

A  
4549

На правах рукописи



Умняшкин Сергей Владимирович

УДК 004.932 : 004.421 : 519.722

**Математические методы и алгоритмы цифровой  
компрессии изображений с использованием  
ортогональных преобразований**

Специальность 05.13.11 - "Математическое и программное обеспечение  
вычислительных машин, комплексов и компьютерных сетей"

Автореферат диссертации на соискание ученой степени  
доктора физико-математических наук

Научная библиотека Уральского Государственного Университета
--

Москва – 2001

Работа выполнена в Московском государственном институте электронной техники (техническом университете)

Научный консультант: доктор физико-математических наук  
профессор Поспелов А.С.

Официальные оппоненты:

Доктор физико-математических наук профессор Ососков Г.А.

Доктор физико-математических наук профессор Селищев С.В.

Доктор технических наук профессор Коекин А.И.

Ведущая организация: ФГУП НИИ Радио (г.Москва)

Защита состоится “ \_\_\_\_\_ ” \_\_\_\_\_ 2002 года в “ \_\_\_\_\_ ” на заседании диссертационного совета Д.212.134.02 в Московском государственном институте электронной техники по адресу: 103498, Москва, г. Зеленоград, МИЭТ (ТУ).

С диссертацией можно ознакомиться в библиотеке МИЭТ (ТУ).

Автореферат разослан “ 10 ” января 2002 года

Ученый секретарь  
диссертационного  
совета профессор



/Воробьев Н.В./

## Общая характеристика работы

**Актуальность темы.** Хранение и передача изображений при непосредственном цифровом представлении в виде матрицы пикселей (точек изображения) вызывает необходимость обработки колоссальных объемов данных. Однако непосредственное представление изображения является неэффективным: вследствие значительной коррелированности элементов матрицы независимое кодирование пикселей порождает избыточные коды. Поэтому особую актуальность среди прочих задач цифровой обработки изображений приобретает задача *сжатия изображений*, которая заключается в поиске путей реализации эффективного кодирования визуальных данных.

**Цель работы.** Сложность алгоритмов, используемых для компрессии изображений, неуклонно растет – сказанное касается не только объема вычислений, но и идейных основ построения алгоритмов, большинство которых основано на использовании дискретных ортогональных преобразований для предварительной обработки данных. Вместе с тем, задача сжатия изображений ставится практикой, что требует при ее решении постоянного внимания к возможностям реальной аппаратуры. Целью работы являлись исследование теоретических вопросов эффективного кодирования изображений с использованием ортогональных преобразований, а также разработка соответствующих алгоритмов сжатия, пригодных для практического применения на базе универсальных вычислительных средств общего назначения.

**Направление исследований.** Проведенные в диссертационной работе исследования включали в себя рассмотрение следующих вопросов:

1. Исследование и разработка методов теоретического анализа и синтеза дискретных преобразований для схем компрессии коррелированных данных;
2. Разработка новых быстрых алгоритмов вычисления дискретного преобразования Крестенсона-Леви (ДПКЛ) и алгоритма сжатия полутоновых изображений, основанного на статистическом кодировании спектров ДПКЛ;

3. Исследование специфики и формализация общей схемы компрессии, использующей поблочную обработку фрагментов изображения при помощи ортогонального преобразования с последующим квантованием и статистическим кодированием коэффициентов-трансформант;
4. Разработка алгоритмов вейвлет-компрессии изображений и изучение возможностей фрактального кодирования в области вейвлет-спектра;
5. Разработка алгоритма сжатия видеопоследовательностей (динамических изображений), пригодного для использования в виде программной реализации на базе универсальных вычислительных средств общего назначения (мультимедийных персональных компьютеров).

**Методы исследования.** В качестве основного теоретического инструмента исследований использовались методы математического и функционального анализа, линейной алгебры, теории вероятностей и математической статистики, теории информации. Значительную часть исследований представляли собой также компьютерные эксперименты по обработке реальных неподвижных и динамических изображений, направленные на получение необходимых статистических данных и определение характеристик итоговых алгоритмов сжатия. Проведенные эксперименты подтвердили верность теоретических решений и эффективность предложенных алгоритмов компрессии.

**Научная новизна.** В результате выполнения диссертационной работы получены новые методы анализа эффективности ортогональных преобразований, предназначенных для сжатия коррелированных данных; специально для сжатия данных введено в рассмотрение (впервые построено) дискретное псевдосинусное преобразование (ДПКП). Разработаны новые быстрые алгоритмы вычисления ДПКЛ, на базе которого впервые получена схема компрессии статических изображений, имеющая аналогичные методу JPEG характеристики. Для обработки неподвижных и динамических изображений предложены как новые алгоритмы, так и общие теоретические подходы, формализующие

процедуры анализа и синтеза схем компрессии цифровых изображений на основе дискретных ортогональных преобразований.

**На защиту диссертации выносятся следующие основные результаты:**

- Метод оценки декоррелирующей эффективности ортогональных преобразований и основанные на нем алгоритмы кластеризации коррелированных данных;
- ДПКП и быстрый алгоритм его вычисления;
- Новый быстрый алгоритм ДПКЛ и его модификация – алгоритм с неполным вычислением; алгоритм совмещенных вычислений ДПКЛ для обработки вещественных массивов в базисе  $(1, \exp(-2\pi i/3))$ ;
- Метод компрессии изображений, основанный на специальном способе арифметического кодирования спектров ДПКЛ блоков изображения;
- Детерминированные и вероятностные оценки коэффициентов дискретного косинусного преобразования (ДКП);
- Алгоритм контекстного кодирования спектров ДКП изображений;
- Общая схема компрессии изображений на основе адаптивного векторного квантования в области ортогональных преобразований;
- Алгоритмы вейвлет-компрессии статических изображений;
- Алгоритм поиска перемещенных блоков изображения;
- Экспериментальная методика построения разбиения спектров на области независимого кодирования;
- Алгоритм видеокомпрессии.

**Практическая ценность.** В целом содержание работы носит прикладную направленность, поэтому полученные теоретические результаты также служат достижению целей, связанных с разработкой конкретных алгоритмов и схем компрессии цифровых изображений. Применение полученных алгоритмов сжатия изображений возможно для широкого класса систем хранения и

передачи визуальной информации, прежде всего, в мультимедийных и сетевых компьютерных приложениях. Разработанные алгоритмы, как подтверждают эксперименты, обладают высокими характеристиками по скорости, качеству обработки и сжатию данных, которые соответствуют современному мировому уровню.

**Реализация результатов работы.** Теоретические результаты работы и алгоритмы сжатия видеоизображений внедрены в ГУ НПК «Технологический Центр» МИЭТ (<http://www.tcep.ru>) и использованы в научно-производственной деятельности НПП «Технология» (Москва) [29,30].

**Апробация работы.** Основные результаты работы докладывались и обсуждались на следующих научных конференциях и совещаниях:

1. VII саратовская зимняя школа по теории функций и приближений (СГУ, январь 1994 г.) [1].
2. Междунар. конференция по теории функций и приближений, посвященная 90-летию акад. С.М.Никольского (Москва, МИ РАН, май 1995 г.) [2].
3. Всероссийские научно-технические конференции “Электроника и информатика” (Москва, МИЭТ, 1995-2000 г.) [3,6,22].
4. Междунар. конференция по теории приближения функций, посвященная памяти проф. П.П.Коровкина (Калуга, КГПУ, 26-29 июня 1996 г.) [4].
5. Международные конференции «Методы оптимизации вычислений» (Киев, 1997, 2001 г.) [7,26].
6. Международная конференция «Проблемы математического образования», посв. 75-летию чл.-корр. РАН проф. Л.Д.Кудрявцева (1998 г.) [9].
7. Международная конференция «Теория приближений и гармонический анализ» (Тула, 26-29 мая 1998 г.) [10].
8. Международная конференция «Информационные технологии в инновационных проектах» (Ижевск, 20-22 апреля 1999 г.) [13].

9. VII Международная конференция «Математика. Экономика. Экология. Образование» – Международный симпозиум «Ряды Фурье и их приложения» (Новороссийск, 1999 г.) [13,14].
10. VII Международная конференция «Математика. Компьютер. Образование» (Дубна, ОИЯИ, 24-29 янв. 2000 г.) [17].
11. Международная конференция, посвященная 80-летию со дня рождения С.Б.Стечкина (Екатеринбург, 28 февраля – 3 марта 2000 г.) [18].

В целом содержание работы обсуждалось на научном семинаре лаборатории информационных технологий в Объединенном институте ядерных исследований (г.Дубна, декабрь 2001 года).

**Публикации.** Основное содержание диссертации отражено в 30 работах.

**Структура и объем диссертации.** Диссертационная работа содержит 382 страницы (из которых 26 страниц – приложения) и состоит из введения, шести глав, заключения и 6 приложений. Библиографический список включает в себя 178 наименований. В приложениях приведены численные результаты ряда экспериментов по обработке изображений, а также справочная информация и копии документов об использовании результатов диссертационной работы.

## **Содержание работы**

**Во введении** (28 страниц) обосновываются актуальность, научная новизна, практическая ценность исследований. Кратко излагается содержание глав.

**В первой главе** (35 страниц) приводятся необходимые для дальнейшего изложения предварительные сведения, дается краткий обзор и классификация основных подходов к реализации эффективного кодирования изображений.

Использование алгоритмов сжатия с потерями данных для полутоновых изображений носит повсеместный характер: допустив наличие ошибки в восстановленном изображении, можно добиться намного более высокого уровня компрессии данных. Чаще всего качество обработки изображений оценивают

по среднеквадратичной ошибке (СКО):  $CKO = \sqrt{\frac{1}{NM} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (x_{i,j} - \hat{x}_{i,j})^2}$ , где

$X=(x_{i,j})$  – матрица исходного изображения,  $\hat{X}=(\hat{x}_{i,j})$  – матрица изображения, полученного после обработки (сжатия и восстановления данных). Для логарифмической величины СКО используется общепринятая мера PSNR (peak signal to noise ratio – отношение пикового значения сигнала к шуму),

$$PSNR = 20 \lg \frac{255}{CKO} \text{ [дБ]}.$$

Методы сжатия изображений удобно рассматривать в виде общей схемы, состоящей из трех основных этапов: снижение межэлементной корреляции данных, квантование элементов данных, статистическое кодирование. Квантование является основным инструментом, используемым при сжатии с потерями данных. По сути дела, квантование есть выделение из входных данных некоторой основной части информации, когда ее менее значимая часть опускается. Применяется как *скалярное*, так и *векторное* квантование.

Перевод изображения в обобщенную спектральную область при помощи линейного преобразования  $F$  может значительно снизить межэлементную корреляцию в матрице-трансформанте  $Y=F(X)$  по сравнению с корреляцией элементов в матрице дискретного изображения  $X$ . Тогда независимое покомпонентное кодирование матрицы  $Y$ , а не матрицы  $X$ , становится более эффективным. Можно также дать энергетическую трактовку цели использования преобразований, которая в таком понимании заключается в концентрации максимальной части энергии исходного дискретного сигнала (матрицы  $X$ ) в минимальном количестве спектральных коэффициентов (элементов матрицы  $Y$ ). Между распределением энергии в обобщенном спектре и декоррелирующими свойствами преобразований имеется определенная связь. Изучение эффективности декоррелирующих свойств, таким образом, является важной задачей при выборе преобразования для применения в схеме компрессии.

Реальные фотографические изображения представляют собой двумерные сигналы, имеющие неоднородности (особенности) в областях контуров объектов, поэтому базис функций, используемых для разложения, должен обладать на исходном изображении хорошей локализацией. Однако в фоновых областях изображение может рассматриваться как реализация стационарного сигнала, что делает предпочтительным использование для разложения частотно-локализованного базиса (хорошо известно, что коэффициенты Фурье разложения по тригонометрической системе стационарного сигнала некоррелированы). Добиться одновременно высокого разрешения в частотной и во временной областях невозможно в силу принципа неопределенности Гейзенберга. Выходом является использование функциональных базисов *вейвлетов* (всплесков), которые обладают переменным время-частотным разрешением. Подходы, связанные с использованием всплесков, на сегодняшний день являются доминирующими в обработке неподвижных изображений, постепенно вытесняя традиционный инструмент декорреляции – дискретное косинусное преобразование.

В первой главе отмечается, что для оптимизации алгоритмов сжатия данных с потерями часто используется подход, основанный на минимизации RD-функции Лагранжа. Пусть  $X$  – некоторый входной набор данных, которому в результате выполнения процедуры сжатия-восстановления ставится в соответствие выходной набор данных той же природы,  $Y=F(X,u)$ , где  $u=(u_1, \dots, u_n)$  – набор управляющих параметров алгоритма сжатия  $F$ . Считаем  $X, Y$  элементами некоторого пространства  $\Omega$  с метрикой  $D(X,Y)$ , множество всех возможных значений управляющего вектора  $u$  обозначим  $U$ . Задача оптимизации кодирования состоит в том, чтобы для заданного набора входных данных  $X$  и максимально допустимых битовых затрат  $R_b$  найти такие параметры  $u^*=(u_1^*, \dots, u_n^*)$  алгоритма  $F$ , чтобы ошибка кодирования данных  $D(X,Y)=D(X,F(X,u))$  принимала бы минимальное значение. То есть

$$\begin{aligned} D(X, F(X, \mathbf{u}^*)) &= D(X, \mathbf{u}^*) = \min_{\mathbf{u} \in U} D(X, \mathbf{u}), \\ R(X, \mathbf{u}^*) &\leq R_b, \end{aligned} \quad (1)$$

где  $R(X, \mathbf{u})$  – число бит, необходимое для кодирования набора данных  $X$  с параметрами  $\mathbf{u}$ .

