

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

3/2007

НАУЧНО - ТЕХНИЧЕСКИЙ ЖУРНАЛ

ЦИФРОВАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ

ТЕМАТИЧЕСКИЙ ВЫПУСК

- ВЕЙВЛЕТ- ПРЕОБРАЗОВАНИЕ
- АРИФМЕТИЧЕСКОЕ КОДИРОВАНИЕ
- СЖАТИЕ И ВОССТАНОВЛЕНИЕ
- ФИЛЬТРАЦИЯ И ОБНАРУЖЕНИЕ
- ОЦЕНКА КАЧЕСТВА

ISSN 1684-2634



9 771684 263005 >

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Научно-технический журнал
№ 3/2007

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:

107031 Москва, Рождественка, 6/9/20, стр. 1
Тел.: (495) 621-71-08, (495) 621-16-39
Факс: (495) 621-16-39
E-mail: nto.popov@mtu.-net.ru
tor@rgta.ryazan.ru
http://www.dspsa.ru

Для писем:

129090 Москва, а/я 48.

Издатель:

Российское научно-техническое общество
радиотехники, электроники и связи им. А.С. Попова
Ответственный редактор: П.И. Зудков
Компьютерная верстка: В.В. Андреехина
Дизайн: И.Е. Артюхина

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

Подписано в печать 30.10.07
Формат 60x90/8.

Гарнитура «Agial». Печать ризографическая.
Бумага офсетная. Печ.л. 6,5. Тираж 700 экз.

Заказ № 6303. 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:

Бехтин Ю.С.

**Вейвлет-кодирование зашумленных изображений
преследованием базиса по когерентным
структурам** 2

Казарян М.Л.

О корректности задач сжатия сигналов 9

Приоров А.Л., Саутов Е.Ю., Хрящев В.В.

Неэталонная оценка качества JPEG изображений . 15

Беляев Е.А., Тюрликов А.М.

**Оценка вероятности появления символа при
адаптивном двоичном арифметическом кодировании
в задачах сжатия видеoinформации** 20

Мачихин А.С.

**Автоматизированный синтез линейных фильтров
для восстановления смазанных изображений** 25

Макаров Д.Г.

**Цифровая обработка телевизионных
измерительных сигналов** 30

Костров Б.В., Конкин Ю.В.

**Алгоритмическое обеспечение системы автономной
коррекции погрешностей навигационной системы
маневренных летательных аппаратов** 37

Иртюга В.А.

**АСС - прогрессивный метод цифрового кодирования
звука** 41

Фадеев А.С. Кочегурова Е.А.

**Метод преобразования форматов музыкальной
информации** 46

Милов А.Н.

**Оптимизированные алгоритмы преобразований
Харра для отечественной платформы DSP
«ELCORE» и их применение в задачах графической
обработки данных** 52

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

ВЕЙВЛЕТ-КОДИРОВАНИЕ ЗАШУМЛЕННЫХ ИЗОБРАЖЕНИЙ ПРЕСЛЕДОВАНИЕМ БАЗИСА ПО КОГЕРЕНТНЫМ СТРУКТУРАМ

Бехтин Ю.С.

Введение

Специфика формирования некоторых классов изображений оптико-электронными приборами обуславливает наличие искажений. Очевидным (но не лучшим) решением задачи сжатия таких зашумленных изображений является применение предварительной фильтрации. Однако из-за несогласованности алгоритмов фильтрации с последующим кодированием возможны ситуации, когда ошибки квантования будут превалировать над остаточным шумом после фильтрации. Следовательно, для согласования фильтрации и компрессии необходимо использовать единый механизм обработки изображения. Такой механизм можно построить на основе пороговой (грубой или мягкой) вейвлет-обработки сигналов (*hard thresholding* или *soft thresholding*) [1, 2], если величину нулевой, или «мертвой» зоны (*dead zone*) кодера (увеличенный интервал квантования около нуля [2, 3]) сделать адаптивной к интенсивности шума. Данный подход применяется в работах [4 – 7], где рассматриваются варианты выбора порогов и интервалов квантования, приводящие к субоптимальным оценкам оригинального изображения. Метод поиска оптимальной величины порога и оптимального интервала квантования, исходя из общего критерия качества, предложен автором в [8].

Вышеуказанные работы используют один базис для вейвлет-преобразования. Однако разные базисы по-разному аппроксимируют сигнал, поэтому выбор наилучшего в смысле некоторого критерия базиса из известной библиотеки базисов обеспечивает эффективное решение поставленной задачи. В работе [9] в качестве критерия отбора базисов используется принцип «минимума длины описания» (*minimum description length*), согласно которому наилучшим является базис, обеспечивающий минимальную суммарную длину кода для значимых вейвлет-коэффициентов, интервалов и уровней квантования, а также названия базиса. Однако данный подход трудно применить на практике, поскольку полученная степень сжатия не гарантирует выполнение ограничений на скорость кодирования (квоту бит). В другой работе [10] базис отбирается исходя из верхней границы минимального риска, который обеспечивает эффективное шумоподавление, но не сжатие. Для целей восстановления искаженных звукозаписей (одномерных сигналов) в работе [11] был разработан

Рассматривается решение задачи вейвлет-компрессии зашумленного изображения путем выбора последовательности «лучших» базисов из библиотеки по предложенной функции стоимости. На каждой итерации в оптимальном базисе алгоритм выделяет в зашумленном изображении когерентные структуры. Алгоритм минимизирует среднеквадратическую погрешность восстановления оригинала с учетом ошибок квантования при заданной квоте бит. Приводятся результаты моделирования, подтверждающие эффективность предложенного алгоритма работы кодера по сравнению с другими методами компрессии изображений, в частности с SPIHT.

метод преследования базиса по когерентным структурам, то есть путем отбора тех частей сигнала, которые сильно коррелированы с базисом по сравнению с белым шумом. В статье данный метод распространяется на случай сжатия зашумленных изображений (двумерных сигналов), при этом одновременно решается задача фильтрации и квантования при заданной скорости кодирования (квоте бит).

Теоретические аспекты совмещения фильтрации и квантования.

Многомасштабный анализ позволяет получить декомпозицию зашумленного изображения

$$Y = X + Z, \quad (1)$$

где Y – наблюдаемое изображение, X – неизвестный оригинал, Z – гауссовский шум с нулевым средним, при заданном числе уровней Q в виде:

$$W_Y = \mathbf{W}Y = \mathbf{W}(X + Z) = \mathbf{W}X + \mathbf{W}Z = W_X + W_\xi, \quad (2)$$

где \mathbf{W} – символ прямого вейвлет-преобразования [1, 2]. Таким образом, требуется выполнить сжатие данных зашумленного изображения (1), кодируя вейвлет-коэффициенты (2) таким образом, чтобы ошибка восстановления была минимальной в смысле среднего квадрата евклидовой нормы:

$$E \left\{ \|W_X - \hat{W}_X\|^2 \right\} \rightarrow \min, \quad (3)$$

где \hat{W}_X – вейвлет-коэффициенты восстановленного после компрессии изображения \hat{X} . В работах [1, 3] показано, что среднеквадратическая ошибка (СКО) восстановления (3), которая вычисляется в области вейвлет-преобразования (трансформанты), эквивалентна СКО $E \left\{ \|X - \hat{X}\|^2 \right\}$ в пространственной области.

Оценки вейвлет-коэффициентов \hat{W}_X для сепаративных вейвлет-базисов могут быть найдены с помощью нелинейной аппроксимации, которая строит-

ся по M первым вейвлет-коэффициентам наибольшей амплитуды. Для того чтобы выполнить нелинейную аппроксимацию, необходимо произвести сортировку всех I вейвлет-коэффициентов по абсолютной амплитуде:

$$|w_{Y_k}| \geq |w_{Y_{k+1}}|, \forall k \in [1, \dots, I]. \quad (4)$$

Сортированные вейвлет-коэффициенты имеют быстрый спад, причем, как показали проведенные исследования, в одном базисе у изображений одного класса поведение кривых спада практически одинаково. Для подбора значения M можно использовать также аппроксимацию кривой спада. Ф.Фальзон и С.Малла в работе [12] определили эту аппроксимацию как $|w_{Y_k}| \sim Ck^{-\gamma}$, где $C > 0$ и $\gamma > \frac{1}{2}$ – некоторые константы при условии, что существуют две другие константы $A > 0, B > 0$, не зависящие от C, γ и I , такие что

$$ACK^{-\gamma} \leq |w_{Y_k}| \leq BCK^{-\gamma}, \forall k \in [1, \dots, I].$$

При действии шума, как показали проведенные эксперименты, удастся подобрать показатель степени γ , при котором ошибка аппроксимации кривой спада всех вейвлет-коэффициентов относительно невелика. Рис.1 иллюстрирует подбор аппроксимации кривой спада с показателем степени $\gamma=0,53$ для зашумленного изображения Lena ($\sigma_{ш}=30$), показанного на рис.2.

Разные базисы обеспечивают и разные кривые спада вейвлет-коэффициентов. На рис.3 приведены графики части нормированных упорядоченных вейвлет-коэффициентов изображения Lena, полученные соответственно с помощью вейвлетов CDF 9.7 (Cohen-Daubeshies-Feauveau) и db4 (Daubeshies 4) при трех уровнях декомпозиции ($Q=3$). Нормировка необходима потому, что величина первого сортированного вейвлет-коэффициента $|w_{Y_1}|$, принадлежащего НЧ субполосе декомпозиции, изменяется при выборе базиса. Для зашумленного изображения Lena (рис. 2) при базисах CDF 9.7 и db4 абсолютная разность между первыми вейвлет-коэффициентами составила 504,3 (для db7 $|w_{Y_1}|=6955,8$, а для CDF9.7 $|w_{Y_1}|=6451,5$). Из рис.3 видно, что основные отличия между кривыми проявляются в начальной области быстрого спада сортированных вейвлет-коэффициентов. При этом нормированные кривые идут параллельно друг другу, не пересекаясь.

Нелинейная аппроксимация является основой для получения оценок вейвлет-коэффициентов $W_{\hat{X}}$, поскольку отбрасываемые вейвлет-коэффициенты малой амплитуды соответствуют высокочастотным субполосам, где сосредотачивается энергия шума.

Сжатие изображения после вейвлет-преобразования также сопровождается отбрасыванием части вейвлет-коэффициентов малой амплитуды из-за нулевой зоны. Таким образом, выбор числа M обеспечивает компромисс между степенью сжатия и эффективностью шумоподавления. Квантованные, или значимые вейвлет-коэффициенты (не попавшие в нулевую зону) кодируются, расходуя заданный пользователем бюджет бит R_C . Выбор базиса и числа M значимых вейвлет-коэффициентов влияет на эффективность использования квоты бит R_C .

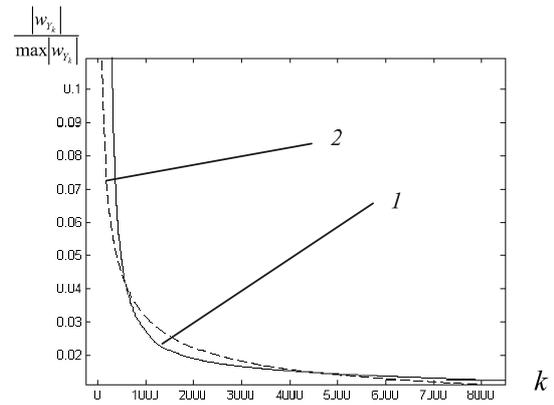


Рис.1. Фрагменты нормированной кривой спада сортированных вейвлет-коэффициентов зашумленного изображения Lena (1) и ее аппроксимации (2) при $\gamma=0,53$; вейвлет db7.



Рис.2. Искаженное изображение Lena при $\sigma_{ш}=30$.

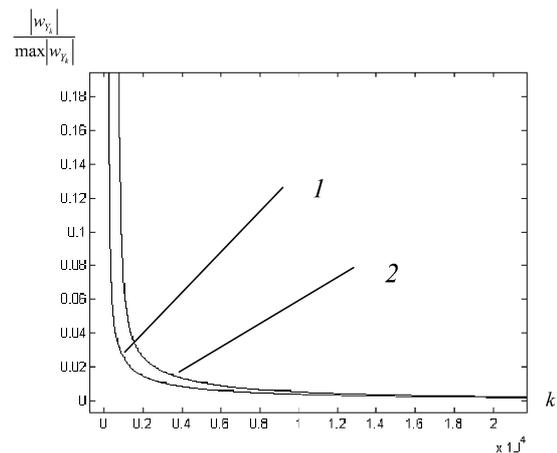


Рис.3. Кривые спада части вейвлет-коэффициентов изображения Lena при двух базисах CDF 9.7 (1) и db4 (2), $Q=3$.

Различают два варианта кодирования и, следовательно, два варианта расходования полного бюджета бит R_C [1, 3]. При высоких скоростях кодирования (свыше 1 бит/пиксель, $R_C > I$) принимается гипотеза о равномерном квантовании с высоким разрешением. Тогда можно доказать, что при кодировании значимых вейвлет-коэффициентов с помощью двоичных слов постоянной длины дисперсия ошибки квантования составит:

$$\sigma_{кв}^2 = \frac{C^2}{3} (1 - M^{-\gamma})^2 2^{-2\bar{R}_C}, \text{ где } \bar{R}_C = R_C / M. \quad (5)$$

Действительно, при равномерном квантовании с высоким разрешением дисперсия ошибки квантования есть $\sigma_{кв}^2 = \frac{\Delta^2}{12}$, где Δ – интервал квантования.

Число интервалов (бинов) квантования K задается квотой бит $\bar{R}_C = \log_2 K$. Поскольку диапазон изменения значимых вейвлет-коэффициентов составит $2(|w_{Y_1}| - |w_{Y_M}|)$, то интервал квантования равен $\Delta = 2 \frac{C}{K} (1 - M^{-\gamma})$, откуда выводится (5).

При малых скоростях кодирования (меньше 1 бит/пиксель, $R_C \leq I$) M квантованных значимых вейвлет-коэффициентов зашумленного изображения кодируются с помощью двоичных слов переменной длины. Дисперсия ошибки квантования определяется аналогичным образом, но для квантования с низким разрешением, поскольку число M сильно ограничено:

$$\sigma_{кв}^2 = C^2 (1 - M^{-\gamma})^2 2^{-2\bar{R}_1}, \quad (6)$$

где $\bar{R}_1 = (R_C - R_0) / M$.

Здесь R_0 – число бит, необходимое для кодирования карты существенности (координат) вейвлет-коэффициентов. Координаты квантованных вейвлет-коэффициентов $\langle w_k \rangle$, $\forall k \in [1, \dots, I]$, где скобки $\langle \cdot \rangle$ обозначают операцию квантования, записываются в виде двоичной карты:

$$s(k) = \begin{cases} 0, & \text{если } \langle w_k \rangle = 0; \\ 1, & \text{если } \langle w_k \rangle \neq 0. \end{cases} \quad (7)$$

Для числа M ($M \geq 1$) значимых вейвлет-коэффициентов относительные количества 0 и 1 в карте существенности соответственно равны $p_0 = (I - M) / I = 1 - \varepsilon$ и $p_1 = M / I = \varepsilon$. При адаптивном кодировании вначале находится число M , которое кодируется числом бит, равным $\log_2 I$. Затем вычисляется величина ε , и полный расход квоты бит на кодирование карты существенности, таким образом, составит:

$$R_0 = \log_2 I - I[(1 - \varepsilon) \log_2 (1 - \varepsilon) + \varepsilon \log_2 \varepsilon]. \quad (8)$$

Для кодирования типа базиса также требуется некоторое число бит R_B из заданной квоты R_C . Тогда целесообразно использовать распределение вероятности символов, описывающих базисы, чтобы код с переменной длиной уменьшил среднее значение расхода бит R_B и слабо влиял на дисперсию ошибки квантования (5) и (6).

Для вычисления полной ошибки, вызванной нелинейной аппроксимацией и квантованием, вначале рассматривается сумма квадратов отклонений оценок вейвлет-коэффициентов

$$w_{\bar{X}_k} = \begin{cases} \langle w_{Y_k} \rangle, & \text{если } |w_{Y_k}| \geq |w_{Y_M}|, \\ 0, & \text{если } |w_{Y_k}| < |w_{Y_M}|, \end{cases} \quad \forall k \in [1, \dots, I], \quad (9)$$

от их истинных значений:

$$\sum_{i=1}^I (w_{X_i} - w_{\bar{X}_i})^2 = \sum_{i=1}^I w_{Y_i}^2 - \sum_{i=1}^M w_{Y_i}^2 + \sum_{i=1}^I w_{\xi_i}^2 + \sum_{i=1}^M \sigma_{\hat{e}i}^2. \quad (10)$$

Для удобства вводятся следующие обозначения:

$$\sigma^2 = \frac{1}{I} \sum_{i=1}^I w_{Y_i}^2 - \text{средняя сумма квадратов всех вейвлет-коэффициентов зашумленного изображения};$$

$$\tilde{\sigma}^2 = \frac{1}{M} \sum_{i=1}^M w_{Y_i}^2 - \text{средняя сумма квадратов значимых вейвлет-коэффициентов};$$

$$\sigma_{W_{\xi}}^2 = \frac{1}{I} \sum_{i=1}^I w_{\xi_i}^2 - \text{дисперсия вейвлет-коэффициентов шума с нулевым средним};$$

$$\sigma_{кв}^2 = \frac{1}{M} \sum_{i=1}^M \sigma_{квi}^2 - \text{выборочная дисперсия ошибки квантования значимых вейвлет-коэффициентов.}$$

Тогда среднее квадрата нормы (3) с учетом (10) и условных обозначений принимает вид:

$$E\{\|W_X - W_{\bar{X}}\|^2\} = \frac{1}{I} \sum_{i=1}^I (w_{X_i} - w_{\bar{X}_i})^2 = \sigma^2 - \varepsilon \tilde{\sigma}^2 + \sigma_{W_{\xi}}^2 + \varepsilon \sigma_{\hat{e}a}^2. \quad (11)$$

Уравнение (11) показывает, что для качественного шумоподавления при сжатии (кодировании) зашумленного изображения желательно иметь как можно большую среднюю сумму квадратов значимых вейвлет-коэффициентов и как можно меньшую дисперсию ошибки квантования. Данные требования являются противоречивыми, поскольку с увеличением числа M значимых вейвлет-коэффициентов одновременно увеличивается и ошибка квантования. Тогда выбор базиса необходимо проводить при фиксированных значениях M , поскольку базис определяет характер кривых спада упорядоченных вейвлет-коэффициентов.

Будем полагать, что базис β^a лучше, чем базис β^{ω} при кодировании зашумленных изображений, если при всех $\varepsilon \in [I^{-1}, 1]$ выполняется неравенство

$$\mathcal{E}^{(a)}(\tilde{\sigma}_a^2(\mathcal{E}^{(a)}) - \sigma_{\hat{e}a}^2(\mathcal{E}^{(a)})) \geq \mathcal{E}^{(\omega)}(\tilde{\sigma}_\omega^2(\mathcal{E}^{(\omega)}) - \sigma_{\hat{e}\omega}^2(\mathcal{E}^{(\omega)})). \quad (12)$$

Чтобы избавиться от неопределенности при парном сравнении базисов, необходимо иметь некоторый критерий, однозначно показывающий преимущество одного базиса перед другим. Для неискаженных сигналов (изображений) в теории вейвлет-преобразования рассматриваются критерии, построенные на вогнутых функциях стоимости Шура [1].

Из выражений (5), (6) следует, что с ростом показателя степени γ возрастает дисперсия ошибки квантования и одновременно убывает сумма квадратов значимых вейвлет-коэффициентов. Проведенные исследова-

ния показали, что характер изменения величины $u\sigma^2 = \varepsilon(\tilde{\sigma}^2 - \sigma_{\hat{e}\hat{a}}^2)$ при низких и высоких скоростях кодирования примерно одинаков. На рис.4 показано изменение теоретической оценки величины $\varepsilon(\tilde{\sigma}^2 - \sigma_{\hat{e}\hat{a}}^2)$ при разных показателях степени γ . Из рис.4 видно, что при выборе базисов при фиксированном значении M выигрывает базис с наименьшим значением γ (кривая 1). Расчеты сделаны без учета расходов R_B на кодирование названия базиса.

Необходимо отметить, что фактические дисперсии ошибок квантования никогда не превосходят верхних границ (5), (6). Следовательно, если величина $\tilde{\sigma}^2$ доминирует над теоретической величиной $\sigma_{\hat{e}\hat{a}}^2$ (5) или (6), то такое соотношение сохранится и при расчете фактических дисперсий ошибок квантования. Для зашумленного изображения, декомпозированного при двух разных базисах, величина $\tilde{\sigma}^2$ отражает точность оценки оригинала (нелинейной аппроксимации), следовательно, наибольшая из двух величин $\varepsilon(\tilde{\sigma}^2 - \sigma_{\hat{e}\hat{a}}^2)$ соответствует наилучшему базису.

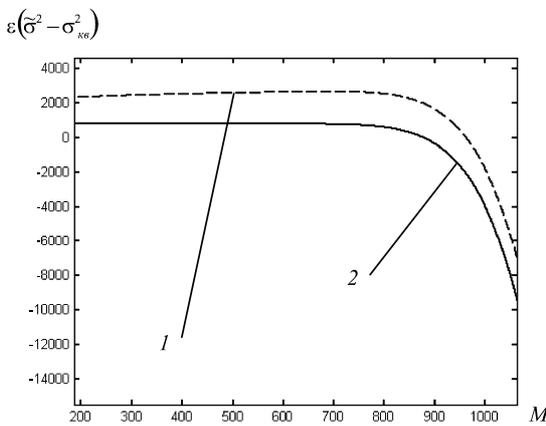


Рис.4. Поведение величины $\varepsilon(\tilde{\sigma}^2 - \sigma_{\hat{e}\hat{a}}^2)$ при $\gamma=0,55$ (1) и $\gamma=1$ (2), $l=20000$; $R_c = 0,2I$ (0,2 бит/пиксель); $C=5000$.

Таким образом, можно доказать, что базис β^α лучше, чем базис β^ω при кодировании зашумленных изображений тогда и только тогда, когда при всех вогнутых функциях $\Phi(u)$, $u>0$, справедливо неравенство

$$\sum_{i=1}^l \Phi\left(\frac{1}{\sigma^2 I} \left(|W_{Y_i}^{(\alpha)}|^2 - \sigma_{\hat{e}\hat{a}i}^{(\alpha)2} \right)\right) \leq \sum_{i=1}^l \Phi\left(\frac{1}{\sigma^2 I} \left(|W_{Y_i}^{(\omega)}|^2 - \sigma_{\hat{e}\hat{a}i}^{(\omega)2} \right)\right) \tag{13}$$

Из практических соображений необходимо использовать какую-либо одну вогнутую функцию $\Phi(u)$. На основании (13) в рассмотрение вводится функция стоимости:

$$C(\beta^l) = \sum_{i=1}^l \Phi\left(\frac{1}{\sigma^2 I} \left(|W_{Y_i}^{(l)}|^2 - \sigma_{\hat{e}\hat{a}i}^{(l)2} \right)\right), \tag{14}$$

где $l=1, \dots, L$, L – число базисов в библиотеке Λ :

$$\Lambda = \{\beta^1, \beta^2, \dots, \beta^L\} \tag{15}$$

Тогда наилучший базис β^α минимизирует стоимость аппроксимации [1]:

$$\tilde{N}(\beta^\alpha) = \min_l \tilde{N}(\beta^l) \tag{16}$$

Выбор наилучшего базиса зависит от вида вогнутой функции $\Phi(u)$. В качестве функций стоимости в статье используется энтропия вида

$$\Phi(u) = -u \ln u, \quad u \geq 0. \tag{17}$$

Библиотека базисов. Библиотека базисов Λ представляет собой коллекцию L ортонормированных базисов (15). Состав библиотеки не является постоянным, а может изменяться в зависимости от класса кодируемого зашумленного изображения. В библиотеку могут, например, входить: стандартный базис евклидова пространства \mathbf{R}^l , базисы Хаара-Уолша, различные вейвлет-базисы семейства Добеши (Daubeshis), биортонормальные вейвлет-базисы (семейство bior), в том числе базисы, эффективные для сжатия, CDF 9.7, LeGall 5.3 и т.п. На их основе строятся соответствующие вейвлет-пакетные базисы и их менее асимметричные версии – куафлеты (coiflet). В состав библиотеки можно добавить локальные тригонометрические базисы. Такая коллекция базисов является достаточно адаптируемой для эффективного представления сигналов любой размерности, в том числе 2D сигналов (изображений).

Вычислительная сложность применения того или иного базиса из библиотеки определяется размером изображения I и составляет:

- для процессов декомпозиции (анализа) и реконструкции (синтеза) по ортонормированным 2D базисам, включая 2D вейвлеты – $O(I)$ операций;
- при нахождении лучшего базиса при 2D вейвлет-пакетной обработке – $O(\text{Log}_4 I)$ операций;
- при нахождении лучшего 2D локально-тригонометрического базиса - $O(I[\log_4 I]^2)$ операций.

В библиотеку базисов также можно добавить базис, вычисляемый по преобразованию Карунена-Лозва, который дает минимальную энтропию [1, 3]. Однако вычислительная сложность такого преобразования равна $O(I^3)$ уже в 1D варианте из-за проблемы нахождения собственных чисел.

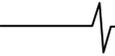
В результате неизвестный сигнал f может быть полностью представлен M элементами ($M < l$) базиса β^l , $l=1, \dots, L$:

$$f = \mathbf{B}^{(l)} W_M^{(l)}, \tag{18}$$

где $\mathbf{B}^{(l)} \in \mathbf{R}^l$ – ортогональная матрица, столбцы которой являются элементами базиса β^l ; а $W_M^{(l)}$, $l=1, \dots, L$, – матрица с M ненулевыми коэффициентами преобразования.

Таким образом, задача совместной фильтрации и сжатия данных искаженного изображения представляется как задача поиска модели из набора моделей, которые формируются базисами β^l , $l=1, \dots, L$.

Подбор числа значимых вейвлет-коэффициентов через когерентные структуры. Вейвлет-декомпозиция зашумленного изображения в библиотеке базисов является его рассмотрением в разных масштабах и определением наилучшего совпадения с выбираемой моделью (базисом). Шум представляет такие компоненты декомпозированного изображения, которые не имеют сильной корреляции ни с одним из базисов библиотеки. Тогда для фильтрации следует выделить когерентные структуры [1, 12], то есть те компоненты двумерного сигнала,



которые сильно коррелированы с базисом. Для сортированных по (4) вейвлет-коэффициентов формируется следующее решающее правило отделения когерентной структуры в базисе β^l , $l = 1, \dots, L$:

$$\rho(Y_k) = \frac{|w_{Y_k}^{(l)}|}{\|Y\|} > \rho_l, \quad (19)$$

где ρ_l - максимум нормированной корреляции шума в l вейвлет-коэффициентах зашумленного изображения. При любом M можно определить остаток Y_M , вычисляемый как разность между исходным зашумленным изображением и изображением, полученным после обратного вейвлет-преобразования (восстановления) над всеми когерентными структурами:

$$Y_M = Y - \sum_{k=1}^M \mathbf{W}^{-1} \{w_{Y_k}\} = \sum_{k=M+1}^L \mathbf{W}^{-1} \{w_{Y_k}\}, \quad (20)$$

где символом \mathbf{W}^{-1} обозначена операция вычисления обратного вейвлет-преобразования. Остаток Y_M является ортогональной проекцией Y на пространство размерности $l - M$. Остаток Y_M будет считаться шумом, если при сравнении нормированной корреляции этого остатка с базисом β^l , $l = 1, \dots, L$, с нормированной корреляцией шума ρ_{l-M} справедливо неравенство [1, 12]:

$$\rho^2(Y_M) = \frac{|w_{Y_M}|^2}{\sum_{k=M+1}^L |w_{Y_k}|^2} > \rho_{l-M}^2. \quad (21)$$

Оценка оригинального сигнала является тогда суммой M когерентных структур:

$$\hat{X} = \sum_{k=1}^M \mathbf{W}^{-1} \{w_{Y_k}\}. \quad (22)$$

Таким образом, из (21) следует, что удаление когерентных структур внешне похоже на грубую пороговую обработку вейвлет-коэффициентов (2) с величиной порога:

$$\tau = \rho_{l-M} \sqrt{\sum_{k=M+1}^L |w_{Y_k}|^2}. \quad (23)$$

Соотношения (21), (22) нуждаются в уточнении относительно ошибок квантования. Поскольку минимизация СКО восстановления зашумленного изображения после компрессии $E\{\|w_x - w_{\hat{x}}\|^2\}$ эквивалентна максимизации величины $\varepsilon(\tilde{\sigma}^2 - \sigma_{\text{ка}}^2)$, то с учетом (2), (21) и (23) выводится следующий алгоритм оценки числа значимых вейвлет-коэффициентов M по когерентным структурам:

$$f(M) = \varepsilon(\tilde{\sigma}^2 - \sigma_{\text{ка}}^2) \rightarrow \max_M, \quad (24)$$

$$\forall |w_{Y_i}|^2 \geq \rho_{l-M}^2 \sum_{i=M+1}^L |w_{Y_i}|^2 = \rho_{l-M}^2 I(\sigma^2 - \varepsilon \tilde{\sigma}^2). \quad (25)$$

Из соотношений (24), (25) видно, что алгоритм имеет итерационный характер. Пусть произведена сортировка всех l вейвлет-коэффициентов по убыванию их абсолютных значений. Тогда алгоритм содержит следующую последовательность шагов.

1. Инициализация алгоритма начинается с установки $M = l - 1$.

2. Вычисляется кумулятивная сумма квадратов вейвлет-коэффициентов и по (23) находится величина

порога при известной нормированной корреляции шума ρ_{l-M} .

3. Проверяется выполнение условия (25), которое дает новое, скорректированное значение числа M .

4. Проверяется достижение максимума целевой функции $f(M)$ по (24).

5. Если максимум достигнут, то останов алгоритма. Иначе переход к п.2.

