С.И., Муравьев В.С. Пространственный алгоритм выделения объектов на основе адаптивной пороговой обработки // Цифровая обработка сигналов и ее применения. Мат. 8-й междунар. конф. Том 2. – М.: 2006. – С. 445-448.
7. Алпатов Б.А. Алгоритм обнаружения и выделения движущегося фрагмента изображения // Техника средств связи. Серия Техника телевидения. – 1991. – №2. – С. 72-76.
8. Алпатов Б.А., Блохин А.Н. Модели и алгоритмы обнаружения и выделения движущихся фрагментов изображений // Автометрия. – 1995. – №4. – С. 100-104.
9. Алпатов Б.А., Бабаян П.В. Выделение движущихся объектов в условиях геометрических искажений изображения // Цифровая обработка сигналов. – 2004. – №4. – С. 9-14.
10. Алпатов Б.А., Балашов О.Е., Степашкин А.И. Прогнозирование управления приводами в видеокomпьютерной системе сопровождения объектов // Цифровая обработка сигналов и ее применения. Мат. 7-й международной конференции. Том 2. – М.: 2005. – С. 333-336.
11. Zitova B., Flusser J. Image registration methods: a survey // Image and Vision Computing 21 – 2003, pp. 977–1000.
12. Алпатов Б.А., Бохан К.А. Алгоритм автоматического обнаружения, выделения и оценки динамических объектов, возникающих в последовательности телевизионных кадров // Цифровая обработка сигналов и её применение. Тез. докл. 3-й междунар. науч. конф. – М.: МЦНТИ, 2000. – С. 105-109.

## ФУНКЦИОНАЛЬНЫЙ МЕТОД ТРАССИРОВКИ ГРАНИЦ РАЗДЕЛА ДВУХ СРЕД, ИМЕЮЩИХ ЦИЛИНДРИЧЕСКУЮ ГЕОМЕТРИЮ

Аринин В. А.

### Введение

Для исследования свойств материалов в области давлений мегабарного диапазона и для уменьшения эффектов необратимого нагрева используется техника квазиизэнтропического нагружения в экспериментальных устройствах цилиндрической геометрии [1, 2], использующих энергию взрывчатых веществ. Наиболее распространенным методом обработки получаемых в этих опытах рентгено снимков были экспертные оценки. Но, из-за неизбежного субъективизма, получаемые результаты не считались слишком надежными. Применение в качестве альтернативы градиентного метода не обеспечивает приемлемую точность из-за трудноконтролируемых регулярных ошибок [3].

Предлагаемый метод трассировки цилиндрических границ на статических рентгено снимках показал его высокую точность.

*Природа цилиндрической геометрии позволяет существенно улучшить статистику изображения за счет суммирования строк, перпендикулярных оси симметрии экспериментального устройства, что делает эффективным применение метода функциональной трассировки. Для его реализации в рамках данной работы для цилиндрического случая решены следующие задачи: определение направления оси симметрии для правильной ориентации изображения; определение положения оси симметрии для измерения асимметрии в работе устройства; контекстное устранение неравномерности экспозиции изображения, минимизирующее ошибки функциональной трассировки.*

**Исходные данные.** Типичные рентгено снимки работы устройств цилиндрической геометрии приведены на рис. 1: а) снимок хорошего контраста с протяженными участками цилиндрических границ и ярко выраженными границами разделов; б) снимок низкого контраста с плохо проработанными границами разделов, участки цилиндрических границ протяженные; в) снимок с неопределенными границами (внутренняя), границами малого контраста (легкая вставка на внешней границе внутреннего цилиндра), с короткими участками цилиндрических границ.