Поиск решения задачи (1) в большинстве случаев сводится к громоздким численным процедурам итерационного характера. Если не задаваться ограничением  $R(X, \mathbf{u}) \leq R_b$ , то для определения оптимальных параметров кодирования  $\mathbf{u}^*$ , соответствующих решению задачи (1) для некоторого (заранее неизвестного) значения  $R_b$ , используется упрощенный вариант минимизации RD-функции Лагранжа:

$$J(\mathbf{u}^*) = \min_{\mathbf{u}} [J(\mathbf{u}) = D(\mathbf{u}) + \lambda R(\mathbf{u})], \quad (2)$$

где  $\lambda$  – неотрицательный параметр, задаваемый внешне. Параметр  $\lambda$  в функции  $J(\mathbf{u})$  устанавливает баланс между качеством и уровнем сжатия данных. Значение  $\lambda=0$  соответствует наименьшей ошибке кодирования, увеличивая значение  $\lambda$ , получаем при оптимизации параметров алгоритма  $F$  по (2) меньшую длину кода, но большую ошибку. Тем самым можно настраивать алгоритм кодирования  $F$  на необходимые характеристики. Для поиска решения задачи (1) минимизация (2) повторяется итерационно, с различными значениями  $\lambda$  – данная процедура носит название *RD-оптимизации*<sup>1</sup>.

В первой главе также кратко отмечаются особенности, связанные с обработкой (сжатием-восстановлением) динамических изображений. Основным используемым преобразованием для видеокомпрессии пока остается ДКП, как более простое по объему вычислений в сравнении с вейвлет-преобразованиями. Так же как и для статической компрессии, алгоритмы кодирования видео чаще всего являются более сложными по сравнению с алгоритмами декодирования. Реализация программной компрессии видео в реальном масштабе времени, та-

<sup>1</sup> Berger T. Rate Distortion Theory. – Englewood Cliffs, NJ: Prentice Hall, 1971.

ким образом, накладывает существенные ограничения на допустимую сложность вычислений.

**Вторая глава** (52 страницы) посвящена изучению эффективности и синтезу ортогональных преобразований, предназначенных для использования в сжатии данных. Предложенный новый метод анализа эффективности основан на следующих рассуждениях. Пусть известна ковариационная матрица  $K_X$  исходного вектора данных  $X = (x_0, x_1, \dots, x_{N-1})^T$ , вектор-спектр  $Y$  получен в результате некоторого ортогонального преобразования с матрицей  $W$ :  $Y = WX$ . Средняя безусловная энтропия коэффициента вектор-спектра может быть записана в виде [5]:

$$\begin{aligned} H_{cp} &= -\frac{1}{N} \sum_{k=0}^{N-1+\infty} \int f_k(m_k, \sigma_k, x) \log f_k(m_k, \sigma_k, x) dx = \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \log \sigma_k - \frac{1}{N} \sum_{k=0}^{N-1+\infty} \int f_k^0(x) \log f_k^0(x) dx, \end{aligned} \quad (3)$$

где  $f_k(m_k, \sigma_k, x)$  – функция плотности распределения вероятностей для спектральной характеристики  $y_k$  ( $k$ -ой компоненты вектора  $Y$ ),  $m_k$  – математическое ожидание,  $\sigma_k$  – среднее квадратичное отклонение,  $f_k^0(x) = f_k(0, 1, x)$ . Чем меньше средняя энтропия (3), тем эффективнее будет последующее независимое кодирование компонент спектра. Наложив ограничение, при котором для класса

рассматриваемых преобразований  $\frac{1}{N} \sum_{k=0}^{N-1+\infty} \int f_k^0(x) \log f_k^0(x) dx = Const$  и учитывая,

что для оптимального преобразования Карунена-Лозва (когда матрица  $W = W_{opt}$  составлена из собственных векторов  $K_X$  и матрица  $K_Y = WK_X W^T$  имеет диагональный вид)

$$\frac{1}{2N} \sum_{k=0}^{N-1} \log \sigma_k^2 = \frac{1}{2N} \log \prod_{k=0}^{N-1} \sigma_k^2 = \frac{1}{2N} \log \det K_X,$$

декоррелирующей эффективности будем рассматривать величину *средней избыточной энтропии*  $\Delta H(W, K_X) = H_{cp}(W, K_X) - H_{cp}(W_{opt}, K_X)$ , которая выра-

жается через элементы матриц  $\mathbf{K}_X = (\text{cov}(x_i, x_j))_{i,j=0}^{N-1}$  и  $\mathbf{W} = (w_{i,j})_{i,j=0}^{N-1}$  следующим образом [12]:

$$\Delta H(\mathbf{W}, \mathbf{K}_X) = \frac{1}{2N} \log \left( \frac{1}{\det \mathbf{K}_X} \prod_{k=0}^{N-1} \left( \sum_{m=0}^{N-1} w_{k,m} \sum_{j=0}^{N-1} w_{k,j} \text{cov}(x_m, x_j) \right) \right). \quad (4)$$

Чем больше значение  $\Delta H(\mathbf{W}, \mathbf{K}_X)$ , тем меньше эффективность декоррелирующего преобразования с матрицей  $\mathbf{W}$ . Численные расчеты величины (4) для различных преобразований и видов ковариационных матриц показали результаты, полностью согласующиеся с известными данными, которые получены другими методами, например, по Пэрл (*Pearl J. On coding and filtering stationary signals by discrete Fourier transforms // IEEE Trans. Inf. Theory. - 1973. - Vol. IT-19. - P. 229-232.*).

Большой интерес для анализа представляет модель дискретного сигнала (вектора  $\mathbf{X}$ ), имеющего статистику дискретного марковского процесса первого порядка, когда ковариационная матрица имеет следующий вид:

$$\mathbf{K}_X(\rho) = \begin{pmatrix} 1 & \rho & \dots & \rho^{N-1} \\ \rho & 1 & \dots & \rho^{N-2} \\ \dots & \dots & \dots & \dots \\ \rho^{N-1} & \rho^{N-2} & \dots & 1 \end{pmatrix}. \quad (5)$$

Данная модель часто используется для описания межстрочной и межстолбцовой корреляции в дискретных изображениях. При  $\rho=1$ , когда все компоненты в исходном векторе  $\mathbf{X}$  одинаковы (для любых двух отсчетов вектора коэффициент корреляции равен единице), расчет введенного критерия (4) для матрицы (5) невозможен, т.к. в этом случае имеем  $\det \mathbf{K}_X = 0$ . Вместе с тем, на фоновых областях изображения  $\rho \rightarrow 1$ . В главе 2 доказана следующая теорема.

**Теорема 2.1.** Для любой ортогональной матрицы  $\mathbf{W}$  ( $N \times N$ ) такой, что  $\forall j = 0, 1, \dots, N-1: w_{0,j} = \frac{1}{\sqrt{N}}$  (базисная функция с нулевым индексом есть нормированная постоянная составляющая) и ковариационной матрицы (5)

рованная постоянная составляющая) и ковариационной матрицы (5)

$$\lim_{\rho \rightarrow 1} \Delta H(\mathbf{W}, \mathbf{K}_X) = \frac{1}{2N} \log \left[ N \prod_{k=1}^{N-1} \left( \frac{1}{2} \sum_{m=0}^{N-1} w_{k,m} \sum_{j=0}^{N-1} w_{k,j} |m-j| \right) \right].$$

Различные исследования, в том числе проведенные в главе 2, показывают, что среди дискретных преобразований, имеющих быстрые алгоритмы вычислений (для размерности  $N$  реализуемых за  $\sim \text{Mlog}N$  арифметических операций) наиболее близкие к оптимальному преобразованию Карунена-Лозва характеристики декорреляции для марковского процесса (5) дает использование ДКП. Однако несмотря на наличие хорошо проработанных быстрых алгоритмов вычисления, ДКП принципиально требует для своей реализации выполнения операций умножения и по объему вычислений заметно уступает, например, преобразованиям Хаара, Уолша, Крестенсона-Леви. Отдельным вопросом, рассмотрению которого уделено значительное внимание в главе 2, является построение (синтез) нового преобразования, имеющего как высокие характеристики декорреляции для модели (5), так и значительно более быстрые, чем для ДКП, алгоритмы вычисления. Полученное в результате дискретное псевдокосинусное преобразование определено для векторов размерности  $N$ , которая допускает разложение  $N=N_1 \dots N_n$ , причем  $\forall k N_k \in \{2,3,4\}$ . Представление  $N=N_1 \dots N_n$  необходимо записать минимальным количеством множителей  $N_k$ , располагая их без убывания, т.е.  $\forall k, \forall m > k: N_k \leq N_m$ . Например, для  $N=8$  имеем  $N_1=2, N_2=4$  (но не  $N_1=4, N_2=2$  и не  $N_1=N_2=N_2=2$ ). Тогда матрица ДПКП  $\mathbf{W}_N$  (в данном случае нижний индекс указывает на размерность преобразования) строится как прямое произведение<sup>2</sup>  $\mathbf{W}_N = \mathbf{W}_{N_1} \otimes \dots \otimes \mathbf{W}_{N_n}$  элементарных матриц ДПКП  $\mathbf{W}_{N_k} \in \{\mathbf{W}_2, \mathbf{W}_3, \mathbf{W}_4\}$ ,  $k=1, \dots, n$ , где элементарные матрицы орто-

<sup>2</sup> Под тензорным (прямым) произведением матриц  $\mathbf{D} = \{d_{l,m}\}$  ( $l=0, \dots, \eta-1$ ;  $m=0, \dots, \mu-1$ ) и

$$\mathbf{G} = \{g_{k,j}\} \text{ понимаем матрицу } \mathbf{S} = \mathbf{D} \otimes \mathbf{G} = \{s_{\alpha,\beta}\} = \begin{pmatrix} d_{0,0} \mathbf{G} & \dots & d_{0,\mu-1} \mathbf{G} \\ \dots & \dots & \dots \\ d_{\eta-1,0} \mathbf{G} & \dots & d_{\eta-1,\mu-1} \mathbf{G} \end{pmatrix}.$$

гональны и получены в результате определенных модификаций из матриц ДКП соответствующей размерности. Элементарные матрицы представимы в виде произведения некоторой диагональной матрицы  $D$  на матрицу  $C$ , а структура  $C$  позволяет реализовать умножение на произвольный вектор  $U$  только при помощи операций сложения и вычитания чисел (умножение на 2 эквивалентно сложению,  $2x=x+x$ ). Именно:

$$\begin{aligned}
 W_2 &= \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} = \underbrace{\begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} \end{bmatrix}}_{D_2} \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{C_2}, \\
 W_3 &= \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{6} & -\sqrt{2}/3 & 1/\sqrt{6} \end{bmatrix} = \underbrace{\begin{bmatrix} 1/\sqrt{3} & 0 & 0 \\ 0 & 1/\sqrt{2} & 0 \\ 0 & 0 & 1/\sqrt{6} \end{bmatrix}}_{D_3} \underbrace{\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & -2 & 0 & 1 \end{bmatrix}}_{C_3} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}}_{C_3}, \\
 W_4 &= \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 2/\sqrt{10} & 1/\sqrt{10} & -1/\sqrt{10} & -2/\sqrt{10} \\ 1/2 & -1/2 & -1/2 & 1/2 \\ 1/\sqrt{10} & -2/\sqrt{10} & 2/\sqrt{10} & -1/\sqrt{10} \end{bmatrix} = \\
 &= \underbrace{\begin{bmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 1/\sqrt{10} & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/\sqrt{10} \end{bmatrix}}_{D_4} \underbrace{\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & -2 & 1 \end{bmatrix}}_{C_4} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}}_{C_4}.
 \end{aligned}$$

Из свойств тензорного произведения следует представление  $W_N = D_N C_N$ , где диагональная матрица  $D_N = D_{N_1} \otimes \dots \otimes D_{N_n}$ , а матрица  $C_N = C_{N_1} \otimes \dots \otimes C_{N_n}$ . Матрицы  $C_2, C_3, C_4, D_2, D_3, D_4$  приведены выше. Таким образом, реализация ДПКП  $Y = W_N X = D_N C_N X$  заключается в реализации умножения матрицы  $C_N$  на вектор,  $\tilde{Y} = C_N X$ , и последующей нормировке полученного вектора  $\tilde{Y}$ ,  $Y = D_N \tilde{Y}$ . Для вычисления ДПКП удобно использовать быстрые алгоритмы, основанные на факторизованном представлении для мат-

рицы  $C_N$  в виде произведения слабозаполненных матриц<sup>3</sup>:  $C_N = T_N^{(n)} \cdot T_N^{(n-1)} \cdot \dots \cdot T_N^{(1)}$ , где  $T_N^{(i)} = I_{N_1} \otimes \dots \otimes I_{N_{i-1}} \otimes C_{N_i} \otimes I_{N_{i+1}} \otimes \dots \otimes I_{N_n}$ ,  $I_{N_j}$  – единичная матрица размерности  $N_j \times N_j$ . Поскольку матрицы  $T_N^{(j)}$  состоят из определенным образом разреженных матриц-блоков  $C_{N_j}$ , умножение матрицы  $T_N^{(j)}$  на вектор также сводится только к операциям сложения и вычитания чисел. Быстрые алгоритмы обратного ДПКП строятся аналогично, т.к. в силу ортогональности ДПКП:  $W_N^{-1} = W_N^T = C_N^T D_N = (T_N^{(1)})^T (T_N^{(2)})^T \cdot \dots \cdot (T_N^{(n)})^T \cdot D_N$ .