Для гауссовского белого процесса Z размером l и с дисперсией σ_w^2 при больших значениях l существует вероятность, стремящаяся к 1, что для любого ортонормированного базиса выполняется [1]:

$$\frac{\max_{1 \leq k \leq l} |w_{\xi}|}{\|Z\|} \leq \frac{\sqrt{2 \ln l} \sigma_w}{\sqrt{l} \sigma_w} = \frac{\sqrt{2 \ln l}}{\sqrt{l}} = \rho_l. \quad (26)$$

Поскольку отделение когерентных структур производится с помощью механизма грубой пороговой обработки вейвлет-коэффициентов, то в результате работы алгоритма (24), (25) автоматически получаются оценки вейвлет-коэффициентов

$$w_{\hat{x}_i} = \begin{cases} w_{Y_i}, & i = \overline{1, M}, \\ 0, & i = \overline{M+1, l}. \end{cases}$$

Предложенный алгоритм может быть модифицирован для мягкой пороговой обработки:

$$w_{\hat{x}_i} = \begin{cases} w_{Y_i}, & i = \overline{1, t}, \\ \text{sign}(w_{Y_i}) \left(|w_{Y_i}| - \rho_{l-M} \sqrt{I(\sigma^2 - \varepsilon \tilde{\sigma}^2)} \right), & \\ i = \overline{t+1, M}, & \\ 0, & i = \overline{M+1, l}. \end{cases} \quad (27)$$

При этом целевая функция (24) изменяется и алгоритм поиска оценки M принимает вид:

$$f(M) = \varepsilon(\tilde{\sigma}^2 - \sigma_{\text{ка}}^2) - (\varepsilon - \theta) \rho_{l-M}^2 I(\sigma^2 - \varepsilon \tilde{\sigma}^2) \rightarrow \max_M, \quad (28)$$

$$\forall |w_{Y_i}|^2 \geq \rho_{l-M}^2 I(\sigma^2 - \varepsilon \tilde{\sigma}^2).$$

Здесь t - число вейвлет-коэффициентов НЧ субполосы (аппроксимации), которые обычно не подвергаются мягкой пороговой обработке [5], $\theta = t/l$.

Преследование базисов по когерентным структурам. Изображения являются сигналами сложной структуры, для которых практически невозможно подобрать один, «идеальный» базис. Кроме того, при действии шума одна часть изображения может оказаться когерентной структурой при одном базисе, а другая - иметь сильную корреляцию с другим базисом. При использовании выше-рассмотренных алгоритмов получения оценок значимых вейвлет-коэффициентов и их числа M (24)-(28) может оказаться, что для одного базиса β^α из библиотеки базисов Λ (15) остаток Y_M вида (20) содержит когерентную структуру для другого базиса β^ω , $\alpha, \omega \in [1, \dots, L]$. Таким образом, качество восстановления оригинального изображения по одному базису не будет высоким из-за потери информативной части изображения.

С учетом вышесказанного алгоритм обработки изображения на основе преследования базиса по когерентным структурам содержит следующую последовательность шагов.

1. Инициализация алгоритма начинается с установки $Y_M = Y$, $l = 1$.

2. Выполняется вейвлет-преобразование (2) над изображением Y по базису β^l и формируется массив вейвлет-коэффициентов W_Y .

3. Все вейвлет-коэффициенты w_{Y_i} , $i = \overline{1, I}$, сортируются по убыванию их абсолютных значений. Находится аппроксимация кривой спада сортированных вейвлет-коэффициентов (не обязательно, опционально).

4. В зависимости от величины R_C производится распределение квоты бит по одному из двух вариантов кодирования (на низких или высоких скоростях). Если аппроксимация кривой спада упорядоченных вейвлет-коэффициентов получена на предыдущем шаге, то рассчитывается теоретическая дисперсия ошибки квантования. Если аппроксимация кривой спада не определена, тогда вычисляется фактическая дисперсия ошибки квантования.

5. Составляется вогнутая функция стоимости Шура (14) и вычисляется ее значение.

6. Если библиотека базисов не исчерпана, то есть $\iota < L$, то $\iota = \iota + 1$ и переход к п.2. В противном случае переход к п.7.

7. Наилучший базис β^α является минимумом функции стоимости Шура $\tilde{N}(\beta^\alpha) = \min \tilde{N}(\beta^l)$. Вейвлет-коэффициенты, полученные по наилучшему базису, автоматически сохраняются в буферной памяти и извлекаются оттуда для следующего шага.

8. Производится отделение когерентных структур алгоритмами по поиску оценки M числа значимых вейвлет-коэффициентов, описанных в предыдущем разделе. Вычисляются оценки значимых вейвлет-коэффициентов каким-либо способом пороговой обработки (грубой или мягкой) с учетом квантования, которые сохраняются в буферной памяти.

9. Если $M=0$, то останов алгоритма. В противном случае выполняется обратное вейвлет-преобразование над отброшенными (незначимыми) вейвлет-коэффициентами, формируется остаток Y_M , корректируется квота бит и осуществляется переход к п.2.

Таким образом, в результате работы алгоритма в буферной памяти остаются несколько множеств квантованных оценок вейвлет-коэффициентов, соответствующих набору «лучших» базисов из библиотеки Λ . Эти коэффициенты кодируются либо кодом постоянной, либо переменной длины в зависимости от величины R_C .

Для восстановления изображения необходимо найти сумму изображений, полученных после обратного вейвлет-преобразования с соответствующими базисами над

множествами квантованных оценок вейвлет-коэффициентов после декодирования.

Результаты моделирования. Чтобы продемонстрировать эффективность алгоритма при ограниченности объема статьи, ниже приводятся результаты, полученные для относительно «слабой», но «быстрой» библиотеки базисов, построенной из 36 типов вейвлетов различных семейств, содержащихся в среде Matlab, а также вейвлетов CDF 9.7, LeGall 5.3 и Coiflet 22.14. Библиотеки базисов на основе пакетного разложения дают блестящие результаты, но требуют больше времени на вычисления. Моделирование зашумленного изображения Lena (рис.2) проводилось при разных интенсивностях шумов и низких скоростях кодирования для грубой пороговой обработки вейвлет-коэффициентов. Обработанные пакетным алгоритмом SPIHT [1-3] и предложенным алгоритмом изображения помещены на рис.5 и рис.6 соответственно. В результате экспериментов вычислялись СКО и пиковое отношение сигнал/шум (ПОСШ), изменения которых по итерациям соответственно показаны на рис.7 и рис.8. Последовательность выбора базиса иллюстрируется табл.1. Из табл.1 и рис.8, 9 видно, что сходимость достигается на 15-й итерации при $M=36$. Достигнутое СКО составило 38,45 против 270,27, рассчитанное при сжатии данного изображения пакетным алгоритмом SPIHT. Величина ПОСШ составила 32,4 дБ против 22,5 дБ у SPIHT. Из рис.5 и рис.6 видно, что искажения оригинала изображения, обработанного по SPIHT, значительно сильнее, чем искажения при обработке предложенным алгоритмом при скорости 0,2 бит/пиксель, на которой все кодеки, как показано в [5], обеспечивают наилучшее шумоподавление у искаженного изображения.

Кроме того, при разных интенсивностях шумов было проведено моделирование для ряда тестовых изображений и сравнение с другими методами кодирования. Результаты моделирования, помещенные в табл.2, показывают преимущество предложенного алгоритма по сравнению с известными алгоритмами кодирования по СКО и ПОСШ для изображений разных классов при 0,2 бит/пиксель. Из табл.2 видно, что для сильно текстурированных изображений (типа Lena и Goldhill) предложенный метод сжатия обеспечивает значительный выигрыш по СКО и ПОСШ по сравнению с методами JPEG и вейвлет-пакетным методом SPIHT. Однако для изображений с большим числом однородных текстур, например для изображения самолета F16, предложенный метод лишь немного выигрывает у пакетного SPIHT.



Рис.5. Обработанное пакетным SPIHT изображение Lena при $\sigma_{ш} = 30$, 0,2 бит/пиксель, $Q=3$.



Рис.6. Обработанное предложенным алгоритмом изображение Lena при $\sigma_{ш} = 30$; 0,2 бит/пиксель, $Q=3$.

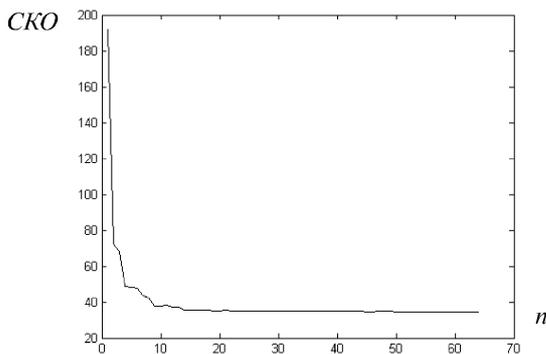


Рис.7. Изменение СКО по итерациям.

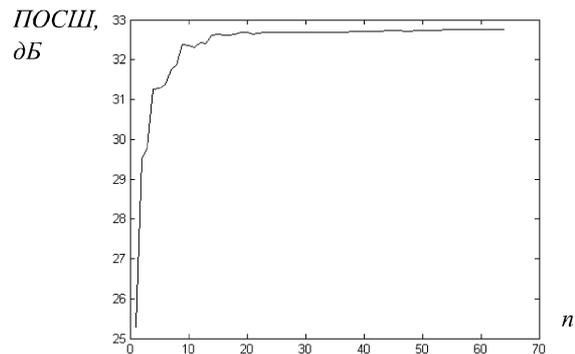


Рис.8. Изменение ПОСШ по итерациям.

Таблица 1.

Номера итераций	1, 5, 7	3, 4, 6, 8-11, 13, 14, 16-36	12	2	15
Название базиса	bior3.5	bior3.9	bior2.2	db6	db3

Таблица 2.

Метод сжатия	JPEG		SPIHT		Преследование базиса		
	СКО	ПОСШ, дБ	СКО	ПОСШ, дБ	СКО	ПОСШ, дБ	
«Лена» (Lena)							
$\sigma_{ш}$	10	96,52	24,56	85,43	27,63	16,78	32,85
	20	285,34	21,67	116,38	25,47	21,13	32,67
	30	390,56	18,32	270,27	22,52	38,45	32,41
«Золотой холм» (Goldhill)							
$\sigma_{ш}$	10	112,87	23,47	92,23	27,12	23,64	30,04
	20	312,45	19,34	135,58	24,12	68,34	28,57
	30	560,34	15,73	310,32	21,57	107,35	25,23
«F16g»							
$\sigma_{ш}$	10	87,33	26,87	26,57	31,93	15,34	32,8
	20	215,34	23,52	86,41	27,44	79,32	28,35
	30	456,24	17,88	176,35	25,67	84,23	25,34

Заключение

Данный алгоритм сжатия зашумленного изображения, как показали проведенные исследования, обладает достаточной эффективностью даже при относительно небольшой и «примитивной» библиотеке базисов. Наилучшие результаты следует ожидать при использовании вейвлет-пакетных и локально-косинусных словарей. Данное обстоятельство приводит к увеличению вычислительных и временных затрат, несмотря на быструю сходимость, продемонстрированную в примере. Таким образом, предложенный подход можно рекомендовать для приложений, связанных с архивацией зашумленных изображений различных классов в базах данных.

Литература

1. С.Малла. Вейвлеты в обработке сигналов. М.: Мир, 2005.
2. Р.Гонсалес, Р.Вудс. Цифровая обработка изображений. М.: Техносфера, 2006.
3. В.И.Воробьев, В.Г.Грибунин. Теория и практика вейвлет-преобразования. С.-Петербург, ВУС, 1999.
4. S.G.Chang, B.Yu, M.Vetterli, "Adaptive wavelet thresholding for image de-noising and compression", *IEEE Trans. on Image Processing*, Vol. 9, No. 9, 2000, pp.1532-1546.
5. O.-K. Al-Shaykh, R.M.Mercereau, "Lossy Compression of Noisy images", *IEEE Trans. on Image Processing*, Vol.7, Nr.12, 1998, pp.1641-1652.

6. M.L. Mittal, V.K.Singh, R. Krishnan, "Wavelet Transform based Technique for Speckle Noise Suppression and Data Compression for SAR Images". *Proc. of the Fifth International Symposium on Signal Processing and Applications*, 1999, pp. 781-784.
7. D.Wei, J.E.Odegard, H. Guo, M. Lang, C.S.Burrus, "Simultaneous Noise Reduction and SAR Image Data Compression Using best Wavelet Packet Basis", *Proceedings of International Conference on Image Processing*, Vol. 3, 1995, pp. 200-203.
8. Ю.С.Бехтин. Поиск оптимального распределения бит при сжатии данных зашумленных изображений на основе вейвлет-преобразования. *Цифровая обработка сигналов*, № 1, апрель, 2007. 10 с.
9. N. Saito, "Simultaneous Noise Suppression and Signal Compression using a Library of Orthonormal Bases and the Minimum Description Length Criterion." In *Wavelets in Geophysics*, E.Foufoula-Georgiou, P. Kumar, Eds. New York: Academic Press, 1994, pp. 299-324.
10. D.Donoho, I.Johnstone, "Ideal De-noising in an Orthonormal Basis chosen from a Library of Bases", *C.R. Acad. Sci. Paris, Série I*, 319, 1994. pp. 1317-1322.
11. J.Berger, R.Coifman, and M.Goldberg, "Removing Noise from Music using Local Trigonometric bases and Wavelet Packets", *J. Audio Eng. Sci.*, 42(10), 1994. pp. 808-818.
12. S.Mallat and F.Falzon, "Analysis of low bit rate image transform coding", *IEEE Trans. Signal Process.*, vol.46, no. 4, pp.1027-1042, Apr. 1998.

УДК 621.397.2

О КОРРЕКТНОСТИ ЗАДАЧИ СЖАТИЯ СИГНАЛОВ

Казарян М.Л.

Введение

Кодирование с использованием различных дискретных ортогональных преобразований являются эффективным методом сокращения избыточности изображений [1-3]. Сигнал подвергается унитарному преобразованию с дальнейшим отбором спектральных коэффициентов, используемых при решении задачи сжатия. Задачу восстановления сигнала можно рассматривать как обратную (некорректно поставленную) задачу [4]. Методика исследования некорректных задач достаточно известна [4-8]. А.Н. Тихонов предложил метод регуляризации, который является наиболее общим в настоящее время методом решения некорректных задач [4, 6, 7].

Поставленная задача $z=R(u)$ на паре метрических пространств (F,U) является корректно поставленной, если удовлетворяются следующие условия:

- 1) для всякого элемента $u \in U$ существует решение z из пространства F ;
- 2) решение определяется однозначно;
- 3) задача устойчива на пространствах (F,U) , т.е. для $\forall \varepsilon > 0 \exists \delta(\varepsilon) > 0$ так, что из неравенства $\rho_U(u_1, u_2) \leq \delta(\varepsilon)$ следует $\rho_F(z_1, z_2) \leq \varepsilon$.

При этом доказывается, что для дискретных преобразований Фурье и Уолша регуляризация спектра коэффициентами вида

$$\frac{1}{1 + \alpha k^{1+\varepsilon}}, \quad 0 \leq \alpha \leq 1, \quad \varepsilon > 0, \quad k = 1, 2, \dots, n-1.$$

позволяет произвести устойчивое суммирование рядов этих преобразований.

Найдена оценка погрешностей при сжатии сигналов с регуляризацией, получены оценки погрешностей для ДПФ (дискретного преобразования Фурье), преобразования Хартли, ДКП (дискретного косинусного преобразования), ДСП (дискретного синусного преобразования), ДПУА (дискретного преобразования Уолша – Адамара) с регуляризацией и без регуляризации. Проведено сравнение соответствующих оценок и найдены значения регуляризирующего параметра α для этих преобразований. Из данного сравнения следует, что восстановленный сигнал с применением регуляризации ближе к оригиналу, чем сигнал восстановленный без применения метода регуляризации.

Математическая постановка задачи

Пусть $x = (x_0, \dots, x_{n-1})$ исходный вектор данных размерности n , рассматриваемый как реализация некоторого случайного процесса с определенными свойствами и

Исследуется применение метода регуляризации Тихонова в задаче сжатия сигналов посредством дискретных ортогональных преобразований. Данная задача исследуется с точки зрения некорректно поставленной. Находятся параметры регуляризации для различных дискретных ортогональных преобразований (ДОП). В работе определяются матрицы выбора спектральных компонент для разных ДОП. Получены оценки погрешностей восстановления сжатого сигнала кодированного посредством ортогональных преобразований с регуляризацией и без регуляризации. Из данных оценок следует, что сигнал, восстановленный с применением метода регуляризации Тихонова, ближе к исходному сигналу нежели сигнал, восстановленный без применения метода регуляризации.

ковариационной матрицей \sum_x ; F – ортогональное преобразование, задаваемое невырожденной $n \times n$ - матрицей отсчетов базисных функций некоторой ортогональной системы $\{\varphi_k(t)\}_{k \geq 0}$:

$$F = \begin{bmatrix} \varphi_0(0) & \varphi_0(1) & \dots & \varphi_0(n-1) \\ \varphi_1(0) & \varphi_1(1) & \dots & \varphi_1(n-1) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{N-1}(0) & \varphi_{N-1}(1) & \dots & \varphi_{N-1}(n-1) \end{bmatrix}$$

F^{-1} – обратное преобразование; S – матрица выбора размерности $m \times n$ ранга m , $1 \leq m \leq n$; W – матрица восстановления размерности $n \times m$.

Задача состоит в выборе F, S, W таких, что

$$\rho(x, F^{-1} W S F x) \rightarrow \min, \quad (1)$$

где ρ – заданная метрика.

При решении более простой задачи зафиксируем базис F и положим $W = S^T$. Тогда в (1) минимизация производится только по S . В такой постановке это решение известно как задача отыскания оптимального метода зонного кодирования при сжатии данных посредством преобразования F [9]. Алгоритм зонного кодирования при сжатии данных описывается в следующем виде.

Пусть x – исходный вектор пространства (X, ρ) векторов размерности N . F – ортогональная матрица размерности $N \times N$.

1 шаг. Вектор – сигнал x подвергается преобразованию F :

$$y = Fx = (y_0, \dots, y_{n-1})^T.$$

2 шаг. Вектор – спектральных компонент y заменяется посредством «оператора выбора» S на меньший по размерности \hat{y} :

$$\hat{y} = Sy = (y_0, y_1, \dots, y_{m-1}),$$

который и подлежит передаче по каналам связи, хранению и т.д. Отметим, что матрица S размерности $m \times n$ имеет вид

$$S = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \end{pmatrix}$$

Величину $k = \frac{n}{m}$ назовем коэффициентом сжатия.

3 шаг. На приемной стороне полученный вектор \hat{y} дополняется до размерности n (все компоненты, кроме отобранных, полагаются равными нулю), т.е.

$$\tilde{y} = S^T \hat{y} = (y_0, y_1, \dots, y_{m-1}, 0, \dots, 0).$$

4 шаг. Вектор \tilde{y} подвергается преобразованию F^{-1} , т.е. определяется вектор $\hat{x} = F^{-1} \tilde{y}$, который восстанавливает исходный вектор данных с погрешностью

$$\varepsilon = \rho(x, \hat{x}) \quad (2)$$

Задача состоит в определении оптимального способа зонного кодирования, т.е. такого выбора заменяемых нулями компонент, который гарантирует при заданном k минимум

$$\varepsilon^* = \sup_{x \in (X, \rho)} \varepsilon.$$

Очевидно, что заменять нулями целесообразно те компоненты вектора y , которые малы по абсолютной величине и, следовательно, вносят меньший вклад при восстановлении исходного вектора. Кроме того, следует иметь в виду, что при построении программ реализации дискретных ортогональных преобразований (ДОП) используется внутренняя структура преобразования. Матрица F разбивается на однотипные блоки, благодаря чему возможно распараллеливание вычислений. Поэтому целесообразно заменять нулями спектральные компоненты, соответствующие целиком блоку (или нескольким блокам) матрицы F .

Более детально остановимся на процессе кодирования и декодирования двумерных сигналов (изображений) посредством ортогональных преобразований.

Поступающие на вход системы исходные изображения размерности $n \times n$ представим в виде их разверток, а именно $m = n^2$ элементным вектором, т.е. рассматриваем одномерный сигнал.

Процесс кодирования производится в два этапа.

1. Посредством преобразования F m -мерное пространство исходных векторов f отображается в m -мерное спектральное пространство векторов \hat{f} .

2. Переходом в k -мерное ($k < m$) пространство посредством оператора S понижается размерность вектора \hat{f} , т.е. осуществляется отбор наиболее информативных спектральных компонент вектора \hat{f} .

Отметим, что наиболее информативными понимаются те спектральные компоненты, передача которых позволяет в декодере получить оценку g исходного вектора f с минимально возможными искажениями, определяемыми выбранной метрикой $\rho(f, g)$.

Процесс декодирования производится также в два этапа:

1. Посредством оператора W осуществляется фильтрация зашумленного усеченного вектора \hat{f} и переход из k -мерного в m -мерное пространство спектральных компонент.

2. Посредством оператора F^{-1} m -мерное пространство спектральных компонент отображается в m -мерное пространство исходных векторов.

Приведенный выше анализ показывает, что в процессе кодирования и декодирования изображений посредством ортогональных преобразований, важное значение имеет решение задачи определения матрицы выбора для различных типов ортогональных дискретных преобразований.

Итак, в процессе сжатия и восстановления сигнала важное место занимает задача определения вида матрицы выбора для различных ортогональных преобразований. В соответствии с этим удобно ввести следующее обозначение:

$$N = r^n; \text{ вектор } y_s = (0, \dots, 0, y_{r^s}, y_{r^{s+1}}, \dots, 0, \dots, 0), \\ s = 0, \dots, n-1 \text{ называется } s\text{-й пачкой вектора } y; \\ X_\delta = \{x = (x_0, \dots, x_{N-1}) : \max_{1 \leq k \leq n} |x_{k-1} - x_k| \leq \delta\}. \quad (3)$$

где x_i - действительные числа, $\delta > 0$.

Перейдем к вопросу оценки спектров различных ортогональных преобразований.

Пусть $F = (f_{i,j})_{i,j=0}^{N-1}$ матрица дискретного преобразования Фурье, $x = (x_0, x_1, \dots, x_{N-1}) \in X_\delta$ - исходный сигнал. Компоненты спектрального вектора $y = Fx$ имеют вид

$$y_l = \frac{1}{N} \sum_{k=0}^{N-1} x_k \exp\left\{\frac{-2\pi i k l}{N}\right\}.$$

В работе [10] было доказано, что

$$\max_{x \in X_\delta} |y_l| \geq \frac{\delta}{2 \sin \frac{\pi l}{N}}, \quad l=1,2,\dots,n-1. \quad (4)$$

Дискретная функция $\sin \frac{\pi l}{N}$, $l=1,2,\dots,n-1$ имеет вид:

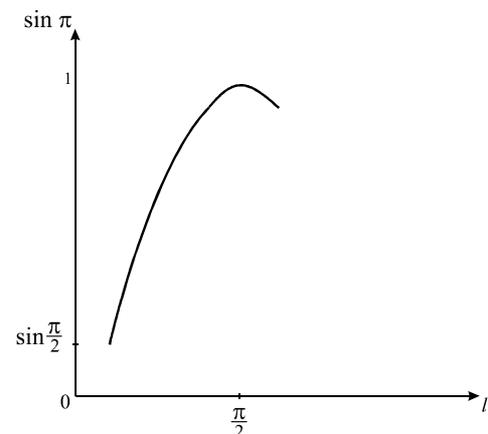


Рис.1. Поведение дискретной функции $\sin \frac{\pi l}{N}$, $l=1,2,\dots,n-1$.

Из рис.1 и формулы (4) заключаем, что центральные компоненты спектра $y = Fx$, $y = (y_0, \dots, y_{N-1})$ менее информативны. Следовательно, при сжатии данных рекомендуется использовать матрицу выбора следующего вида:

$$\begin{pmatrix} I & N_0 & N_0 \\ N_0 & N_0 & N_0 \\ N_0 & N_0 & I \end{pmatrix},$$

где N_0 - нулевые матрицы соответствующих порядков.

Например, пусть длина исходного вектора x равна $n=64$, а коэффициент сжатия равен 4, т.е. при сжатии необходимо сохранить 16 компонент спектрального вектора y . В этом случае матрица выбора имеет следующий вид

$$\begin{matrix} I_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 \\ 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 \\ 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 \\ 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 \\ 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 \\ 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 \\ 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 \\ 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & 0_8 & I_8 \end{matrix}$$

Пользуясь той же методикой и оценками, приведенными в [10], можно прийти к следующему утверждению.

При сжатии сигнала необходимо заменять нулями:

- центральные компоненты спектра для преобразования Фурье и Хартли;

- последние компоненты спектра для дискретного косинусного преобразования (ДКП), дискретного синусного преобразования (ДСП), дискретного преобразования Уолша-Пэли (ДПУП), дискретного преобразования Уолша (ДПУ);

- компоненты спектра с номерами

$$2^{r-1} \cdot (2^i - 1), \quad r = 1, 2, \dots, k; \quad i = 1, 2, \dots, 2^{n-r}$$

для дискретного преобразования Уолша – Адамара (ДПУА), где 2^n - длина сигнала, 2^k - коэффициент сжатия.

Оценим погрешности восстановления сжатого сигнала кодированного посредством ДОП с регуляризацией и без регуляризации.

Пусть F – матрица дискретного преобразования Фурье. Если x – исходный сигнал, а \bar{x} - восстановленный после сжатия сигнал, то, согласно равенства Парсевала и линейности преобразования Фурье, имеет место равенство:

$$\frac{1}{n} \cdot \rho_{l_2}(x, F^{-1}WSFx) = \rho_{l_2}(Fx, WSFx) = \rho_{l_2}(y, \bar{y}),$$

где ρ_{l_2} - среднеквадратическая метрика.

Следовательно, для оценки погрешности при сжатии сигналов достаточно иметь оценку вида

$$\varepsilon(n, k) = \sup_{x \in X_\delta} \rho_{l_2}(y, \bar{y}). \quad (5)$$

Если спектр сигнала рассматривать как исходный сигнал $y = Fx$, то после сжатия, передачи и экстраполяции этот сигнал отличается от исходного сигнала (спектра), и это отличие можно с некоторой точностью устранить при помощи регуляризации исходного сигнала. Ошибка восстановления с регуляризацией имеет вид:

$$\varepsilon(n, k, \alpha) = \sup_{x \in X_\delta} \rho_{l_2}(y_r, \bar{y}_r), \quad (6)$$

где k – заданный коэффициент, y_r – исходный регуляризованный сигнал, \bar{y}_r - соответствующий сжатый сигнал, α -параметр регуляризации.

Возникает задача нахождения оценки погрешности восстановления исходного сигнала x без регуляризации и с регуляризацией, т.е. сравнение (5) с равенством

$$\varepsilon_r(n, k, \alpha) = \sup_{x \in X_\delta} \rho_{l_2}(x, x_r), \quad (7)$$

где x_r - восстановленный сигнал, что согласно (3) эквивалентно задаче определения $\rho_{l_2}(y, y_r)$.

Последующее утверждение посвящено получению оценок вида (5) – (7) для различных дискретных ортогональных преобразований и их сравнению.

Справедливы следующие оценки для:

- дискретного преобразования Фурье (ДПФ)

$$\varepsilon(n, k) \leq \frac{\delta}{2} \cdot \sqrt{1 + \frac{4}{\pi^2}} \cdot \sqrt{n - \frac{n}{k}},$$

$$\varepsilon_r(n, k, \alpha) \leq \frac{\alpha \cdot (n-1)^{1+\varepsilon} \cdot \delta}{2 \cdot (1+\alpha)} \cdot \sqrt{1 + \frac{4}{\pi^2}} \cdot \sqrt{n - \frac{n}{k}},$$

$$\alpha \leq \frac{1}{(n-1)^{1+\varepsilon} - 1};$$

- дискретного преобразования Хартли (ДПХ)

$$\varepsilon(n, k) \leq \frac{\delta}{2} \cdot \sqrt{n - \frac{n}{k}},$$

$$\varepsilon_r(n, k, \alpha) \leq \frac{\alpha \cdot (n-1)^{1+\varepsilon} \cdot \delta}{2 \cdot (1+\alpha)} \cdot \sqrt{n - \frac{n}{k}},$$

$$\alpha \leq \frac{1}{(n-1)^{1+\varepsilon} - 1};$$

- дискретного косинусного преобразования (ДКП)

$$\varepsilon(n, k) \leq \frac{\delta \cdot \sqrt{2 \cdot n}}{\pi} \cdot \sqrt{n - \frac{n}{k}},$$

$$\varepsilon_r(n, k, \alpha) \leq \frac{\alpha \cdot \sqrt{2 \cdot n} \cdot \left(\frac{n}{k} + 1\right)^{1+\varepsilon} \cdot \delta}{\pi \cdot (1+\alpha)} \cdot \sqrt{n - \frac{n}{k}},$$

$$\alpha \leq \frac{1}{\left(\frac{n}{k} + 1\right)^{1+\varepsilon} - 1};$$

- дискретного синусного преобразования (ДСП)

$$\varepsilon(n, k) \leq \frac{\delta \cdot \sqrt{n+1}}{\sqrt{2}} \cdot \sqrt{n - \frac{n}{k}},$$

$$\varepsilon_r(n, k, \alpha) \leq \frac{\alpha \cdot \sqrt{2 \cdot n} \cdot \left(\frac{n}{k} - 1\right)^{1+\varepsilon} \cdot \delta}{\pi \cdot \sqrt{2} \cdot (1+\alpha)} \cdot \sqrt{n - \frac{n}{k}},$$

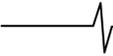
$$\alpha \leq \frac{1}{\left(\frac{n}{k} - 1\right)^{1+\varepsilon} - \pi};$$

- дискретного преобразования Уолша-Адамара (ДПУА), Уолша – Пэли (ДПУП), Уолша (ДПУ) (здесь коэффициент сжатия равен 2^k)

$$\varepsilon(n, k) \leq \frac{\delta \cdot \sqrt{2^n} \cdot 2^t}{\pi} \cdot \sqrt{2^n - 2^{n-t}},$$

$$\varepsilon_r(n, k, \alpha) \leq \frac{\alpha \cdot 2^{(n-t) \cdot (1+\varepsilon) - 1} \cdot \sqrt{2^n} \cdot 2^t \cdot \delta}{(1+\alpha)} \cdot \sqrt{2^n - 2^{n-t}},$$

$$\alpha \leq \frac{1}{2^{(n-t) \cdot (1+\varepsilon) + 1} - 1}.$$



Для доказательства последних утверждений докажем следующую теорему 1.