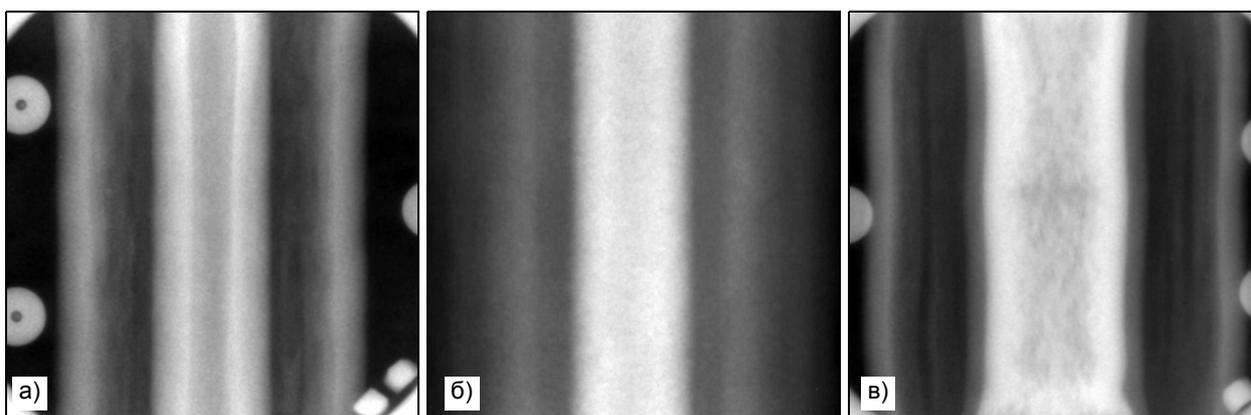


Рис. 1. Примеры рентгено снимков работы устройств цилиндрической геометрии

**Сущность метода.** Предлагаемый метод основан на следующей гипотезе о функциональных особенностях протяженных цилиндрических границ на рентгено снимках.

1. Из-за экспоненциальной природы поглощения рентгеновского излучения логарифм интенсивности прошедшего излучения  $D = \ln(I)$ , при отсутствии искажающих факторов, пропорционален просвечиваемой толщине материала, следовательно,

слева и справа от изображения цилиндрической границы функция  $D$  должна хорошо аппроксимироваться эллипсом.

2. Искажения передаточной характеристики и взаимное влияние границ друг на друга в коаксиальной геометрии несущественно исказят приграничную функцию.

3. Функция, аппроксимирующая изображение ближе к его оси, при пересечении на границе с функцией, аппроксимирующей изображение с дру-

гой стороны границы, в точке пересечения должна иметь бесконечную производную. Точка пересечения является искомой координатой границы.

4. Размытие изображения компенсируется экстраполяцией данных в окрестность границы.

5. Наличие шумов на изображении не вызывает систематической ошибки.

Проверка гипотезы на многочисленных математических моделях и на статических рентгенограммах показала ее применимость.