Отметим, что необходимая при вычислении ДПКП и обратного ДПКП нормировка (умножение на матрицу  $D_N$ ) для схемы сжатия со скалярным квантованием коэффициентов преобразования не влечет усложнения вычислений. Нормировка может быть объединена при сжатии данных с этапом скалярного квантования  $\tilde{y}_j = \text{round}(y_j / q)$  путем выбора для каждого элемента  $y_j$  преобразованного вектора  $Y$  индивидуального шага квантования  $q_j = q/d_{ij}$  (где  $d_{ij}$  – элемент диагональной нормировочной матрицы  $D_N$ ). При деквантовании  $\hat{y}_j = m_j \tilde{y}_j$  множитель для элемента  $\hat{y}_j$  следует выбирать в виде  $m_j = q \cdot d_{ij}$ .

Как показали расчеты величины средней избыточной энтропии (4) и остаточной корреляции по Пэрл, для данных, имеющих статистику марковского процесса первого порядка (5), ДПКП обладает большей эффективностью в декорреляции по сравнению с другими быстрыми преобразованиями, реализация которых также сводится только к операциям сложения и вычитания чисел.

**Третья глава** (48 стр.) посвящена изучению вопросов применения для сжатия изображений дискретного преобразования Крестенсона-Леви (ДПКЛ) и является развитием исследований кандидатской работы автора.

<sup>3</sup> Для обоснования справедливости данного представления см. с.84-85 из монографии «Абстрактные алгебраические системы и цифровая обработка сигналов» / Вариченко Л.В., Лабунец В.Г., Раков М.А. - Киев: Наукова думка, 1986. – 248 с.

Система функций Крестенсона-Левы  $\{\chi_m(x)\}_{m=0}^{\infty}$  на полуинтервале  $x \in [0; 1)$  в нумерации Пэли определяется следующим образом:

$$\chi_m(x) = \prod_{k=1}^{l(m)} \exp\left(i \frac{2\pi}{p} x_k \cdot m_{-k}\right) = \exp\left(i \frac{2\pi}{p} \sum_{k=1}^{l(m)} x_k \cdot m_{-k}\right), \quad (6)$$

где  $p \geq 2$  (целое число),  $x_k \equiv [x \cdot p^k] \pmod{p}$ ,  $m_{-k} \equiv [m / p^{k-1}] \pmod{p}$ , а число  $l(m)$  находится из условия:  $p^{l(m)-1} \leq m \leq p^{l(m)} - 1$ . Система функций Крестенсона-Левы является обобщением на комплексную плоскость хорошо известной системы функций Уолша, которая получается из (6) при  $p=2$ .

Предварительно в главе 3 изучаются вопросы построения быстрых алгоритмов ДПКЛ. В матричном виде ДПКЛ можем записать:

$$Y = \frac{1}{\sqrt{p^n}} W_{(n)} \cdot X, \quad (7)$$

где  $X = (x_0, x_1, \dots, x_{p^n-1})^T$  - исходный вектор,  $Y = (y_0, y_1, \dots, y_{p^n-1})^T$  - вектор-спектр ДПКЛ. Нижний индекс  $n$  матрицы  $W_{(n)}$  указывает на размерность преобразования,  $N=p^n$ . Элементами симметрической матрицы  $W_{(n)}$  являются отсчеты функций системы (6):  $w_{k,l} = \bar{\chi}_k(l/p^n)$ . Быстрые алгоритмы ДПКЛ основаны на факторизованном представлении матрицы  $W_{(n)}$  в виде произведения слаботолпленных матриц. Один из возможных способов факторизации матрицы ДПКЛ предложен автором и может быть сформулирован в виде следующей теоремы [1].

**Теорема 3.1.** Матрица ДПКЛ представима в виде  $W_{(n)} = W_{(n)}^{(1)} W_{(n)}^{(2)} \dots W_{(n)}^{(n)}$ ,

$$\text{где } W_{(n)}^{(j)} = \begin{pmatrix} W_{(n-1)}^{(j)} & & 0 \\ & \ddots & \\ 0 & & W_{(n-1)}^{(j)} \end{pmatrix}, \quad j = 1, \dots, n-1$$

(на диагонали матрицы  $W_{(n)}^{(j)}$  размерности  $p^{n-1} \times p^{n-1}$  находятся  $p$  матриц  $W_{(n-1)}^{(j)}$  размера  $p^{n-1} \times p^{n-1}$ , остальные элементы матрицы  $W_{(n)}^{(j)}$  - нулевые). Матрица

$W_{(n)}^{(n)} = \{w_{j,k}\}$  ( $j, k=0, 1, \dots, p^n-1$ ) имеет следующий вид:  $w_{j,k} = \delta_{j \bmod p}^{[k/p]} \cdot q^{j \cdot n - k - 1}$ , где  $k_{-1} \equiv k \pmod{p}$ ,  $j_{-n} \equiv [j/p^{n-1}] \pmod{p}$ ,  $q = \exp(-2\pi i/p)$ ,  $\delta_j^m = \begin{cases} 1, & \text{при } m = j \\ 0, & \text{при } m \neq j \end{cases}$ .

Дискретные (цифровые) полутоновые изображения описываются в виде вещественной матрицы  $X$ . При обработке вещественных векторов (матриц) спектр ДПКЛ распадается на пары комплексно сопряженных элементов, поэтому часть элементов спектра можно не вычислять. Учет этих особенностей позволил предложить для обработки вещественных массивов алгоритмы ДПКЛ и обратного ДПКЛ с *неполным вычислением*. При обработке вещественных наборов данных возможно также применение «совмещенных» преобразований, аналогично тому, как это делается при вычислениях дискретного преобразования Фурье. При этом из двух вещественных векторов  $X^{(1)}$  и  $X^{(2)}$  формируют комплексный вектор  $X = X^{(1)} + iX^{(2)}$  и выполняют преобразование.

Если использовать для комплексных чисел  $z = x + iy$  представление в базисе  $(1, q)$ :  $z = \alpha + \beta q$  ( $q = e^{-2\pi i/3}$ ), то вычисление ДПКЛ (при  $p=3$ ) может быть выполнено без использования операций умножения (нормирующий множитель  $1/\sqrt{p^n}$ , см. (7) – не учитывается) при помощи операций сложения и вычитания (данный факт установлен А.В.Ефимовым). Использование в алгоритмах с *неполным вычислением* базиса  $(1, q)$  влечет за собой ряд особенностей, которые рассматриваются в главе 3. При размерности матрицы изображения  $3^n \times 3^r$  вычисление ДПКЛ или обратного ДПКЛ по алгоритмам с *неполным вычислением* требует выполнения  $7(n+r)3^{n+r-1}$  операций сложения (вычитания), операции умножения и деления не используются. Использование базиса  $(1, q)$  в вычислительном плане является более эффективным и для алгоритмов совмещенных вычислений. Специфика использования ДПКЛ и базиса  $(1, q)$  рассмотрена в главе 3, там же получены необходимые для реализации совмещенных ДПКЛ и обратного ДПКЛ расчетные формулы.

Для обработки неподвижных изображений в кандидатской работе автора был предложен алгоритм компрессии, в котором исходное изображение разбивается для обработки на элементарные фрагменты размером  $9 \times 9$  точек, и каждая матрица-фрагмент обрабатывается при помощи ДПКЛ. В главе 3 приводится краткое описание алгоритма компрессии [20], который подтвердил обоснованность использования дискретного преобразования Крестенсона-Леви для сжатия изображений. Характер вносимых в процессе сжатия-восстановления ошибок различен для основанного на ДКП метода JPEG и для предложенной схемы: при использовании JPEG имеет место “размывание”, в то время как новая схема сжатия приводит к “зубчатости” изображения. Тем не менее, субъективное качество восприятия при использовании обоих рассмотренных способов сжатия примерно одинаково. Оценки величины искажений по величине PSNR, выполненные для ряда тестовых изображений размерности  $720 \times 504 = 362880$  точек, также дали близкие результаты: для одних изображений может наблюдаться незначительное преимущество основанного на ДКП стандарта JPEG, для других изображений лучшее качество восстановления возможно при использовании нового метода. Различия невелики, особенно при невысоких уровнях сжатия. Оценки вычислительных затрат показывают, что алгоритм на основе ДПКЛ не уступает JPEG и в части, касающейся вычислительной сложности реализации. Алгоритм компрессии изображений [20], использующий ДПКЛ и имеющий сравнимые с JPEG характеристики, представляет собой результат, впервые полученный автором.

В четвертой главе (54 страницы) рассматриваются вопросы разработки схем сжатия, использующих пофрагментную обработку изображений. Обычно для подобных схем наиболее эффективным является использование ДКП, поэтому изучению его свойств, необходимых для построения алгоритмов компрессии, посвящена значительная часть четвертой главы. Т.к. двумерное ДКП является разделимым преобразованием (сводится к одномерным преобразова-

ниям вдоль строк и вдоль столбцов обрабатываемой матрицы), то многие свойства двумерного ДКП вытекают из свойств одномерного.

Для коэффициентов одномерного ДКП, определяемого формулой

$$y_k = \sqrt{\frac{2}{N}} \cdot c(k) \sum_{j=0}^{N-1} x_j \cos\left(\frac{\pi k}{N} \left(j + \frac{1}{2}\right)\right),$$

$$k = 0, 1, \dots, N-1,$$

получено следующее соотношение ( $k \neq 0$ ):

$$y_k = \frac{-1}{\sqrt{2N} \sin \frac{\pi k}{2N}} \sum_{j=1}^{N-1} \Delta_j \sin \frac{\pi k}{N} j, \quad (8)$$

где  $\Delta_j = x_j - x_{j-1}$  – первые разности отсчетов исходного вектора данных  $X = (x_0, x_1, \dots, x_{N-1})^T$ . Формула (8) показывает, что разности  $\Delta_j$  имеют различный характер влияния на спектр ДКП. Так, разность  $\Delta_{N/2} = x_{N/2} - x_{N/2-1}$  (при четном  $N$ ) входит во все коэффициенты  $y_{2k+1}$  с максимальными по модулю (единичными) весами. Разность между отсчётами  $x_0$  и  $x_{N-1}$  вообще не входит в (8), что принципиально отличает ДКП от его «прародителя» – дискретного преобразования Фурье, амплитудный спектр которого инвариантен к циклическим сдвигам ( $x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_{N-1} \rightarrow x_0$ ) компонент вектора данных  $X$ .

В предположении некоррелированности разностей<sup>4</sup>  $\Delta_j$  и равенства нулю их математических ожиданий  $E(\Delta_j) = 0$  исследовались также вероятностные оценки, в частности, величина суммарной дисперсии переменных составляющих спектра ДКП:  $\tilde{E} = \sum_{k=1}^{N-1} D(y_k)$ . Данная сумма представима через дисперсии

первых разностей вектора данных следующим образом:  $\tilde{E} = \frac{1}{2N} \sum_{j=1}^{N-1} S(j) D(\Delta_j)$ ,

где  $S(j)$  – некоторый весовой коэффициент.

<sup>4</sup> Для марковской модели (5) данное допущение не верно.

Доказана теорема 4.1: для коэффициентов  $S(j)$  верно соотношение:  $S(j)=2j(N-j)$ .

Полученное соотношение вновь подтверждает, что для использовавшейся вероятностной модели вектора  $X$  разность  $\Delta_{N/2}$  (в данном случае, ее дисперсия) вносит наибольший вклад. Чем ближе номер разности к  $N/2$ , тем больше значение веса  $S(j)$  и тем больший вклад вносит разность  $\Delta_j$  в энергию переменных составляющих. Исследование спектров ДКП для некоторых характерных сигналов также подтвердило, что при использовании приемов кодирования спектров из JPEG (в частности, кодирование серий нулей – run-length encoding) эффективность кодирования будет ухудшаться в случае, когда информационная насыщенность дискретного сигнала будет приходиться на центральные отсчеты вектора.

Помимо изучения общих свойств ДКП, в четвертой главе исследуется возможность оптимизации кодирования по схеме JPEG, сохраняющей формат выходных данных. В развитие идей работы Crouse-Ramchandran<sup>5</sup> предложен алгоритм [21], осуществляющий дополнительную оптимизацию скалярного квантования коэффициентов ДКП, что влечет некоторое повышение характеристик сжатия по стандарту JPEG за счет незначительного усложнения процедуры оптимизации.

Для сжатия статических изображений с использованием ДКП  $8 \times 8$  в четвертой главе разработан новый алгоритм [24], отличающийся от JPEG этапом статистического кодирования, для которого используется несколько моделей и алгоритм арифметического кодирования. Обозначим  $Z = (z_0, \dots, z_{63})$  – вектор, полученный в результате зигзагообразного считывания по JPEG матрицы скалярно проквантованных коэффициентов ДКП в одномерную последователь-

---

<sup>5</sup> *M.Crouse and K.Ramchandran*. Joint thresholding and quantizer selection for transform image-coding: entropy-constrained analysis and applications to baseline JPEG // IEEE Trans. on Image Processing. – 1997. - Vol. 6. - №2 - P. 285-297.