Теорема 1. Если \hat{y} -вектор, который получается из вектора $y = Fx$ заменой нулями $2m$ центральных компонент ($m < \frac{n}{2}$), т.е. коэффициент сжатия имеет вид

$k = \frac{n}{n-2m}$.., то справедливо следующее соотношение

$$\frac{\delta}{2} \cdot \sqrt{n - \frac{n}{k}} \leq \varepsilon(n, k) \leq \frac{\delta}{2} \cdot \sqrt{1 + \frac{4}{\pi^2}} \cdot \sqrt{n - \frac{n}{k}}. \quad (7)$$

Доказательство. Согласно теореме [10], для $j = 1, 2, \dots, n-1$ имеем

$$\frac{\delta}{2 \cdot \sin \frac{\pi \cdot j}{n}} \leq \max_{x \in X_\delta} |y_j| \leq \frac{\delta \cdot \sqrt{1 + \frac{4}{\pi^2}}}{2 \cdot \sin \frac{\pi \cdot j}{n}}. \quad (8)$$

Обозначим

$$y_1(j) = \frac{\delta}{2 \cdot \sin \frac{\pi \cdot j}{n}}, \quad y_2(j) = \frac{\delta \cdot \sqrt{1 + \frac{4}{\pi^2}}}{2 \cdot \sin \frac{\pi \cdot j}{n}}. \quad (9)$$

Очевидно, что и для вектора \hat{y} также выполняется соотношение (8). Следовательно, используя равенство Парсеваля, из (8) находим

$$\left[\sum_{j=\frac{n}{2}-m}^{\frac{n}{2}+m-1} y_1^2(j) \right]^{\frac{1}{2}} \leq \varepsilon(n, k) \leq \left[\sum_{j=\frac{n}{2}-m}^{\frac{n}{2}+m-1} y_2^2(j) \right]^{\frac{1}{2}}. \quad (10)$$

Сначала оценим выражение в правой части (10). Из формул (8), (9) имеем

$$\left[\sum_{j=\frac{n}{2}-m}^{\frac{n}{2}+m-1} y_2^2(j) \right]^{\frac{1}{2}} = \frac{\delta \cdot \sqrt{1 + \frac{4}{\pi^2}}}{2} \cdot \left[\sum_{j=\frac{n}{2}-m}^{\frac{n}{2}+m-1} \frac{1}{\sin^2 \frac{\pi \cdot j}{n}} \right]^{\frac{1}{2}} \quad (11)$$

Так как $\sin\left(\frac{\pi}{2} \pm \frac{\pi \cdot m}{n}\right) = \cos \frac{\pi \cdot m}{n}$, то при $j \neq \frac{n}{2}$ находим

$$\begin{aligned} \sum_{j=\frac{n}{2}-m}^{\frac{n}{2}+m-1} \frac{1}{\sin^2 \frac{\pi \cdot j}{n}} &= \frac{1}{\sin^2\left(\frac{\pi}{2} - \frac{\pi \cdot m}{n}\right)} + \frac{1}{\sin^2\left(\frac{\pi}{2} - \frac{\pi \cdot (m-1)}{n}\right)} + \dots \\ &+ \frac{1}{\sin^2\left(\frac{\pi}{2} - \frac{\pi}{n}\right)} + \frac{1}{\sin^2\left(\frac{\pi}{2} + \frac{\pi}{n}\right)} + \dots + \frac{1}{\sin^2\left(\frac{\pi}{2} + \frac{\pi \cdot (m-1)}{n}\right)} = (12) \\ &= 2 \cdot \sum_{j=1}^m \frac{1}{\cos^2 \frac{\pi \cdot j}{n}} - \frac{1}{\cos^2 \frac{\pi \cdot m}{n}}. \end{aligned}$$

В силу монотонности на интервале $\left[0, \frac{\pi}{2}\right]$ функции $\frac{1}{\cos^2 x}$ находим

$$2 \cdot \sum_{j=1}^m \frac{1}{\cos^2 \frac{\pi \cdot j}{n}} \leq \frac{2 \cdot n}{\pi} \cdot \int_{\frac{\pi}{n}}^{\frac{\pi \cdot (m+1)}{n}} \frac{dx}{\cos^2 x} = \frac{2 \cdot n}{\pi} \cdot \left(\operatorname{tg} \frac{\pi \cdot (m+1)}{n} - \operatorname{tg} \frac{\pi}{n} \right). \quad (13)$$

Подставляя значения (13) в (11), получим

$$\varepsilon(n, k) \leq \frac{\delta \cdot \sqrt{1 + \frac{4}{\pi^2}}}{2} \cdot \sqrt{\frac{2 \cdot n}{\pi} \cdot \left(\operatorname{tg} \frac{\pi \cdot (m+1)}{n} - \operatorname{tg} \frac{\pi}{n} \right)}.$$

Это неравенство для достаточно больших значений n можно записать в виде

$$\varepsilon(n, k) \leq \frac{\delta \cdot \sqrt{1 + \frac{4}{\pi^2}}}{2} \cdot \sqrt{2 \cdot m} \quad (14)$$

Аналогичными вычислениями получается также оценка

$$\varepsilon(n, k) \geq \frac{\delta}{2} \cdot \sqrt{\frac{2 \cdot n}{\pi} \cdot \operatorname{tg} \frac{\pi \cdot m}{n}} \quad (15)$$

Подставляя в (14) и (15) значения m ,

$m = \frac{n \cdot k - n}{2 \cdot k}$, при больших значениях n окончательно находим

$$\frac{\delta}{2} \cdot \sqrt{n - \frac{n}{k}} \leq \varepsilon(n, k) \leq \frac{\delta}{2} \cdot \sqrt{1 + \frac{4}{\pi^2}} \cdot \sqrt{n - \frac{n}{k}}.$$

Следовательно, теорема доказана.

Проанализируем полученный результат с точки зрения сжатия и восстановления изображений. При помощи матрицы выбора S вместо исходного вектора $y = (y_i)_{i=0}^{n-1}$ рассматриваем вектор $z = (y_i + \gamma_i)_{i=0}^{n-1}$ вида

$$z = (z_0, z_1, \dots, z_{\frac{n}{2}-m-1}, 0, \dots, 0, z_{\frac{n}{2}+m}, \dots, z_{n-1}),$$

где

$$\gamma_i = \begin{cases} 0, & i \text{ дè } i = 0, 1, \dots, \frac{n}{2} - m - 1, \\ -y_i, & i \text{ дè } i = \frac{n}{2} - m, \frac{n}{2} - m + 1, \dots, \frac{n}{2} + m - 1, \\ 0, & i \text{ дè } i = \frac{n}{2} + m, \frac{n}{2} + m + 1, \dots, n - 1 \end{cases}$$

В этих условиях, согласно теореме 1, ошибка восстановления имеет вид (7),

где $k = \frac{n}{n-2m}$ - коэффициент сжатия.

Регуляризацию суммирования с приближенными коэффициентами можно произвести при помощи регулирующих множителей вида:

$$R(k, \alpha) = \frac{1}{1 + \alpha \cdot k^{1+\varepsilon}},$$

$k = 1, 2, \dots, n-1, 0 < \alpha \leq 1, R(0, \alpha) = 1, \varepsilon > 0$.

Использование метода регуляризации при сжатии и восстановлении сигналов означает, что вместо исходного вектора y рассматривается вектор

$$\hat{y} = (\hat{y}_i)_{i=0}^{n-1}, \text{ где}$$

$$\hat{y}_0 = y_0, \quad \hat{y}_j = \frac{y_j}{1 + \alpha \cdot j^{1+\varepsilon}}, \quad j = 1, 2, \dots, n-1 \quad (16)$$

При помощи матрицы выбора производится сжатие вектора \hat{y} , т.е. рассматриваем вектор $z = (z_i)_{i=0}^{n-1}$, где

$$z_j = \begin{cases} y_0, & i \text{ дè } j = 0 \\ \frac{y_j}{1 + \alpha \cdot j^{1+\varepsilon}}, & i \text{ дè } j = 1, 2, \dots, \frac{n}{2} + m - 1, \\ \frac{y_j}{1 + \alpha \cdot j^{1+\varepsilon}}, & i \text{ дè } j = \frac{n}{2} + m, \frac{n}{2} + m + 1, \dots, n - 1. \end{cases} \quad (17)$$

Так как имеет место неравенство (8), то очевидно также следующее соотношение

$$\frac{\delta}{2 \cdot (1 + \alpha \cdot j^{1+\varepsilon}) \cdot \sin \frac{\pi \cdot j}{n}} \leq \max_{x \in X_\delta} |z_j| \leq \frac{\delta \cdot \sqrt{1 + \frac{4}{\pi^2}}}{2 \cdot (1 + \alpha \cdot j^{1+\varepsilon}) \cdot \sin \frac{\pi \cdot j}{n}} \quad (18)$$

Условия (18) имеют место также для вектора \hat{y} . Следовательно, используя равенство Парсеваля и, учитывая формулы (16), (17), из (18) находим

$$\left[\sum_{j=\frac{n}{2}-m}^{\frac{n}{2}+m-1} z_1^2(j) \right]^{\frac{1}{2}} \leq \varepsilon(n, k, \alpha) \leq \left[\sum_{j=\frac{n}{2}-m}^{\frac{n}{2}+m-1} z_2^2(j) \right]^{\frac{1}{2}} \quad (19)$$

где

$$z_1(j) = \frac{\delta}{2 \cdot (1 + \alpha \cdot j^{1+\varepsilon}) \cdot \sin \frac{\pi \cdot j}{n}},$$

$$z_2(j) = \frac{\delta \cdot \sqrt{1 + \frac{4}{\pi^2}}}{2 \cdot (1 + \alpha \cdot j^{1+\varepsilon}) \cdot \sin \frac{\pi \cdot j}{n}}.$$

Оценим сверху выражение в правой части (19).

$$\left[\sum_{j=\frac{n}{2}-m}^{\frac{n}{2}+m-1} z_2^2(j) \right]^{\frac{1}{2}} = \frac{\delta \cdot \sqrt{1 + \frac{4}{\pi^2}}}{2} \left[\sum_{j=\frac{n}{2}-m}^{\frac{n}{2}+m-1} \frac{1}{(1 + \alpha \cdot j^{1+\varepsilon})^2 \cdot \sin^2 \frac{\pi \cdot j}{n}} \right]^{\frac{1}{2}} \leq \frac{\delta \cdot \sqrt{1 + \frac{4}{\pi^2}}}{2 \cdot (1 + \alpha (\frac{n}{2} - m)^{1+\varepsilon})} \left[\sum_{j=\frac{n}{2}-m}^{\frac{n}{2}+m-1} \frac{1}{\sin^2 \frac{\pi \cdot j}{n}} \right]^{\frac{1}{2}}.$$

Используя формулы (12) и (13), окончательно находим

$$\varepsilon(n, k, \alpha) \leq \frac{\delta \cdot \sqrt{1 + \frac{4}{\pi^2}}}{2 \cdot (1 + \alpha (\frac{n}{2} - m)^{1+\varepsilon})} \cdot \sqrt{2 \cdot n \cdot \left(\operatorname{tg} \frac{\pi(m+1)}{n} - \operatorname{tg} \frac{\pi}{n} \right)}. \quad (20)$$

Аналогичными вычислениями получаем оценку снизу

$$\varepsilon(n, k, \alpha) \geq \frac{\delta}{2 \cdot (1 + \alpha (\frac{n}{2} + m - 1)^{1+\varepsilon})} \cdot \sqrt{\frac{2 \cdot n}{\pi} \cdot \operatorname{tg} \frac{\pi \cdot m}{n}}. \quad (21)$$

Подставляя в (20) и (21) значения $m = \frac{n \cdot k - n}{2 \cdot k}$ при больших числах n , находим

$$\frac{\delta}{2 \cdot (1 + \alpha (n - \frac{n}{2k})^{1+\varepsilon})} \cdot \sqrt{n - \frac{n}{k}} \leq \varepsilon(n, k, \alpha) \leq \frac{\delta \cdot \sqrt{1 + \frac{4}{\pi^2}}}{2 \cdot (1 + \alpha (\frac{n}{2k})^{1+\varepsilon})} \cdot \sqrt{n - \frac{n}{k}}. \quad (22)$$

Сравнивая оценки (7) и (22), можно заключить, что при фиксированном n ошибка восстановления с использованием ДПФ уменьшается, если применять метод регуляризации.

Теперь вычислим погрешность между вектором y и регуляризованным и сжатым вектором z . Учитывая, что

$$y = \left(y_0, y_1, \dots, y_{\frac{n}{2}-m-1}, y_{\frac{n}{2}-m}, \dots, y_{\frac{n}{2}+m-1}, y_{\frac{n}{2}+m}, \dots, y_{n-1} \right)$$

$$z = \left(y_0, \frac{y_1}{1 + \alpha}, \frac{y_2}{1 + \alpha \cdot 2^{1+\varepsilon}}, \dots, \frac{y_{\frac{n}{2}-m-1}}{1 + \alpha \cdot (\frac{n}{2} - m - 1)^{1+\varepsilon}}, 0, \dots, 0, \frac{y_{\frac{n}{2}+m}}{1 + \alpha \cdot (\frac{n}{2} + m)^{1+\varepsilon}}, \dots, \frac{y_{n-1}}{1 + \alpha \cdot (n - 1)^{1+\varepsilon}} \right)$$

Вычислим

$$\varepsilon_r(n, k, \alpha) = \left[\sum_{j=1}^{\frac{n}{2}-m-1} \left(y_j - \frac{y_j}{1 + \alpha \cdot j^{1+\varepsilon}} \right)^2 + \sum_{j=\frac{n}{2}+m}^{n-1} \left(y_j - \frac{y_j}{1 + \alpha \cdot j^{1+\varepsilon}} \right)^2 + \sum_{j=\frac{n}{2}-m}^{\frac{n}{2}+m-1} y_j^2 \right]^{\frac{1}{2}} = \left[\sum_{j=1}^{\frac{n}{2}-m-1} \left(\frac{\alpha \cdot j^{1+\varepsilon}}{1 + \alpha \cdot j^{1+\varepsilon}} \right)^2 \cdot y_j^2 + \sum_{j=\frac{n}{2}+m}^{n-1} \left(\frac{\alpha \cdot j^{1+\varepsilon}}{1 + \alpha \cdot j^{1+\varepsilon}} \right)^2 \cdot y_j^2 + \sum_{j=\frac{n}{2}-m}^{\frac{n}{2}+m-1} y_j^2 \right]^{\frac{1}{2}} \quad (23)$$

Отсюда, используя (8), находим

$$\varepsilon_r(n, k, \alpha) \leq \frac{\alpha(n-1)^{1+\varepsilon} \cdot \delta}{2 \cdot (1 + \alpha)} \cdot \sqrt{1 + \frac{4}{\pi^2}} \cdot \sqrt{n - \frac{n}{k}}. \quad (24)$$

Теперь сравним оценки (7) и (24) и определим α , требуя выполнение соотношения

$$\varepsilon_r(n, k, \alpha) \leq \varepsilon(n, k).$$

Вычислим

$$\frac{\varepsilon_r(n, k, \alpha)}{\varepsilon(n, k)} \leq \frac{\alpha \cdot (n-1)^{1+\varepsilon}}{1 + \alpha} \leq 1.$$

Отсюда находим

$$\alpha \leq \frac{1}{(n-1)^{1+\varepsilon} - 1}.$$

Аналогичными рассуждениями доказываются оценки для преобразований Хартли, ДКП, ДСП, ДПУА, ДПУП и ДПУ.

Для наглядности ниже приводится график, на котором прослеживается поведение ошибки сжатия с регуляризацией и без регуляризации сигналов с применением вышерассмотренных ДОП.

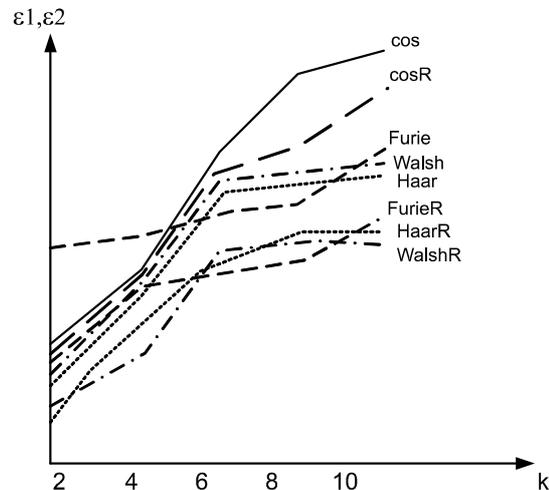


Рис.2. Ошибки сжатия с регуляризацией и без регуляризации.

Здесь ε_1 - ошибка сжатия без регуляризации; ε_2 - ошибка сжатия с регуляризацией; k – коэффициент сжатия.

Таким образом, получены оценки погрешностей восстановления сжатого сигнала кодированного посредством ортогональных преобразований с регуляризацией и без регуляризации.

Проведено сравнение оценок погрешностей и найдены значения регуляризирующего параметра α для ДПФ, преобразования Хартли, ДКП, ДСП и ДПУА при которых восстановленный сигнал более близок к оригиналу, чем сигнал восстановленный без применения метода регуляризации.

Литература

1. Н. Ахмед, К.Р. Рао. Ортогональные преобразования при обработке цифровых сигналов. – М.: Связь, 1980.
2. У. Прэтт. Цифровая обработка изображений. – М.: Мир, 1982.
3. П. Уинтц. Кодирование изображения посредством преобразований. // ТИИЭР, 1972, т.60, стр.69-83.
4. А.Н. Тихонов, В.Я. Арсенин. Методы решения некорректно поставленных задач. – М.: Наука, 1979.
5. В.К. Иванов. О линейных некорректных задачах. // Докл. АН СССР, 1962, т.145, №2, с.270-272.
6. А.Н. Тихонов. О регуляризации некорректно поставленных задач. // Докл. АН СССР, 1963, т.153, №1, с.49-52.
7. А.Н. Тихонов. Об устойчивых методах суммирования рядов Фурье. // Докл. АН СССР, 1964, т.156, №2, с.268-271.
8. М.М. Лаврентьев, В.Г. Романов, С.П. Шишатский. Некорректные задачи математической физики и анализа. – М.: Наука, 1980, 288 с.
9. Б.И. Голубов, А.В. Ефимов, В.А. Скворцов. Ряды и преобразования Уолша. – М.: Наука, 1987.
10. С.С. Агаян, А.А. Петросян. Оптимальные методы зонного кодирования посредством дискретных ортогональных преобразований. – Препринт №89-3, 1989.
11. С.С. Агаян, С.Б.Алавердян. Разработка автоматизированных систем цифровой обработки многомерных сигналов (изображений). В сб. «Системные и математические вопросы создания автоматизированных научных исследований (АСНИ) коллективного пользования» Ереван, 1984, с.61-66.
12. Тесчер А.Г. Кодирование изображений с преобразованием. В кн. Методы передачи изображений. Сокращение избыточности. Под ред. У.К. Прэтта. М., Радио и связь, 1983г. с. 103-143

ПРЕСС-РЕЛИЗ

Предприятия радиоэлектронного комплекса России намерены принять участие в подготовке и проведении Олимпиады-2014 в г. Сочи

Радиоэлектронный комплекс России выступил с инициативой по оснащению олимпийских объектов отечественным радиоэлектронным оборудованием. Федеральным агентством по промышленности (Роспром) создан оргкомитет по подготовке необходимых предложений, в который вошли руководители следующих компаний: ОПК «Оборонпром», ОАО «Телеком», РТОС, ОАО «Концерн «Созвездие», ОАО «ЭКОС», ОАО «НИИВК им. М.А. Карцева», ОАО «Концерн «РТИ Системы», ОАО «ЦНИИ «Электроника», ОАО «Компания «Импульс», ФГУП «НТЦ «Промтехазро», ОАО «Концерн ПВО «Алмаз-Антей», ОАО «Концерн «Вега», ЗАО «ЧипЭКСПО».

Для оснащения олимпийских объектов и инфраструктуры может быть использована следующая продукция радиоэлектронной промышленности: аппаратура цифрового телерадиовещания; связанное оборудование; оборудование для управления воздушным движением; системы и средства обеспечения охраны, безопасности и борьбы с терроризмом; навигационная аппаратура; информационные системы; оборудование для коммунального хозяйства; системы для управления дорожным движением, городским транспортом, электронные системы слежения; оборудование для оснащения офисов, гостиниц, корреспондентских пунктов, культурных центров, спортивных комплексов и других объектов инфраструктуры; медицинская аппаратура и др.

Участие предприятий радиоэлектронного комплекса России в подготовке и проведении Олимпиады-2014 в г.Сочи обеспечит оснащение инфраструктуры олимпийских объектов высококачественным отечественным оборудованием и даст динамичный толчок в научно-техническом и производственном развитии радиоэлектронной промышленности.

Оргкомитет по подготовке предложений РЭК к Олимпиаде 2014 года
Председатель оргкомитета – заместитель руководителя Роспрома Борисов Ю.И.

УДК 621.397

НЕЭТАЛОННАЯ ОЦЕНКА КАЧЕСТВА JPEG ИЗОБРАЖЕНИЙ

Приоров А.Л., Саутов Е.Ю., Хрящев В.В.

Введение

Для оценки искажений цифровых изображений наиболее часто используются: пиковое отношение сигнал/шум (ПОСШ) и среднеквадратическая ошибка (СКО). Однако эти метрики не всегда хорошо определяют визуально воспринимаемое качество [1-4]. Они задаются следующими соотношениями:

$$\text{ПОСШ} = 20 \log_{10} \frac{255}{\sqrt{\text{СКО}}}, \quad \text{СКО} = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2,$$

где x_i и y_i – значения пикселей исходного и восстановленного изображений соответственно.

В последнее время предпринимается большое количество попыток разработать новые метрики объективной оценки качества изображений, учитывающие свойства зрительной системы человека [5-8].

Следует отметить, что большинство из предлагаемых методов оценки качества используют исходное (эталонное) изображение. Разработка алгоритма объективного неэталонного измерения качества – сложная задача. В [9] указывается, что эффективная неэталонная модель оценки качества изображения возможна только при наличии априорной информации о типах его искажений.

Анализ работ [10-14] показывает, что при сжатии изображений с использованием стандарта JPEG при дискретном косинусном преобразовании (ДКП) наиболее отчетливо проявляются два вида искажений – блочность и размытие границ. В определении степени влияния каждого из этих искажающих факторов на изображение и заключается идея предложенного ниже неэталонного алгоритма оценки качества.

Схема предлагаемого алгоритма неэталонной оценки качества JPEG изображений приведена на рис. 1. Последовательность вычисления такой оценки состоит из пяти операций.

1. Обозначим исследуемое декодированное изображение, имеющее размеры $M \times N$ пикселей, как $x(m, n)$, где $m \in [1, M]$, $n \in [1, N]$ и вычислим разницу между соседними пикселями в каждой строке $d_h(m, n) = x(m, n+1) - x(m, n)$ и в каждом столбце изображения $d_v(m, n) = x(m+1, n) - x(m, n)$.

2. Определим среднюю разницу между блоками размером 8×8 по формуле:

$$B_h = \frac{1}{M([N/8]-1)} \sum_{i=1}^M \sum_{j=1}^{[N/8]-1} |d_h(i, 8j)|.$$

Описывается алгоритм неэталонной оценки качества сжатых изображений формата JPEG. Приведенные результаты моделирования показывают хорошую коррелированность получаемых оценок с визуально воспринимаемым качеством декодированных изображений. Описана процедура оптимизации этапа кодирования JPEG с использованием предложенной неэталонной оценки.

3. Оценим энергию сигнала изображения. Оценка проводится в два этапа, т.к. два фактора характеризуют энергию изображения:

а) Первый фактор – отклонение средней абсолютной разности на границе блоков от той же величины, подсчитанной для всего изображения

$$A_h = \frac{1}{7} \left[\frac{8}{M(N-1)} \sum_{i=1}^M \sum_{j=1}^{N-1} |d_h(i, j)| - B_h \right].$$

б) Второй фактор – число переходов через нуль. Пусть функция $z_h(m, n)$ принимает значение равно единице, если функция $d_h(m, n)$ пересекает нуль по горизонтали. В противном случае значение $z_h(m, n)$ равно нулю. Находим значения функции $z_h(m, n)$ для $n \in [1, N-2]$. Тогда количество переходов через нуль по горизонтали определяется как

$$Z_h = \frac{1}{M(N-2)} \sum_{i=1}^M \sum_{j=1}^{N-2} z_h(m, n).$$

Аналогичным способом вычисляются соответствующие величины B_v , A_v и Z_v по вертикали.

4. Производим усреднение полученных результатов:

$$B = \frac{B_h + B_v}{2}; \quad A = \frac{A_h + A_v}{2}; \quad Z = \frac{Z_h + Z_v}{2}.$$

5. Для получения конечной оценки качества сжатых изображений - JQ (JPEG Quality) используется следующая комбинация величин A, B и Z:

$$JQ = \alpha + \beta B^{\gamma_1} A^{\gamma_2} Z^{\gamma_3},$$

где $\alpha, \beta, \gamma_1, \gamma_2, \gamma_3$ – параметры, которые подбирались таким образом, чтобы результат как можно точнее учитывал визуально воспринимаемое качество декодированных изображений. Субъективные визуальные оценки изображений, сжатых алгоритмом JPEG, использованы для оптимизации параметров алгоритма. Экспериментально подобраны следующие значения параметров:

$$\alpha = -246, \quad \beta = 262, \quad \gamma_1 = -0.024, \\ \gamma_2 = 0.016, \quad \gamma_3 = 0.064$$

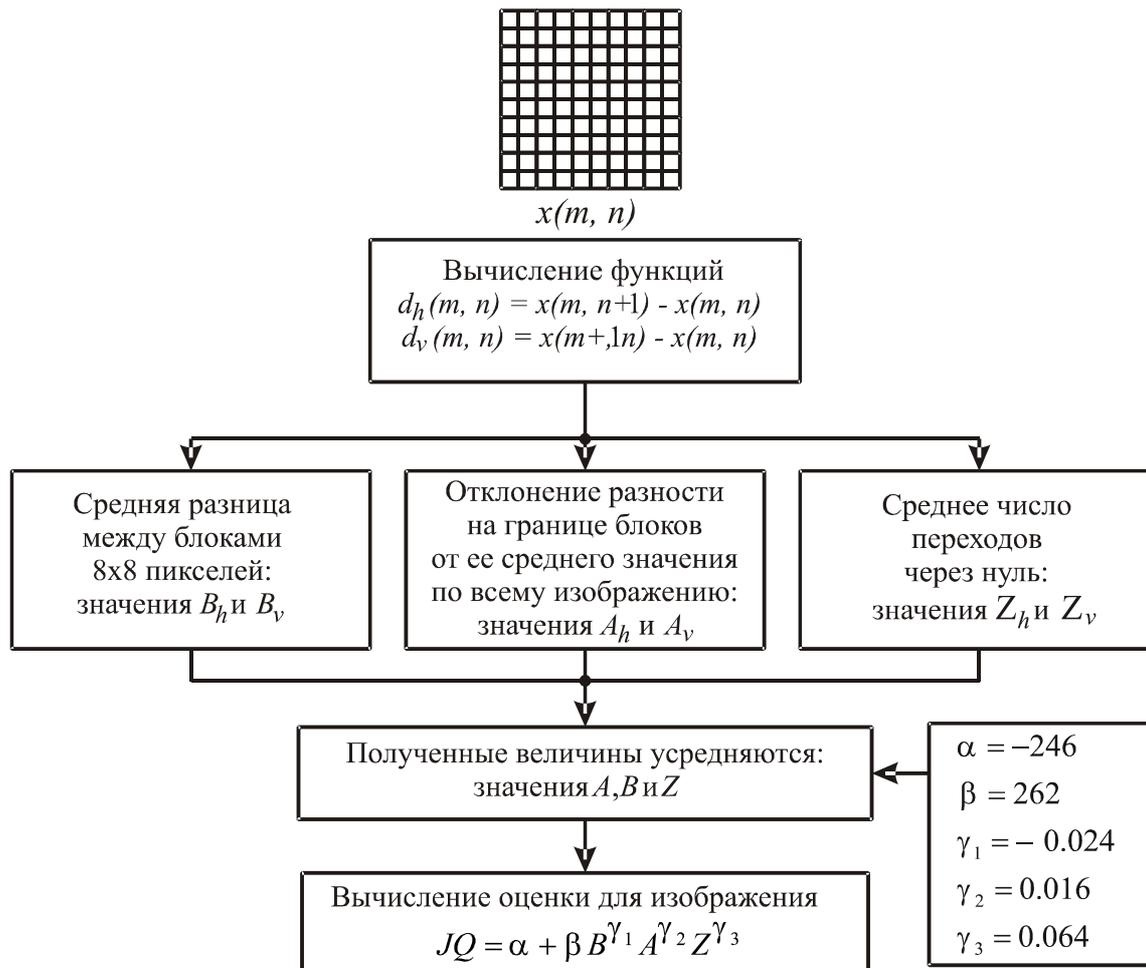


Рис. 1. Алгоритм неэталонной оценки качества JPEG изображений

Этот алгоритм неэталонной оценки качества JPEG изображений реализован в среде математического моделирования Matlab. Для тестирования алгоритма выбраны три полутоновых изображения разрешением 512×512 пикселей с разной степенью детализации: «Перцы», «Лена» и «Бабуин».

Результаты работы алгоритма для изображения «Лена» приведены на рис. 2.

Здесь K и JQ обозначают коэффициент сжатия исходного изображения и полученную неэталонную оценку. При низких коэффициентах сжатия (2-10) JPEG изображение практически полностью соответствует своему оригиналу (искажения незначительные). Как видно из рис. 2б, оценка JQ для таких изображений достаточно высока (находится в области цифры 10). При увеличении степени сжатия качество JPEG изображения ухудшается, проявляются искажающие факторы: блочность и размытие границ. В этом случае оценка JQ уменьшается, отражая воспринимаемое качество изображения (рис. 2в, 2г).

На рис. 3 представлены зависимости оценки JQ от коэффициента сжатия K , полученные для трех указанных тестовых изображений. Отметим, что эти зависимости носят монотонно спадающий, практически линейный характер. Причиной этому служит проявление на изображении искажений, вносимых алго-

ритмом JPEG с ростом коэффициента сжатия.

Для разных изображений зависимости различны, что подчеркивает тот факт, что искажения с различной степенью «эффективности» воздействуют на изображения с разным уровнем детализации. Так, например, из рис. 3 видно, что при одинаковом коэффициенте сжатия качество изображения «Бабуин» на 1,5-2 балла ниже по шкале JQ , чем для изображений «Лена» и «Перцы». Подобные зависимости получены и для других типов тестовых изображений.

Проведенные эксперименты позволили составить примерную шкалу перевода оценок JQ в пятибалльную шкалу визуальных оценок, согласно рекомендациям ITU-R BT.500-11 (табл. 1).

Рекомендательный характер ступеней кодирования позволяет разработчикам оптимизировать алгоритм JPEG с целью получения, например, более высокой скорости кодирования или декодирования, большей степени сжатия или улучшения качества восстановленного изображения. Так, например, в работах [15-17] описывается процедура выбора оптимальной матрицы квантования ДКП; рекурсивный алгоритм, учитывающий локальную статистику изображения, предлагается в [18]; а оптимизация этапа кодирования JPEG с помощью векторного квантования – в [19].