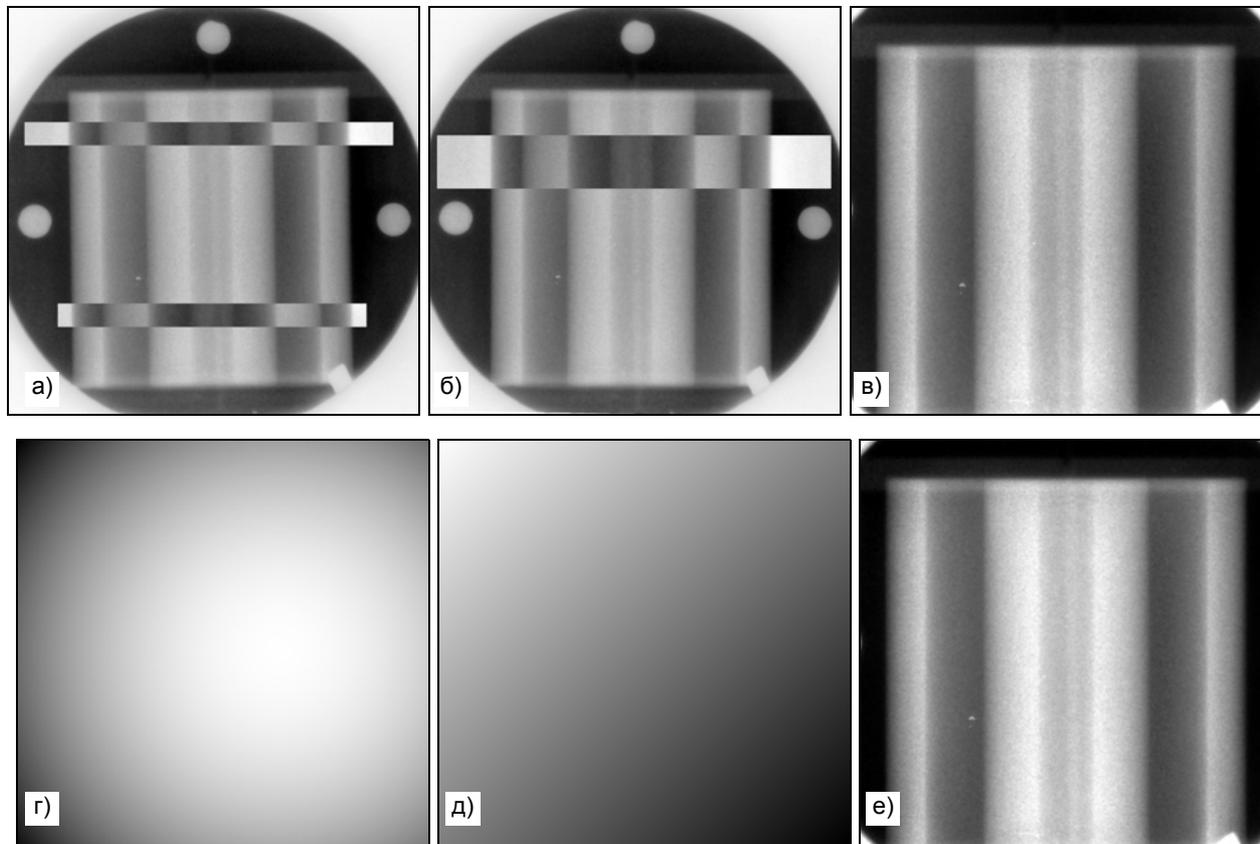


Рис. 2. Подготовка изображений к трассировке границ

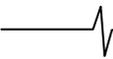
**Подготовка изображений: ориентация оси симметрии.** В общем случае при проведении рентгено съемки ось симметрии устройства образует с вертикалью некоторый угол. Это не позволяет напрямую суммировать строки изображения с целью получения качественного функционального профиля. Для определения угла наклона оси воспользуемся корреляционным подобием относительно нее рентгеновского изображения цилиндра. Выберем две полосы из верхней и нижней частей изображения, причем базовая полоса (допустим, верхняя) должна быть более длинной. Для каждой полосы просуммируем столбцы с целью получения одномерных дискретных функций. Пусть длина базовой функции -  $M$  элементов, а второй  $N$  элементов, причем  $M > N$ . Тогда, делая выборки из базовой функции длиной  $N$  слева направо в шаг в один дискрет, можно получить  $M - N + 1$  коэффициентов корреляции, используя соотношение (1), где  $a_i$  – приведенный элемент базовой выборки длины  $N$  (после вычитания из него среднего арифметического выборки),  $b_i$  – приведенный элемент второй функции.

$$k_{\text{корр}} = \frac{\sum_{i=1}^N a_i b_i}{\sqrt{\sum_{i=1}^N a_i^2 \cdot \sum_{i=1}^N b_i^2}} \quad (1)$$

Для субпиксельного определения положения максимума функции сдвиговой корреляции применим квадратичную интерполяцию. Пусть максимум находится в точке с координатой  $i$ , тогда уточненное положение максимума –  $i + \Delta i$ , где  $\Delta i$  определяется следующим соотношением:

$$\Delta i = \frac{1}{2} \cdot \frac{k_{i-1} - k_{i+1}}{k_{i+1} - 2k_i + k_{i-1}}$$

Положение максимума функции корреляции определяет сдвиг нижней полосы относительно верхней, что, с учетом расстояния между полосами, позволяет найти угол поворота. На рис. 2, а полосы, выбранные для корреляции, показаны негативом. Определенный по ним угол поворота, ориентирующего ось симметрии вертикально, составил  $-1,087^\circ$ . Результат поворота показан на рис. 2, б.