ность. В алгоритме сжатия используются следующие потоки данных, каждому из которых соответствует отдельная статистическая модель (накопленная гистограмма частот появления символов).

- Поток  $D$ , соответствующий постоянным составляющим яркости  $z_0 = \tilde{Y}_{0,0}$ .
- $N_0$  потоков  $\{S_i\}_{i=1}^{N_0}$ , в которых кодируются длины нулевых серий.
- $N_1$  потоков  $\{C_i\}_{i=1}^{N_1}$ , в которых кодируются ненулевые компоненты.

Сначала для каждого вектора  $Z$  в потоке  $D$  постоянная составляющая яркости  $z_0$  кодируется при помощи дифференциальной импульсно-кодовой модуляции. Затем выполняется следующая обработка компонент  $z_1, \dots, z_{63}$ .

**Шаг 0.** (Подготовительный).

- 0.1.** Определить индекс  $j_{\max}$  последнего ненулевого элемента  $z_{j_{\max}} \neq 0$ ,  
 $\forall k > j_{\max} : z_k = 0$ . (Если  $z_1 = \dots = z_{63} = 0$ , то полагаем  $j_{\max} = 0$ .)
- 0.2.** Индекс текущей обрабатываемой компоненты  $z_j$ :  $j \leftarrow 0$ .
- 0.3.** Установить текущий поток (модель)  $T \leftarrow S_{N_0}$ .
- 0.4.** Длина нулевой серии:  $n_z \leftarrow 0$ .

**Шаг 1.**  $j \leftarrow j+1$ . Если  $j > j_{\max}$ , то перейти на шаг 3.

**Шаг 2.** Если  $z_j \neq 0$  то

**2.1.** Длина нулевой серии:  $n_z \leftarrow n_z + 1$ . Перейти на шаг 1.

**иначе**

**2.2.** В текущем потоке  $T$  закодировать  $n_z$ ; положить  $n_z \leftarrow 0$ .

**2.3.** Вычислить контекстный прогноз  $p = p_f(z_1, \dots, z_{j-1})$ .

**2.4.** Установить текущий поток  $T \leftarrow C_k$ . Номер  $k$ ,  $1 \leq k \leq N_1$ ,

находится по условию:  $p \in [t_{k-1}^C, t_k^C)$ , где  $\{t_m^C\}_{m=0}^{N_1}$  –  
некоторые пороги,  $0 = t_0^C < t_1^C < \dots < t_{N_1}^C = \infty$ .

**2.5.** В потоке  $T$  закодировать компоненту  $z_j$ .

**2.6.** Установить текущий поток (модель)  $T \leftarrow S_k$ . Номер  $k$ ,

$1 \leq k \leq N_0$ , находится по условию:  $p \in [t_{k-1}^S, t_k^S)$ , где

$\{t_m^S\}_{m=0}^{N_0}$  – пороги,  $0 = t_0^S < t_1^S < \dots < t_{N_0}^S = \infty$ .

## 2.7. Перейти на шаг 1.

**Шаг 3. Если  $j < 64$  то** закодировать в потоке  $T$  признак конца блока ЕОВ.

### **Конец**

Ключевую роль в работе алгоритма играет правило построения прогнозной величины  $p = p_j(z_1, \dots, z_{j-1})$  по контексту уже закодированных компонент, которое имеет вид

$p_j = \sum_{k=1}^{j-1} \alpha_{j,k} |z_k|$ , где весовые коэффициенты  $\alpha_{j,k} \geq 0$ . Другими

важными параметрами являются значения порогов  $\{c_m\}_{m=0}^{N_1}$  и  $\{s_m\}_{m=0}^{N_0}$ , используемые на шагах 2.4 и 2.6. Универсальные, пригодные для обработки различных изображений пороги были определены по результатам экспериментов, так же, как и весовые коэффициенты  $\alpha_{j,k}$ . Описанный алгоритм контекстного кодирования повышает сжатие данных на 10% по сравнению со стандартной схемой JPEG (эталоном служила программа JPEG Optimizer™ v.4.0, см. <http://xat.com>).

Значительное внимание в четвертой главе уделено изучению теоретических и практических аспектов применения векторного квантования (ВК) для компрессии изображений в области преобразований. Для этого спектры фрагментов изображения необходимо разбить на определенные зоны, каждая из которых соответствует отдельному потоку данных, подвергаемому ВК. Сформулируем задачу разбиения коррелированного набора данных, который представим в виде вектора  $\mathbf{Y}$ , на кластеры (классы) [13,14]. Исходными параметрами являются ковариационная матрица  $\mathbf{K}_Y$  вектора  $\mathbf{Y}$  и ограничение на максимальную неопределённость (энтропию)  $H_{\max}$  для каждого кластера в разбиении. Требуется найти такое разбиение множества случайных компонент  $\mathbf{Y} = \{Y_0, \dots, Y_{N-1}\}$  на подмножества  $\mathbf{Y}^{(k)} = \{Y_0^{(k)}, \dots, Y_{N_k-1}^{(k)}\}$ ,  $k=1, \dots, M$ , чтобы:

$$\begin{aligned} \sum_{k=1}^M H(\mathbf{Y}^{(k)}) &\rightarrow \min, \\ H(\mathbf{Y}^{(k)}) &\leq H_{\max}, \quad k=1, \dots, M. \end{aligned} \quad (9)$$

Минимум ищется по всем возможным разбиениям. Решение задачи определяет

правило разбиения спектра на зоны (кластеры) которые соответствуют  $M$  потокам данных, каждый из которых обрабатывается отдельным векторным квантователем. По полученной в результате статистической обработки фрагментов изображения ковариационной матрице коэффициентов преобразования  $K_Y$  для построения разбиения спектра на зоны ВК могут быть использованы два квази-оптимальных способа [13,14]: алгоритм «Выращивания кластеров» и алгоритм «Выделения сильных связей». Оба предложенных способа численного поиска решения (9) основаны на минимизации энтропии межкластерных связей  $H_{св}(Y^{(1)}, \dots, Y^{(M)}) = \sum_{k=1}^{M-1} H(Y^{(k)}) - H(Y)$ . Более близкие к оптимальному (удовлетворяющему (9)) разбиению показал второй из названных алгоритмов.

В четвертой главе предложена общая схема компрессии, использующая адаптивное RD-оптимизированное ВК в области спектров фрагментов изображения. Пусть разбиение спектра  $Y$  фрагмента изображения на зоны для векторного квантования (кластеры)  $\{Y^{(j)}\}$  задано; для каждой зоны некоторым образом сформированы (например, в результате обучающей обработки тестовых изображений) кодовые книги  $\{C^{(j)}\}$ ,  $C^{(j)} = \{c_0^{(j)}, \dots, c_{L_j-1}^{(j)}\}$ , где каждому вектору  $c_k^{(j)}$  поставлен в соответствие вес  $p_k^{(j)}$ , имеющий смысл вероятности появления соответствующего вектора в  $j$ -ом входном потоке,  $\sum_{k=0}^{L_j-1} p_k^{(j)} = 1$ , причем векторы кодовых книг  $\{C^{(j)}\}$  упорядочены по весам,  $\forall j, \forall k > m: p_m^{(j)} \geq p_k^{(j)}$ . Для обеспечения возможности адаптации кодовой книги необходимо предусмотреть процедуру альтернативного «невекторного» квантования и кодирования кластеров (назовем такую процедуру *непосредственным кодированием*), для которой возможно применение различных вариантов, рассмотренных в диссертационной работе. В частности, предложена экспериментальная методика поиска оптимального разбиения дискретных спектров на области для независимого стати-

стического кодирования, когда каждой такой области соответствует отдельная статистическая модель арифметического кодера. Непосредственно закодированный вектор помещается в кодовую книгу на место вектора, имевшего наименьший вес, а если в кодовой книге имеется несколько таких векторов, то удаляется вектор, который попал в кодовую книгу раньше других.

Пусть  $D(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|^2$  – квадрат евклидовой нормы разности векторов  $\mathbf{x}$  и  $\mathbf{y}$ ;  $D_{HK} = D(\mathbf{Y}^{(j)}, \hat{\mathbf{Y}}^{(j)})$  – ошибка ( $\hat{\mathbf{Y}}^{(j)}$  – декодированный кластер), а  $R_{HK} = R(\mathbf{Y}^{(j)})$  – битовые затраты непосредственного кодирования кластера. Предлагаемый алгоритм RD-оптимизированной обработки фрагмента изображения сводится к повторению для каждого кластера  $\mathbf{Y}^{(j)}$  ( $j=1, \dots, M$ ) следующей процедуры.

**Шаг 1.** Вычислить RD-функцию векторного кодирования

$$J_{BK}^{(j)} = \min_{\mathbf{c}_k^{(j)} \in \mathcal{C}^{(j)}} (D(\mathbf{c}_k^{(j)}, \mathbf{Y}^{(j)}) - \lambda \log_2 p_k^{(j)}).$$

**Шаг 2.** Вычислить RD-функцию непосредственного кодирования

$$J_{HK}^{(j)} = D_{HK} + \lambda R_{HK}.$$

**Шаг 3.** Кодирование по наилучшему варианту.

**Если**  $J_{HK}^{(j)} - \lambda \log_2 v_{HK}^{(j)} > J_{BK}^{(j)} - \lambda \log_2 v_{BK}^{(j)}$ ,

где  $v_{HK}^{(j)}$  и  $v_{BK}^{(j)}$  – соответственно частоты непосредственно и векторно кодируемых кластеров в  $j$ -ом потоке ( $v_{HK}^{(j)} + v_{BK}^{(j)} = 1$ ),

**то** /\* кодировать векторно \*/

**3.1.** Закодировать битовый признак BK;

**3.2.** Закодировать индекс  $i$  наилучшего вектора

$$\mathbf{c}_i^{(j)} = \arg \min_{\mathbf{c}_k^{(j)} \in \mathcal{C}^{(j)}} (D(\mathbf{c}_k^{(j)}, \mathbf{Y}^{(j)}) - \lambda \log_2 p_k^{(j)})$$

и изменить веса векторов

$$p_i^{(j)} \leftarrow p_i^{(j)} + \Delta, \quad \forall m \neq i: p_m^{(j)} \leftarrow p_m^{(j)} - \Delta_m;$$

**3.3.** Выполнить сортировку кодовой книги  $\mathcal{C}^{(j)}$ :

– определить индекс  $m$  новой позиции вектора  $\mathbf{c}_i^{(j)}$ :  $m = \min_{p_k^{(j)} = p_i^{(j)}} k$ ,

– выполнить циклическую перестановку векторов и их весов:

$$\begin{aligned} \mathbf{c}_m^{(j)} &\rightarrow \mathbf{c}_{m+1}^{(j)} \rightarrow \dots \rightarrow \mathbf{c}_{i-1}^{(j)} \rightarrow \mathbf{c}_i^{(j)} \rightarrow \mathbf{c}_m^{(j)}, \\ p_m^{(j)} &\rightarrow p_{m+1}^{(j)} \rightarrow \dots \rightarrow p_{i-1}^{(j)} \rightarrow p_i^{(j)} \rightarrow p_m^{(j)}; \end{aligned}$$

**иначе** /\* кодировать непосредственно \*/

**3.4.** Закодировать битовый признак непосредственного кодирования;

**3.5.** Закодировать скалярно проквантованный вектор  $\hat{\mathbf{Y}}^{(j)}$ ;

**3.6.** Обновить кодовую книгу  $\mathbf{C}^{(j)}$ :

– определить индекс начала  $m$  векторов с минимальным весом,

$$m = \min_{p_k^{(j)} = p_{\min}^{(j)}} k, \text{ где } p_{\min}^{(j)} = \min_{l=0, \dots, L_j-1} p_l^{(j)},$$

$$- \forall k = m+1, \dots, L_j-1: \quad \mathbf{c}_k^{(j)} \leftarrow \mathbf{c}_{k-1}^{(j)}; \quad \mathbf{c}_m^{(j)} \leftarrow \hat{\mathbf{Y}}^{(j)}.$$

**Конец**

Сделаем пояснения к приведенному алгоритму адаптивного ВК. Для каждого потока векторов, помимо кодовой книги, изначально задается некоторое распределение  $\{p_0^{(j)}, \dots, p_{L_j-1}^{(j)}\}$ , которое является внутренней статистической моделью адаптивного арифметического кодера, используемой для кодирования индексов кодовой книги. Изменение веса векторов при кодировании индекса на шаге 3.2 представляет собой адаптацию соответствующей модели к входным данным и является внутренней функцией арифметического кодера; конкретное значение величин  $\Delta$ ,  $\Delta_m$  меняется в зависимости от объема уже обработанных данных и во многом определяется спецификой статистического моделирования кодера. После изменения на шаге 3.2 весов векторов выполняется шаг 3.3, который представляет собой «погружение» только что закодированного вектора «вглубь» кодовой книги на позицию с минимально допустимым значением индекса, при котором не нарушается упорядоченность весов,  $\forall j, \forall k > m: p_m^{(j)} \geq p_k^{(j)}$ . Тем самым внутри серий индексов, соответствующих равным весам (и, соответственно, равным битовым затратам на кодирование), производится сортировка векторов по времени появления, которая гарантиру-

ет, что при пополнении кодовой книги новым вектором (шаг 3.6) удален будет наиболее давний из самых редко встречающихся векторов,  $c_{L_j-1}^{(j)}$ .