Таблица 1

Шкала соответствия оценок JQ и визуально воспринимаемого качества декодированных JPEG изображений

Оценка JQ	Визуально воспринимаемое качество
8 – 11	отличное
6 – 8	хорошее
4 – 6	удовлетворительное
0 – 4	неудовлетворительное
отрицательная	плохое



а) оригинальное изображение



б) JPEG изображение ($K = 6.92, JQ = 8.993$)



в) JPEG изображение ($K = 19.46, JQ = 5.741$)



г) JPEG изображение ($K = 32.78, JQ = 3.9585$)

Рис. 2. Результаты тестирования предложенного алгоритма на изображении «Лена»

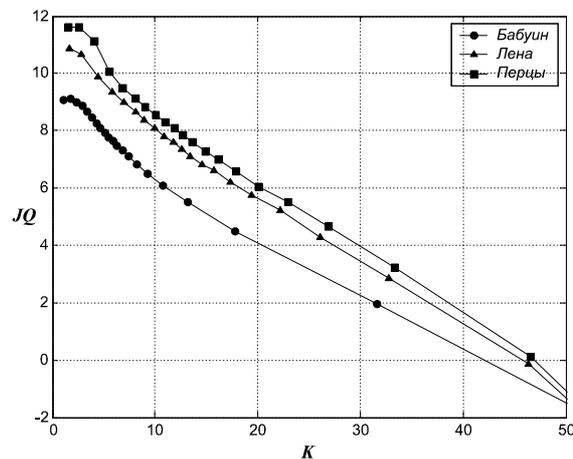


Рис. 3. Зависимость оценки JQ от коэффициента сжатия K

Возможны два способа оптимизации алгоритма JPEG цифровой обработки изображений с использованием оценки JQ .

1 способ. Выбирается единая для всего изображения матрица квантования с учетом значений оценки JQ . На первом этапе применяется наиболее «мягкая» матрица квантования и вычисляется оценка JQ_1 . Затем применяются более «грубые» матрицы квантования и контролируется отклонение оценки JQ_i от JQ_1 . Если значение оценки превышает первоначально заданный порог, повторяется алгоритм кодирования (рис. 4а).

2 способ. Этот способ учитывает тот факт, что различные участки изображения имеют различную детализацию. Учет этого факта позволяет применять к участкам с меньшей детализацией более «грубые» матрицы квантования и, как следствие, сжимать такие участки сильнее. Эта идея лежит в основе второго способа оптимизации.

Матрица квантования для каждого блока размером 8×8 пикселей подбирается индивидуально на основе анализа оценки JQ окрестности сжимаемого блока размером $N \times N$ пикселей. На первом этапе к k -ому блоку применяется наиболее «мягкая» матрица квантования и вычисляется оценка JQ_k^1 . Затем к кодируемому блоку размером 8×8 пикселей применяется матрица квантования с «грубыми» коэффициентами и вычисляется оценка JQ_k^2 . Если отклонение JQ_k^2 от JQ_k^1 не выходит за пределы заданного порога, применяется более «грубая» матрица квантования. В противном случае к блоку применяется матрица квантования, полученная на предыдущем этапе, работа алгоритма для данного блока завершается и происходит переход к следующему блоку. С целью умень-

шения вычислительных затрат, число итераций алгоритма кодирования для каждого из блоков ограничивалось пятью (рис. 4б).

Исследования показали, что наиболее предпочтительным размером области анализа является размер блока 24×24 пикселей. При дальнейшем увеличении размеров указанной области эффективность второго способа оптимизации снижалась и приближалась к значениям, полученным при первом способе оптимизации.

Представленная на рис. 5. зависимость отображает поведение оценки JQ восстановленного изображения «Перцы» от степени сжатия K для обычного (реализация в Matlab) и оптимизированного первым способом JPEG кодирования. Как видно из графика, оптимизация позволяет в среднем на единицу улучшить значения оценок JQ практически во всем диапазоне коэффициентов сжатия K .

На рис. 6 представлены усредненные по трем тестовым изображениям зависимости SKO восстановленного изображения от коэффициента сжатия при стандартной и оптимизированной схемах JPEG кодирования.

Отметим, что второй способ оптимизации позволяет улучшить качество изображения в значениях SKO преимущественно при низких степенях сжатия, при больших коэффициентах сжатия зависимости совпадают. Следовательно, учет участков с различной детализацией эффективнее использовать при небольших коэффициентах сжатия изображения ($K \in [2, 10]$).

Таким образом, рассмотрев возможность оптимизации этапа кодирования стандарта JPEG на основе неэталонной оценки JQ , показано, что процесс оптимизации позволяет улучшить качество изображения в значениях оценки JQ на 1-2 балла, а в значениях SKO восстановленного изображения на 10-15 единиц.

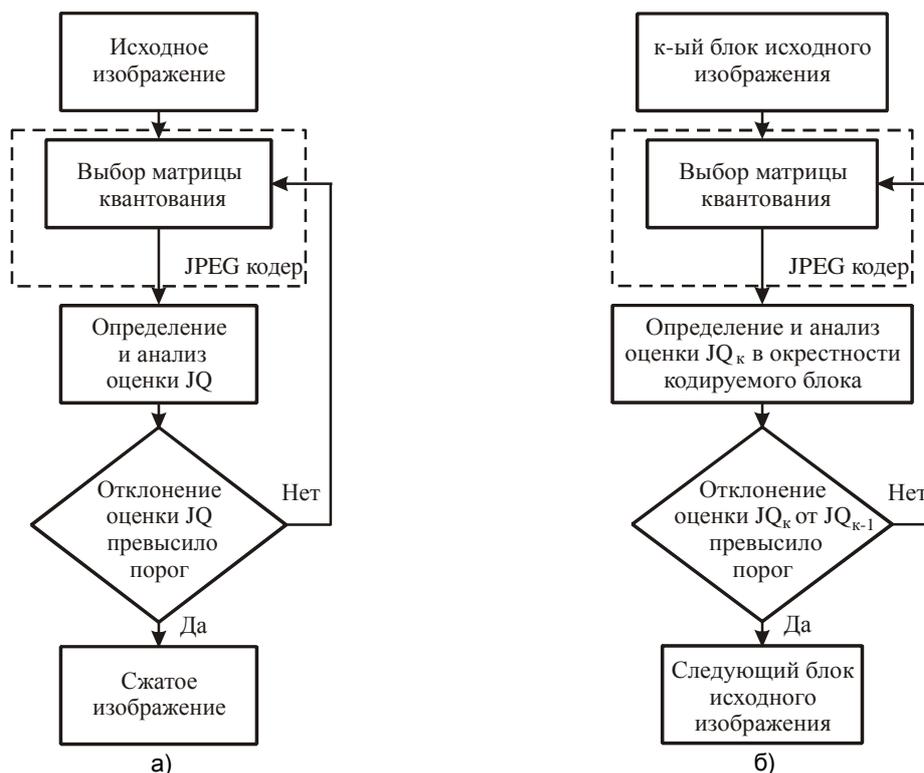


Рис. 4. Оптимизация этапа кодирования алгоритма JPEG: а) 1 способ; б) 2 способ

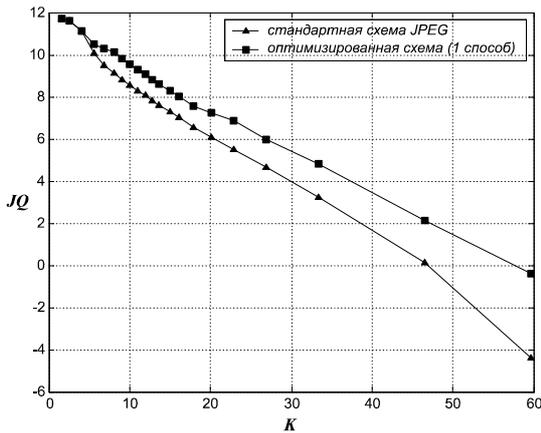


Рис. 5. Зависимость оценки JQ от степени сжатия K для стандартной и оптимизированной схемы JPEG кодирования

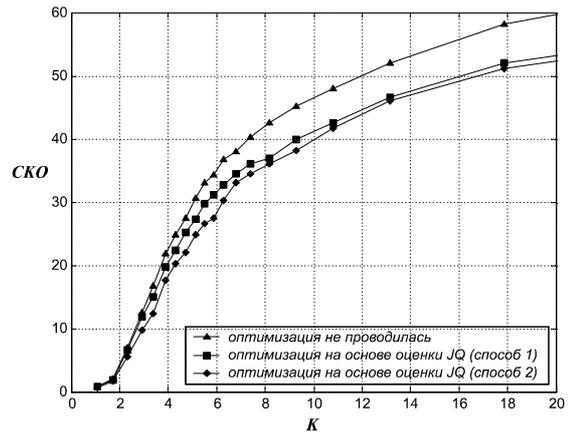


Рис. 6. Зависимость СКО восстановленного изображения от степени сжатия K для различных схем JPEG кодирования

Литература

1. Сэлмон Д. Сжатие данных, изображений и звука. – М.: Техносфера, 2004.
2. Wang, Z., Bovik A. Modern image quality assessment // Synthesis lectures on image, video & multimedia processing. Morgan & Claypool. 2006.
3. Eskicioglu A., Fisher P. Image quality measures and their performance // IEEE Trans. communications. 1995. V. 43, № 12. P. 2959-2965.
4. Mrak M., Grgic S., Grgic M. Picture quality measures in image compression system // Proc. IEEE region 8 int. conf. on computer as a tool. EUROCON 2003. 2003. V. 1, № 9. P. 233-236.
5. Egiazarian K., Jaakko A., Ponomarenko N., Lukin V., Battisti F., Carli M. Two new full-reference quality metrics based on HVS // Proc. of the 2nd int. workshop on video process. and quality metrics for consumer electronics. VPQM. 2006.
6. Ilushkina N., Tchobanou M. Three full reference quality measures for compressed images // In proc. of the 9th int. conf. "Digital Signal Processing and its Applications". 2007. V. 2, P. 497-499.
7. Wang Z., Bovik A., Sheikh H., Simoncelli E. Image quality assessment: from error visibility to structural similarity // IEEE Trans. on image processing. 2004. V. 13, № 4. P. 600-612.
8. Арляпов С.А., Приоров А.Л., Хрящев В.В. Модифицированный критерий оценки качества восстановленных изображений // Цифровая обработка сигналов. 2006. № 2. С. 27-33.
9. Бекренев В.А., Саутов Е.Ю., Куйкин Д.К. Алгоритм неэталонной оценки качества JPEG изображений // Докл. 9-й межд. конф. «Цифровая обработка сигналов и ее применение» (DSPA-2007). Москва. 2007. Т. 2, С. 312-315.
10. Li X. Blind image quality assessment // Proc. IEEE int. conf. image processing. 2002. V. 1, P. 449-452.
11. Wang Z., Bovik A., Evans B. Blind measurement of blocking artifacts in images // Proc. IEEE int. conf. on image processing. Vancouver, Canada. 2000. V. 3, P. 981-984.
12. Bovik A., Liu S. DCT-domain blind measurement of blocking artifacts in DCT-coded images // Proc. IEEE int. conf. acoust., speech, and signal processing. 2001. V. 3, P. 1725-1728.
13. Boev A., Foi A., Egiazarian K., Katkovnik V. Adaptive scales as a structural similarity indicator for image quality assessment // Proc. of the 2nd int. workshop on video processing and quality metrics for consumer electronics. VPQM. 2006.
14. Babu R.V., Perki A. An HVS-based no-reference perceptual quality assessment of JPEG coded images using neural networks // Proc. IEEE int. conf. on image processing. 2005. V. 1, P. 433-436.
15. Ratnakar V., Livny M. Efficient algorithm for optimizing DCT quantization // IEEE Trans. on image processing. 2000. V. 9, № 2. P. 267-270.
16. Watson A. Visually optimal DCT quantization matrices for individual images // Proc. data compression conf. 1993. P. 178-187.
17. Monro D., Sherlock B. Optimum DCT quantization // Proc. data compression conf. 1993. P. 188-194.
18. Wu S., Gersho A. Rate-constrained picture-adaptive quantization for JPEG baseline coders // Proc. int. conf. acoustics, speech, signal processing. 1993. V. 5, P. 389-392.
19. Crouse M., Ramchandran K. JPEG optimization using an entropy-constrained quantization framework // Proc. data compression conf. 1995. P. 342-351.

ОЦЕНКА ВЕРОЯТНОСТИ ПОЯВЛЕНИЯ СИМВОЛА ПРИ АДАПТИВНОМ ДВОИЧНОМ АРИФМЕТИЧЕСКОМ КОДИРОВАНИИ В ЗАДАЧАХ СЖАТИЯ ВИДЕОИНФОРМАЦИИ

Беляев Е.А., Тюрликов А.М.

Введение

В современных алгоритмах сжатия информации широко применяется адаптивное арифметическое кодирование данных с заранее неизвестной статистикой. Предполагается, что эти данные можно разбить на последовательность интервалов так, что статистические характеристики данных внутри каждого интервала постоянны, но могут существенно отличаться для разных интервалов. Степень сжатия таких данных во многом зависит как от точности оцениваемой внутри каждого интервала статистики, так и от скорости адаптации к изменяющейся между интервалами статистике.

В настоящей работе рассматриваются алгоритмы адаптивной оценки вероятности появления символа на выходе источника с двоичным алфавитом, сочетающие в себе как точность оценки, так и скорость адаптации.

Ниже показано, что описанные в работах [1,3-6] алгоритмы адаптивной оценки требуют высоких затрат как памяти так и вычислительных ресурсов кодера и декодера. Один из подходов, позволяющих понизить эти затраты, заключается в использовании конечного автомата при оценке вероятности [9-15]. Такой подход применяется, например, в стандарте сжатия видеоинформации H.264/AVC [16]. Конечный автомат хранится в памяти кодера и декодера и используется при оценке вероятности для всех контекстных моделей, имеющих в стандарте. Подобная реализация не требует операций умножения и деления при вычислении оценок вероятности. Недостаток такого подхода состоит в том, что при использовании одного и того же конечного автомата не учитываются отличия статистических свойств двоичных источников, соответствующих различным контекстным моделям. Один из способов, позволяющих учесть статистические отличия источников, заключается в использовании нескольких конечных автоматов [18], что приводит к дополнительному увеличению затрат памяти.

Второй способ, позволяющий учесть различия статистических свойств двоичных источников может быть реализован путем применения алгоритма оценки вероятности с периодическим масштабированием счетчиков [2]. Такой способ оценки позволяет повысить степень сжатия, но при этом существенно увеличивает требования к вычислительным ресурсам кодера и декодера, так как в нем используются операции умножения и деления.

В данной работе предложен целочисленный алгоритм «виртуального скользящего окна», который, с одной стороны, не использует операции умножения и деления при вычислении оценки вероятности и является

Представлено описание ряда известных алгоритмов адаптивной оценки вероятности появления символа на выходе двоичного источника. Предложен алгоритм «виртуального скользящего окна». Приведены результаты практического использования алгоритма для стандарта сжатия видеоинформации H.264/AVC.

наиболее предпочтительным с точки зрения степени сжатия, с другой стороны. Схожая целочисленная реализация адаптивной оценки вероятности описана в работе [19]. Однако она является более сложной как с вычислительной точки зрения (используется операция умножения), так и с точки зрения количества параметров алгоритма, которые необходимо назначить перед началом кодирования.

Алгоритм адаптивной оценки с периодическим масштабированием счетчиков. Пусть на вход кодера подается последовательность двоичных символов x_1, x_2, \dots, x_N . Для оценки вероятности появления символа $x_{t+1} \in \{0, 1\}, 1 \leq t \leq N$, на выходе двоичного источника может быть использована оценка Кричевского-Трофимова [1]. Согласно этой оценке, вероятность того, что символ x_{t+1} будет равен единице

$$\hat{p}_{t+1} = \frac{n_t^1 + \frac{1}{2}}{n_t^0 + n_t^1 + 1} \quad (1)$$

где n_t^1 и n_t^0 - число единиц и нулей в последовательности x_1, x_2, \dots, x_t .

Оценка (1) имеет следующие недостатки. Во-первых, данная оценка неэффективна при сжатии данных с изменяющейся статистикой. Во-вторых, при реализации на практике существует вероятность переполнения счетчиков, содержащих число единиц и нулей в последовательности поступивших символов. Описанные недостатки (см., например, работу [2]) могут быть устранены путем периодического масштабирования счетчиков. Если выполняется условие $\min\{n_t^0, n_t^1\} > N_{\min}$ или $\max\{n_t^0, n_t^1\} > N_{\max}$, то значения счетчиков масштабируются следующим образом:

$$\begin{cases} n_0(t) = \beta n_0(t) \\ n_1(t) = \beta n_1(t) \end{cases}$$

где $\beta \in (0...1)$ - коэффициент масштабирования, а N_{\min} и N_{\max} - пороговые значения счетчиков.

«Скользящее окно» и его аппроксимации. Другой способ устранения описанных выше недостатков заключается в использовании конструкции так называемого «скользящего окна» [3], в котором вероятность появления очередного символа x_{t+1} источника определяется за счет анализа содержимого окна, то есть последова-

тельности символов $x_{t-W+1}x_{t-W+2}\dots x_t$, где $W \geq 1$ - длина окна. После кодирования очередного символа содержащее окно сдвигается на одну позицию, новый символ x_{t+1} заносится в освободившуюся ячейку, а последний символ x_{t-W+1} удаляется. Оценка вероятности появления единицы на выходе двоичного источника при использовании «скользящего окна» определяется по формуле (1) с учетом того, что n_t^0 и n_t^1 - число нулей и единиц в окне соответственно и $n_t^0 + n_t^1 = W$. Основным недостатком данного подхода является необходимость хранения последних W закодированных символов в памяти кодера и декодера.

При аппроксимации «скользящего окна» в памяти хранится только число символов в окне, которое вычисляется по некоторому правилу. Одно из таких правил, названное «мнимым скользящим окном», предложено в [4,5] для двоичного источника и в [3] для недвоичного источника. Согласно этому правилу, после кодирования символа x_t , из окна удаляется не последний, а случайный символ $y_t \in \{0,1\}$, и число единиц в окне $n_{t+1}^1 = n_t^1 - y_t + x_t$, где y_t - случайная величина, которая генерируется со следующими вероятностями:

$$\begin{cases} \Pr\{y_t = 1\} = \frac{n_t^1}{W}, \\ \Pr\{y_t = 0\} = 1 - \frac{n_t^1}{W}. \end{cases}$$

В алгоритме «мнимого скользящего окна» требуется генерировать случайную величину y_t , значение которой должно поступать на вход кодера и декодера. Поэтому в работе [3] предлагается использовать специальным образом формируемую псевдослучайную последовательность, что затрудняет реализацию алгоритма.

В [6-8] показано, как можно отказаться от генерирования случайной величины. Для этого случайная величина y_t заменяется ее математическим ожиданием. При этом правило пересчета числа единиц в окне изменяется следующим образом:

$$n_{t+1}^1 = n_t^1 - \frac{n_t^1}{W} + x_t = \left(1 - \frac{1}{W}\right)n_t^1 + x_t. \quad (3)$$

В соответствии с (3), вероятность появления единицы

$$\hat{p}_{t+1} = \left(1 - \frac{1}{W}\right)\hat{p}_t + \frac{1}{W}x_t. \quad (4)$$

Выражение (4) можно получить и исходя из других соображений. В [9] предлагается использовать геометрически убывающую оценку вероятности:

$$\hat{p}_{t+1} = \gamma(x_t + (1-\gamma)x_{t-1} + (1-\gamma)^2x_{t-2}\dots), \quad (5)$$

где γ - действительное число, $\gamma \in [0..1]$.

После записи в рекуррентной форме выражение (5) примет следующий вид:

$$\hat{p}_{t+1} = (1-\gamma)\hat{p}_t + \gamma x_t \quad (6)$$

Аналогичное правило вычисления приведено в [6,10]. При значении $\gamma = \frac{1}{W}$ выражения (4) и (6) совпадают. Это означает, что проведение процедуры «дерандомизации» «мнимого скользящего окна» приводит к частному случаю геометрически убывающей оценки

вероятности (подразумевается, что W принимает целые значения).

Реализация алгоритмов оценки вероятности при помощи конечного автомата. Во многих случаях при реализации ранее рассмотренных алгоритмов оценки вероятности целесообразно воспользоваться конечным автоматом, заданным в виде таблиц состояний и переходов. Каждое состояние автомата соответствует некоторой оценке вероятности. В зависимости от значения входного символа x_t происходит переход из одного состояния автомата в другое. Подобная реализация не требует операций умножения и деления при вычислении оценок вероятности. Кроме того, фиксированный набор оценок вероятностей позволяет отказаться от операции умножения при выполнении собственно арифметического кодирования. Этот способ оценки вероятности используется в таких реализациях арифметического кодера как Q-coder [11], его модификациях QM-coder [12] и MQ-coder [13], а также в Quasi-arithmetic coder [10], Z-coder [14], ELS-coder [9]. Табличная реализация алгоритма оценки с периодическим масштабированием счетчиков рассмотрена в работе [15].

Покажем, каким образом происходит переход от алгоритма оценки вероятности к конечному автомату на примере стандарта H.264/AVC [16]. Для кодирования недвоичных данных в этом стандарте может использоваться контекстный адаптивный двоичный арифметический кодер (Context-Based Adaptive Binary Arithmetic Coder). Перед кодированием очередной недвоичный символ при помощи процедуры бинаризации отображается в некоторую двоичную последовательность. Для каждого двоичного символа этой последовательности из заранее определенного фиксированного множества по некоторому правилу выбирается контекстная модель. Каждая контекстная модель содержит описание двоичного источника (фактически текущую оценку вероятности). Двоичный символ кодируется с использованием оценок вероятностей, задаваемых выбранной контекстной моделью. По завершении кодирования производится модификация оценок вероятностей соответствующей контекстной модели. Для кодирования двоичной последовательности используется M-coder, рассмотренный в работе [17]. Алгоритм оценки вероятности в M-coder основан на выражении (6), реализованном путем использования конечного автомата (см. рис.1), состоящего из 64-х состояний. Каждое состояние этого конечного автомата определяет фиксированную оценку вероятности входного символа. При этом входные символы делятся на наиболее вероятные символы (Most Probable Symbol – MPS) и наименее вероятные символы (Least Probable Symbol – LPS). Множество значений оценок вероятности $\{\hat{p}_0, \hat{p}_1, \dots, \hat{p}_{63}\}$ задается следующим образом: $\hat{p}_i = (1-\gamma)\hat{p}_{i-1}$, где $i = 1, \dots, 63$, $\hat{p}_0 = 0.5$, $\gamma = 1 - \left(\frac{p_{\min}}{0.5}\right)^{\frac{1}{63}}$, $\hat{p}_{\min} = 0.01875$.

При этом оценка вероятности

$$\hat{p}_{t+1} = \begin{cases} (1-\gamma)\hat{p}_t + \gamma, & \text{если } x_t \text{ - наименее вероятный символ,} \\ \max\left[(1-\gamma)\hat{p}_t, \hat{p}_{62}\right], & \text{в противном случае.} \end{cases}$$

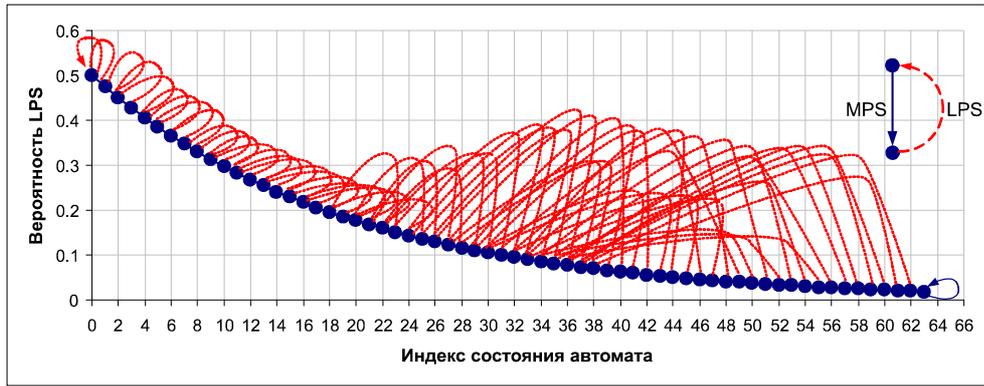


Рис.1 Оценка вероятности в стандарте H.264/AVC

В схеме кодирования, предложенной в [17], параметр адаптации γ одинаков для источников, соответствующих различным контекстным моделям. Однако можно достичь лучшего сжатия, если каждому источнику ставить в соответствие собственное значение γ . Один из способов, позволяющих реализовать этот подход, заключается в использовании нескольких значений γ , каждому из которых соответствует конечный автомат, заданный в виде таблиц состояний и переходов (подобный метод описан в работе [18]).

«Виртуальное скользящее окно». В настоящей работе предлагается целочисленная реализация (6) без использования таблиц состояний и переходов. Можно дать следующую интерпретацию работы алгоритма «мнимого скользящего окна». Пусть имеется «скользящее окно» из W ячеек. При поступлении очередного символа случайным образом выбирается одна ячейка. Символ, находящийся в этой ячейке заменяется на поступивший символ. Тогда правило пересчета (6) при целочисленной реализации допускает следующую интерпретацию. Имеется «скользящее окно» из cW ячеек, где c - параметр алгоритма. Значение поступившего символа заносится в c ячеек, выбранных случайным образом. При этом из окна удаляется среднее число единиц в выбранных c ячейках. Тогда число единиц s_{t+1} в окне из cW ячеек после кодирования очередного символа x_t можно пересчитывать по следующему правилу:

$$s_{t+1} = \begin{cases} s_t + \left\lfloor \frac{cW - s_t + \frac{W}{2}}{W} \right\rfloor, & \text{при } x_t = 1, \\ s_t - \left\lfloor \frac{s_t + \frac{W}{2}}{W} \right\rfloor, & \text{при } x_t = 0. \end{cases} \quad (7)$$

Далее в тексте правило пересчета (7) называется «виртуальным скользящим окном». Оценка вероятности появления единицы для алгоритма «виртуального скользящего окна»

$$\hat{p}_t = \frac{s_t}{cW}. \quad (8)$$

При аппроксимации алгоритма «скользящего окна» алгоритмом «виртуального скользящего окна», параметр c необходимо выбрать так, чтобы минимальные и максимальные оценки вероятностей для этих алгоритмов совпадали. Рассмотрим сначала, каким образом

будет обеспечено совпадение минимальных оценок вероятностей. Для этого укажем следующее множество наборов входных данных, для которых оценка вероятности принимает минимальное значение. Сначала подается произвольная конечная последовательность длины k . Затем подается последовательность из t нулей. Далее эти последовательности будем называть начальной и заключительной последовательностями соответственно. В этом случае, при любой начальной последовательности должно выполняться равенство следующих пределов:

$$\lim_{t \rightarrow \infty} \frac{n_{k+t}^1 + \frac{1}{2}}{W + 1} = \lim_{t \rightarrow \infty} \frac{s_{k+t}}{cW}, \quad \text{при } \forall k < \infty. \quad (9)$$

Левая и правая части (9) соответствуют оценке вероятности при использовании «скользящего окна» и «виртуального скользящего окна» для случая, когда заключительная последовательность имеет неограниченную длину.

Рассмотрим сначала левую часть равенства (9). При поступлении на вход кодера, использующего «скользящее окно», заключительной последовательности длины W и более количество единиц в окне станет равным нулю, поэтому левая часть (9) не зависит от начальной последовательности и равна:

$$\lim_{t \rightarrow \infty} \frac{n_{k+t}^1 + \frac{1}{2}}{W + 1} = \frac{1}{2(W + 1)}, \quad \text{при } \forall k < \infty. \quad (10)$$

Теперь рассмотрим правую часть равенства (9). Если перед началом кодирования $s_1 \in \{0, 1, \dots, \frac{W}{2} - 2\}$, то всегда найдутся две такие начальные последовательности, что значения правой части (9) будут различны при заключительной последовательности любой длины. Например, если начальная последовательность состоит из нулей, то значение $s_{t+k} = s_1$ и не меняется в процессе работы кодера, так как значение

$$\left\lfloor \frac{s_t + \frac{W}{2}}{W} \right\rfloor = 0, \quad \text{при } s_t < \frac{W}{2}.$$

Если начальная последовательность содержит хотя бы одну единицу, то $s_{t+k} > s_1$. Поэтому необходимым условием выполнения равенства (9) является выбор начального значения $s_1 \geq \frac{W}{2} - 1$. Очевидно, что при таком выборе для любой начальной последовательности правая часть (9) равна:

$$\lim_{t \rightarrow \infty} \frac{s_{k+t}}{cW} = \frac{w/2 - 1}{cW}, \text{ при } \forall k < \infty. \quad (11)$$

Приравнявая (10) и (11), получим:

$$\frac{1}{2(W+1)} = \frac{w/2 - 1}{cW}, \text{ тогда } c = W - 1 - \frac{2}{W}.$$

Аналогичным образом можно показать, что при выборе начального значения $s_1 \leq cW - W/2 + 1$, и параметре $c = W - 1 - 2/W$, будут совпадать и максимальные оценки вероятностей. При $W \gg 1$, можно считать, что $c = W - 1 - \frac{2}{W} \approx W$.

Таким образом, для совпадения соответствующих оценок вероятности алгоритмов «скользящего окна» и «виртуального скользящего окна» необходимо и достаточно, чтобы перед началом кодирования значение $s_1 \in \{w/2 - 1, \dots, cW - w/2 + 1\}$ и параметр алгоритма $c = W$.

Применение предложенного алгоритма «виртуального скользящего окна» позволяет отказаться от использования таблиц состояний и переходов. При выборе $W = 2^i$, где i - положительное целое число, вместо операции деления можно использовать операцию побитового сдвига. Таким образом, при вычислении оценки вероятности используются только операции сложения и сдвига. Характеристики алгоритма легко изменяются путем выбора различных значений длины окна W . Поэтому каждому источнику можно назначить собственное значение параметра $\gamma = 1/W$ без использования таблиц.

Практическое сравнение алгоритмов. Сравнение практической эффективности алгоритмов оценки вероятности проводилось на примере стандарта сжатия видеoinформации H.264/AVC. Алгоритмы сжатия видеoinформации сравниваются по двум основным параметрам: визуальное качество декодированной видеопоследовательности и битовая скорость (число бит на выходе кодера в единицу времени). Для получения практических результатов алгоритм оценки вероятности, используемый в кодере и декодере H.264/AVC, был заменен сначала алгоритмом с периодическим масштабированием счетчиков, а затем алгоритмом «виртуального скользящего окна». Для этого использовался открытый JVT кодек, версии JM. 10.2 (FRExt), поддерживающий данный стандарт. Результаты были получены для тридцати первых кадров известных тестовых HDTV видеопоследовательностей ("riverbed", "rush hour", "station", "sunflower", "tractor") разрешением 1920x1080. Алгоритм с периодическим масштабированием счетчиков был взят

из работы [2], в которой $N_{\max} = \infty$, и коэффициент масштабирования

$$\beta = \frac{N_{\min} + \Delta - 1}{N_{\min} + \Delta}, \text{ где } \Delta = 0.4.$$

При реализации алгоритма «виртуального скользящего окна» каждой контекстной модели с номером i необходимо назначить параметр кодирования $W_{opt}(i)$. Для этого в процессе кодирования тестовой видеопоследовательности (в данном случае использовалась последовательность "tractor") для каждого двоичного символа с номером t источника, соответствующего контекстной модели с номером i , и для каждой длины окна $W_l = 2^l, l = 3, 4, \dots, 10$, вычислялась оценка вероятности появления единицы

$$\hat{p}_t(W_l, i) = \frac{s_t(i)}{W_l^2},$$

где $s_t(i)$ - число единиц в «виртуальном скользящем окне» для контекстной модели с номером i перед кодированием символа $x_t(i)$.