**Подготовка изображений: определение положения оси симметрии снимка.** Эта операция не обязательна, но весьма полезна, так как позволяет определить степень асимметрии границ. По взрывным снимкам можно оценить качество срабатывания экспериментального устройства, а по статическим – точность метода. Для определения положения оси симметрии также применим корреляционный метод. Максимум сдвиговой корреляции между базовой полосой и ее участком, отраженным относительно вертикали, определяет наиболее вероятное положение оси. На рис. 2б выбранная полоса показана негативом. Изображение в) представляет собой фрагмент изображения б), вырезанный так, чтобы ось симметрии проходила точно по его середине.

**Подготовка изображений: выравнивание поля экспозиции.** Данная операция обеспечивает равные веса суммируемых строк при получении функционального профиля. Ее влияние на точность трассировки границ в большинстве случаев ничтожно, тем не менее, ее следует выполнять, да и восприятие снимков с выровненным полем существенно лучше. Выравнивание поля производится с помощью

мультипликативного маскирования, то есть, для устранения неравномерности экспозиции, изображение поэлементно умножается на маску. Используется два типа масок: сферическая и плоская (если смотреть на трехмерное представление изображения маски). Контроль производится по гистограмме фона: чем уже гистограмма, тем лучше выравнивание. При подборе масок варьируется шесть параметров, поэтому процесс маскирования является в некотором роде искусством. Его автоматизация сложна и требует больших вычислительных ресурсов. На рис. 2,гд приведены сферическая и линейная маски, использовавшиеся при коррекции неравномерности поля экспозиции снимка, представленного на рис. 2,в. Диапазон значений сферической маски составил  $0,801 \div 1$ , линейной –  $0,962 \div 1$ .

**Трассировка.** На рис. 3 приведен результат трассировки опыта с цилиндрической геометрией, границы имеют большую протяженность. Приведенный снимок соответствует области  $99,34 \times 99,34$  мм и представлен дискретным изображением размером  $1000 \times 1000$  пикселей. Функциональный профиль получен по 400 строкам: с 311-й по 710-ю (показаны белой стрелкой).

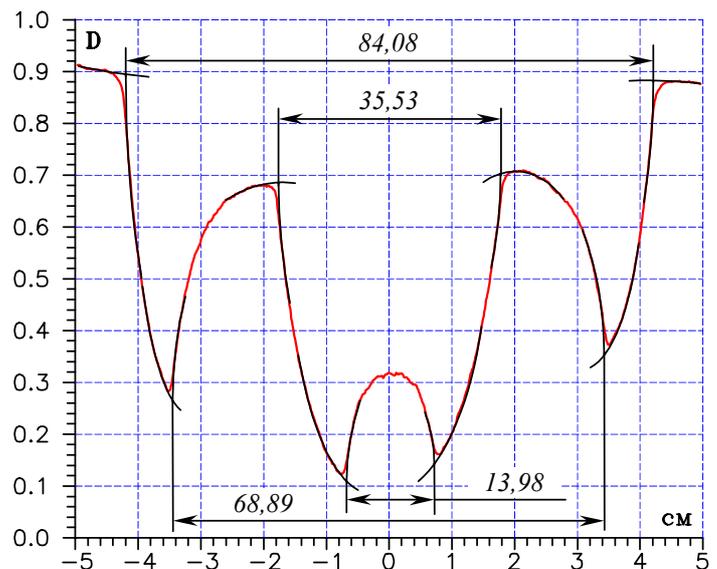
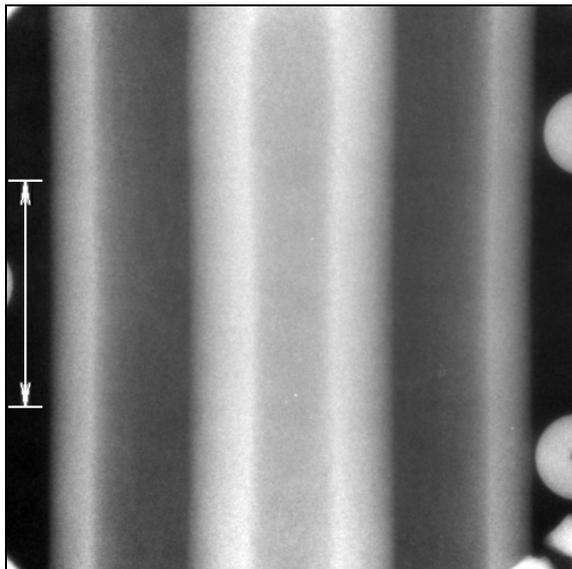