Для бинарного признака непосредственного или векторного кодирования необходимо использование отдельной модели арифметического кодера, алфавит которой состоит лишь из двух символов; в приведенном описании алгоритма подразумевается, что необходимое изменение значений частот  $v_{HK}^{(j)}$  и  $v_{BK}^{(j)}$  (адаптация модели) производится арифметическим кодером при выполнении шага 3.1 или 3.4.

Таким образом, в четвертой главе формализована общая схема сжатия статических изображений с использованием векторного квантования в области ортогональных преобразований, которая ориентирована на пофрагментную обработку.

**Пятая глава** (65 страниц) посвящена изучению и разработке алгоритмов компрессии неподвижных изображений в области вейвлет-преобразований. Везде далее  $W = \{w_{k,m}\}$  обозначает матрицу дискретного вейвлет-спектра, полученную в результате двумерного  $n$ -шагового вейвлет-преобразования матрицы дискретного изображения. Количество шагов вейвлет-преобразования определяет число частотных уровней в спектре, для  $n$  шагов имеем  $n+1$  уровень. При этом коэффициенты вейвлет-разложения можно упорядочить в виде совокупности деревьев, корнями которых являются элементы, лежащие в самом низкочастотном диапазоне спектра (*саббэнде*). На рис. 1 приведен пример структуры спектра трехшагового преобразования ( $n=3$ ).

Дерево, включающее в себя все свои спектральные коэффициенты, обозначим  $T$ . Строчными буквами  $i$  будем обозначать индексы узлов дерева, каждый из которых соответствует спектральному коэффициенту. Под  $T_i$  понимаем дерево, начинающееся из узла  $i$  и включающее в себя этот узел. Таким образом,  $T_i$  – это определенная ветвь полного дерева  $T=T_0$ . *Подрезанным* деревом  $S_i$

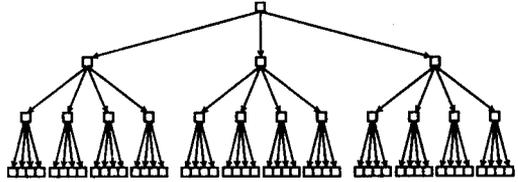
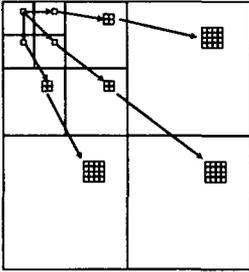


Рисунок 1. Структура связей «родитель-потомки» между коэффициентами вейвлет-спектра

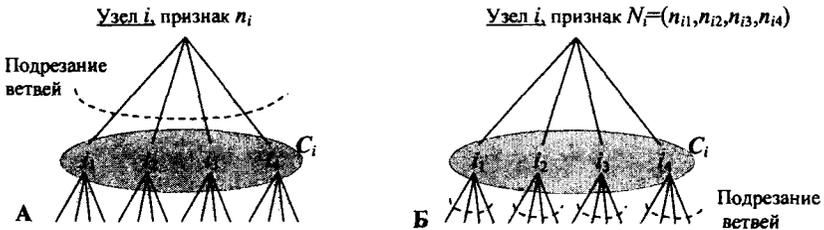
назовем любое поддерево дерева  $T_i (S \subseteq T_i)$ , содержащее узел  $i$ . Аналогично,  $S$  – это дерево, полученное в результате подрезания ветвей полного дерева  $T$ , а все возможные варианты топологии дерева характеризуются множеством  $\{S | S \subseteq T\}$ . Деревом остатка  $U_i$  будем называть множество всех прямых потомков узла  $i$ , т.е.  $U_i = T_i \setminus \{i\}$ . Множество непосредственных потомков узла  $i$  обозначаем через  $C_i$ . Тогда можем записать также, что  $U_i = \bigcup_{j \in C_i} T_j$ . Уровнями дерева называем

горизонтальные слои узлов (см. рис. 1) и обозначаем их как  $L_k$ . Для  $n$ -шагового преобразования в исходном дереве  $T$  имеем  $n+1$  уровень  $\{L_0, \dots, L_n\}$ , при этом нулевой уровень состоит из единственного узла. Таким образом,  $L_0 = \{0\}$ ,  $L_1 = \{1, 2, 3\}$ ,  $L_2 = \{4, \dots, 15\}$ , ...,  $L_n = \{2^{2(n-1)}, \dots, 2^{2n} - 1\}$  (слой с номером  $n$  состоит из листьев). Ограничим рассмотрение случаем, когда единственным параметром кодирования коэффициентов из дерева  $S$  является шаг  $q$  скалярного квантования вейвлет-коэффициентов. Основная задача, стоящая перед алгоритмом компрессии, состоит в поиске RD-оптимальной топологии при фиксированных значениях  $\lambda, q$ :  $J(S^*) = \min_{S \subseteq T} [D(S) + \lambda R(S)]$ .

Идея предложенного в пятой главе алгоритма сжатия [25,27] основана на работах Lewis-Knowles<sup>6</sup> (LK) и Xiong-Ramchandran-Orchard<sup>7</sup> (XRO). При коди-

<sup>6</sup> Lewis A.S., Knowles G. Image Compression Using the 2-D Wavelet Transform // IEEE Trans. Image Proc. – 1992. – Vol. 1. - №2. – P.244-250.

ровании топологии  $S$  в XRO для всех узлов дерева, кроме листьев, использовалась бинарная карта  $\{n_i\}$ : если  $n_i=0$ , то дерево в данном узле подрезается, а если  $n_i=1$ , то по крайней мере непосредственные потомки сохраняются. Сказанное иллюстрируется рис. 2А. Статистическое кодирование признаков  $\{n_i\}$  в алгоритме XRO не используется. Вместе с тем, признаки  $\{n_i\}$  соседних (по положению внутри суббэнда) узлов являются коррелированными величинами. Чтобы учесть эту корреляцию, в разработанном алгоритме признаки для соседних узлов  $\{n_i\}_{i \in C_j}$  предлагается группировать в единый элемент данных, так что карта подрезания ветвей (топология дерева) описывается новым алфавитом данных с символами  $N_j = (n_{i_1}, n_{i_2}, n_{i_3}, n_{i_4}) = n_{i_1} + 2n_{i_2} + 4n_{i_3} + 8n_{i_4}$ , которые, кроме того, кодируются статистически. Новый расширенный признак  $N_j$  оказывается связанным уже с узлом  $j$  более высокого уровня, см. рис. 2Б.



**Рисунок 2.** Способ подрезания ветвей при послойном просмотре узлов: А – алгоритм XRO, Б – предлагаемое кодирование топологии.  $C_i = \{i_1, i_2, i_3, i_4\}$

Идея, которая восходит к работе LK и в том же виде используется в XRO, заключается в следующем: чем больше абсолютная величина вейвлет-коэффициента  $|w_i|$  (или энергия,  $w_i^2$ ) узла-родителя  $i$ , тем менее вероятно появление у данного узла нулевой (т.е. подрезанной) ветви. Более точное предсказание появления нулевой ветви можно составить, если использовать в

<sup>7</sup> Xiong Z., Ramchandran K. and Orchard M.T. Space-frequency Quantization for Wavelet Image Coding // IEEE Trans. Image Proc. – V.6 – May 1997, P. 677-693.

качестве прогнозной величины  $P_i$  сумму, включающую в себя, помимо  $w_i^2$ , также квадраты значений вейвлет-коэффициентов, соседних в саббэнде к узлу  $i$ . Для прогнозной величины узла  $i$  предлагается использовать следующую сумму абсолютных величин вейвлет-коэффициентов:

$$P_i = \frac{1}{16} \left( 4|\tilde{w}_i| + 2 \sum_{m \in I_1(i)} |\tilde{w}_m| + \sum_{m \in I_2(i)} |\tilde{w}_m| \right), \quad (10)$$

где множества для индексов суммирования определяются среди соседей узла  $i$  в саббэнде в соответствии с рис. 3. Весовые коэффициенты, входящие в сумму, были получены в результате статистической обработки ряда тестовых изображений с целью выявления максимума выборочного значения для коэффициента корреляции между прогнозной величиной (10) и энергией проквантованных

вейвлет-коэффициентов непосредственных потомков  $C_i$ :  $\rho \left( P_i, \sum_{j \in C_i} \tilde{w}_j^2 \right) \rightarrow \max$ .

Предлагаемый алгоритм вейвлет-компрессии использует несколько статистических моделей. Функцию арифметического кодера, которая оценивает по внутренним статистическим моделям количество бит, необходимое для кодирования символа  $c$  в  $k$ -ом потоке, обозначим  $H(k, c)$ . Обозначение  $H_{spec}$  относится к потокам, в которых кодируются проквантованные вейвлет-коэффициенты,  $H_{map}$  – к потокам, в которых кодируются признаки начала нулевой ветви. Деревья спектра обрабатываются последовательно; после оптимизации топологии очередного дерева его необходимо закодировать, и тем самым произвести адаптацию статистических моделей арифметического кодера.

**Шаг 0.** /\* инициализация \*/

$S \leftarrow T$ .

$l \leftarrow n-2$ . /\*  $l$  – просматриваемый уровень дерева \*/

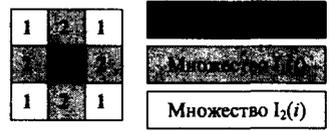


Рисунок 3

**Шаг 1.** /\*подготовительный \*/

$\forall i \in L_{n-1}$ : /\* просматриваются все узлы предпоследнего уровня \*/

**1.1.** /\*корректировка проквантованных коэффициентов\*/

$$\forall j \in C_i : \tilde{w}_j \leftarrow \arg_k \left( J(k) = \min_{k \in \{\tilde{w}_j, \tilde{w}_j \pm 1, 0\}} \left[ (kq - w_j)^2 + \lambda H_{spec}(\text{IndSpec}(j), k) \right] \right).$$

**1.2.** /\* расчет RD-функций сохранения и подрезания листьев \*/

$$J_{U_i}^0 \leftarrow \sum_{j \in U_i} w_j^2; \quad J_{U_i}^1 \leftarrow \sum_{j \in U_i} \left( (\hat{w}_j - w_j)^2 + \lambda H_{spec}(\text{IndSpec}(j), \tilde{w}_j) \right)$$

**Шаг 2.**  $\forall i \in L_i$ : /\*просмотр текущего уровня с попыткой подрезания ветвей \*/

**Если**  $i \neq 0$  **то** /\* не достигли начала дерева \*/

**1.1.** /\* определение оптимальной топологии ветвей \*/

$$J_i^* \leftarrow J_{\sum_{j \in C_i} U_j}^* = \min_{\substack{n_j \in \{0,1\} \\ j \in C_i}} \left( \sum_{j \in C_i} J_{U_j}^{n_j} + \lambda H_{map}(\text{IndMap}(\bar{P}_i), N_i) \right),$$

$$\text{где } \bar{P}_i = \frac{1}{4} \sum_{j \in C_i} P_j \text{ (см. (10)), } N_i = 8n_{i_4} + 4n_{i_3} + 2n_{i_2} + n_{i_1},$$

$$\text{здесь: } i_4 > i_3 > i_2 > i_1, \quad \{i_1, i_2, i_3, i_4\} = C_i.$$

**1.2.**  $\forall j \in C_i$ : **Если**  $n_j=0$ , **то**  $U_j \leftarrow \emptyset$ . /\* подрезать ветвь \*/

**1.3.** /\*корректировка проквантованных коэффициентов\*/

$$\forall j \in C_i : \tilde{w}_j \leftarrow \arg_k \left( J(k) = \min_{k \in \{\tilde{w}_j, \tilde{w}_j \pm 1, 0\}} \left[ (kq - w_j)^2 + \lambda H_{spec}(\text{IndSpec}(j), k) \right] \right).$$

**1.4.** /\* подготовка для просмотра следующего уровня \*/

$$J_{U_i}^0 \leftarrow \sum_{j \in C_i} \left( w_j^2 + J_{U_j}^0 \right) \quad J_{U_i}^1 \leftarrow J_i^* + \sum_{j \in C_i} \left( (\hat{w}_j - w_j)^2 + \lambda H_{spec}(\text{IndSpec}(j), \tilde{w}_j) \right).$$

**1.5.**  $l \leftarrow l-1$ . Перейти на шаг 2.

**иначе** /\*  $i=0$ , достигли начала дерева \*/

**1.6.** /\* определение оптимальной топологии дерева \*/

$$J_0^* \leftarrow J_{\sum_{j \in C_0} U_j}^* = \min_{\substack{n_j \in \{0,1\} \\ j \in C_0}} \left( \sum_{j \in C_0} J_{U_j}^{n_j} + \lambda H_{map}(0, 4n_3 + 2n_2 + n_1) \right), \text{ где } C_0 = \{1, 2, 3\}.$$

**2.7.**  $\forall j \in C_0$  : **Если**  $n_j=0$ , **то**  $U_j \leftarrow \emptyset$ . /\* подрезать ветвь \*/

Перейти на шаг 3.