По завершении кодирования вычислялись оценки битовых затрат при адаптивном арифметическом кодировании

$$\hat{R}(W_l, i) = \sum_t \hat{r}_t(W_l, i), \text{ где}$$

$$\hat{r}_t(W_l, i) = \begin{cases} -\log_2 \hat{p}_t(W_l, i), & \text{если } x_t(i) = 1, \\ -\log_2 (1 - \hat{p}_t(W_l, i)), & \text{если } x_t(i) = 0. \end{cases}$$

Для контекстной модели с номером i параметр кодирования $W_{opt}(i)$ устанавливался равным

$$W_{opt}(i) = \arg \min_{W_l} \hat{R}(W_l, i).$$

Начальное значение $s_1(i)$ определялось следующим образом:

$$s_1(i) = \min \left\{ W_{opt}^2(i) - \frac{W_{opt}(i)}{2} + 1, \max \left\{ \frac{W_{opt}(i)}{2} - 1, \left\lfloor W_{opt}^2(i) \cdot \hat{p}(i) \right\rfloor \right\} \right\},$$

где $\hat{p}(i)$ - определенная в стандарте H.264/AVC начальная оценка вероятности единицы для контекстной модели с номером i .

В представленных ниже таблицах приведено уменьшение битовой скорости (в процентах) относительно оригинальной версии кодека, при фиксированном качестве, в зависимости от значения номера шага квантования (QP). Таблица №1 относится к случаю использования алгоритма с периодическим масштабированием счетчиков. Таблица №2 относится к случаю использования алгоритма «виртуального скользящего окна».

Таблица №1.

Уменьшение битовой скорости при использовании алгоритма с периодическим масштабированием счетчиков

QP	10	20	30	40	50
riverbed	0.72	0.67	0.40	0.23	0.65
rush hour	0.58	0.47	0.57	0.17	1.04
station	0.77	0.57	-0.15	0.35	1.43
sunflower	0.25	0.19	0.05	0.22	1.66
tractor	0.53	0.61	0.79	0.89	1.45

Уменьшение битовой скорости при использовании алгоритма «виртуального скользящего окна»

QP	10	20	30	40	50
riverbed	0.98	0.93	0.93	0.90	0.90
rush hour	0.86	0.73	0.92	0.50	1.23
station	1.08	0.98	0.27	0.63	1.43
sunflower	0.68	0.50	0.45	0.58	1.58
tractor	0.84	0.90	1.10	1.10	1.62

Как видно из таблиц 1-2, назначение каждому двоичному источнику индивидуального параметра кодирования приводит к уменьшению битовой скорости (для таблицы №1 близкие результаты получены в [20]) при фиксированном качестве. В случае практической реализации такого подхода с использованием оценки с периодическим масштабированием счетчиков необходимо либо использовать операции умножения и деления при вычислении оценки вероятности либо вводить большое количество заранее вычисленных таблиц. Алгоритм «виртуального скользящего окна» может быть реализован как без использования таблиц, так и без операций умножения и деления. При этом он является наиболее предпочтительным с точки зрения эффективности кодирования.

Литература

1. R. E. Krichevski and V. E. Trofimov, "The performance of universal encoding" IEEE Trans. Inform. Theory, vol. IT-27, pp. 199–207, Mar. 1981.
2. D.L. Duttweiler and C. Chamzas, "Probability estimation in arithmetic and adaptive-Huffman entropy coders" IEEE Transactions on Image Processing Vol. 4, pp. 237-246 Mar. 1995
3. Рябко Б.Я., Фионов А.Н., Эффективный метод арифметического кодирования для источников с большими алфавитами // Проблемы передачи информации Т.35, №4. С.95-108, 1999.
4. T. Leighton and R. L. Rivest, "Estimating a probability using finite memory" IEEE Trans. Inform. Theory, vol. IT-32, pp. 733–742, Nov. 1986.
5. A. Zandi, G.G. Langdon, "Adaptation for non-stationary binary sources for data compression", p.224, 29th Asilomar Conference on Signals, Systems and Computers (2-Volume Set), 1996.
6. E. Meron and M. Feder, "Finite-Memory Universal Prediction of Individual Sequences" IEEE Trans. Inform. Theory, vol. 50-7, pp. 1506–1523, July 2004.
7. E. Belyaev, M. Gilmutdinov and A. Turlikov, "Binary Arithmetic Coding System with Adaptive Probability Estimation by "Virtual Sliding Window" // Proc. of the 10th IEEE International Symposium on Consumer Electronics – ISCE'06, St.-Petersburg, Russia, pp. 194–198, 2006.
8. Беляев Е.А. Использование «виртуального скользящего окна» при адаптивном арифметическом кодировании// Научная сессия ГУАП, посвященная всемирному Дню авиации и космонавтики и 65-летию ГУАП: Сб. докл./ ГУАП. СПб., 2006.
9. D. Withers, "The ELS-coder: a rapid entropy coder", Data Compression Conference, p. 475,1997.
10. P. G. Howard and J. S. Vitter, "Practical implementations of arithmetic coding", in Image and Text Compression, Storer, Ed., pp. 85–112, Kluwer Academic,1992.
11. Pennebaker W. B., Mitchell J. L., Langdon G. G., Arps R. B. "An Overview of the Basic Principles of the Q-Coder Adaptive Binary Arithmetic Coder" // IBM J. Research and Development. Vol. 32, N.6. P. 717-726, 1988.
12. ITU-T and ISO/IEC JTC 1, ITU-T Recommendation T.81 and ISO/IEC 10918-1 Digital Compression and Coding of Continuous-Tone Still Images, ITU-T Recommendation T.81 and ISO/IEC 10918-1 (JPEG), Sep. 1992.
13. ITU-T and ISO/IEC JTC 1, JPEG2000 Image Coding System: Core Coding System, ITU-T Recommendation T.800 and ISO/IEC 15444-1 (JPEG2000 Part 1), 2000.
14. L. Bottou, P. G. Howard, and Y. Bengio, "The Z-coder adaptive binary coder", Proc. IEEE Data Compression Conference, Snowbird (USA), pp. 13–22, 1998.
15. Detlev Marpe, Heiko Schwarz, Gabi Blättermann, Guido Heising, and Thomas Wiegand, "Context-Based Adaptive Binary Arithmetic Coding In JVT/H.26L", Image Processing Department, Heinrich-Hertz-Institute, Berlin, Germany.
16. ITU-T and ISO/IEC JTC 1, Advanced video coding for generic audiovisual services, ITU-T Recommendation H.264 and ISO/IEC 14496-10 (AVC), May 2003.
17. Detlev Marpe, Heiko Schwarz, and Thomas Wiegand, "Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard", IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 620–636, July 2003.
18. Eeckhaut, H.; Schrauwen, B.; Christiaens, M.; Van Campenhout, J. "Tuning the M-coder to improve Dirac's Entropy Coding". WSEAS transactions on Information Science and Applications. WSEAS. Vol. 2 (10). pp. 1563-1571, 2005.
19. Семенюк В.В., Метод адаптивного кодирования двоичной информационной выборки. Известия вузов. Приборостроение.-Т.47, Вып. 5.-С.36-41, 2004.
20. Detlev Marpe and Thomas Wiegand, "A Highly Efficient Multiplication-Free Binary Arithmetic Coder and its Application in Video Coding", Proceedings IEEE International Conference on Image Processing (ICIP '03), vol. II, pp. 263–266, Oct. 2003.

УДК 621.396.965.8

АВТОМАТИЗИРОВАННЫЙ СИНТЕЗ ЛИНЕЙНЫХ ФИЛЬТРОВ ДЛЯ ВОССТАНОВЛЕНИЯ СМАЗАННЫХ ИЗОБРАЖЕНИЙ

Мачихин А.С.

Введение

Формирование любого изображения сопровождается искажениями, которые вносят все звенья формирующей оптико-электронной системы (ФОЭС). Помимо линейных и шумовых искажений, свойственных реальным физическим устройствам, при регистрации могут возникать дополнительные факторы, негативно влияющие на качество изображения.

Один из таких факторов – скоростной сдвиг (смаз) изображения, причиной появления которого является относительное перемещение объекта и ФОЭС в процессе экспонирования (в дальнейшем, не нарушая общности, будем говорить о движущемся объекте относительно неподвижной ФОЭС). Действие данного фактора проявляется в наложении различных участков изображения, расположенных вдоль траектории этого перемещения. Восприятие и анализ смазанных изображений затрудняются, а эффективность их автоматической обработки значительно снижается.

Устранение смаза является разновидностью задачи восстановления изображения, которая формулируется как задача воссоздания исходного изображения $f(x, y)$ по наблюдаемому искаженному изображению $g(x, y)$.

Для пространственно инвариантной системы регистрации формирование изображения с учетом аддитивного шума $n(x, y)$ этой системы описывается выражением типа свертки

$$g(x, y) = f(x, y) \otimes h(x, y) + n(x, y) \quad (1)$$

где $h(x, y)$ - импульсный отклик ФОЭС.

Таким образом, задача восстановления изображения $f(x, y)$ сводится к задаче деконволюции, т.е. решению интегрального уравнения (1) типа свертки. Известно несколько способов решения задачи деконволюции: методы линейной фильтрации, итерационные алгоритмы, методы некорректных задач оптимизации и др. [1]. Для получения удовлетворительных результатов при использовании этих методов необходима априорная информация о характеристиках искажения и шума.

Возможности оценивания статистических характеристик шума как на основе экспериментальных исследований ФОЭС, так и непосредственно по искаженному изображению достаточно изучены [1]. Поэтому основной проблемой восстановления изображений является оценивание функции $h(x, y)$. В случае же, когда изображение искажено смазом, описываемым функцией $h_{см}(x, y)$, в дальнейшем называемой функцией смаза точки (ФСТ), произвольного вида, даже при наличии ФОЭС ни один из известных методов не позволяет эффективно оцени-

Рассмотрена проблема синтеза линейных фильтров для восстановления смазанных изображений, предложен метод моделирования функций смаза точки, приведены примеры восстановления изображений с помощью синтезированных фильтров.

вать параметры искажения в автоматизированном режиме. Поэтому, как правило, ФСТ определяют на основе визуального анализа искаженного изображения. Необходимость в этой длительной и трудоемкой процедуре, заключающейся в интерактивном подборе параметров ФСТ, существенно ограничивает возможности применения алгоритмов устранения смаза.

Данная работа посвящена разработке автоматизированного метода оценки и моделирования ФСТ произвольного вида при известных параметрах ФОЭС и условиях съемки.

Формирование смазанного изображения

Для определения ФСТ представим смазанное изображение $g_{см}(x, y)$ как свертку изображения $g(x, y)$, регистрируемого ФОЭС в отсутствие смаза, и ФСТ

$$g_{см}(x, y) = g(x, y) \otimes h_{см}(x, y). \quad (2)$$

В пространственно-частотной области соотношение (2) преобразуется к виду

$$G_{см}(v_x, v_y) = G(v_x, v_y) H_{см}(v_x, v_y). \quad (3)$$

С другой стороны, формирование смазанного изображения можно описать интегральным уравнением [2]

$$g_{см}(x, y) = \frac{1}{T} \int_0^T g(x - x_0(t), y - y_0(t)) dt, \quad (4)$$

где T – время экспонирования; $x_0(t)$, $y_0(t)$ - закон движения изображения объекта (ЗДИО).

После перехода в пространственно-частотную область и осуществления несложных преобразований выражение (4) принимает вид

$$G_{\bar{m}}(v_x, v_y) = \frac{1}{T} G(v_x, v_y) \int_0^T \exp(-i2\pi(v_x x_0(t) + v_y y_0(t))) dt \quad (5)$$

Из (3) и (5) следует, что функция $H_{см}(v_x, v_y)$, далее называемая функцией передачи смаза (ФПС), может быть записана как

$$H_{см}(v_x, v_y) = \frac{1}{T} \int_0^T \exp(-i2\pi(v_x x_0(t) + v_y y_0(t))) dt. \quad (6)$$

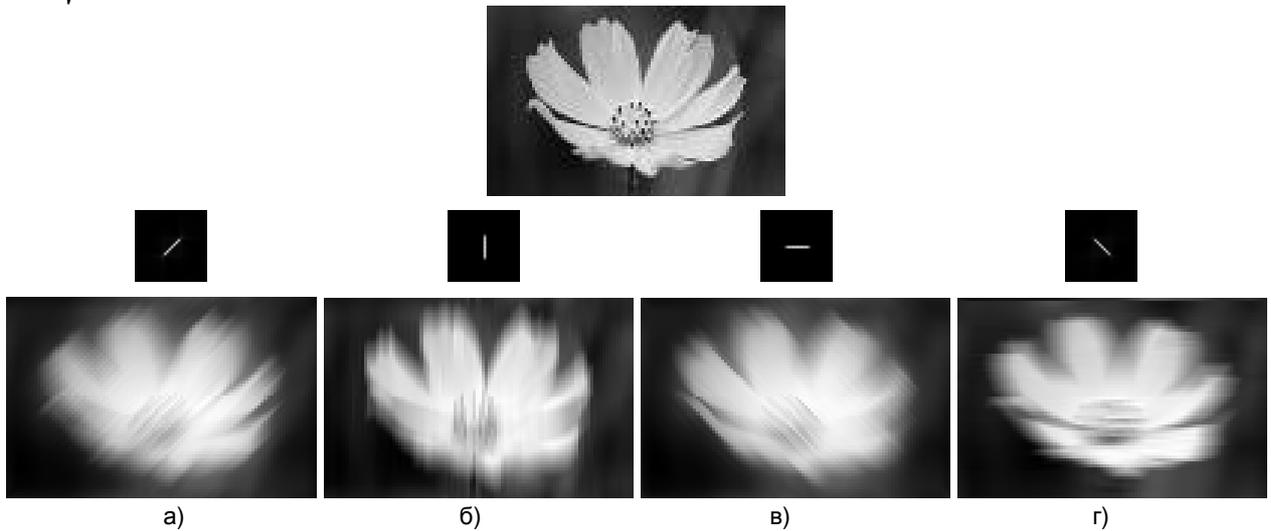


Рис. 1. Примеры моделирования ФСТ (в средней строке) для случая линейного равномерного смаза величиной 10пк с различными направлениями:

- а) $\theta = -45^0$ ($l_x = 7$ пк, $l_y = -7$ пк); б) $\theta = 90^0$ ($l_x = 0$, $l_y = 10$ пк); в) $\theta = 0^0$ ($l_x = 10$ пк, $l_y = 0$);
 г) $\theta = 45^0$ ($l_x = 7$ пк, $l_y = 7$ пк).

В нижней строке представлены примеры искажения изображения (в верхней строке) соответствующими ФСТ.

Взяв обратное преобразование Фурье от ФПС, получим выражение для ФСТ

$$h_{см}(x, y) = \frac{1}{T} \int_0^T \delta(x - x_0(t), y - y_0(t)) dt \quad (7)$$

где $\delta(x, y)$ – дельта-функция.

В отличие от моделей формирования смазанного изображения, представленных в [1,4] и справедливых только в случае равномерного прямолинейного смаза, выражения (6) и (7) корректно описывают смаз изображения при любом ЗДИО. Кроме того, общим недостатком моделей [1,4] является то, что они никак не учитывают параметры ФОЭС и условия съемки. Между тем, как вид ФСТ в каждом конкретном случае определяется в том числе и этими факторами.

Используя (6), получим аналитические выражения для ФПС в практически важных случаях смаза при равномерном прямолинейном и равноускоренном движениях объекта.

Равномерное прямолинейное движение объекта

В этом случае ЗДИО на цифровом изображении имеет вид

$$\begin{cases} x_0(t) = v_x t \\ y_0(t) = v_y t \end{cases}, \quad (8)$$

где v_x, v_y [пк/с] – проекции вектора \vec{v} скорости изображения объекта на координатные оси.

Подставляя (8) в (6), получаем

$$H_{см}(v_x, v_y) = \frac{1}{T} \int_0^T \exp(-i2\pi(v_x v_x + v_y v_y) t) dt, \quad (9)$$

откуда после несложных преобразований имеем

$$\begin{aligned} H_{см}(v_x, v_y) &= \text{sinc}(\pi(v_x v_x + v_y v_y) T) \exp(-i\pi(v_x v_x + v_y v_y) T) = \\ &= \text{sinc}(\pi(v_x l_x + v_y l_y)) \exp(-i\pi(v_x l_x + v_y l_y)) = \\ &= \text{sinc}(\pi l (v_x \cos \theta + v_y \sin \theta)) \exp(-i\pi l (v_x \cos \theta + v_y \sin \theta)) \end{aligned} \quad (10)$$

где $l_x = v_x T$, $l_y = v_y T$ – проекции траектории смаза, представляющей собой отрезок прямой линии, на координатные оси, $l = \sqrt{l_x^2 + l_y^2}$ – величина смаза; $\theta = \arctg(l_y / l_x)$ – угол, образуемый линией смаза с горизонтальной осью.

Выражение (10) позволяет задавать функцию передачи равномерного прямолинейного смаза его величиной и направлением - параметрами, которые могут быть оценены непосредственно по искаженному изображению. На практике это оказывается важным, т.к. далеко не всегда доступна априорная информация об условиях съемки и параметрах ФОЭС.

На рис. 1 представлены примеры моделирования ФСТ в случае прямолинейного равномерного смаза с различными параметрами, а также искажения реального изображения полученными ФСТ. Искажение производилось согласно (3) с последующим обратным преобразованием Фурье.

Равноускоренное движение объекта

В этом случае ЗДИО на цифровом изображении имеет вид

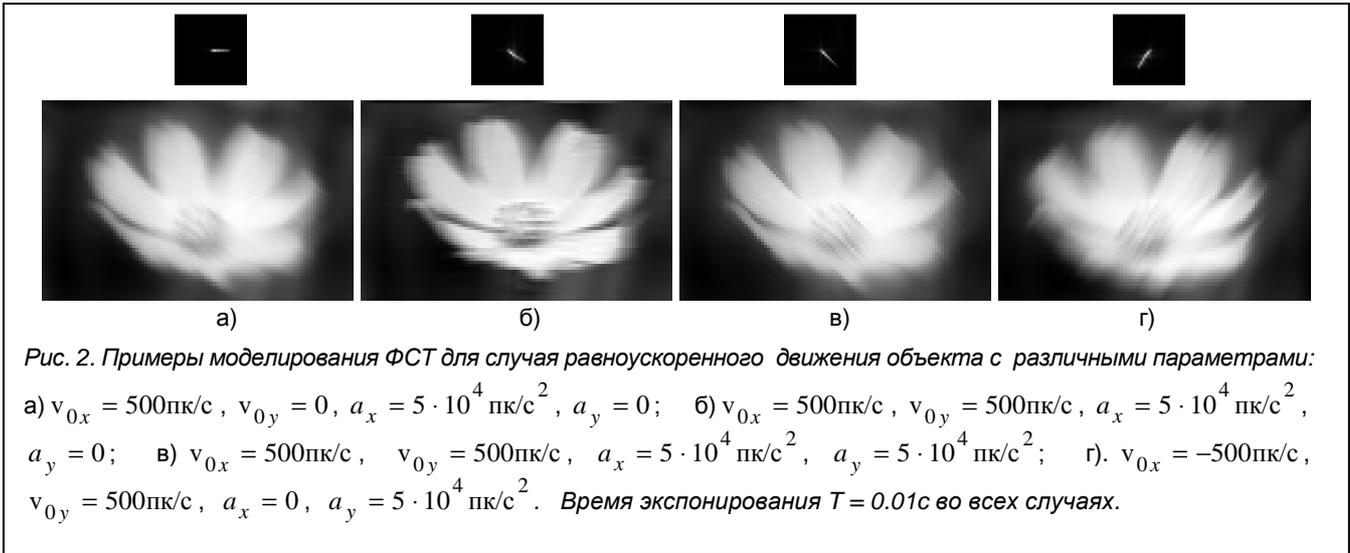
$$\begin{cases} x_0(t) = v_{0x} t + \frac{a_x t^2}{2} \\ y_0(t) = v_{0y} t + \frac{a_y t^2}{2} \end{cases}, \quad (11)$$

где v_{0x}, v_{0y} [пк/с] – проекции вектора \vec{v}_0 начальной скорости изображения объекта на координатные оси; a_x, a_y [пк/с²] – проекции вектора \vec{a} ускорения изображения объекта на координатные оси.

Подставляя выражения (11) в (6), после перегруппировки слагаемых под экспонентой получим

$$H_{см}(v_x, v_y) = \frac{1}{T} \int_0^T \exp(A(v_x, v_y) t - B(v_x, v_y) t^2) dt, \quad (12)$$

где для сокращения дальнейших выкладок и наглядности введены следующие обозначения:



$$A(v_x, v_y) = -i2\pi(v_{0x}v_x + v_{0y}v_y), \quad B(v_x, v_y) = i\pi(a_x v_x + a_y v_y).$$

Выделив в (12) под экспонентой полный квадрат и введя функцию ошибок

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-p^2) dp,$$

после несложных преобразований приведем (12) к виду

$$H_{\bar{m}}(v_x, v_y) = \frac{\sqrt{\pi}}{2\sqrt{B(v_x, v_y)}} \cdot \left(\text{erf} \left(\frac{1}{2} \frac{A(v_x, v_y)}{\sqrt{B(v_x, v_y)}} \right) - \text{erf} \left(\frac{1}{2} \frac{A(v_x, v_y)}{\sqrt{B(v_x, v_y)}} - T\sqrt{B(v_x, v_y)} \right) \right) \cdot \exp \left(\frac{1}{4} \frac{A(v_x, v_y)^2}{B(v_x, v_y)} \right). \quad (13)$$

Следует отметить, что выражение (13) справедливо только в случае $B(v_x, v_y) \neq 0$, т.е. $|\vec{a}| \neq 0$. В случае $|\vec{a}| = 0$, выражение (11) переходит в (8), и вместо (13) необходимо использовать (10).

Функция erf, представляющая собой бесконечный ряд, является стандартной для большинства современных математических пакетов (MathCad, Matlab и др.). Известны численные методы быстрого вычисления ее значений с заданной точностью. Поэтому выражение (13) можно считать окончательным и применять без дальнейших упрощений.

На рис. 2 представлены примеры моделирования ФСТ в случае равноускоренного смаза с различными параметрами, а также искажения реального изображения полученными ФСТ. Искажение производилось согласно (3) с последующим обратным преобразованием Фурье.

Синтез линейных частотных восстанавливающих фильтров

Известен ряд методов восстановления изображений и, в частности, устранения смаза [1-5]. Область применимости каждого из них определяется прежде всего требованиями к скорости и качеству восстанов-

ления. На практике, несмотря на невозможность экстраполяции спектра и учета ограничений на получаемые решения, наиболее широкое распространение приобрели линейные методы. Это объясняется простотой их реализации и высокой скоростью вычислений.

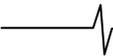
Суть линейных методов заключается в синтезе передаточной функции $H_{восст}(v_x, v_y)$ восстанавливающего линейного фильтра, так что восстановленное изображение $g(x, y)$ представляется в виде

$$g(x, y) = F^{-1} \{ G_{см}(v_x, v_y) H_{восст}(v_x, v_y) \}. \quad (14)$$

В работах [1–3] приведено множество линейных частотных восстанавливающих фильтров. Передаточная функция $H_{восст}(v_x, v_y)$ каждого из них зависит от ФПС и характеристик шума, который в большинстве практически важных случаев считают не коррелированным с изображением. Разумный компромисс между разрешением и зашумленностью восстановленного изображения достигается, как правило, интерактивным вмешательством пользователя в процесс восстановления посредством изменения некоторых числовых параметров.

Алгоритм (7) моделирования ФСТ допускает задание смаза несколькими способами: зависимостями $x_0(t)$ и $y_0(t)$ координат от времени и функциональной зависимостью координат $y_0(x_0)$ или массивом координат траекторных точек. Задание непосредственно $H_{см}(v_x, v_y)$, т.е. пространственно-частотного спектра функции $h_{см}(x, y)$, избавляет от необходимости осуществлять преобразование Фурье, быстрому варианту которого свойственны погрешности.

Использование полученных выражений (10) и (13), в которые не входит интеграл по времени, приводит к существенной экономии времени. Как показали исследования, в большинстве практически важных случаев время экспонирования можно разбить на интервалы, в течение которых движение объекта можно считать либо равномерным либо равноускоренным, т.е. разбить интеграл (6) на сумму интегралов вида (9) и (12).



Кроме того, на практике зачастую имеются данные об ЗДИО. Так, например, при аэрофотосъемке, как правило, с высокой точностью известны параметры используемой ФОЭС и параметры полета: высота, скорость носителя и др. Эти данные позволяют определить ЗДИО и использовать предлагаемый метод моделирования ФПС.

Указанные обстоятельства обеспечивают удобство и автоматизацию синтеза частотных восстанавливающих фильтров на основе описанного метода.

Программная реализация

Предлагаемый подход к синтезу частотных восстанавливающих смазанных изображения фильтров реализован в виде законченного программного модуля. Пользователю предоставляется возможность задавать ЗДИО, условия съемки, параметры ФОЭС, характеристики шума, а также вид используемой фильтрации: инверсная, винеровская, среднегеометрическая, эволюционная или регуляризация Тихонова. Возможно задание и использование других фильтров. Предусмотрено интерактивное вмешательство пользователя в ходе многовариантного анализа в процесс подбора передаточной функции восстанавливающего фильтра за счет изменения параметров, определяющих отношение сигнал-шум.

На рисунке 3 представлен пример восстановления изображения, искаженного смазом, траектория и неравномерность которого аппроксимировались равноускоренным движением (см. (11)) с параметрами: $v_{0x} = -1000$ пк/с; $v_{0y} = 0$; $a_x = 0$; $a_y = 10^5$ пк/с². Время экспонирования полагалось равным 0.01с. Шумовые

характеристики, необходимые для использования восстанавливающих фильтров, предварительно оценивались на основе специального экспериментального исследования ФОЭС. На основе моделирования ФПС (см. (13)) производилось восстановление изображения согласно (14) с помощью нескольких синтезированных фильтров:

1). Фильтра Винера (рис. 3б)

$$H_V(v_x, v_y) = \frac{H_{cm}^*(v_x, v_y)}{|H_{cm}(v_x, v_y)|^2 + \frac{P_n(v_x, v_y)}{P_f(v_x, v_y)}},$$

где $P_n(v_x, v_y)$ – энергетический спектр аддитивного шума ФОЭС, $P_f(v_x, v_y)$ – энергетический спектр исходного неискаженного изображения.

2). Фильтра, реализующего регуляризацию Тихонова (рис. 3в)

$$H_T(v_x, v_y) = \frac{H_{cm}^*(v_x, v_y)}{|H_{cm}(v_x, v_y)|^2 + \alpha Q(v_x, v_y)},$$

где

$$Q(v_x, v_y) = 10^{-12} + 10^{-13}(v_x^2 + v_y^2) + 10^{-14}(v_x^2 + v_y^2)^2,$$

$\alpha = 0.005$.

3). Эволюционного фильтра (рис. 3г)

$$H_E(v_x, v_y) = \left(\frac{1}{|H_{\bar{n}}(v_x, v_y)|} \right)^\sigma \left(\frac{|H_{\bar{n}}(v_x, v_y)|^2}{|H_{\bar{n}}(v_x, v_y)|^2 + \alpha Q(v_x, v_y)} \right)^\mu,$$

где $\sigma = 0.92$, $\mu = 0.4$.



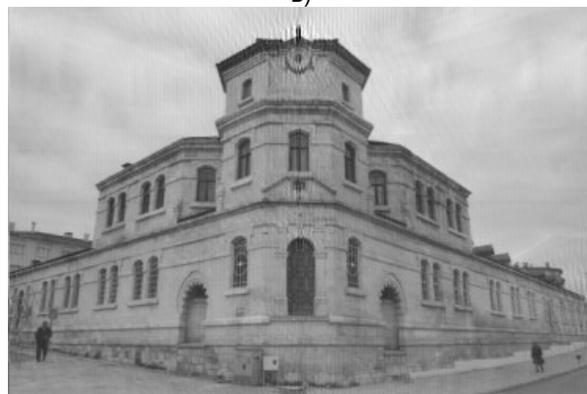
а)



в)



б)



г)

Рис. 3. Пример восстановления изображения (а), искаженного смазом, с помощью б). фильтра Винера, синтезированного автоматически; в). регуляризации Тихонова (значение α подбиралось интерактивно); г). эволюционного фильтра (значения α , σ , μ подбирались интерактивно).

Заключение

Предложен метод моделирования смаза и синтеза восстанавливающих смазанные изображения линейных частотных фильтров. Средством достижения удовлетворительного результата является использование максимума априорной информации о характере смаза и условиях регистрации. Описанный метод обладает рядом важных преимуществ:

- строгость аналитического решения и удобство использования в теоретических исследованиях;
- возможность моделирования и учета произвольного (криволинейного, неравномерного) смаза;
- учет условий съемки и параметров ФОЭС;
- высокая скорость вычислений.

Использование данного метода на этапе предобработки позволяет расширить возможности современных систем обработки и распознавания изображений. При

достаточной априорной информации об условиях регистрации и параметрах ФОЭС возможно его применение в режиме, близком к режиму реального времени.

Литература

1. Василенко Г.И., Тараторин А.М. Восстановление изображений. – М.: Радио и связь, 1986.
2. Прэтт У. Цифровая обработка изображений. – М.: Мир, 1982.
3. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005.
4. Сизиков В.С. Устойчивые методы обработки результатов измерений. – С.-П.: Специальная литература, 1999.
5. Колючкин В.Я., Мачихин А.С. Модифицированный итерационный алгоритм восстановления изображений. //Вестник МГТУ. Сер. Приборостроение. – 2007. – №1. С. 114-121.



ИТОГОВЫЙ ПРЕСС-РЕЛИЗ 5-ой международной выставки по электронике, микроэлектронике и компонентам

С 3 по 5 октября 2007 года в Москве в Экспоцентре на Красной Пресне успешно прошла 5-я международная выставка по электронике, оборудованию, компонентам, технологиям «ChipEXPO-2007», организованная компанией «ЧипЭКСПО» при поддержке Министерства промышленности и энергетики РФ, Правительства Москвы, Федерального агентства по промышленности и Московской торгово-промышленной палаты.