Рис. 3. Трассировка границ в устройствах цилиндрической геометрии

**Погрешность метода.** Контекстная оценка точности состоит в сравнении результата обработки по выбранному участку с обработкой по одинаковым выборкам, покрывающим весь участок. Представительная оценка может быть получена по  $7 \div 10$  выборкам, для устранения влияния возможного дрейфа радиусов по вертикали на искомую погрешность, выборки следует делать чересстрочно. Пусть, например, необходимо обработать  $n=8$  выборок, тогда в первую войдут строки участка с номерами  $1, 9, \dots, 8n+1$ , во вторую –  $2, 10, \dots, 8n+2$  и так далее. Пусть среднеквадратичное отклонение границ, найденных по выборкам,

от значения, найденного по всему участку, равно  $\sigma_{pt}$ , тогда искомое значение среднеквадратичного отклонения результата  $\sigma = \sigma_{pt} \sqrt{2/n}$ . Отметим, что усреднение результатов обработки по выборкам не совпадает с результатом обработки всего участка (ошибки не эргодичны). Оценка точности по модели состоит в проведении многократной рентгено съемки теста, близкого по характеристикам к состоянию экспериментального устройства на момент регистрации и при тех же условиях съемки. Полученное значение  $\sigma_{ist}$  определяет функцию нормального распределения ошибок в статическом эксперименте. Эта функция должна

быть свернута с функцией зависимости интенсивности излучения рентгеновской установки от времени с учетом скорости движения границы, что и определит рабочее значение  $\sigma$  (в процессе проведения рентгенографических экспериментов зависимость интенсивности излучения от времени все-

гда фиксируется). Увеличение погрешности из-за динамической нерезкости связано с необходимостью экстраполяции приграничных функций на большее расстояние по оси ординат. В таблице приведены значения  $\sigma$ , полученные контекстной оценкой по 8-ми выборкам.

Таблица

Результаты обработки опыта по квазиизэнтропическому нагружению дейтерия

Код	R <sub>1</sub>			R <sub>2</sub>			R <sub>3</sub>			R <sub>4</sub>		
	R, мм	$\sigma$ , мм	$\sigma$ , %	R, мм	$\sigma$ , мм	$\sigma$ , %	R, мм	$\sigma$ , мм	$\sigma$ , %	R, мм	$\sigma$ , мм	$\sigma$ , %
13/2	6,99	0,12	1,72	17,77	0,09	0,51	34,45	0,12	0,35	42,04	0,10	0,24

### Заключение

Предложенный функциональный метод определения местоположения границ раздела двух сред в экспериментальных устройствах цилиндрической геометрии по их рентгенограммам основан на простой и эвристически понятной гипотезе. Получаемые с помощью него результаты наглядны и не содержат систематических ошибок. Точность метода исключительно высока: в статических экспериментах при использовании рентгенопленки составляет приблизительно 0,3%, при использовании современных регистраторов на основе пакетов фотохромных экранов достигает 0,05%. Более того, оказалось, что регистрируется ошибка в определении коэффициента рентгенопоглощения в четвертом знаке после запятой, что соответствует ошибке измерения расстояния от источника излучения до объекта в 3 мм при характерном значении этого расстояния в 4 м. Подобная точность не свойственна классической технической

рентгенографии, использующей установки с жестким спектром излучения ( $E_{\text{тп}} \approx 50$  МэВ) [4].