**Шаг 3.** /\* формирование и выдача результата \*/

$$J^* \leftarrow J_0^* + \sum_{j \in C_0} ((\tilde{w}_j - w_j)^2 + \lambda H_{spec}(\text{IndSpec}(j), \tilde{w}_j)); \quad S^* \leftarrow S.$$

**Конец**

Просмотр узлов дерева проводится от листьев к корню. На первом (подготовительном) шаге просматриваются те узлы  $i$ , которые имеют только непосредственных потомков ( $C_i=U_i$ ), формируются массивы из значений RD-функций, соответствующих вариантам подрезания ( $J_{U_i}^0$ ) и сохранения ( $J_{U_i}^1$ ) листьев. Предварительно, на шаге 1.1, для каждого узла-листа анализируется возможность дополнительной минимизации RD-функции Лагранжа, характеризующей скалярное квантование вейвлет-коэффициентов. Данная процедура основана на известных свойствах распределения вероятностей вейвлет-коэффициентов, в соответствии с которыми битовые затраты  $R$  на кодирование коэффициента априори тем меньше, чем меньше его абсолютное значение (по этой причине для случая  $\tilde{w}_j = 0$  процедура дополнительной минимизации не применяется). На втором шаге, который выполняется для всех узлов  $i$  следующих уровней,  $i \in L_l$  ( $l=n-2, \dots, 0$ ), производится выбор RD-оптимального способа подрезания ветвей (шаги 2.1-2.2), начинающихся из узлов  $j \in C_i$ . Шаги 2.3 и 2.4 несут тот же смысл, что и шаги 1.1, 1.2. Отметим, что введение дополнительной оптимизации скалярного квантования (шаги 1.1 и 2.3) позволяет при том же уровне сжатия дополнительно повысить PSNR на 0,02-0,03 дБ.

Для всех узлов, за исключением корневого, выбор модели для кодирования признаков  $N_i$  (на шаге 2.1) производится с использованием правила, определяемого функцией  $\text{IndMap}(i)$ . Для кодирования признака  $N_0$ , связанного с корнем дерева, используется отдельный поток данных, которому условно присвоен номер 0 (см. шаг 2.6).

Как следует из приведенного описания алгоритма оптимизации, важнейшую роль в его работе играют функции  $\text{IndMap}(\bar{P}_i)$  и  $\text{IndSpec}(i)$ , которые задают правила выбора потоков для кодирования данных. Первая функция производит выбор модели кодирования признака  $N_i = (n_{i1}, n_{i2}, n_{i3}, n_{i4})$  по среднему значению прогнозных величин  $P_i(10)$ ,  $i = i_1, i_2, i_3, i_4$ , и имеет следующий вид:

$$\text{IndMap}(P) = \begin{cases} 1, & \text{если } P < 0.3 \\ 2, & \text{если } 0.3 \leq P < 1.1 \\ 3, & \text{если } 1.1 \leq P < 4 \\ 4, & \text{если } 4 \leq P \end{cases}$$

Пороги  $(t_1, t_2, t_3) = (0.3; 1.1; 4.0)$  для выбора моделей находились в результате минимизации длины выходного битового кода  $R(t_1, t_2, t_3)$  при обработке известных тестовых изображений *Lena*, *Barbara*, *Goldhill*. Эксперименты показали также, что введение большего числа моделей для признаков лишено смысла.

Вторая ключевая функция алгоритма вейвлет-сжатия,  $\text{IndSpec}(i)$ , представляет собой правило выбора модели для кодирования не попавших в нулевые ветви вейвлет-коэффициентов. Для кодирования спектральных коэффициентов более эффективным оказывается использование прогнозной величины, полученной не только по родительскому узлу, но и по вейвлет-коэффициентам узлов, которые находятся в том же саббэнде, рядом с обрабатываемым<sup>8</sup>. В рассматриваемом алгоритме последовательность кодирования и декодирования узлов вейвлет-спектра определяется схемой рисунка 4 (цифры обозначают порядок обработки саббэндов). Для использования в функции  $\text{IndSpec}(j)$  прогнозная величина текущего узла  $j$  (помечен черным) формируется в виде (весовые коэффициенты получены экспериментально):

$$s_j = 0.36P_i + 1.06 \left( |\tilde{w}_{j_y}| + |\tilde{w}_{j_x}| + 0.4|\tilde{w}_{j_d}| \right), \text{ где } j_y - \text{узел-сосед по вертикали, } j_x -$$

<sup>8</sup> Идея использования контекста соседних вейвлет-коэффициентов предложена в работе *Chrysafts C., Ortega A. Efficient Context-Based Entropy Coding for Lossy Wavelet Image Compression // Proc. Data Compression Conference. – Snowbird (Utah), 1997. – P. 241-250.*

узел-сосед по горизонтали,  $j_d$  – узел-сосед по диагонали (которые уже обработаны, см. рис. 4),  $P_i$  определяется для узла-родителя  $i$  ( $j \in C_i$ ) по (10). При этом сама функция, определяющей номер модели кодирования вейвлет-коэффициентов, имеет вид:

$$\text{IndSpec}(j) = \begin{cases} 5, & \text{если } j > 3 \text{ и } s_j < 1.72 \\ 4, & \text{если } j > 3 \text{ и } 1.72 \leq s_j < 4.10 \\ 3, & \text{если } j > 3 \text{ и } 4.10 \leq s_j < 9.80 \\ 2, & \text{если } j > 3 \text{ и } 9.80 \leq s_j < 26 \\ 1, & \text{если } 0 < j \leq 3 \text{ или } 26 \leq s_j \\ 0, & \text{если } j = 0 \end{cases}$$

Пороги (1.72;4.10;9.80;26) также уточнялись по результатам обработки тестовых изображений.

Нулевая модель относится к кодированию спектральных коэффициентов при масштабирующих функциях и вновь обособлена. Первая модель включает в себя самые низкочастотные вейвлет-коэффициенты (уровня  $L_1$ ), а также коэффициенты, для которых прогноз  $s_j$  – наибольший.

Характеристики, которые дает предложенный алгоритм компрессии для стандартных тестовых изображений, приведены в табл. 1. В экспериментах применялись пятишаговые вейвлет-преобразования из работы *Wei D., Pai H.-T., and Bovik A. C., Antisymmetric Biorthogonal Coiflets for*

Image Coding // Proc. IEEE International Conference on Image Processing. – Chicago, 1998. – V. 2. – P. 282-286. Сравнение достигнутых характеристик с результатами применения других известных алгоритмов показывает, что предложенный алгоритм имеет весьма высокие показатели.

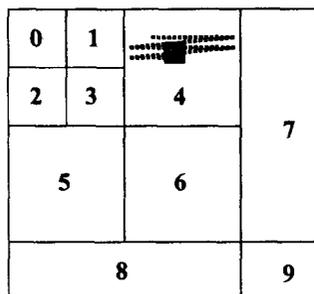


Рисунок 4

Таблица 1. Величина PSNR (в дБ), полученная при сжатии тестовых изображений по предложенному алгоритму

BPP	Lena	Barbara	Goldhill
0.250	34.59	28.37	30.81
0.500	37.66	32.36	33.49
1.000	40.81	37.36	36.99

Завершающие исследования главы 5 относятся к построению гибридной схемы кодирования вейвлет-спектра, когда в дополнение к описанному выше способу подрезания ветвей вейвлет-коэффициентов допускается также возможность векторного «самоквантования» ветвей, которое может трактоваться как фрактальное кодирование в области вейвлет-преобразований<sup>9</sup>. Полученный в результате гибридный алгоритм [26] требует существенно больших вычислений, однако фрактальная составляющая кодирования при этом оказалась практически полностью подавленной базовой схемой вейвлет-компрессии, основанной на подрезании ветвей. Необходимо отметить, однако, что соединение двух подходов в гибридном алгоритме было произведено несложным способом, и возможности дальнейшего развития оставляют здесь обширное поле для исследований.

**Шестая глава** (45 стр.) посвящена исследованию алгоритмов сжатия динамических изображений с целью построения схемы видеокompрессии, пригодной для программной реализации, которая обеспечивает обработку в реальном масштабе времени на базе персональных компьютеров.

Кадром видеопоследовательности назовем матрицу пикселей из  $M_1$  строк и  $M_2$  столбцов:  $\mathbf{V}=(b_{k,l})$ ,  $k=0,1,\dots,M_1-1$ ,  $l=0,1,\dots,M_2-1$ , а под *видеопоследовательностью* будем понимать упорядоченный набор кадров  $\mathbf{V}^0, \mathbf{V}^1, \dots, \mathbf{V}^i, \dots$ . Назовем  $(y,x)$ -блоком кадра  $\mathbf{V}$  ( $y, x$  – целочисленные координаты) некоторую подматрицу  $\mathbf{V}_{y,x}=(b_{k,l})$ , где  $k=y,y+1,\dots,y+N_1-1$ ,  $l=x,x+1,\dots,x+N_2-1$ . В разработанном алгоритме каждый кадр видеопоследовательности разбивается при обработке на смежные матрицы-блоки  $\{\mathbf{V}_{m,n}\}$  размером  $8 \times 8$ ,  $m,n=0,8,16\dots$ . Если какой-либо блок  $\mathbf{V}_{y,x}^{i-1}$  видеопоследовательности оказался в определенном смысле «похож» на оригинальный блок  $\mathbf{V}_{m,n}^i$ , считаем, что блок  $\mathbf{V}_{m,n}^i$  – это *перемещенный*

<sup>9</sup> См., например, *Davis G.M., A wavelet-based analysis of fractal image compression // IEEE Trans. Image Proc. – 1998. – V.7 – №2. – P.141-154.*

фрагмент  $\mathbf{V}_{y,x}^{i-1}$  предыдущего кадра, и для кодирования  $(m,n)$ -блока изображения достаточно задать координаты блока в предыдущем кадре,  $y$  и  $x$ , либо изменения координат  $y$ - $m$  и  $x$ - $n$ . Частным случаем перемещенного блока является *неподвижный* блок, когда  $m=y$ ,  $n=x$ . Если блоку  $\mathbf{V}_{m,n}^i$  нельзя найти похожий в предыдущем кадре, блок должен кодироваться как *новый*. Для выбора способа кодирования очередного обрабатываемого блока  $\mathbf{V}_{m,n}^i$  вновь руководствуемся критерием минимума функции Лагранжа  $J(b)=D(b)+\lambda R(b)$ . Положим, что аргумент  $b=0$  соответствует кодированию перемещенного (неподвижного) блока, а  $b=1$  – нового. Т.е. если  $J(1)>J(0)$ , то блок кодируется как перемещенный, и как новый – в противном случае.

При использовании RD-оптимизации задача поиска перемещенных блоков формулируется следующим образом. Для заданного  $(m,n)$ -блока  $\mathbf{V}_{m,n}^i$   $i$ -ого кадра найти в предыдущем восстановленном кадре такой  $(y,x)$ -блок  $\hat{\mathbf{V}}_{y,x}^{i-1}$ , чтобы минимальное значение принимала RD-функция Лагранжа

$$J^* = \|\mathbf{V}_{m,n}^i - \hat{\mathbf{V}}_{y,x}^{i-1}\| + \lambda R(y-m, x-n) = \min_{(u,v) \in \Omega} \left\{ \|\mathbf{V}_{m,n}^i - \hat{\mathbf{V}}_{v,u}^{i-1}\| + \lambda R(u-m, v-n) \right\}. \quad (11)$$

Здесь учтено, что координаты найденного блока будут кодироваться как относительные, т.е. вектором перемещений  $r=(y-m, x-n)$ . Для того чтобы поиск можно было осуществить в реальном масштабе времени, в качестве области  $\Omega$  необходимо рассматривать лишь точки  $(v,u)$ , достаточно близко расположенные к точке  $(m,n)$ . Повышение эффективности поиска за счет расширения области  $\Omega$  достигается применением различных алгоритмов направленного поиска, ориентированных на минимизацию ошибки представления перемещенного блока  $\|\mathbf{V}_{m,n}^i - \hat{\mathbf{V}}_{y,x}^{i-1}\|$ , что соответствует частному случаю (11) при  $\lambda=0$ .

Для того, чтобы учесть вклад битовых затрат в RD-функцию  $J^*$ , будем проводить минимизацию (11) поэтапно, на каждом из этапов приближенно полагая, что рассматриваемые векторы перемещений влекут одинаковые затраты на ста-

тистическое кодирование. Кроме того, для повышения универсальности итерационных алгоритмов поиска малые перемещения будем искать более точно. Действительно, если некоторый блок изображения переместился на значительное расстояние по сравнению с предыдущим кадром, то соответствующий фрагмент изображения воспринимается человеческим глазом смазано, и в точном определении вектора перемещения нет необходимости. Малые же перемещения блоков не только являются преобладающими, но и должны точно отслеживаться в силу специфики визуального восприятия. Предлагаемый алгоритм RD-поиска имеет следующий вид.