Международная выставка «ChipEXPO-2007» является ведущей российской выставкой по электронике. Выставка объединила в своей работе представителей министерств и федеральных агентств, руководителей и специалистов российских и зарубежных предприятий, общественных организаций, журналистов ведущих отраслевых и массовых изданий. В этом году она собрала 275 предприятий, работающих в области микроэлектроники, электронных компонентов, технологий, измерительной техники и оборудования для электронной промышленности из России, Украины, Белоруссии, Нидерландов, Финляндии, Австрии, Великобритании, Германии, Гонконга, США, Чехии и Тайваня. В числе участников выставки ведущие компании отрасли, производители оборудования и компонентов, дистрибьюторы, средства массовой информации. За 3 дня работы выставка привлекла более 12 000 специалистов. Более 60 СМИ освещали работу выставки.

В церемонии торжественного открытия юбилейной выставки «ChipEXPO-2007» приняли участие представители Министерства промышленности и энергетики Российской Федерации, Министерства обороны Российской Федерации, Управления радиоэлектронной промышленности и систем управления Федерального агентства по промышленности, Департамента науки и промышленной политики города Москвы, Префектур административных округов города Москвы, а также руководители ведущих предприятий отрасли. Участники церемонии торжественного открытия выставки отметили ведущую роль мероприятия, как традиционной площадки для профессионального общения специалистов и создания новых отношений с зарубежными партнерами.

Центральным событием деловой программы выставки стала Пленарная сессия Министерства промышленности и энергетики Российской Федерации «Годовые итоги и перспективы развития электронной промышленности России».

В дни работы выставки были проведены семинары «Российский рынок электронных компонентов» и «Как снизить риск потерь времени и денег при разработке бренда для радиоэлектронной промышленности», а также конференция «Современная наноэлектроника» и ряд технических семинаров.

Состоялся традиционный конкурс на присуждение отраслевой премии «Золотой Чип». Компетентное Жюри определило победителей конкурса. В номинации «За развитие российской электроники» Диплом первой степени получил ФГУП «НПП Исток». В номинации «Лучшее технологическое решение» первое место получил ОАО «Авангард», а в номинации «Лучшее изделие микроэлектроники 2006-2007 гг.» ЗАО «НТЦ Модуль».

Выставки «ChipEXPO» - это всегда яркие, интересные мероприятия, насыщенные событиями, важными для отрасли, эффективными деловыми встречами и переговорами. Высокий интерес к мероприятию вызван тем, что выставка «ChipEXPO» в полном объеме отражает пути развития электронной промышленности России и способствует созданию условий для формирования рынка перспективной наукоемкой продукции и интеграции отечественных предприятий в мировой рынок электроники.

ЗАО «ЧипЭКСПО», Тел: (495) 221-5015
info@chipexpo.ru , sa@chipexpo.ru
<http://chipexpo.chipexpo.ru/>



ЦИФРОВАЯ ОБРАБОТКА ТЕЛЕВИЗИОННЫХ ИЗМЕРИТЕЛЬНЫХ СИГНАЛОВ

Макаров Д.Г.

Введение

В условиях значительного роста сети ТВ вещания, разработки и внедрения перспективных телевизионных систем все большее значение приобретают действенные методы контроля и измерений телевизионного оборудования в процессе его разработки, в условиях массового производства и эксплуатации. Основная идея алгоритмов обработки ИС, состоящих из последовательностей прямоугольных импульсов, серий синусоидальных колебаний, импульсных сигналов специальной формы и др., заключается в анализе искажений этих сигналов, зависящих от линейных и нелинейных характеристик ТВ каналов, наличия в них флуктуационных, синусоидальных и импульсных помех.

Основной задачей при разработке алгоритмов обработки ТВ измерительных сигналов является минимизация погрешности оценки параметров, характеризующих искажения ТВ ИС, которые разделяют на:

методические (теоретические), заложенные в самих алгоритмах (методе) измерения данной величины;

инструментальные (аппаратные), характеризующиеся схемой, конструкцией, состоянием измерительных приборов в процессе эксплуатации;

внешние, связанные с внешними условиями измерений, а также возникающие из-за изменения состояния объекта в процессе измерений.

Разрабатываемая система измерений параметров ТВ каналов должна обеспечить минимизацию всех указанных погрешностей оценки параметров измерительных сигналов.

Измерение постоянных уровней ИС

Для часто встречающихся на практике флуктуационных помех с нормальным законом распределения энтропийное значение погрешности определяется соотношением:

$$Entr \cong \pm 2,07k\sigma_n$$

где σ_n - эффективное напряжение помехи, k - коэффициент, характеризующий изменение уровня помехи при обработке сигналов.

Чем меньше величина k , тем эффективнее фильтрация сигналов и цифровая обработка их дискретных отсчетов и результатов измерений.

Так, при оценке постоянных уровней сигналов фильтрация, наилучшим образом используя различные корреляционные свойства сигнала и помехи, в классе линейных оценок наиболее эффективно уменьшает влияние помех в k_α раз:

$$k_\alpha = \sqrt{\frac{\int_0^{f_{\max}} K^2(f)G(f)df}{\int_0^{f_{\max}} G(f)df}}$$

Данная статья посвящена вопросам реализации эффективных алгоритмов цифровой обработки измерительных сигналов (ИС), используемых для оценки качественных показателей телевизионных (ТВ) систем формирования и передачи ТВ изображения.

где $G(f)$ - спектральная плотность мощности помехи, $K(f)$ - АЧХ фильтра, f_{\max} - граничная частота спектра помехи.

Минимальное значение k_α может быть получено при использовании оптимального фильтра, характеристика которого определяется допустимой длительностью переходного процесса, спектральной плотностью мощности помехи, а также свойствами сигнала. Для обеспечения высокой точности оценки постоянного уровня сигналов при наличии флуктуационных помех специальные ИС в области измерения должны обладать максимально возможным числом равных нулю четных производных [1]. При помехе с равномерным спектром частотная характеристика фильтра должна иметь вид:

$$K(f) = 1 + \sum_{m=1}^M (-1)^m \left(\frac{f}{f_{\max}} \right)^{2m}, \quad f \leq f_{\max},$$

где M - число равных нулю четных производных,

$$\sum_{m=1}^M (-1)^{m-1} \frac{b_{2m}}{2m+2k+1} = \frac{1}{2k+1}, \quad 1 \leq k \leq M \text{ и}$$

$$k_\alpha = 1 + \sum_{m=1}^M (-1)^m \frac{b_{2m}}{2m+1}.$$

В Таблице 1 приведены несколько значений коэффициента уменьшения эффективного напряжения помехи k_α в зависимости от числа M .

Таблица 1

M	0	1	2	3	4
k_α	1	0,444	0,284	0,209	0,165

Для оценки характерных уровней ТВ ИС и результатов измерений параметров ТВ каналов при наличии флуктуационных и импульсных помех весьма эффективным средством существенного уменьшения их влияния является использование порядковых статистик [2].

При этом накапливающаяся последовательность результатов измерений одной и той же величины X_1, X_2, \dots, X_n преобразуется в порядковую статистику $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$. Такое упорядочивание выборки не изменяет объема информации, однако часто позволяет сократить объем данных для вычисления и уменьшения мешающих факторов за счет оптимального выбора данных при обработке выборочных значений.

В алгоритм функционирования системы анализа данных заложено использование уинсоризованного среднего [2]:

$$\bar{X}_n(p) = \frac{1}{n} \left[(r+1)(X_{r+1} + X_{n-r}) + \sum_{i=r+2}^{n-r-1} X_i \right], \quad 0 < r < (n-1)/2,$$

При нормальном законе распределения результатов измерений и $n \geq 10$ достоверность такой оценки весьма высока (более 0,95). В некоторых случаях при больших объемах выборки другой удобной оценкой, обладающей значительной робастностью относительно аномальных результатов, является медианная оценка:

$$\bar{X}_M = \begin{cases} X_{\frac{n+1}{2}}, & n - \text{нечетное}; \\ (X_{\frac{n}{2}} + X_{\frac{n}{2}+1})/2, & n - \text{четное}. \end{cases}$$

При нормальном законе распределения эффективность этой оценки асимптотически стремится к величине $2/\pi \approx 0,637$.

Основную информацию об энтропийном отклонении результатов измерений несут крайние порядковые статистики, и отбрасывание экстремальных значений, обеспечивая устранение «загрязнения» выборки, снижает эффективность оценки энтропии. В этом плане при нормальном законе распределения особый интерес имеет оценка [2, 3]:

$$Entr(\bar{X}) = \bar{\sigma} \sqrt{\frac{\pi e}{2}},$$

$$\text{где } \bar{\sigma} = \frac{\sqrt{\pi}}{n(n-1)} \sum_{i=1}^{[n/2]} (n-2i+1)(X_{n+1-i} - X_i).$$

Эта оценка весьма эффективна (более 0,9779) и в меньшей степени подвержена влиянию аномальных результатов изменений, чем стандартная оценка [3].

Фильтрация элементов ИС и оценка их параметров

Фильтрация компонент ТВ сигнала, используемых при анализе его параметров и отдельных элементов ТВ ИС, осуществляется путем перемножения формируемого с использованием быстрого преобразования Фурье (БПФ) спектра сигнала на спектральное окно Тьюки [4] с характеристикой, определяемой следующим соотношением:

$$K(f) = \begin{cases} \sin^2 \frac{\pi(f-f_1)}{2(f_2-f_1)}, & f_1 \leq f \leq f_2; \\ 1, & f_2 \leq f \leq f_3; \\ \cos^2 \frac{\pi(f-f_3)}{2(f_4-f_3)}, & f_3 \leq f \leq f_4. \end{cases}$$

Вне пределов указанных интервалов $K(f) = 0$. Если используется фильтр нижних частот, то величины

$f_1 = f_2 = 0$. При измерении постоянных уровней сигнала применяется фильтр, в котором также $f_3 = 0$, что соответствует равенству нулю второй и четвертой производных сигнала в области измерения ($M=2$).

Отфильтрованный сигнал формируется с помощью обратного преобразования Фурье (ОБПФ).

Выделение огибающей синусоидального сигнала $u(t) = U \cos(\omega t + \varphi)$ производится путем определения его Гильбертовой составляющей $\Gamma[u(t)] = U \sin(\omega t + \varphi)$ и вычисления амплитуды по формуле:

$$U = \sqrt{(u(t))^2 + (\Gamma[u(t)])^2}.$$

Для получения огибающей синусоидальных сигналов ТВ ИС применяется алгоритм цифровой обработки, блок-схема которого приведена на рис. 1.

Для выделения гильбертовой составляющей сигнала его спектр умножается на величину $\exp(i\pi/2)$, что означает трансформацию действительной составляющей БПФ в мнимую, а мнимой – в инвертированную действительную составляющую. Затем производится ОБПФ.

Преобразование Гильберта используется также при оценке фазовых соотношений в синусоидальных сигналах – дифференциальной фазы цветовой поднесущей. Первоначально в каждой дискретной точке определяется значение фазы:

$$\varphi(k) = \text{Arctg} \frac{\Gamma[u(k)]}{u(k)}, \quad 1 \leq k \leq N,$$

затем рассчитывается фазовая характеристика:

$$\bar{\varphi}(k) = \varphi(k) + 2\pi r,$$

причем целое число r подбирается таким образом, чтобы $\bar{\varphi}(k) > \bar{\varphi}(k-1)$.

Предположим, что изменение фазы должно оцениваться относительно фазы поднесущей в области дискретного отсчета $k = K$. Тогда значение фазы и среднее значение изменения фазы в этой точке определяются в области $K - k_0 \leq k \leq K + k_0$ формулами:

$$\hat{\varphi}(K) = \frac{1}{k_0 + 1} \left[\bar{\varphi}(K) + \sum_{n=1}^{k_0} \frac{\varphi(K-n) + \varphi(K+n)}{2} \right],$$

$$\Delta\varphi = \frac{1}{2k_0} \sum_{n=K-k_0+1}^{K+k_0} [\varphi(n) + \varphi(n-1)],$$

а характеристика дифференциальной фазы определяется по формуле:

$$\varphi_d(k) = \bar{\varphi}(K) - \hat{\varphi}(K) - \Delta\varphi(k - K), \quad K \leq k \leq N.$$

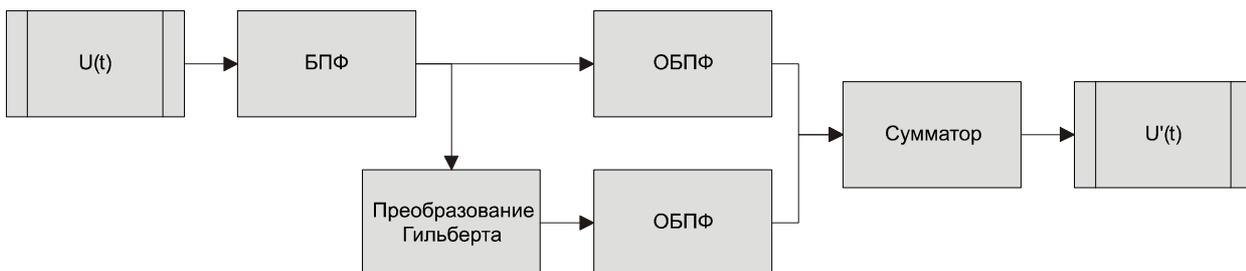
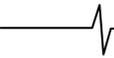


Рис. 1. Блок-схема формирования «пакета» синусоидальных колебаний



При вычислении частоты сигнала помимо его преобразования по Гильберту анализируются также спектры производных сигнала и составляющей Гильберта. Оценка частоты после ОДПФ осуществляется по формуле:

$$f(k) = \frac{1}{2\pi} \sqrt{\frac{(u'(k))^2 + (\Gamma[u'(k)])^2}{(u(k))^2 + (\Gamma[u(k)])^2}},$$

где $u'(k)$, $\Gamma[u'(k)]$ - производная сигнала и ее преобразование по Гильберту соответственно.

Поиск экстремальных значений сигналов

Поиск экстремальных значений и значений сигналов в иных характерных точках ТВ ИС, и их местоположения осуществляется с использованием передискретизации – увеличения числа дискретных отсчетов на заданном интервале. С этой целью производится БПФ, спектральные компоненты в области частот выше $F(N/2)$ дополняются нулевыми компонентами, так что общее их число $N_1 \gg N$, а затем выполняется операция ОБПФ.

Дальнейшее уточнение положений и значений уровней ТВ ИС характерных точек производится путем соответствующих аппроксимаций непрерывной функции между близко расположенными дискретными отсчетами.

В ряде случаев удобно анализировать экстремальные значения сигналов непосредственно по его дискретным отсчетам. Этот вариант особенно эффективен при обработке серий синусоидальных сигналов с малым числом периодов, располагаемых на постоянном уровне.

В данном случае ИС можно представить в виде:

$$u(t) = A \cos(\omega_0 t + \varphi) + B, \quad t \geq 0,$$

где A и ω_0 - соответственно амплитуда и круговая частота синусоидального сигнала, $\varphi = \omega_0 t_0$ - начальная фаза синусоидальных колебаний, t_0 - начальная точка отсчета сигнала, B - постоянный уровень сигнала.

Если дискретизация сигнала осуществляется через интервал Δt , то при указанной синхронизации импульсов дискретизации от измерительного сигнала его дискретные отсчеты можно представить в виде:

$$x_n = A \cos\left[(n-1)\Delta x + \varphi - \frac{\xi_{1n}}{A}\right] + B + \xi_{2n},$$

где $\Delta x = \omega_0 \Delta t$, ξ_{1n} - дискретное значение помехи, определяющей изменение положения импульсов дискретизации за счет ее проникновения в канал синхронизации, ξ_{2n} - дискретное значение аддитивной помехи в точке дискретизации. Величины дискретных значений помех ξ_{1n} и ξ_{2n} могут содержать как флуктуационную, так и импульсную составляющие.

Алгоритм оценки размаха синусоидального колебания можно разделить на два этапа:

- поиск положений экстремальных значений синусоидального сигнала с одновременной оценкой его размаха;
- оценка разности экстремальных значений сигнала в характерных точках, определенных на первом этапе, с одновременной оценкой изменений измеряемого сигнала.

Алгоритм поиска положений экстремальных значений сигнала с одновременной оценкой его размаха заключается в вычислении преобразованных значений сигнала:

$$\bar{X}_N = \sum_{n=1}^N a_n x_n,$$

где a_n - коэффициенты взвешивания отсчетов, обеспечивающие максимальное подавление помехи при выполнении некоторых условий, характеризующих точную оценку экстремумов сигнала.

В случае, если дискретные отсчеты сигнала располагаются симметрично относительно экстремума, условиями безыскаженной его оценки являются:

$$\sum_{n=1}^N a_n = 1 \quad \text{и} \quad \sum_{n=1}^N a_n \cos\left[\left(n - \frac{N-1}{2}\right) \cdot \Delta x\right] = 1.$$

Первое из этих условий гарантирует безыскаженную оценку постоянной составляющей сигнала B , а второе - экстремума синусоидальной составляющей.

Минимизация дисперсии помехи обеспечивается при

$$\sum_{n=1}^N a_n^2 \Rightarrow \min$$

и соответствует значениям

$$a_n = \alpha_1 + \alpha_2 \cos\left(n - \frac{N+1}{2}\right) \cdot \Delta x, \quad 1 \leq n \leq N,$$

где коэффициенты α_1 и α_2 соответственно равны

$$\alpha_1 = \frac{\sin N\Delta x}{N \sin \Delta x}; \quad \alpha_2 = \frac{\sin(N\Delta x/2)}{N \sin(\Delta x/2)}.$$

При этом коэффициент уменьшения дисперсии помехи составляет

$$\mu_N = \lambda_1 + \lambda_2 = \frac{1}{N} \cdot \theta_N,$$

$$\text{где } \lambda_1 = \frac{1 + \alpha_1 - 2\alpha_2}{N(1 + \alpha_1 - 2\alpha_2^2)}, \quad \lambda_2 = \frac{2(1 - 2\alpha_2)}{N(1 + \alpha_1 - 2\alpha_2^2)},$$

$$\theta_N = \frac{3 + \alpha_1 - 4\alpha_2}{1 + \alpha_1 - 2\alpha_2^2} - \text{эффективность оценки величины}$$

преобразованного сигнала \bar{X}_N в точке

$$t_N = t_0 + \frac{N+1}{2} \cdot \Delta t.$$

Очевидно, если эта точка совпадает с положением t^* максимума сигнала (1), то, учитывая соотношения (2) и (3), при отсутствии помех $\bar{X}_N = A + B$. Если же эта точка совпадает с положением минимума, то $\bar{X}_N = B - A$.

Из соотношения для a_n следует, что

$$a_n = a_{N-n+1}, \quad 1 < n < \frac{N+1}{2}.$$

Важным вопросом является выбор величины шага дискретизации Δt , определяющего методическую погрешность оценки экстремума. Учитывая, что максимальный сдвиг точки t_N относительно экстремума составляет

$$|t_N - t^*| \leq \Delta t / 4,$$

максимально возможная относительная погрешность оценки величины A составляет

$$\Delta_A \leq \frac{(\Delta x)^2}{16} \cdot 100\%.$$

Задавшись величиной $\Delta x = 0,2$, получим величину относительной погрешности анализа размаха (амплитуды) сигнала $\Delta_A \leq 0,25\%$.

Величины коэффициентов

$$a_1, a_2, \dots, a_{\left[\frac{N+1}{2}\right]} \quad \text{для } 5 \leq N \leq 16$$

приведены в таблице 2, в которой приведена также величина коэффициента подавления помехи μ_N .

Таблица 2

Коэффициенты взвешивания отсчетов при поиске экстремумов

N	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	μ_N
5	-0,0865	0,3427	0,4876						0,4878
6	-0,0948	0,2177	0,3771						0,3972
7	-0,0965	0,1409	0,2872	0,3368					0,3368
8	-0,0951	0,0910	0,2193	0,2848					0,2930
9	-0,0923	0,0571	0,1683	0,2368	0,2602				0,2602
10	-0,0888	0,0334	0,1296	0,1960	0,2298				0,2991
11	-0,0852	0,0163	0,0999	0,1620	0,2003	0,2134			0,2134
12	-0,0816	0,0038	0,0766	0,1340	0,1735	0,1937			0,1937
13	-0,0781	-0,0055	0,0583	0,1107	0,1498	0,1738	0,1820		0,1820
14	-0,0746	-0,0125	0,0436	0,0914	0,1290	0,1550	0,1681		0,1681
15	-0,0714	-0,0179	0,0316	0,0752	0,1110	0,1376	0,1540	0,1598	0,1598
16	90,0683	-0,0219	0,0219	0,0615	0,0953	0,1218	0,1402	0,1495	0,1495

Таблица 3

Эквивалентное число отсчетов алгоритма поиска экстремума синусоидального сигнала

N	5	6	7	8	9	10	11	12	13	14	15	16
$N_{\text{экв}}$	4,5	5,6	6,6	7,6	8,6	9,7	10,4	11,5	12,2	13,2	13,9	14,9

Начиная с числа $N = 7$, производится расчет последовательности трех отсчетов преобразованного сигнала $\bar{X}_N, \bar{X}_{N-1}, \bar{X}_{N-2}$ и осуществляется расчет относительных величин изменения этой последовательности:

$$F_{N1} = \frac{\bar{X}_N}{\bar{X}_{N-1}} - 1 \text{ и } F_{N2} = \frac{\bar{X}_{N-1}}{\bar{X}_{N-2}} - 1.$$

В случае $F_{N1} \geq 0, F_{N2} < 0$ максимум сигнала соответствует величине:

$$\bar{A}_{\text{max}} = \bar{X}_{N-1} \approx A + B.$$

При $F_{N1} \leq 0, F_{N2} > 0$ имеет место минимум сигнала:

$$\bar{A}_{\text{min}} = \bar{X}_{N-1} \approx A - B.$$

Эквивалентное число отсчетов, определяющее эффективность поиска экстремумов, приведено в таблице 3.

Из этой таблицы следует, что $N_{\text{экв}} \sim N - 1$.

Положение максимума или минимума сигнала может быть уточнено по формуле:

$$x^* = \frac{1}{2} \left[N - 2 \pm \left(\frac{1}{4} - \frac{2F_{N1}}{\Delta x^2} \right) \right] \Delta x,$$

где знак «+» соответствует случаю, когда при поиске максимума $\bar{X}_N > \bar{X}_{N-2}$, а при поиске минимума $\bar{X}_N < \bar{X}_{N-2}$; в противном случае используется знак «-».

Анализ параметров флуктуационных и синусоидальных помех

Измерение эффективного напряжения флуктуационных помех и размахов синусоидальных помех производится также с помощью БПФ уровня одной из строк кадра, в которой измерительный сигнал либо отсутствует, либо плавно изменяется, занимая спектр в низкочастотной области (менее 50 кГц). При анализе флуктуационных помех и синусоидальных помех полоса частот ниже 50 кГц отфильтровывается. На рис.2 приведены примерный вид ТВ ИС для анализа и спектр данного участка ИС с выделенными помехами.

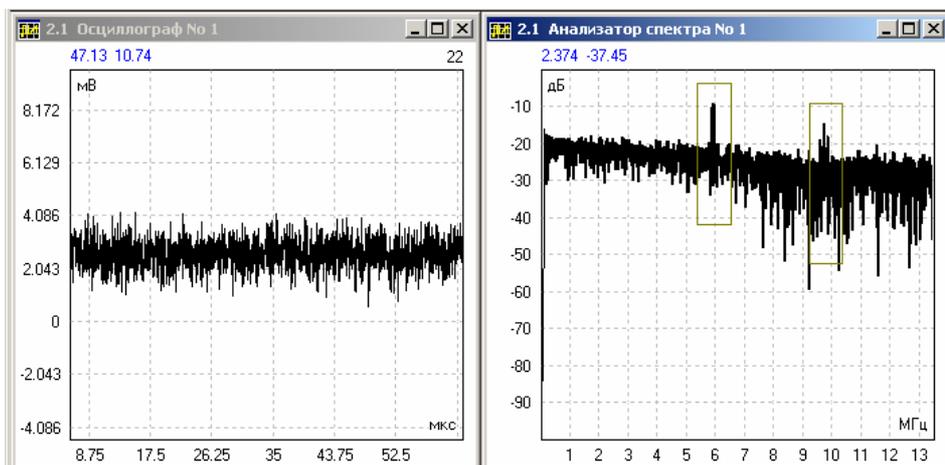


Рис. 2. Форма помехи на постоянном уровне сигнала и ее спектр

Для анализа спектра помех при преобразовании используется весьма эффективная оконная функция [5]:

$$Y(t) = a_0 + 2 \sum_{n=1}^2 (-1)^n a_n \cos \frac{2\pi n t}{T}, \quad 0 \leq t \leq T,$$

где T – интервал обработки,

$$a_0 = \frac{1}{\sqrt{6}}, \quad a_1 = \frac{1}{4}, \quad a_2 = \frac{1}{4} - \frac{1}{2\sqrt{6}}.$$

Выделение модулей спектральных компонент синусоидальных сигналов $F(n_1), F(n_2), \dots, F(n_N)$ осуществляется путем оценки величин, превышающих уровень $\bar{F} = F_{cp} + 3,5\sigma$,

$$\text{где } F_{cp} = \frac{1}{K_2 - K_1 + 1} \sum_{k=K_1}^{K_2} F(k);$$

$$\sigma = \sqrt{\frac{1}{(K_2 - K_1 + 1)(K_2 - K_1)} \sum_{k=K_1}^{K_2} [F(k) - F_{cp}]^2},$$

K_1 и K_2 – числа, определяющие анализируемую область частот.

Точная оценка параметров синусоидальной помехи осуществляется по уровню максимальной спектральной компоненты, например, $F(n_1)$ и уровням соседних с ней спектральных компонент $F(n_1 - 1)$ и $F(n_1 + 1)$. Расчет частоты f_{s1} и амплитуды U_{s1} синусоидальной помехи производится по формулам:

$$f_{s1} = (n_1 + \Delta) \cdot f_{дл} / N,$$

$$U_{s1} = F(n_1) + [F(n_1 + 1) - F(n_1 - 1)] \cdot \Delta / 4,$$

$$\text{где } \Delta = \frac{F(n_1 + 1) - F(n_1 - 1)}{2[F(n_1) - F(n_1 + 1) - F(n_1 - 1)]}.$$

При оценке мощности флуктуационных помех производится суммирование мощностей всех спектральных компонент (невзвешенных или взвешенных) в анализируемой области частот за вычетом компонент, соответствующих выявленным синусоидальным помехам. При расчете учитывается коэффициент пропорциональности, определяемый используемой оконной функцией и равный

$$R = \frac{a_0^2 + 2(a_1^2 + a_2^2)}{[a_0 + 2(a_1 + a_2)]^2}.$$

Результаты разработки алгоритмов цифровой обработки ТВ ИС и анализа качества передаваемой информации позволили реализовать прецизионные цифровые измерения качественных показателей каналов аналогового, аналого-цифрового и цифрового телевидения в видеоанализаторах компьютерных ВК-1 и ВК-2 [6-9].

Видеоанализаторы зарегистрированы в Государственном реестре средств измерений и обеспечивают создание информационно-измерительных систем аппаратно-студийных комплексов ТВ центров и систем трансляции ТВ программ. Прибор ВК-2 реализует функции 12 виртуальных измерителей.

На рис. 4 приведен протокол результатов измерений телевизионного канала непосредственно в процессе передачи программ по сигналам испытательных строк. На рис.5 изображены диаграммы виртуального осциллографа и спектрального анализатора, а на рис.6 приводятся диаграммы измерений цветных полос систем SECAM и PAL.

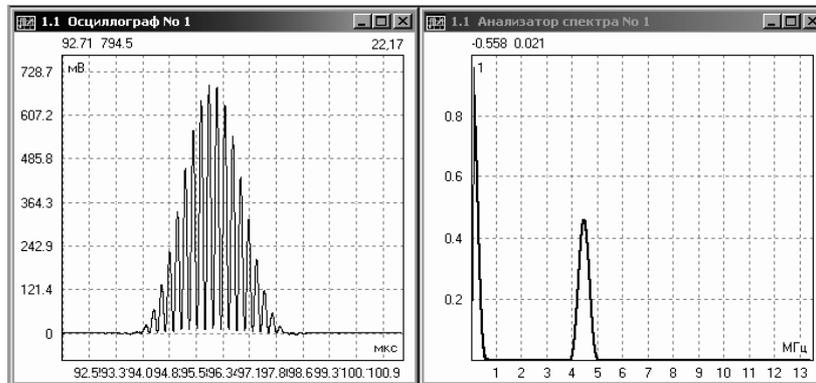


Рис. 5. Диаграммы осциллографа и спектрального анализатора.

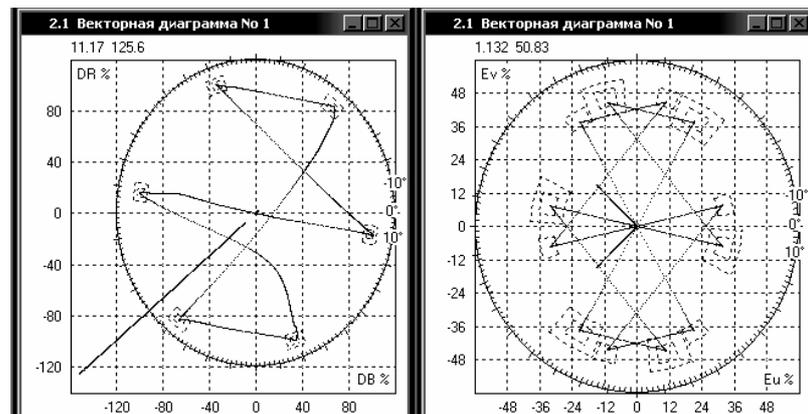


Рис. 6. Векторные диаграммы сигналов цветных полос по стандартам SECAM и PAL.

Измерительное устройство:	№ 1	Время измерения:	12 час. 24 мин.
Номер входа устройства:	1	Дата измерения:	6 июля 2006 г.
Тип устройства:	КИ-ТВ	Дата поверки:	6 июля 2006 г.
Серийный номер устройства:	083	Очередная дата поверки:	6 июля 2008 г.