Автор благодарит сотрудников Института физики взрыва Российского Федерального ядерного центра ВНИИ экспериментальной физики М. А. Мочалова и В. Д. Орлова за любезно предоставленный иллюстративный материал.

### Литература

1. В. Д. Урлин, М. А. Мочалов, О. Л. Михайлова, ЖЭТФ 111 вып. 6, 2099 (1997).
2. О. Л. Михайлова, М. А. Мочалов, А. И. Соколова, В. Д. Урлин, Теплофизика высоких температур, 38, №2, 227 (2000).
3. У. Претт. «Цифровая обработка изображений», М.: Мир, 1982.
4. А. Роуз. «Зрение человека и электронное зрение», М.: Мир, 1977.

## НОВЫЕ КНИГИ!

**Гонсалес Р., Вудс Р. Цифровая обработка изображений** \ М.: Техносфера, 2005. – 1072 с.

Монография раскрывает базовые понятия и методологию компьютерной обработки изображений, дает основы для дальнейшего изучения этой многогранной и быстро развивающейся области. Является одним из наиболее популярных, известных в мире и полных учебников в области теории и методов цифровой обработки видеoinформации.

Рассмотрены все основные направления обработки и анализа изображений, включая основы теории восприятия и регистрации видеoinформации, методы фильтрации, вейвлет-преобразования, улучшения, восстановления и сжатия черно-белых и цветных изображений. Обсуждаются также вопросы сегментации, распознавания образов, описания и представления деталей, морфологического анализа изображения. Все разделы сопровождаются большим количеством примеров и иллюстраций.

Книга рассчитана на научных работников и профессиональных программистов, специалистов по компьютерному дизайну, студентов и преподавателей.



**9-я Международная Конференция и Выставка**  
**ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ И ЕЕ ПРИМЕНЕНИЕ**  
*28-30 марта 2007 г., Москва, Россия*  
*The 9<sup>th</sup> International Conference and Exhibition*  
**DIGITAL SIGNAL PROCESSING AND ITS APPLICATIONS**

**ИНФОРМАЦИОННОЕ ПИСЬМО**

*Уважаемый коллега!*

**Приглашаем Вас принять участие в работе 9-ой Международной конференции и выставки «Цифровая обработка сигналов и ее применение – DSPA'2007», которая состоится в Москве 28-30 марта 2007 года**

**ОРГАНИЗАТОРЫ:**

- Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова
- IEEE Signal Processing Society
- Российская секция IEEE
- Институт радиотехники и электроники РАН
- Институт проблем управления РАН
- Институт проблем передачи информации РАН
- Московский научно-исследовательский телевизионный институт
- Компания AUTEX Ltd. (АВТЭКС)

**ОСНОВНЫЕ НАПРАВЛЕНИЯ РАБОТЫ КОНФЕРЕНЦИИ:**

- Теория сигналов и систем
- Теория и методы ЦОС
- Цифровая обработка многомерных сигналов
- Цифровая обработка речевых и звуковых сигналов
- Цифровая обработка изображений
- ЦОС в системах телекоммуникаций
- ЦОС в радиотехнических системах
- ЦОС в системах управления
- Цифровая обработка измерительной информации
- Нейрокомпьютерная обработка сигналов и изображений
- Цифровое телевидение и телевидение
- Цифровое радиовещание
- ЦОС в системах защиты информации
- Проектирование и техническая реализация систем ЦОС
- Проблемы подготовки специалистов в области ЦОС

**ОРГАНИЗАЦИИ И ФИРМЫ** приглашаются представить свои разработки и продукцию на **ВЫСТАВКЕ** или провести **СЕМИНАРЫ-ПРЕЗЕНТАЦИИ**, освещающие их коммерческие проекты, продукцию и услуги.