**Шаг 0.** Вычисление значения RD-функции неподвижного блока,  $r=(0,0)$ :

$$J_0 = \left\| \mathbf{B}_{m,n}^i - \hat{\mathbf{B}}_{m,n}^{i-1} \right\| + \lambda R(0,0).$$

**Шаг 1.** Близкий к перебору точный поиск небольших перемещений.

**1.1.** Среди девяти  $(y,x)$ -блоков предыдущего кадра,

$$(y,x) = (m + \Delta_{y1}, n + \Delta_{x1}), \Delta_{y1} \in \{-2,0,2\}, \Delta_{x1} \in \{-2,0,2\},$$

находим такой блок  $\hat{\mathbf{B}}_{y_1, x_1}^{i-1}$ , что

$$\left\| \mathbf{B}_{m,n}^i - \hat{\mathbf{B}}_{y_1, x_1}^{i-1} \right\| = \min_{(\Delta_{y1}, \Delta_{x1})} \left\| \mathbf{B}_{m,n}^i - \hat{\mathbf{B}}_{m+\Delta_{y1}, n+\Delta_{x1}}^{i-1} \right\|.$$

**1.2.** Среди девяти  $(y,x)$ -блоков предыдущего кадра,

$$(y,x) = (y_1 + \Delta_{y2}, x_1 + \Delta_{x2}), \Delta_{y2}, \Delta_{x2} \in \{-1,0,1\},$$

находим такой блок  $\hat{\mathbf{B}}_{y_1^*, x_1^*}^{i-1}$ , что

$$\left\| \mathbf{B}_{m,n}^i - \hat{\mathbf{B}}_{y_1^*, x_1^*}^{i-1} \right\| = \min_{(\Delta_{y2}, \Delta_{x2})} \left\| \mathbf{B}_{m,n}^i - \hat{\mathbf{B}}_{y_1+\Delta_{y2}, x_1+\Delta_{x2}}^{i-1} \right\|.$$

**1.3.** Вычисление значения RD-функции

$$J_1 = \left\| \mathbf{B}_{m,n}^i - \hat{\mathbf{B}}_{y_1^*, x_1^*}^{i-1} \right\| + \lambda R(y_1^* - m, x_1^* - n).$$

**Шаг 2.** Грубый поиск больших перемещений.

**2.1.** Среди восьми  $(y,x)$ -блоков предыдущего кадра,

$$(y,x) = (m + \Delta_{y3}, n + \Delta_{x3}),$$

$$(\Delta_{y3}, \Delta_{x3}) \in \{(8,0), (0,8), (-8,0), (0,-8), (6,6), (-6,6), (-6,-6), (6,-6)\},$$

находим такой блок  $\hat{\mathbf{B}}_{y_2, x_2}^{i-1}$ , что

$$\left\| \mathbf{B}_{m,n}^i - \hat{\mathbf{B}}_{y_2, x_2}^{i-1} \right\| = \min_{(\Delta_{y3}, \Delta_{x3})} \left\| \mathbf{B}_{m,n}^i - \hat{\mathbf{B}}_{m+\Delta_{y3}, n+\Delta_{x3}}^{i-1} \right\|.$$

**2.2.** Среди девяти  $(y, x)$ -блоков предыдущего кадра,

$$(y, x) = (y_2 + \Delta_{y4}, x_2 + \Delta_{x4}),$$

$$(\Delta_{y4}, \Delta_{x4}) \in \{(0,0), (2,2), (-2,2), (-2,-2), (2,-2), (3,0), (0,3), (-3,0), (0,-3)\}$$

находим такой блок  $\hat{\mathbf{B}}_{y_2^*, x_2^*}^{i-1}$ , что

$$\|\mathbf{B}_{m,n}^i - \hat{\mathbf{B}}_{y_2^*, x_2^*}^{i-1}\| = \min_{(\Delta_{y4}, \Delta_{x4})} \|\mathbf{B}_{m,n}^i - \hat{\mathbf{B}}_{y_2 + \Delta_{y4}, x_2 + \Delta_{x4}}^{i-1}\|.$$

**2.3.** Вычисление значения RD-функции

$$J_2 = \|\mathbf{B}_{m,n}^i - \hat{\mathbf{B}}_{y_2^*, x_2^*}^{i-1}\| + \lambda R(y_2^* - m, x_2^* - n).$$

**Шаг 3.** Выбор наилучшего варианта перемещения блока.

$$J^* = \min\{J_0, J_1, J_2\}, \quad (y^*, x^*) = \arg J^*(y, x).$$

**Конец**

Для вычисления значения функции  $J^0$  необходимо учесть битовые затраты на кодирование признака перемещенного блока:  $J^0 = J^* - \lambda \log_2(v_{mov})$ , где  $v_{mov}$  – частота появления перемещенных блоков в уже обработанных данных. Основной объем вычислений в приведенном алгоритме связан с подсчетом отклонений  $\|\mathbf{B}_{m,n}^i - \hat{\mathbf{B}}_{y,x}^{i-1}\|$ . При вычислениях одна пробная точка переходит с шага 0 на шаг 1.1, одна – с шага 1.1 на 1.2, одна – с шага 2.1 на 2.2. В итоге вычисление отклонения требуется выполнить 33 раза.

Для повышения эффективности приведенного алгоритма поиска для вектора  $(\Delta_y, \Delta_x)$  обрабатываемого блока следует строить прогноз по векторам  $(\Delta_y^v, \Delta_x^v)$  и  $(\Delta_y^h, \Delta_x^h)$  двух уже обработанных соседних блоков (соответственно соседа по вертикали и соседа по горизонтали). Собственно *прогноз* – это относительные координаты  $(y^0, x^0)$ , которые определяют перенос центра области поиска: из точки  $(m, n)$  в точку  $(m', n') = (m + y^0, n + x^0)$ . Эксперименты показывают, что число новых блоков изображения сокращается на 5...25%, если принять следующее правило составления прогноза:

$$(y^0, x^0) = \begin{cases} (0,0), \text{ оба соседних блока - новые} \\ (\Delta_y^v, \Delta_x^v), \text{ по горизонтали - новый, по вертикали - перемещенный} \\ (\Delta_y^h, \Delta_x^h), \text{ по горизонтали - перемещенный, по вертикали - новый} \\ ((\Delta_y^v, \Delta_x^v) + (\Delta_y^h, \Delta_x^h))/2, \text{ оба соседа - перемещенные блоки} \end{cases}$$

Следуя идеологии стандарта MPEG, обработку новых блоков также осуществляем при помощи квантования с последующим статистическим кодированием коэффициентов двумерного ДКП. Пусть  $S = F(B_{m,n})$  – результат ДКП блока  $B_{m,n}$ . Обозначим  $\tilde{S}_Q = \{\tilde{s}_{k,j} = \text{round}(s_{k,j}/q_{k,j})\}_{k,j=0}^Q$ ,  $\hat{S}_Q = \{\hat{s}_{k,j} = q_{k,j}\tilde{s}_{k,j}\}_{k,j=0}^Q$ , где  $Q = \{q_{k,j}\}_{k,j=0}^Q$  – одна из матриц квантования JPEG. Для статистического кодирования матрицы  $\tilde{S}$  применен алгоритм контекстного кодирования, рассмотренный в главе 4, в который введен дополнительный этап RD-оптимизации. Пусть вновь  $Z_Q = (z_0, \dots, z_{63})$  – вектор, полученный в результате зигзагообразного считывания данных из матрицы  $\tilde{S}_Q$  по правилу, определяемому стандартом JPEG ( $\tilde{S}_Q \leftrightarrow Z_Q$ ),  $G(Z_Q) = \{g_k\}_{k=1}^C \subseteq \{0, \dots, 63\}$  – множество индексов ненулевых элементов вектора  $Z_Q$ , т.е.  $z_{g_k} \neq 0$ , если  $g_k \in G$ ;  $\forall g_k \notin G: z_{g_k} = 0$ . RD-оптимизация статистического кодирования возможна за счет удлинения нулевых серий путем сокращения количества элементов во множестве  $G$  (дополнительного обнуления компонент вектора  $Z_Q$ ). Для того, чтобы в результате не произошло заметного усложнения алгоритма кодирования, будем анализировать лишь возможность увеличения финальной нулевой серии, что дает наибольший вклад в минимизацию функции  $J(Z_Q) = D(Z_Q) + \lambda R(Z_Q)$ . Пусть  $Z_Q^m$  – вектор, полученный из вектора  $Z_Q$  в результате обнуления последних компонент  $\{z_k\}_{k=g_m+1}^{63}$ , т.е.  $G(Z_Q^m) = \{g_k\}_{k=1}^m \subseteq G(Z_Q)$ ,  $m=1, \dots, C$ . Тогда простейшая RD-оптимизация состоит в поиске такого индекса  $g_{m^*} \in G(Z_Q)$ , что

$$J_Q = J(\mathbf{Z}_Q^{m*}) = \min_{m \in \{1, \dots, C\}} [J(\mathbf{Z}_Q^m) = D(\mathbf{Z}_Q^m) + \lambda R(\mathbf{Z}_Q^m)]. \quad (12)$$

Рассмотренная выше процедура кодирования нового блока изображения предполагает использование заданной матрицы квантования  $\mathbf{Q}$ . Универсальный же алгоритм кодирования должен оперировать некоторым набором матриц квантования  $\{\mathbf{Q}_j\}$  с возможностью выбора требуемой для конкретных условий. Если набор достаточно большой, то подбор матрицы квантования  $\mathbf{Q}$  по принципу минимума функции  $J_Q$  (12) превращается в громоздкую процедуру, нереализуемую в реальном масштабе времени стандартными средствами. Кроме того, при большом диапазоне возможных значений индекса  $j$  его кодирование для каждого нового блока по отдельности влечет недопустимо высокие дополнительные битовые затраты. Поэтому в качестве исходного набора были выбраны лишь несколько матриц из множества, рекомендуемого JPEG, которые соответствуют наилучшему, наихудшему и некоторым промежуточным уровням качества. В экспериментах выбиралось число матриц  $|\{\mathbf{Q}_j\}|=4$ . Для ускорения выполнения операций деления, которые необходимы при квантовании, элементы матриц  $\{\mathbf{Q}_j\}$  были округлены до ближайшего значения  $2^k$ ,  $k=0,1,\dots$  Такой подход позволяет заменить операции целочисленного деления и умножения сдвигами разрядов двоичного представления чисел, которые обычно выполняются реальной аппаратурой намного быстрее.

С учетом битовых затрат, которые необходимы для кодирования индекса матрицы квантования, функция Лагранжа, соответствующая кодированию нового блока изображения, определяется следующим образом:

$$J^1 = \min_{\mathbf{Q} \in \{\mathbf{Q}_j\}} (J_Q - \lambda \log_2 v_Q) - \lambda \log_2 v_{new},$$

где  $J_Q$  находится в соответствии с (12),  $v_Q$  – частота появления матрицы  $\mathbf{Q}$ ,  $v_{new}$  – частота появления новых блоков в ходе предшествовавшей обработки.

При исследовании характеристик итогового алгоритма видеокompрессии для оценки величины ошибки кодирования в восстановленной последователь-

ности  $\hat{\mathbf{V}}^0, \hat{\mathbf{V}}^1, \dots, \hat{\mathbf{V}}^{K-1}$  использовалось отношение пикового значения сигнала к шуму, которое определялось следующим образом:

$$PSNR = 10 \lg \frac{255^2}{M_1 M_2 K \sum_{k=0}^{K-1} \|\mathbf{V}^k - \hat{\mathbf{V}}^k\|^2} \quad [\text{дБ}], \quad (13)$$

где  $M_1$  и  $M_2$  задают размер кадра в пикселях. Для экспериментов были выбраны широко известные тестовые последовательности *News*, *Container ship*, *Hall monitor*, *Akiyo*, *Claire*. Все они имеют размер кадра  $144 \times 176$  пикселей. Перед проведением экспериментов оригинальные последовательности были прорежены: только каждый третий кадр,  $\mathbf{V}^{3k}$  ( $k=0, \dots, 24$ ), использовался для формирования тех 25-кадровых видеопоследовательностей, которые затем и были подвергнуты обработке. Данное временное прореживание было призвано смоделировать скорость видеозахвата 10 кадров в секунду, вместо оригинальных (для всех вышеназванных последовательностей) 30 кадров в секунду. Обработке подвергалась только яркостная Y-составляющая, и только по яркостной составляющей анализировалась ошибка (13). Программное сжатие видеопоследовательностей при этом удалось осуществить в реальном масштабе времени.

Результаты численных экспериментов, полученные аспирантом Ф.В.Стрелковым [28], отражены в таблице 2. В скобках приведено значение PSNR (13), достигнутое на тех же самых прореженных 25-кадровых тестовых видеопоследовательностях с использованием общедоступного кодера MPEG-2 (`mpeg2encode`, версия 1.1), разработанного группой MSSG, см. <http://www.mpeg.org/MSSG>. Длина файла сжатых данных, полученного в каждом эксперименте, в точности равна произведению величины битового потока данных (приведена в таблице) на множитель 2.5. Во всех тестах предложенная схема компрессии дает высокие результаты, превосходя характеристики указанного кодера MPEG-2, несмотря на то, что кодек `mpeg2encode` использовал

для поиска перемещенных блоков изображения полный перебор в области из 23×23 пикселей.

**Таблица 2.** Характеристики сжатия предложенного алгоритма видеокомпрессии и кодера MPEG-2 (MSSG)

Видеопоследовательность (10 кадров в сек., всего 25 кадров)	Битовый поток, килобит в сек.	PSNR (13), дБ
<i>News</i>	64	34.01 (32.94)
<i>Container ship</i>	64	33.36 (33.35)
<i>Hall monitor</i>	48	35.62 (32.54)
<i>Akiyo</i>	24	34.42 (32.90)
<i>Claire</i>	24	35.87 (34.83)

Описанный алгоритм видеокомпрессии был реализован программно в рамках работ, проводившихся в ГУ НПК «Технологический Центр» Московского государственного института электронной техники и в НПП «Технология» [29,30], внедрение разработанных библиотек видеокомпрессии осуществлено в ряде программных систем, среди которых наибольший практический интерес представляет система видео контроля и регистрации Visual Security (см. <http://www.tcen.ru/vs>).