Измерение по испытательным строкам. SECAM

Параметр		Отклонение	Примечание
<u>Отклонения основных уровней</u>			
Синхроимпульс	СИ	-0.0 %	Норм.
Импульс В2	ИБ	-0.0 %	Норм.
Цветовая поднесущая	ЦВП	-0.0 %	Норм.
Размах немодулированной поднесущей DR	ЦВPR	0.0 %	Норм.
Размах немодулированной поднесущей DB	ЦВPB	-0.0 %	Норм.
<u>Линейные искажения</u>			
Перекося вершины импульса В2	ПБ	-0.0 %	Норм.
Тянущееся продолжение импульса В2	ТП	-0.0 %	Норм.
К-фактор импульса В1	КП	0.3 %	Норм.
Относительное отклонение импульса В1	2Т	0.0 %	Норм.
Различие усиления яркости и цветности	РУ	0.0 %	Норм.
Различие усиления по элементу F	РУ1	-0.0 %	Норм.
Расхождение во времени яркости и цветности	РВ	-0.2 нс	Норм.
<u>Амплитудно-частотная характеристика</u>			
АЧХ 0.5 МГц	П1	0.0 %	Норм.
АЧХ 1.0 МГц	П2	-0.0 %	Норм.
АЧХ 2.0 МГц	П3	0.0 %	Норм.
АЧХ 4.0 МГц	П4	0.0 %	Норм.
АЧХ 4.8 МГц	П5	0.0 %	Норм.
АЧХ 5.8 МГц	П6	0.0 %	Норм.
<u>Нелинейные искажения</u>			
Нелинейность яркости	НЯ	0.0 %	Норм.
Нелинейность цветности	НЦ	0.0 %	Норм.
Влияние цветности на яркость	ЦЯ	-0.0 %	Норм.
Дифференциальное усиление	ДУ	0.0 %	Норм.
Дифференциальная фаза	ДФ	-0.0 °	Норм.
<u>Отношение ТВ сигнала к помехе</u>			
Флуктуационная помеха	ФП	67.2 дБ	Норм.
Взвешенная флуктуационная помеха	ФПв	> 75. дБ	Норм.
Взвешенная флукт. помеха с фильтром 0.33 мкс	ФП33	> 75. дБ	Норм.
Синусоидальная помеха 1	СП1	---	
Частота синусоидальной помехи 1	fСП1	---	
Синусоидальная помеха 2	СП2	---	
Частота синусоидальной помехи 2	fСП2	---	
Фоновая помеха	СФ	> 70. дБ	Норм.
<u>Качество ТВ канала</u>			
Качество ТВ канала	Q	5.0	Норм.

Рис. 4. Протокол результатам измерений параметров ТВ канала.

Литература

1. Дворкович В.П. К вопросу о дискретной и непрерывной линейной фильтрации постоянного уровня видеоимпульсов // Радиотехника и электроника. - 1979. - Т. XXIV, № 11. - С. 2278-2289.
2. Дейвид Г. Порядковые статистики. - М.: Наука, 1979. - 119 с.
3. Дворкович В.П., Кривошеев М.И. Особенности измерения и контроля в цветном телевидении (Учебное пособие). - М.: Изд. ГКТР, 1981. - 131 с.
4. Хэррис Ф.Дж. Использование окон при гармоническом анализе методом дискретного преобразования Фурье // ТИИЭР. - 1978. - Т. 66, №1. - С. 60-96.
5. Патент 1518924 (РФ). Способ измерения линейных характеристик канала связи / В.П. Дворкович // Б.И. - 1993.
6. Дворкович В.П., Дворкович А.В., Макаров Д.Г. Высокоч-

ные измерения параметров и качественных показателей телевизионного канала // Метрология и измерительная техника в связи. - 1998. - № 3. - С. 17-21.

7. Видеоанализатор компьютерный ВК-1. Проспект фирмы НИИР-КОМ. - М., 1998. - 16 с.
8. Видеоанализатор ВК-2 – прибор нового поколения измерений и контроля качественных показателей каналов передачи телевизионных сигналов / В.П. Дворкович, А.В. Дворкович, Д.Г. Макаров и др. // Контрольно-измерительные приборы и системы. - 2006. - № 5. - С. 22-26.
9. Видеоанализатор ВК-2 – прибор нового поколения измерений и контроля качественных показателей каналов передачи телевизионных сигналов. Метрология видеоанализатора / В.П. Дворкович, А.В. Дворкович, Д.Г. Макаров и др. // Контрольно-измерительные приборы и системы. - 2006. - № 6. - С. 32-35.

КРАТКИЕ СООБЩЕНИЯ

УДК 621.372

ЭФФЕКТИВНЫЕ ОКОННЫЕ ФУНКЦИИ НА ОСНОВЕ ФУНКЦИЙ ГАУССА-ЭРМИТА

Вашкевич М.И.

Рассматривается метод расчета эффективных оконных функций, используемых при гармоническом анализе сигналов с применением ДПФ.

Как известно, задача минимизации мощности спектра оконных функций на заданном частотном интервале приводит к использованию вытянутых сфероидальных волновых функций (ВСВФ). Особенность их в том, что они являются собственными (инвариантными) функциями преобразования Фурье на ограниченном интервале. Это означает, что преобразование Фурье переводит каждую ВСВФ в такую же функцию с точностью до константы. Значит, задав во временной области достаточно «узкую» функцию, можно получить функцию ограниченную и в частотной области. Одной из наиболее известных и удачных аппроксимаций ВСВФ являются оконные функции Кайзера. Кроме ВСВФ известен ряд функций, инвариантных к преобразованию Фурье в бесконечных пределах. Функции Гаусса-Эрмита, как один из примеров таких функций, имеют вид:

$$\varphi_n = \omega_n(x) \cdot e^{-x^2/2},$$

где $\omega_n(x)$ - многочлен Эрмита n -го порядка.

Из сказанного выше можно сделать вывод, что функции Гаусса-Эрмита (ФГЭ) могут являться неплохим «материалом» для построения оконных функций. Тем более, что в [1] было показано, что энергия ФГЭ сконцентрирована на ограниченном отрезке, как во временной, так и в частотной области.

Окно, которое исследовалось в данной работе, имеет вид:

$$\omega(k) = e^{\frac{k^2}{2\sigma^2}} \times \frac{H_4\left(\frac{k \cdot \alpha}{N}\right)}{H_4(0)}, \quad k \in \left[-\frac{N}{2}; \frac{N}{2}\right],$$

N – половина ширины окна,

$H_4(x)$ - многочлен Эрмита 4-го порядка.

В этой формуле между α и σ существует определенная эмпирически установленная зависимость: $\sigma(\alpha) = 0.96275 - 0.8375 \cdot \alpha$. Параметр α может принимать значения из интервала $[0; 0.5]$. По сути α - это константа, аналогичная константе β в окне Кайзера. При увеличении α происходит расширение главного лепестка, но при этом падает уровень боковых лепестков.

Сравнение данного окна производилось с окном Кайзера в среде MatLab 6.5 (при помощи утилиты wvtool). Параметр β в окне Кайзера подбирался таким образом, чтобы ширина главного лепестка была одинаковой у обоих окон. Во всех случаях длина окна $N=65$. В таблице 1 приведены наиболее интересные из полученных результатов.

Таблица 1.

Сравнение экспериментального окна с окном Кайзера

Экспериментальное окно		Окно Кайзера			разность уровней боковых лепестков (дБ)	
α	$K_y, \%$	УБЛ (дБ)	β	$K_y, \%$		УБЛ (дБ)
0,30	0,67	25,3	2,91	0,70	23,6	1,7
0,35	0,33	29,0	3,34	0,35	26,3	2,7
0,40	0,11	34,3	3,95	0,13	30,2	4,1
0,45	0,03	39,6	4,72	0,03	35,4	4,2
0,50	0,01	44,6	5,76	0,01	42,3	2,3

Здесь: K_y - коэффициент утечки, который показывает, какая доля общей мощности сосредоточена в боковых лепестках его спектра. УБЛ - уровень боковых лепестков (уровень максимального из боковых лепестков относительно главного лепестка).

Как видно из полученных результатов, предлагаемое окно имеет несколько лучшие показатели по сравнению с окном Кайзера (при $\beta = [3; 6]$), и поэтому может успешно применяться при гармоническом анализе сигналов с применением ДПФ.

Литература

1. Хонина С.Н., Котляр В.В. Вилияние дифракции на изображении, согласованные с вытянутыми сфероидальными волновыми функциями // Компьютерная оптика, ИСОИ РАН, Самара-Москва, 2001, №21, с. 64-69.

УДК 629.7.05:681.3.06

АЛГОРИТМИЧЕСКОЕ ОБЕСПЕЧЕНИЕ СИСТЕМЫ АВТОНОМНОЙ КОРРЕКЦИИ ПОГРЕШНОСТЕЙ НАВИГАЦИОННОЙ СИСТЕМЫ МАНЕВРЕННЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ

Костров Б.В., Конкин Ю.В.

Введение

Для решения задач навигации маневренных летательных аппаратов (МЛА) используется, как правило, географическая система координат, построенная на определении в каждый текущий момент времени широты, долготы и высоты полета. Использование для этой цели инерциальных или инерциально-доплеровских навигационных систем приводит к значительным ошибкам выхода МЛА на навигационную точку, которые могут превышать единицы километров. Таким образом, точность автономного определения координат местоположения летательного аппарата на борту без коррекции погрешностей работы навигационных систем затрудняет действия экипажа по выполнению полетного задания. В качестве источников информации для коррекции текущих координат местоположения МЛА используются спутниковые навигационные системы и радиотехнические системы дальней навигации. Системы подобного уровня нельзя считать абсолютно надежными, особенно в экстремальных ситуациях. Наличие автономного, бортового средства, осуществляющего с приемлемой точностью и быстродействием подстройку системы навигации МЛА, позволяет повысить его живучесть и выполнять поставленные задачи в экстремальных условиях [1].

В статье рассматривается технология реализации метода совмещения текущего радиолокационного изображения, полученного с помощью бортовой радиолокационной станции (РЛС), с эталонным изображением, построенным на основе цифровой карты местности с известными географическими координатами. Это позволяет осуществлять коррекцию текущих координат навигационной системы МЛА без использования внешних источников информации.

Постановка задачи. Источником текущего изображения (РЛИ) является когерентная бортовая радиолокационная станция (БРЛС) миллиметрового диапазона волн. БРЛС осуществляет передний обзор земной поверхности и формирует на выходе РЛИ в прямоугольных координатах, что позволяет использовать полученное изображение в качестве основы для расчета координат местонахождения МЛА. Антенна БРЛС установлена на гиростабилизатор по углам крена и тангажа, что исключает этап геометрической коррекции из предварительной обработки изображений.

В качестве эталонного изображения используется модель РЛИ участка местности, полученная путем предварительного синтеза с помощью геоинформационной системы (ГИС) "Карта 2005" и с учетом парамет-

Рассматривается алгоритм совмещения изображения, полученного от радиолокационной станции, и изображения участка местности, синтезированного по цифровой карте с помощью инвариантных моментов H_i . Совмещение выполняется с целью определения географических координат точки привязки. Приводятся результаты работы алгоритмов, реализующих данный метод.

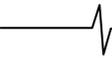
ров БРЛС. Результатом такого синтеза является РЛИ выбранного участка местности, содержащего наиболее информативные объекты. Данные объекты и требуется найти на текущем изображении при совмещении изображений. Алгоритм синтеза требует больших вычислительных затрат, поэтому не рассчитан на использование в реальном времени и не является этапом совмещения изображений. Эта работа выполняется на этапе предполетной подготовки МЛА.

Текущее РЛИ представляет собой монохромное изображение, содержащее в себе информацию о радиолокационном контрасте участка местности. Оно является полутонным дискретизированным изображением, заданным в прямоугольной системе целочисленных координат с функцией яркости $F_{ТН}(x, y)$, где $x = \overline{1, N}$, $y = \overline{1, M}$ - размерность изображения.

Эталонное изображение радиолокационного контраста участка местности представляет собой полутонное дискретизированное изображение в прямоугольной системе целочисленных координат с функцией яркости $F_{ЭИ}(x, y)$, где $x = \overline{1, K}$, $y = \overline{1, L}$ - размерность изображения. Интенсивность текущего и эталонного изображений изменяется в интервале $F_{ЭИ} = \overline{0, 255}$.

Для центрального элемента разложения эталонного изображения $(x_{Ц}, y_{Ц})$ известны прямоугольные географические координаты привязки на местности $(x_{Г}, y_{Г})$. Данные координаты получены по электронной карте местности на этапе синтеза эталонного изображения. Начало системы прямоугольных координат находится на борту МЛА. Для каждого элемента изображения известны размеры сторон участка местности в метрах $\Delta S_x, \Delta S_y$.

Предполагается, что этапу совмещения предшествует этап предварительной обработки изображений, на котором необходимо в максимальной степени подавить шумы в составе РЛИ, обнаружить и подчеркнуть границы полезных сигналов. Решение задач предварительной обработки дано в частности в [2]. Сам этап совмещения выполняется в рамках корреляционно-экстремальной системы навигации (КЭСН).



Отметим основные особенности функционирования КЭСН по полю радиолокационного контраста, влияющие на выбор метода совмещения изображений.

Во-первых, координаты ЛА в пространстве при формировании текущего и эталонного изображений не совпадают. Кроме того, в реальных условиях функционирования КЭСН текущие координаты всегда известны с некоторой ошибкой. Координатное рассогласование неизбежно сопровождается геометрическими искажениями текущего и эталонного РЛИ. Для компенсации указанных искажений необходимо выполнить последовательно операции сдвига, поворота и масштабирования. Эти операции образуют четырехпараметрическую группу аффинных преобразований, которая в декартовой прямоугольной системе координат описывается выражениями:

$$\begin{aligned} x' &= a(x \cos \beta + y \sin \beta) + x_0, \\ y' &= a(x \sin \beta + y \cos \beta) + y_0, \end{aligned} \quad (1)$$

где a - параметр масштабирования; β - угол поворота; x_0, y_0 - параметры сдвига.

Во-вторых, особенность формирования РЛИ заключается в том, что геометрические преобразования изображения обязательно приведут к радиометрическим преобразованиям изображения. Известно, что мощность отраженного радиолокационного сигнала, а, следовательно, и яркость РЛИ сильно зависят от угла падения луча и дальности от источника излучения до отражателя. Выполнить радиометрическую коррекцию эталонного изображения в реальном времени невозможно из-за большого количества вычислительных операций. Радиометрические искажения сглаживаются на этапе предварительной обработки [2].

В-третьих, распознаваемый эталонный объект должен описываться вектором признаков, инвариантных по отношению к его размерам, положению и ориентации. При этом вектор признаков должен обладать малой размерностью, так как количество возможных эталонных изображений может быть большим.

Метод решения поставленной задачи. Критериальные функции для оценки сходства изображений, используемые КЭСН, являются чувствительными к сдвигу, повороту и изменениям масштаба. Поэтому при разработке КЭСН для МЛА, где взаимные искажения изображений являются обычной ситуацией, необходимо использовать критериальные функции, инвариантные к указанным типам искажений. Среди разнообразных алгоритмов выделения признаков изображения лучшие результаты позволяет получить метод моментных инвариантов Ну [3], основанный на выделении статистических закономерностей, описывающих изображение. Математической основой выделения инвариантных признаков является теория алгебраических инвариантов. Эта теория исследует класс алгебраических функций, не изменяющихся при определенных преобразованиях координат.

Сущность метода заключается в следующем. Двумерные степенные моменты $(p+q)$ -го порядка от функции $f(x, y)$ в декартовой прямоугольной системе координат задаются выражениями вида:

$$m_{pq} = \iint_D x^p y^q f(x, y) dx dy, \quad p, q = 0, 1, 2, \dots, \quad (2)$$

где D - область изображения, занятого объектом, для которого вычисляются моменты или некоторый шаблон, если моменты строятся для всего изображения в целом.

Для того чтобы сделать моменты инвариантными относительно сдвига следует использовать центральные моменты:

$$\mu_{pq} = \iint_D (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy, \quad p, q = 0, 1, 2, \dots, \quad (3)$$

где $\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}}$ координаты центра тяжести

(центра масс) области D . Инвариантность в этом случае обеспечивается независимостью степенных произведений под знаком интеграла от сдвига.

При обработке дискретных изображений можно записать:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y). \quad (4)$$

Для того чтобы сделать эти моменты инвариантными относительно масштабирования следует перейти к безразмерным моментам:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\frac{p+q}{2}+1}}. \quad (5)$$

И, наконец, для построения инвариантов относительно поворота строятся собственно инварианты Ну:

$$\begin{aligned} S_1 &= \eta_{20} + \eta_{02} \\ S_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ S_3 &= (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} - 3\eta_{21})^2 \\ S_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2 \\ S_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \\ &+ [(\eta_{30} + \eta_{12})^2 - 3(\eta_{03} + \eta_{21})^2] + \\ &+ (3\eta_{21} - \eta_{03})(\eta_{03} + \eta_{21}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{03} + \eta_{21})^2] \\ S_6 &= (\eta_{20} - \eta_{02}) \\ &+ [(\eta_{30} + \eta_{12})^2 - (\eta_{03} + \eta_{21})^2 + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{03} + \eta_{21})] \\ S_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) \\ &+ [(\eta_{30} + \eta_{12})^2 - 3(\eta_{03} + \eta_{21})^2] - \\ &- (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{21})^2 - (\eta_{03} + \eta_{21})^2] \end{aligned} \quad (6)$$

Таким образом, на основе изложенного выше метода можно выделить вектор инвариантных признаков $P_{ЭИ}$ эталонного РЛИ и вектор $P_{ТИ}$ соответствующей области текущего РЛИ, каждые из которых состоят из значений инвариантных моментов $S_1 - S_7$. На основании данных признаков может быть получена система корреляционной привязки изображений, которая позволит рассматривать не полный растр изображений, а только соответствующий ему вектор признаков, инвариантный к взаимному сдвигу, повороту и масштабированию.

нию изображений в определенных пределах. Малый размер вектора признаков позволит сформировать достаточно большое количество эталонных изображений (ЭИ) под разными углами подхода МЛА и для разных высот.

Для сравнения инвариантных признаков текущего и эталонного изображений в данной работе была использована корреляционная функция (КФ) вида:

$$K(x, y) = \frac{\sum_{i=1}^7 S_i \bar{S}_i(x, y)}{\left[\sum_{i=1}^7 S_i^2 \sum_{i=1}^7 \bar{S}_i^2(x, y) \right]^{1/2}}, \quad (7)$$

где

\bar{S}_i - i -й момент эталонного изображения,

S_i - i -й момент подобласти текущего изображения в положении (x, y) .

КФ вычисляется для всех возможных положений окна эталонного изображения размером $K \times L$ в пределах окна текущего изображения $N \times M$. Точка совмещения (x_m, y_m) определяется максимальным значением КФ.

Корреляционная функция $K(x, y)$ вычисляется для всех взаимных сдвигов окна эталонного изображения относительно текущего изображения в пределах окна (рис. 1). Координаты окна (x_{\max}, y_{\max}) , в которых функция $K(x_{\max}, y_{\max})$ приобретает максимальное значение, и являются координатами точки наилучшего совмещения РЛИ. При размерах изображений $K \times L$ и $N \times M$, $N = M$ количество точек вычисления $K(x, y)$ равно:

$$T_1 = (K - N + 1)(L - N + 1). \quad (8)$$

При больших размерах изображений количество вычислений функции $K(x, y)$ является значительным. В этом случае необходимо искать пути повышения производительности алгоритмов корреляционного совмещения изображений.

Как отмечалось выше, $F_{ЭИ}(n, m)$ в реальном времени сформировать невозможно, поэтому разработка КЭСН выполняется с условием, что ЭИ формируется на этапе предполетной подготовки МЛА. Следовательно, и вектор признаков $P_{ЭИ}$ также формируется заранее.

Таким образом, при корреляционном совмещении будем вычислять только вектор признаков $P_{ТИ}$ соответствующей области ТИ. Моменты можно вычислить как по значениям функции яркости изображения $F_{ТИ}(x, y)$, так и по бинарному изображению $\bar{f}(x, y)$, содержащему границы радиолокационно-контрастных объектов. Значения точек, равные 1, соответствуют границе, а значения точек, равные 0, - области фона. При этом $\bar{f}(x, y)$ содержит только единичные значения и их координаты в изображении, а значения фона исключены, так как в данном случае не несут информации. При этом в несколько раз сокращается размерность изображения, так как количество точек фона превышает количество точек границ. Данную особенность изображения $\bar{f}(x, y)$, прошедшего этап сегментации, можно использовать для вычисления моментов. В этом случае формула для вычисления моментов будет выглядеть следующим образом:

$$\mu_{pq} = \sum_{i=1}^I x_i^p y_i^q, \quad p, q = 0, 1, 2, \dots, \quad (9)$$

где x_i, y_i - координаты единичной точки в $\bar{f}(x, y)$; I - количество единичных точек.

Вычисление моментов по формуле (9) обладает меньшей вычислительной сложностью по сравнению с формулой (3), так как значение I намного меньше общего числа точек изображения (области D). Кроме того, в формуле (9) отсутствует операция умножения на значение функции яркости.

Другим способом повышения производительности алгоритмов совмещения может быть последовательное снижение разрешения большего изображения, с последующим иерархическим вычислением корреляционной функции $K(x, y)$. В обычных алгоритмах совмещения значение $K(x, y)$ вычисляется при всех возможных положениях эталонного изображения относительно текущего. Однако очевидно, что точные вычисления имеет смысл производить только для небольшого числа точек вблизи максимума корреляционной функции. Возможно использование двухэтапного алгоритма получения КФ.

На первом этапе работы алгоритма выполняется снижение разрешения изображения с помощью замены каждого $s \times t$ элементов разложения так называемым блоком, яркость которого вычисляется как средняя яркость $s \times t$ элементов. Для полученного изображения с пониженным разрешением и ЭИ вычисляется корреляционная функция и находятся координаты взаимного положения изображений, соответствующие ее максимуму.

На втором этапе вычисляется корреляционная функция между эталонным и частью текущего изображения, ограниченной областью, смещенной на $\Delta x_1, \Delta y_1$ относительно максимума, найденного на первом этапе в меньшую сторону и $\Delta x_2, \Delta y_2$ в большую сторону. Координаты максимума $K(x, y)$, найденные на втором этапе, и являются координатами точки совмещения изображений.

Объем вычислений J при использовании этого алгоритма можно оценить следующим образом:

$$J = A + P_T B, \quad (10)$$

где A - объем вычислений на первом шаге алгоритма; P_T - вероятность превышения порога на первом этапе (т. е. вероятность нахождения точки, которая будет обрабатываться на втором этапе); B - объем вычислений на втором этапе алгоритма.

С точки зрения быстродействия данный алгоритм обладает недостатком, связанным с необходимостью вычисления средней яркости блока. Поэтому авторами предлагается модифицировать алгоритм следующим образом.

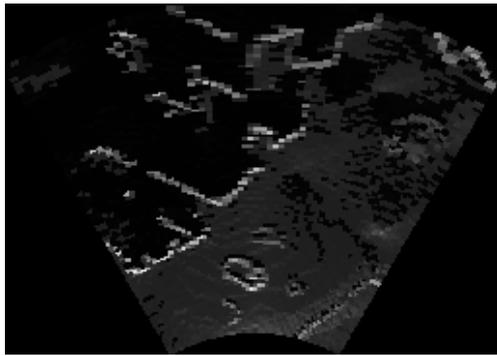
На первом этапе работы алгоритма выполняется вычисление корреляционной функции $K(x, y)$ с шагом по осям x и y больше единицы: $\Delta x = s$, $\Delta y = t$. Второй этап выполняется аналогично методу, рассмотренному выше.

Количество вычислений корреляционной функции в данном алгоритме можно оценить по следующей формуле:

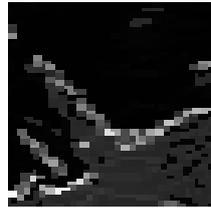
$$T_2 = \left[(K - N + 1) / s \right] \left[(L - N + 1) / t \right] +, \quad (11)$$

$$(\Delta x_1 + \Delta x_2 + 1)(\Delta y_1 + \Delta y_2 + 1)$$

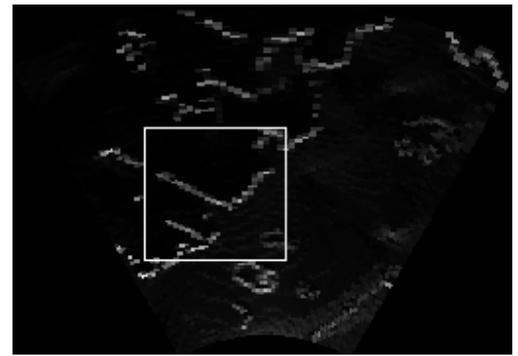
где $K \times L$ - размерность изображения; $s \times t$ - шаг изменения координат для вычисления корреляционной функ-



а) Текущее РЛИ



б) Эталонное РЛИ



в) Результат совмещения

Рис. 1. Корреляционное совмещение изображений

ции первого этапа алгоритма; $\Delta x_1, \Delta x_2, \Delta y_1, \Delta y_2$ - смещение границ окна вычисления корреляционной функции второго этапа относительно максимума первого этапа.

Процент снижения объема вычислений можно вычислить по формуле:

$$k = \left(1 - \frac{T_2}{T_1}\right) \cdot 100 \quad (12)$$

В ходе практических исследований по совмещению РЛИ с помощью корреляционной функции на основе инвариантных моментов возникла необходимость отбраковки ложных результатов совмещения. Для этого предлагается использовать следующий механизм.

По полученным оценкам корреляционной функции $K(x, y)$ вычисляется среднеквадратическое отклонение σ_k и формируется пороговое значение $\Delta = \alpha \cdot \sigma_k$, где α - коэффициент, значение которого устанавливается экспериментально (в ходе практических исследований определено оптимальное значение коэффициента $\alpha = 1.2$). Решение о ложном совмещении принимается при условии:

$$K(x_{\max}, y_{\max}) < \Delta, \quad (13)$$

где $K(x_{\max}, y_{\max})$ - максимум корреляционной функции.

На рис. 1 приводятся примеры эталонного и текущего изображений (а, б) и результат их корреляционного совмещения (в).

Заключение. Таким образом, разработан алгоритм корреляционного совмещения текущего и эталонного РЛИ на основе моментов, инвариантных к сдвигу, повороту и изменению масштаба изображений.

Для данного алгоритма разработаны процедуры повышения производительности на основе вычисления моментов для контурных изображений и двух-

уровневое вычисление корреляционной функции. Кроме того, разработана процедура повышения надежности совмещения информативных участков на текущем и эталонном РЛИ, представленных векторами инвариантных моментов, за счет отбраковки ложных результатов совмещения.

При этом получены следующие основные результаты:

- значительно повышено быстродействие привязки текущих координат местоположения МЛА к географическим координатам;

- обеспечена корреляционная привязка в условиях координатного рассогласования местоположения МЛА в различных точках пространства, которое неизбежно возникает при формировании текущего и эталонного изображений;

- за счет работы разработанного алгоритма компенсируются погрешности измерений датчиков курса, высоты и географических координат МЛА, что повышает живучесть летательного аппарата в экстремальных условиях.

Литература

1. Андреев Г.А., Потапов А.А. Активные системы ориентации по геофизическим полям // Зарубежная радиоэлектроника. 1988. № 9. С. 62-85.
2. Новиков А.И., Конкин Ю.В., Архипов С.А. Применение операторов с симметричной весовой функцией в задачах обработки радиолокационной информации // Математические методы в научных исследованиях: Межвуз. сб. / Рязан. гос. радиотехн. акад. Рязань, 2006. – С. 46-55.
3. Андреев Г.А., Потапов А.А. Алгоритмы обработки навигационной пространственно-временной информации. Часть II // Зарубежная радиоэлектроника. 1989. № 4. С. 3-21.

УДК 621.396.7

ААС – ПРОГРЕССИВНЫЙ МЕТОД ЦИФРОВОГО КОДИРОВАНИЯ ЗВУКА

Иртюга В.А.

Введение

Стандарт звукового сжатия звукового сигнала ААС (Advance Audio Coding – усовершенствованное звуковое кодирование) был разработан группой компаний Fraunhofer IIS, AT&T, Sony и Nokia. По результатам этой работы в апреле 1997 года стандарт кодека был официально опубликован в составе международного стандарта MPEG-2 (ISO/IEC 13818, часть 7 «ААС») [1].

С самого начала ААС позиционировался, как эффективная замена распространенному стандарту звуковой обработки MP3, разработанному в рамках стандарта MPEG-1 (в части «Аудио», уровень 3), что и было с успехом достигнуто благодаря нововведенным алгоритмам и усовершенствованиям, использованным при разработке стандарта кодека. Исходя из анализа результатов сравнительных испытаний, можно утверждать, что для всех скоростей битового потока до 192 кбит/с (на канал) декодер ААС дает более качественный звуковой сигнал, чем декодер MP3. Отметим, что сравнивать качественные характеристики кодеков выше указанной границы скорости становится существенно сложнее, так как при потоке более 192 кбит/с на канал отличия от оригинального сигнала трудно различимы для обоих кодеков и требуют уже тонкого подбора тестового аудиоматериала и наличия большого количества высокопрофессиональных музыкальных экспертов.

Как и во многих стандартах подобного рода, собственно стандартизации подвергнута только декодирующая часть кодека, что оставляет разработчикам и производителям поле для конкуренции и совершенствования выпускаемых ими кодеров. В настоящее время существует несколько реализаций стандарта ААС. В частности, можно упомянуть разработанный компанией Apple ААС-кодер, входящий в состав пакета iTunes, бесплатно распространяемый кодек FAAC/FAAD2, а также ААС-кодек, разработанный в ЗАО «НИИР-КОМ». По результатам испытаний последнего из указанных кодеков можно сказать, что ААС-поток скоростью от 24 кбит/с соответствует качеству FM-стереотрансляции, поток от 32 кбит/с - качеству Audio-CD, а поток 128 кбит/с - сравним по качеству восстановленного сигнала с шестиканальным звуком стандарта "Dolby Digital Surround EX".

В настоящее время ААС получил свое развитие в стандарте MPEG-4 (ISO/IEC 14496, часть 3, подчасть 4 «Общее кодирование звука») [2]. При этом новый стандарт является только дополнением к уже существующим и описанным в MPEG-2 алгоритмам звуковой обработки, а не отменяет его. В частности, профиль ААС-LC (low complexity - упрощенный профиль) описан только в стан-

Рассматривается один из лучших на сегодняшний день метод цифрового сжатия звукового сигнала – ААС (Advance Audio Coding – усовершенствованное звуковое кодирование). Приводятся структурные схемы кодеров и декодеров, соответствующих различным версиям стандарта кодека, при этом особое внимание уделяется кодирующей части кодека. На примере разработанной в ЗАО «НИИР-КОМ» усовершенствованной версии кодека ААС – HE-AAC (High Efficiency) рассматривается вопрос позиционирования данного продукта на рынке мультимедийных услуг.

дарте MPEG-2, являясь при этом базовым для нового высокоэффективного варианта кодека - HE-AAC (High Efficiency).

На настоящий момент существуют две версии расширения кодека ААС: «HE-AAC v.1» и «HE-AAC v.2». Первая версия (v.1) использует технологию спектрального «дублирования» (replication) нижней части спектра звукового сигнала в верхнюю: данный алгоритм носит название SBR – Spectral Band Replication. Вторая версия (v.2) полностью включает в себя всю функциональность первой и расширяет ее с помощью алгоритма параметрического кодирования стереосигнала, этот алгоритм носит название Parametric Stereo (PS).

Структура кодирующего устройства. На рис. 1 приведена структурная схема кодера ААС. Отметим, что здесь и далее под понятиями «звуковой сигнал» и «спектр звукового сигнала» подразумеваются соответственно дискретизированный звуковой сигнал и дискретный спектр дискретизированного звукового сигнала.