По опыту прошедших «DSPA», проведение одновременно Конференции и Выставки, а также тематических семинаров оказалось очень эффективным для всех участников мероприятия. «DSPA» стала местом встречи и продуктивного общения представителей науки и производства, заказчиков и разработчиков, потребителей и поставщиков, преподавателей и студентов, работодателей и соискателей, научно-технических издательств и читателей.

**РАБОТА КОНФЕРЕНЦИИ** будет организована в форме пленарных и проблемно-тематических заседаний, стендовых докладов и заседаний «круглого стола». В качестве пленарных будут заслушаны заказные доклады по основным направлениям работы Конференции.

**Рабочие языки конференции** – русский и английский.

Для участия в работе Конференции с докладами необходимо представить в Оргкомитет следующие материалы (на дискете 3.5 дюйма или по E-mail: [nto.popov@mtu-net.ru](mailto:nto.popov@mtu-net.ru)):

1. Заявка (фамилия, имя, отчество, место работы, должность, почтовый адрес для переписки, телефон, факс, E-mail); в этих сведениях указать название секции.

2. Тезисы на русском и английском языках (по одной полной странице с учетом названия доклада, авторов и организации).

3. Текст доклада (3-4 полные страницы) на русском языке.

4. Одновременно с присылкой материалов оплатить регистрационный взнос. (В случае отклонения доклада взнос возвращается). Оплата платежным поручением или через Сбербанк. При необходимости может быть выставлен счет.

Доклады, включенные в Программу конференции, будут опубликованы в Трудах конференции на русском и английском языках до начала ее работы.

**СРОКИ ПРЕДСТАВЛЕНИЯ МАТЕРИАЛОВ**

– **до 20 декабря 2006** *Тексты тезисов и докладов*

– **после 20 января 2007** *Информация для авторов о приеме докладов на сайте: [www.rntores.ru](http://www.rntores.ru)*

– **15 февраля 2007** *Рассылка приглашений по почте*

**РЕГИСТРАЦИОННЫЙ ВЗНОС** за одного участника (публикацию) от предприятия и организации из России и стран СНГ – **2400 руб.**; для представителей ВУЗов России – **2000 руб.** (за каждый второй доклад взнос увеличивается на **800** рублей) В период после **10.01.07** г. взнос увеличивается на **20%**.

Для членов РНТОРЭС им. А.С. Попова – скидка **30%** (взнос – 1600 рублей). Для аспирантов и студентов – **450 руб.** (за публикацию, без предоставления сборника Трудов конференции).

Для зарубежных участников – **250 €**, для членов IEEE скидка **20%**. После 10.01.2007 года взнос увеличивается до **300 €**.

**Реквизиты для перечисления:** РНТОРЭС имени А.С. Попова, **ИНН 7702021967, КПП 770201001, БИК 044525225, Р/сч. 40703810038090105080** в Сбербанке России г. Москва, **ОСБ 7811 Мещанское, Кор/счет 3010181040000000225**,

Назначение платежа: «Целевой взнос на конференцию ДСПА-2007»

**ДОКЛАДЫ ДОЛЖНЫ НАПРАВЛЯТЬСЯ В ПРОГРАММНЫЙ КОМИТЕТ ПО АДРЕСУ:**

Российское НТОРЭС им. А.С. Попова, Рождественка, 6/9/20, стр. 1, Москва, 107031  
Тел. (495) 921-1639 (**Мусянков Михаил Иванович**), 921-0610, 921-7108; Факс: (495) 921-1639  
E-mail: [nto.popov@mtu-net.ru](mailto:nto.popov@mtu-net.ru); [tor@rgta.ryazan.ru](mailto:tor@rgta.ryazan.ru)  
[http:// www.dspa.ru](http://www.dspa.ru) и [www.rntores.ru](http://www.rntores.ru) (раздел конференции)