Итог исследованиям, проведенным в диссертационной работе, подводится в заключительном разделе – «Основные выводы и заключение» (3 стр.).

### **Основные выводы и заключение**

В представленной диссертационной работе исследованы различные аспекты применения дискретных ортогональных преобразований для цифровой компрессии изображений, как со стороны сугубо формального теоретического анализа, так и со стороны тех требований и ограничений, которые практика на-

лагает на конкретные вычислительные схемы и алгоритмы. В целом содержание работы носит прикладную направленность, поэтому большинство теоретических результатов подкреплено вычислительными экспериментами, результаты которых, в свою очередь, не только служили иллюстрацией или проверкой теории, но часто давали толчок и представляли собой исходный материал для дальнейших изысканий. По результатам проведенных в диссертационной работе исследований можно сделать следующие выводы.

1. Ортогональные преобразования являются основным инструментом, используемым для декорреляции данных при компрессии изображений. В случае, когда математическая модель дискретного сигнала задается ковариационной матрицей, для анализа эффективности декоррелирующей обработки целесообразно применение предложенного в работе критерия средней избыточной энтропии.
2. Специально для сжатия коррелированных данных получено и впервые введено в рассмотрение дискретное псевдосинусное преобразование (ДПКП). В схеме компрессии, предполагающей наличие этапа скалярного квантования коэффициентов преобразования, среди рассмотренных быстрых преобразований, реализация которых сводится только к операциям сложения-вычитания (Уолша, Хаара, псевдосинусное), ДПКП дает наилучшие результаты декорреляции для дискретного сигнала, описываемого марковской моделью.
3. С использованием полученных быстрых алгоритмов ДПКЛ, учитывающих специфику обработки вещественных массивов, предложенная схема компрессии изображений на основе арифметического кодирования коэффициентов ДПКЛ достигает характеристик, которые по качеству обработки и вычислительной сложности близки к варианту JPEG на основе ДКП.

4. При использовании статистического кодирования коэффициентов ДКП по методу JPEG наличие «скачков» дискретного сигнала наименее желательно в центральной области фрагментов обработки.
5. Высокую эффективность применения в различных схемах и алгоритмах сжатия данных имеет метод многомодельного (многопоточкового) арифметического кодирования, а одним из ключевых моментов в разработке схем компрессии является определение правил для выбора текущей модели кодирования по контексту уже обработанных данных. Так, применение в схеме JPEG предложенного в главе 4 алгоритма многомодельного контекстного арифметического кодирования коэффициентов ДКП повышает эффективность сжатия данных на 10%.
6. При сжатии изображений с использованием многомодельного кодирования древовидных структур вейвлет-спектров правило выбора моделей следует строить по комбинированному контексту, который учитывает как окружение самого вейвлет-коэффициента в суббэнде, так и окружение коэффициента-«родителя». Полученный на этой основе новый эффективный алгоритм компрессии цифровых изображений с потерями, который разработан по результатам изучения статистических свойств спектров дискретных вейвлет-преобразований, показывает высокие характеристики сжатия при сложности реализации, приемлемой для широкого круга приложений.
7. Для устранения межкадровой (временной) избыточности видеоданных наиболее предпочтительным для практического применения среди исследованных алгоритмов блочной компенсации перемещений следует признать предложенный гибридный алгоритм направленного поиска, при котором малые перемещения ищутся точно и тщательно, а значительные перемещения – более грубо.

8. При использовании предложенного алгоритма видеокompрессии, который был разработан на основе RD-оптимизационного подхода с учетом требований и специфики программной реализации, достигается сжатие и восстановление видео в реальном масштабе времени на базе современных персональных компьютеров, при высоком качестве обработки.

В целом, в диссертационной работе получены новые научные результаты, теоретические положения которых позволили в значительной степени развить и формализовать процедуры анализа и синтеза схем компрессии цифровых изображений, основанных на использовании дискретных ортогональных преобразований. Выработанные подходы и рекомендации привели к построению конкретных схем и алгоритмов компрессии, многие из которых были реализованы программно и экспериментально подтвердили эффективность своего применения.

### **Список основных работ по теме диссертации**

1. *Ефимов А.В., Умняшкин С.В.* Быстрые алгоритмы вычисления дискретного преобразования Крестенсона-Леви и оценки его спектральных характеристик // Теор. функций и прикл.: Тр. 7-й саратов. зим. шк. (1994 г.). Ч. 2.- Саратов: Изд.-во СГУ, 1995. - С. 9-20.
2. *Ефимов А.В., Поспелов А.С., Умняшкин С.В.* Применение преобразования Крестенсона-Леви в задачах цифровой обработки информации // Междунар. конф. "Функциональные пространства, теория приближений, нелинейный анализ", посв. 90-летию акад. С.М.Никольского (27 апр. - 3 мая 1995 г.): Тез. докл. - М.: Изд.-во МФТИ, 1995. - С.124-125.
3. *Умняшкин С.В.* Применение дискретного преобразования Крестенсона-Леви (ДПКЛ) для кодирования изображений: сравнение с дискретным преобразо-

- ванием Фурье (ДПФ) // Всерос. науч.-тех. конф. "Электроника и информатика" 15-17 нояб. 1995 г.: Тез. докл. - М.: МГИЭТ (ТУ), 1995. - С. 265-266.
4. Умняшкин С.В. Оценка дисперсии элементов спектра дискретного косинусного преобразования стационарного марковского процесса первого порядка // Междун. конф. по теор. прикл. функ., посв. памяти проф. П.П.Коровкина (Калуга, 26-29 июня 1996 г.): Тез. докл. -Т.2.-Тверь: ТГУ, 1996. - С. 217-218.
  5. Умняшкин С.В. Оценка эффективности использования унитарных преобразований для кодирования дискретных сигналов // Информатика и связь: Сб. Научных трудов под ред. В.А.Бархоткина. М.: МИЭТ.- 1997. С.73-78.
  6. Умняшкин С.В. Оценка эффективности применения дискретных преобразований для сжатия данных // «Электроника и информатика – 97». Вторая всероссийская научно-техническая конференция с международным участием (Зеленоград, 25-26 ноября 1997г.): Тез. док. Часть 2. - С.79.
  7. Ефимов А.В., Поспелов А.С., Умняшкин С.В. Теоретические основы и некоторые особенности применения дискретных мультипликативных преобразований в задачах сжатия цифровых изображений // Праці міжнародної конференції "Питання оптимізації обчислень" (6-8 жовтня 1997 р., м. Київ) - Київ: Інститут кібернетики імені В.М.Глушкова, 1997. - С. 108-112.
  8. Ефимов А.В., Поспелов А.С., Умняшкин С.В. Некоторые свойства мультипликативных ортонормированных систем, используемые в цифровой обработке сигналов // Труды математического института им. В.А.Стеклова РАН. - Т.219. - 1997. - С 137-182.
  9. Ефимов А.В., Умняшкин С.В. О некоторых свойствах обобщенной системы Хаара и оценке эффективности применения ортогональных преобразований для сжатия данных // Функциональные пространства. Дифференциальные операторы. Проблемы математического образования: Труды междун. конф., посв. 75-летию чл.-корр. РАН проф. Л.Д.Кудрявцева. - Том 1. - М.: Рос. ун-т дружбы народов, 1998. - С. 70-73.

10. Умняшкин С.В. О модификации дискретного косинусного преобразования // Теория приближений и гармонический анализ: Тез. докл. междунар. конф. (Тула, 26-29 мая 1998 г.). - Тула: ТулГУ, 1998. - С.264-265.
11. Умняшкин С.В. О модификации дискретного косинусного преобразования // Изв. Тул. гос. ун-та. Сер. Математика. Механика. Информатика. Тула: ТулГУ, 1998. Т. 4. Вып. 1. С. 143-147.
12. Умняшкин С.В., Кочетков М.Е. Анализ эффективности использования дискретных ортогональных преобразований для цифрового кодирования коррелированных данных // Известия вузов. Электроника. - №6. - 1998. - С. 79-84.
13. Умняшкин С.В. О кластеризации коррелированных данных. // Информационные технологии в инновационных проектах. Междунар. конф. (г.Ижевск, 20-22 апр. 1999г.): Материалы докладов. - Ижевск, ИжГТУ, 1999. - С. 59-65.
14. Умняшкин С.В. Алгоритм кластеризации коррелированных данных // VII Междунар. конф. Математика. Экономика. Экология. Образование. Междунар. симп. Ряды Фурье и их приложения: Тез. док. – Ростов: Рост. гос. эконом. акад., 1999. – С. 211-212.
15. Ефимов А.В., Постелов А.С., Умняшкин С.В. Некоторые вопросы применения мультипликативных систем и преобразований в задачах цифровой обработки изображений // VII Междунар. конф. Математика. Экономика. Экология. Образование. Междунар. симп. Ряды Фурье и их приложения: Тез. док. – Ростов: Рост. гос. эконом. акад., 1999. – С. 154-155.
16. Умняшкин С.В. Псевдокосинусное преобразование для сжатия дискретных сигналов // Информационные технологии и проблемы микроэлектроники. Сб. науч. тр. – М.: МИЭТ. – 1999. – С.158-170.
17. Умняшкин С.В. Алгоритм компрессии неподвижных изображений на основе дискретного вэйвлет-преобразования // VII международная конференция «Математика. Компьютер. Образование» (Дубна, ОИЯИ, 24-29 янв. 2000): Тез. док. – Москва: Прогресс-Традиция, 2000. – С.327.

18. *Ефимов А.В., Умняшкин С.В.* О структуре некоторых прямых и обратных вейвлет-преобразований // Теория приближений функций и операторов: Тез. докл. Междунар. конф., посв. 80-лет. со дня рожд. С.Б.Стечкина (Екатеринбург, 28 фев. – 3 мар. 2000). -Екатеринб.: УрГУ, 2000. – С.74-75.
19. *Умняшкин С.В., Стрелков Ф.В., Жуков В.Г.* Трехшаговые алгоритмы поиска перемещенных блоков изображения // Информационные технологии и системы управления. Сб. научн. тр. под ред. В.А.Бархоткина. – М: МИЭТ, 2000. – С. 47-55.
20. *Умняшкин С.В.* Цифровая компрессия изображений с использованием дискретного преобразования Крестенсона-Леви // Межотраслевой научно-технический журнал «Оборонный комплекс – научно-техническому прогрессу России», №2, 2000. С.28-39.
21. *Умняшкин С.В., Космач М.В.* Оптимизация кодирования цифровых изображений по методу JPEG //Изв. вуз. Электроника. -№4-5. -2000. -С.139-141.
22. *Умняшкин С.В.* Компенсация перемещения объектов при сжатии видеоданных // «Электроника и информатика – XXI век» Третья международная научно-техническая конференция: Тез. док. – М.: МИЭТ, 2000. – С. 365-366.
23. *Умняшкин С.В.* Алгоритм поиска перемещенных блоков для кодирования цифровых видеоизображений // Межотрасл. н.-т. журнал «Оборонный комплекс – научно-техническому прогрессу России», №3, 2001. – С. 38-41.
24. *Умняшкин С. В.* Использование контекстного арифметического кодирования для повышения сжатия данных по схеме JPEG // Известия вузов. Электроника. - №3. - 2001. – С. 96-99.
25. *Умняшкин С. В.* Вейвлет-компрессия цифровых изображений с прогнозированием статистических моделей //Изв. вузов. Электроника. - №5. - 2001. - С.86-94.
26. *Умняшкин С. В.* Алгоритм фрактального кодирования изображений в области вейвлет-преобразований // Компьютерна математика. Оптимизация об-

числень: Збірник наукових праць. – Том 1. – Київ: Інститут кібернетики ім. В.М.Глушкова, 2001. – С. 385-391.

27. *Umnyashkin S.V.* Image compression based on mixed predictive modeling of wavelet trees // Reports from Växjö University (Sweden) – Mathematics, natural sciences and technology. – Nr.11 (September), 2001. – 18 pages.
28. *Umnyashkin S.V., Strelkov F.V.* An RD-optimized scheme for real-time video compression // Reports from Växjö University (Sweden) – Mathematics, natural sciences and technology. – Nr.12 (September), 2001. – 15 pages.
29. Разработка алгоритмов и программного обеспечения для реализации цифрового анализа и компрессии видеоизображений в реальном масштабе времени на базе программно-аппаратного комплекса общего назначения: Отчет о НИР (закл.) / НПП «Технология»; рук. – Умняшкин С.В. – «Страж»; № гос. рег. 01990011697; Инв. №01077. – Москва, 1999. – 74 с.
30. Исследование и разработка алгоритмов программного сжатия данных с потерями для цифровой обработки видеоизображений: Отчет о НИР (закл.) / НПП «Технология»; рук. – Умняшкин С.В. – «Дозор»; № гос. рег. 01200004624; Инв. №100704. – Москва, 2000. – 48 с.

Подписано в печать: 25 декабря 2001 года

Заказ № 332. Тираж 100 экз. Уч.-изд.л. 2,4. Формат 60×84 1/16