Регулировка усиления (Gain control). Данный модуль не является обязательным в структуре кодека и предназначен для регулировки уровней звукового сигнала. На стороне кодера происходит поблочная регулировка уровня звукового сигнала и передача декодеру информации, необходимой для восстановления исходного соотношения. В состав данного модуля входит 4 полифазных квадратурных фильтров (PQF), обеспечивающих четырехполосное разделение спектра сигнала с последующим контролем уровня в каждой из четырех полос во временной области.

Психоакустическая модель вычисляет максимальные значения ошибок, маскируемых (то есть незаметных для человеческого уха) основным сигналом. Это так называемые пороги ошибок, превышение которых будут заметны при воспроизведении декодированного звукового сигнала. В [1] приведены таблицы (табл. С.1-С.24 в Информационном Приложении С), в соответствии с которыми производятся вычисления указанных порогов. Кроме этого в блоке психоакустической модели осуществляется выбор типа и длины взвешивающего окна, используемого в дальнейшем непосредственно перед МДКП-преобразованием (см. ниже).

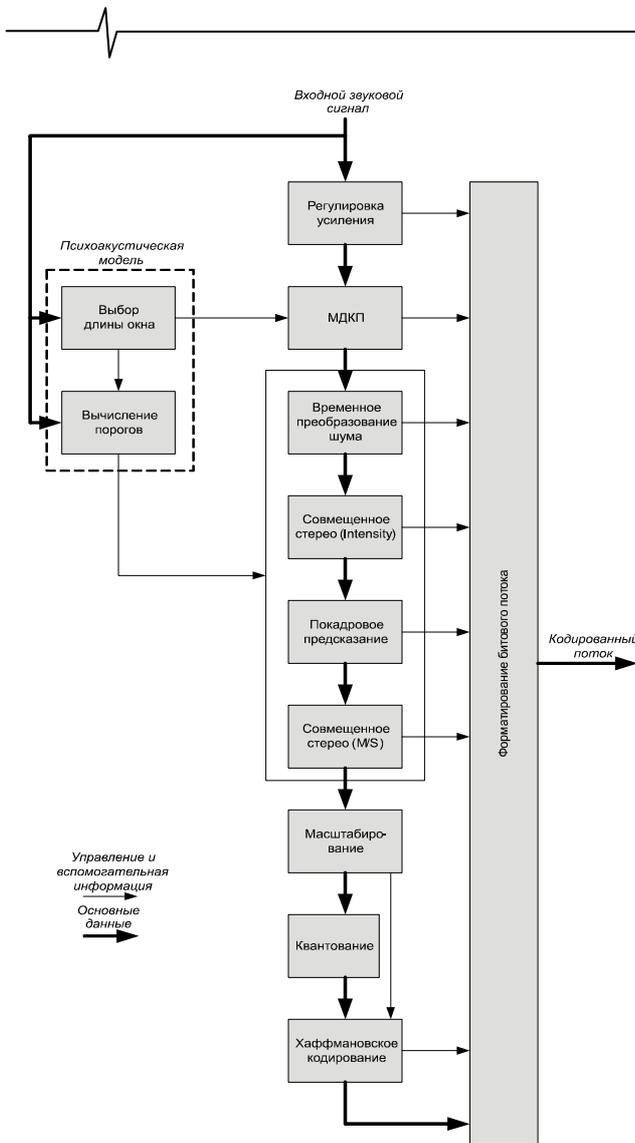


Рис. 1. Схема кодера AAC

Отметим, что выбор осуществляется между двумя типами окон – Кайзера-Бесселя и синусным. Выбор длины взвешивающего окна и, соответственно, длины МДКП для одного кадра осуществляется исходя из следующих альтернатив: 2048 либо 256 при длине кадра 1024 отсчета или 1920 либо 240 при длине кадра 960 отсчетов. Короткие преобразования (256 и 240) осуществляются на том же интервале кадра 8 раз подряд, при этом используются соответствующие короткие взвешивающие окна. При обнаружении участка резкого изменения уровня сигнала с целью локализации этой динамической особенности данный блок выносит решение о применении коротких взвешивающих окон. Здесь же происходит оценка количества бит, необходимых для кодирования очередного анализируемого кадра.

МДКП – метод модифицированного косинусного преобразования (MDCT – Modified Discrete Cosine Transform). МДКП осуществляет преобразование временного звукового сигнала в частотную область. Формула преобразования выглядит следующим образом:

$$X_k = 2 \sum_{n=0}^{2N-1} w_n x_n \cos\left(\frac{\pi}{N} \left(n + \frac{1}{2} + \frac{N}{2}\right) \left(k + \frac{1}{2}\right)\right), \text{ где}$$

X_k -- значения спектральных компонент,
 x_n -- отсчеты входного сигнала,
 w_n -- коэффициенты взвешивающего окна,
 n – временной индекс, $0 \leq n \leq 2N-1$,
 k -- частотный индекс, $0 \leq k \leq N-1$,
 $2N$ -- длина окна преобразования.

Здесь размерность входного сигнала в два раза больше размерности выходного. Это связано с тем, что МДКП в отличие, например, от обычного косинусного преобразования работает с перекрывающимися (на 50%) областями сигнала. Этим решается или, по крайней мере, ослабляется проблема блочности, то есть, появления немаскируемых артефактов на границах преобразования при последовательной обработке сигнала, длительность которого больше, чем длина окна используемого преобразования [3-5]. Кроме этого, для предотвращения разрывов фазы на границах окна преобразования и, как следствие, появления мешающих артефактов в этих областях, в алгоритм AAC вводятся взвешивающие временные окна. Решение о длине окна и типе окна, как уже было сказано, принимается в блоке психоакустической модели кодека.

Метод временного преобразования формы шума квантования (TNS - Temporal Noise Shaping). Данный модуль служит для борьбы с последствиями процедуры квантования сигнала, что проявляется, например, в частом появлении немаскируемых, то есть различных человеческим ухом, пре-эхо сигналов при наличии в сигнале резких фронтов, особенно характерных для речевых сигналов. Суть метода состоит в том, чтобы на основе спектрального анализа звукового сигнала предсказать ошибку квантования (forward prediction или опережающее предсказание) и ввести ее в сигнал (“вычесть” ее). Таким образом, при восстановлении сигнала немаскируемые артефакты, вызванные квантованием, будут частично подавлены [6].

Совмещенное стерео (Joint stereo): Intensity-стерео или M/S-стерео. Данный модуль служит для сокращения информационной избыточности, обусловленной высокой корреляцией между левым и правым каналами стереосигнала. При этом метод Intensity-стерео является алгоритмом сжатия с потерей информации, а M/S – алгоритмом без её потери. Соответствующему блоку декодера передается информация о выбранном типе стереокодирования. Алгоритм Intensity использует корреляционную зависимость между верхними частями спектров стереопары звукового сигнала. При этом в выходной поток передаются нижние части спектров левого и правого каналов в полосе до 6 кГц, а оставшаяся верхняя часть полосы передается в виде масштабированной невозстанавливаемой на стороне декодера суммы спектров левого и правого каналов. Алгоритм M/S (middle/side = среднее/разность) анализирует возможность передачи вместо сигналов левого (L) и правого (R) каналов пары $M=(L+R)/2$ и $S=(L-R)/2$ с точки зрения минимизации выходного битового потока. При этом оцениваются распределения мощностей для обоих вариантов и выносится решение о передаче пары L/R либо M/S.

Блок покадрового предсказания (Prediction). Данный модуль используется для сокращения информационной избыточности, связанной с высокой степенью

корреляции между спектрами последовательных кадров. При этом на стороне кодера используется так называемый алгоритм обратного предсказания (backward prediction), осуществляющий итеративное квантование и восстановление данных с целью предотвращения распространения ошибки предсказания. Добавим, что блок кадрового предсказания включается только тогда, когда при первоначальном анализе звукового сигнала на уровне психоакустической модели принимается решение об отсутствии в данном кадре резких переходов и участков «нестационарности спектра» (так называемые transients).

Блок масштабирования спектральных компонент (Scaling). В этом модуле, на стороне кодера, происходит умножение спектральных компонент на масштабные множители (scalefactors). Данная процедура используется для управления в частотной области формой шумов квантования (во временной области этим занимается блок TNS, см. выше). Сами масштабрующие множители передаются вместе с потоком. Количество коэффициентов определяется разбиением спектра сигнала на так называемые «масштабные полосы» (scalefactor bands), каждой из которых соответствует один коэффициент.

Блок квантования (Quantization) - осуществляет неравномерное квантование спектральных компонент сигнала и разделен на 3 уровня. Верхний уровень занимается вычислением количества бит, в которые нужно «уложить» поток, проверяет не равны ли все спектральные компоненты нулю и, если нет, вызывает процедуру следующего уровня – так называемый внешний итеративный цикл, который в свою очередь вызывает процедуру внутреннего итеративного цикла. Внутренний цикл квантует входной вектор, увеличивая шаг квантования до тех пор, пока выходной вектор все еще укладывается в отведенные ему количество бит. После выполнения внутреннего цикла внешний цикл оценивает искажения, вызванные последней итерацией квантования внутри каждой масштабной полосы (см. описание блока Scaling) и, если ошибка больше пороговой, уменьшает значение масштабующего коэффициента для тех масштабных полос, в которых было допущено превышение порога ошибки.

Блок хаффмановского кодирования (Huffman coding) представляет собой статистический кодер, работающий с парами или квартетами передаваемых квантованных спектральных коэффициентов и с единичными значениями масштабных коэффициентов. Для этого предусмотрено 10 основных кодовых книг кодирования для спектральных компонент и одна для масштабных коэффициентов, различающихся между собой тем, над чем они определены (пары или квартеты значений), максимальным значением входных величин, а также знаковостью самой кодовой книги (знаковая или беззнаковая). Порядок кодирования спектральных коэффициентов – от нижней частоты к верхней.

Блок формирования битового потока (Bitstream formatter). Данный блок компонует выходной битовый поток, который включает в себя следующие основные типы управляющей информации:

- квантованные спектральные коэффициенты звукового сигнала,
- масштабрующие коэффициенты модуля "Scaling",
- информация для восстановления входных уровней звукового сигнала в случае использования на стороне кодера блока "Gain control",
- информация о типе и длине использовавшихся взвешивающих окон,
- информация, необходимая для декодирования временно-го преобразования формы шума (TNS),
- информация об использовавшемся типе объединенного стерео,
- информация об использовавшихся на стороне кодера параметрах хаффмановского кодирования.

Структура декодирующего устройства. Декодирование осуществляется в обратном (относительно кодирования) порядке. Одноименные блоки на рис. 2 представляют собой соответствующие декодеры. На вход декодера поступает упакованный битовый поток, основная часть которого состоит из квантованных и закодированных с помощью хаффмановского кодера спектральных компонент.

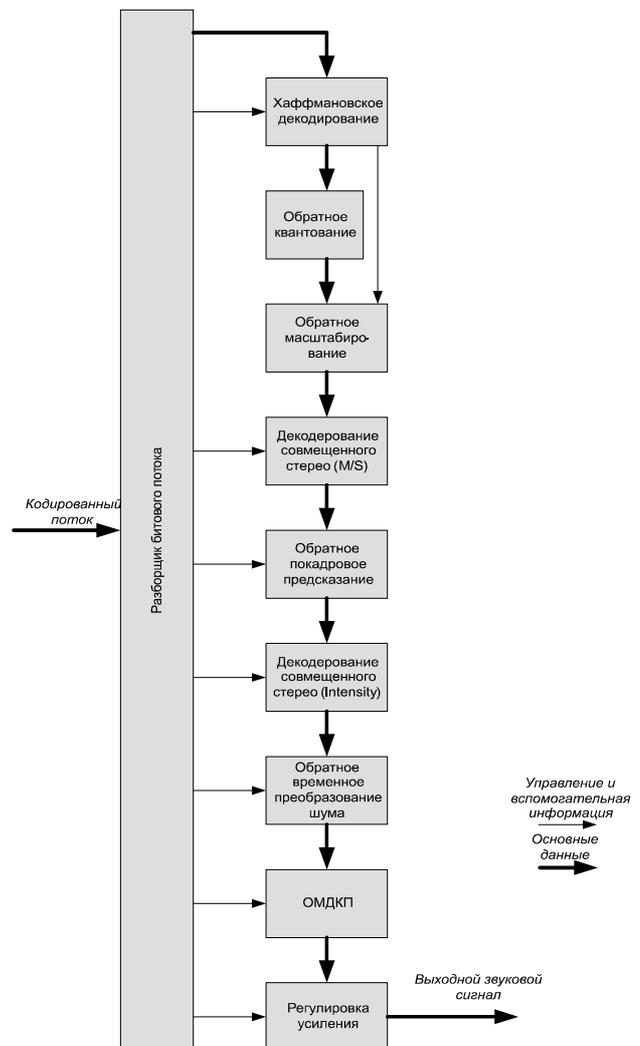


Рис. 2. Схема декодера AAC

Вектор развития стандарта звукового кодера AAC можно обозначить с помощью таких символических формул:

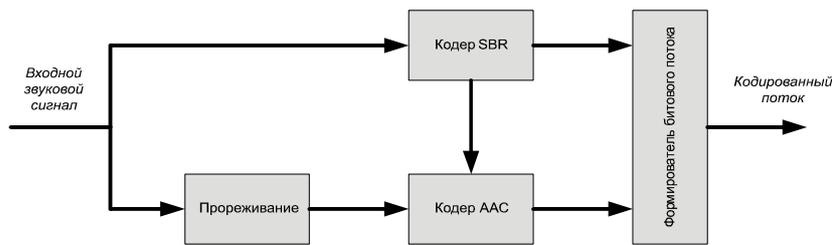


Рис. 3. Общая схема кодера HE-AAC v.1 (SBR)

«HE-AAC v.1» = «AAC» + «SBR»

«HE-AAC v.2» = «AAC» + «SBR» + «PS»

Метод спектрального дублирования или отбраживания (SBR – Spectral Band Replication) основывается на двух вводимых в рамках данной концепции качественных постулатах:

1. В верхней части спектра звукового сигнала содержится в значительной степени избыточная с точки зрения слухового восприятия информация. Для слухового восприятия важны только мощностные характеристики высокочастотных составляющих, а фазовые соотношения могут быть искажены без потери воспринимаемого качества восстановленного звукового сигнала.

2. На некотором временном интервале существует высокая корреляционная зависимость между мощностными характеристиками верхней и нижней половинами спектра звукового сигнала.

На рис. 3 приведена общая схема HE-AAC кодера первой версии.

Суть SBR-алгоритма состоит в том, чтобы не передавать полностью верхнюю половину спектра (коэффициенты МДКП) звукового сигнала. Вместо этого передается информация о мощности спектральных компонент отдельных подполос верхней половины спектра. Кроме этого на стороне передачи анализируется мощность шумовых составляющих этих подполос и их уровень также передается декодеру для возможности восстановления исходного отношения сигнал/шум. Мощность высокочастотных спектральных компонент оценивается линейным предсказателем 2-го порядка с помощью матрицы ковариаций огибающих спектра. Как видно из рис. 3, SBR-кодер работает с непрореженным по времени сигналом, в то время как на AAC-кодер подается уже прореженный в два раза сигнал (блок "Прореживание"). Важно отметить, что блок прореживания должен обладать некоторой "интеллектуальностью", а именно, содержать в себе НЧ-фильтр, предотвращающий зеркальное наложение верхней половины спектра на нижнюю при осуществлении собственно самого прореживания. Помимо перечисленных алгоритмов SBR-кодер отслеживает и оценивает возможность возникно-

вения в высокочастотной части спектра мощных индивидуальных узкополосных сигналов (в стандарте они именуются "синусоидальными компонентами"), образующихся, например, вследствие возникновения в звуковом сигнале участков нестационарности спектра (так называемые "transients"). Информация о таких компонентах спектра также передается декодеру. Таким образом, SBR-декодер получает следующую информацию для восстановления звукового сигнала:

- мощностные коэффициенты подполос верхней половины спектра декодируемого сигнала (так называемые *огигающие спектра*);
- мощностные коэффициенты шума для восстановления соотношения сигнал/шум;
- положение и мощность индивидуальных узкополосных сигналов, не поддающихся линейному предсказанию.

Заметим, что данная информация об огибающих спектра передается не напрямую, а в виде разностей оригинальных и предсказанных значений. Поэтому на стороне декодера вновь осуществляется линейное предсказание по низкочастотной части спектра, получаемой от AAC-декодера, после чего производится восстановление (по мощности) огибающих верхней половины спектра сигнала.

Таким образом, полностью передается только нижняя часть спектра сигнала, а высокочастотные составляющие восстанавливаются по сопроводительной составляющей, занимающей лишь небольшой процент передаваемой информации. Технология SBR получила широкое признание и существует множество источников, кроме самого стандарта [2], в которых можно почерпнуть более подробную информацию о деталях этого алгоритма. Это прежде всего сайт компании Coding Technologies (www.codingtechnologies.com/products/sbr.htm), а также следующий ряд статей по этой теме: [7-10].

Параметрическое стерео кодирование (PS - Parametric Stereo) осуществляет преобразование стереосигнала в моносигнал (downmix), сопровождая его дополнительным набором параметров, позволяющих на стороне декодера восстановить стереобраз. Данное преобразование относится к классу кодирования с информационными потерями. На рис. 4 изображена схема PS-кодера.

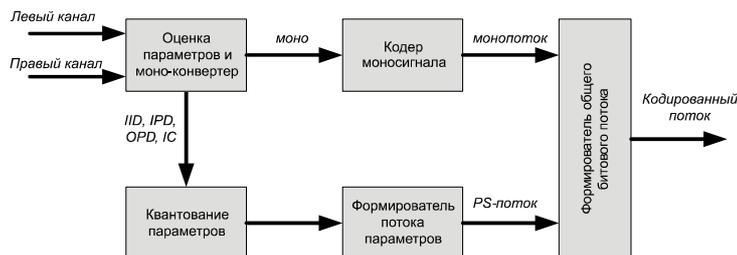


Рис. 4. Схема PS-кодера

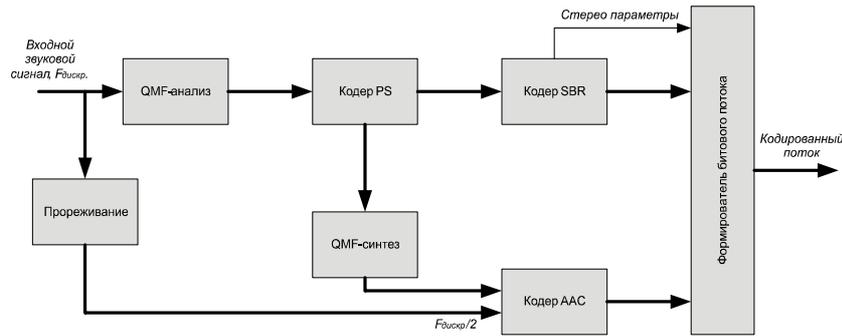


Рис. 5. Схема кодера HE-AAC v.2

В блоке "Parameter estimation" осуществляется оценка следующих четырех типов параметров:

- межканальные разности уровней (IID - Inter-channel Intensity Differences);
- межканальные разности фаз (IPD - Inter-channel Phase Differences);
- общий фазовый сдвиг (OPD - Overall Phase Difference);
- межканальная когерентность (IC - межканальная когерентность).

Последний параметр (IC) измеряется как максимум кросс-корреляции функции фаз или времени. Перечисленных параметров достаточно для того, чтобы восстановить стереообраз на стороне приема. Для восстановления этого стереообраз без заметных на слух искажений под поток этой дополнительной информации достаточно отвести 9-10 кбит/с. При этом в канале будет передаваться только моносигнал (после процедуры преобразования моно-стерео). Перечисленные 4 параметра квантуются (блок "Parameter quantization") и кодируются как дополнительные части монопотока (в блоке "Bit stream multiplexer") таким образом, чтобы обеспечить обратную совместимость при воспроизведении данного потока моно-кодеками, игнорирующими эту дополнительную информацию. Более подробно с технологией параметрического стереокодирования можно ознакомиться, например, в [11-14].

Общая схема HE-AAC-кодека второй версии, включающего в себя кодеры Spectral Band Replication (SBR) и Parametric Stereo (PS) приведена на рис. 5:

Заметим здесь еще раз, что кодеры PS и SBR работают с двойной частотой дискретизации сигнала относительно кодера AAC.

В заключение отметим еще раз, что в ЗАО «НИИР-КОМ» разработан звуковой кодек, совместимый со стандартом HE-AAC v.1. Испытания, проведенные специалистами этой фирмы, подтвердили превосходство стандарта AAC над самым популярным на сегодняшний день форматом сжатия звука MP3 для всех скоростей кодирования и всех типов звуковых сигналов. При этом установлено, что при скорости цифрового потока 24 кбит/с кодек HE-AAC в состоянии обеспечить качество FM-стереовещания, а при скорости 32 кбит/с - качество Audio-CD. Направление дальнейших работ по совершенствованию этого продукта связано с созданием кодека, поддерживающего технологию параметрического стереокодирования (PS) и, следовательно, совместимого с версией HE-AAC v.2, что безусловно позволит добиться повышения качества звука при за-

данной скорости цифрового потока, либо осуществлять передачу звука в каналах с меньшей пропускной способностью, сохраняя при этом качественные характеристики передаваемого звукового сигнала.

Литература

1. ISO/IEC MPEG-2 13818-7 AAC
2. ISO/IEC MPEG-4 14496-3 Audio
3. Henrique S. Malvar, Signal Processing With Lapped Transforms (Artech House: Norwood MA, 1992).
4. John P. Princen and Alan B. Bradley, "Analysis/synthesis filter bank design based on time domain aliasing cancellation", IEEE Trans. Acoust. Speech Sig. Proc. ASSP-34 (5), 1153-1161 (1986).
5. A. W. Johnson and A. B. Bradley, "Adaptive transform coding incorporating time domain aliasing cancellation", Speech Comm. 6, 299-308 (1987).
6. Jurgen Herre, "Temporal Noise Shaping, Quantization and Coding Methods in Perceptual Audio Coding: a Tutorial Introduction", AES 17th International Conference on High Quality Audio Coding
7. Per Ekstrand, "Bandwidth extension of audio signals by spectral band replication", Proc.1st IEEE Benelux Workshop on Model based Processing and Coding of Audio (MPCA-2002), Leuven, Belgium, November 15, 2002
8. Gilbert A. Soulodre and Michel Lavoie, "Subjective Evaluation of MPEG Layer II with Spectral Band Replication", Audio Engineering Society, Convention Paper 6185, Presented at the 117th Convention, 2004 October 28-31 San Francisco, CA, USA
9. O. Kunz, "SBR explained: White paper", Coding Technologies
10. A. Gröschel, M. Schug, M. Beer and F. Henn, "Enhancing audio coding efficiency of MPEG Layer-2 with Spectral Band Replication for DigitalRadio (DAB) in a backwards compatible way", 114th AES Convention, Amsterdam, 2003, March
11. E. Schuijers, W. Oomen, B. den Brinker, and J. Breebaart, "Advances in parametric coding for high-quality audio," in Proc. 114th AES Convention, Amsterdam, The Netherlands, Mar. 2003, Preprint 5852.
12. C. Faller and F. Baumgarte, "Estimation of auditory spatial cues for binaural cue coding," in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP), Orlando, Florida, USA, May 2002.
13. J. Breebaart, S. van de Par, A. Kohlrausch, and E. Schuijers, "High-quality parametric spatial audio coding at low bitrates," in Proc. 116th AES Convention, Berlin, Germany, May 2004.
14. W. Oomen, E. Schuijers, H. Purnhagen, and J. Engdegard, "MPEG4-Ext2: CE on low complexity parametric stereo," ISO/IEC JTC1/SC29/WG11 MPEG2003/M10366, Dec. 2003.

МЕТОД ПРЕОБРАЗОВАНИЯ ФОРМАТОВ МУЗЫКАЛЬНОЙ ИНФОРМАЦИИ

Фадеев А.С. Кочегурова Е.А.

Введение

Быстрый рост информационных технологий сделал возможным программные и аппаратные реализации многих прикладных математических задач, в том числе, связанных с обработкой аудиозаписей. Большое внимание уделяется вопросам сжатия, фильтрации, линейной и нелинейной обработки, а также вопросам преобразования форматов современных мультимедийных систем. Рассмотрение многих вопросов, требующих больших вычислительных мощностей, становится возможным лишь с появлением современных высокопроизводительных ЭВМ. Решаются задачи классификации, распознавания и формирования объектных форматов, в том числе для графических систем. Обработке музыкальной информации уделяется не столь много внимания.

Музыкальной можно назвать воспроизводимую информацию, воспринимаемую человеком на слух, которую можно записать в виде нотных партитур. Музыкальное произведение — это записанный нотами на аналоговых, цифровых или иных носителях конечный фрагмент потока музыкальной информации.

По аналогии с графическими цифровыми форматами любое музыкальное произведение можно представить в виде объектного и амплитудно-временного форматов.

Наиболее популярным объектным форматом музыки является нотная запись. Она представляет собой совокупность объектов (нот, пауз, знаков альтерации и др.) с определенным набором свойств: длительность, высота, уровень громкости, временные значения (Рис.1.). Мультимедийные системы достаточно часто используют формат *MIDI* (*Musical Instrument Digital Interface*), представляющий собой объектное описание партий музыкальных инструментов. Отличительной особенностью объектного формата является чрезвычайно малый размер файлов и гибкость музыкального произведения.

Для воспроизведения музыкального произведения,

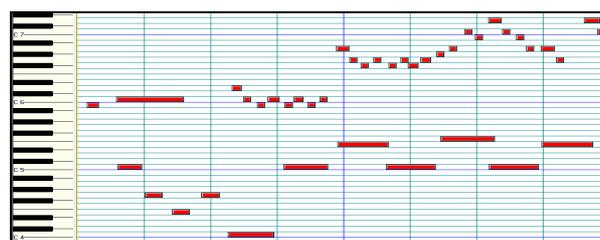
Дано описание объектного и амплитудно-временного музыкальных форматов и методов перевода музыкальной информации из формата в формат. Описан программно-технический комплекс, позволяющий производить соответствующие преобразования музыкальной информации, математической основой которого является аппарат непрерывных вейвлет-преобразований с использованием нестандартных вейвлет-функций. Приводится методика формирования базисных вейвлетов, описан эксперимент и интерпретация полученных результатов.

записанного в *MIDI*-формате, необходимо использовать программно- и/или аппаратно-реализованные комплексы. Разные мультимедийные системы могут воспроизвести одно и то же объектное представление музыкального произведения с отличиями, явно воспринимаемыми на слух.

Амплитудно-временной формат музыкальных произведений не требует сложного программно-аппаратного обеспечения для воспроизведения музыки. При отсутствии нелинейных искажений в системах воспроизведения амплитудно-временного формата все системы воспроизведут одно и то же музыкальное произведение идентично. Наибольшую популярность среди аналоговых носителей информации получили грампластинки и аудиокассеты с магнитной лентой. Среди цифровых носителей информации без сжатия наиболее популярны компакт-диски формата *CD-DA* (*Compact Disk Digital Audio*) и файлы *Wave* (*Microsoft Waveform data*). Музыкальная информация форматов *Wave* и *CD-DA* представлена в виде решетчатой функции стандарта *PCM* (*Pulse Code Modulation* — импульсно-кодовая модуляция). Наиболее часто цифровые музыкальные форматы используют частоту дискретизации сигнала 44100 Гц с количеством уровней квантования 2^{16} . Такие параметры позволяют передавать музыкальные записи без воспринимаемых на слух искажений. На рис.2 приведено амплитудно-временное представление (АВП) фрагментов мелодии фортепианной полифонии В.А.Моцарта.



а)



б)

Рис.1. Визуализированное объектное представление фрагмента мелодии фортепианного произведения В.А. Моцарта: а) нотная запись; б) представление Sakewalk Piano Roll.

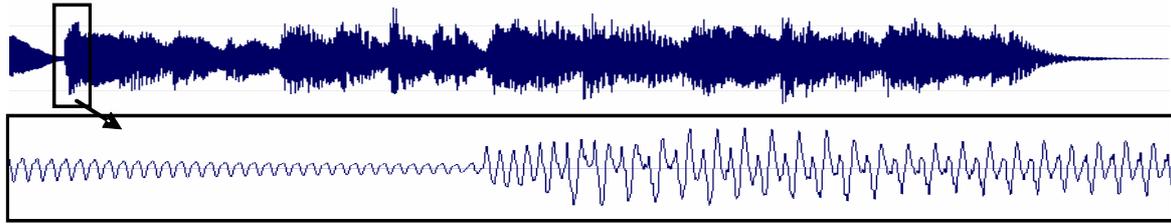


Рис. 2. Амплитудно-временное представление мелодии В.А. Моцарта: фрагмента продолжительностью 2,0 с (сверху); фрагмента продолжительностью 0,3 с (снизу). Вдоль оси ординат отложены значения амплитуды. Ось абсцисс (здесь и далее) представляет собой ось времени.

Одним из недостатков амплитудно-временных форматов является большой объем цифровых данных получаемых при хранении и обработке музыкальных произведений. Так, полифония В.А. Моцарта продолжительностью 11 минут в MIDI формате занимает 43Кб, а в формате Wave - 69,5 Мб.

Соответственно двум форматам представления музыкальных произведений, возможны два типа преобразований: прямое, позволяющее объектный формат музыкального произведения преобразовать в амплитудно-временной, и обратное, позволяющее по амплитудно-временному формату получить объектный.

Задача преобразования объектного формата в амплитудно-временной (прямое преобразование) решена, и активно используется с 80-х годов XX века многими музыкальными синтезаторами.

Однако, актуальна и обратная задача, согласно которой необходимо амплитудно-временной формат звукозаписи преобразовать в объектный. Например, по звукозаписи получить нотные партитуры всех звучащих инструментов. Область применения такого преобразования обширна:

- восстановление архивных звукозаписей;
- нотное представление игры виртуозов-исполнителей и композиторов;
- уменьшение объемов звуковых данных;
- редактирование партий музыкальных инструментов;
- создание караоке-произведений;
- процесс обучения.

На сегодняшний момент рынок программных продуктов не предлагает программных решений, способных выполнить преобразование амплитудно-временного формата музыкального произведения в объектное. А единственной технологией осуществления преобразо-

вания является слуховой анализ музыкального произведения с последующим рутинным подбором музыкантом-аранжировщиком партий каждого инструмента, входящего в состав произведения [1].

В данной работе приведено описание программно-технического комплекса, выполняющего формирование аудио-данных в объектном формате с последующим преобразованием их в амплитудно-временной формат и снова в объектный. Структурная схема комплекса приведена на рисунке (Рис.3). Комплекс включает в себя ряд технических и программных компонентов, обозначенных на схеме отдельными блоками. Стрелками обозначены возможные направления и пути передачи информации. Каждый элемент схемы снабжен подписью с названием программного пакета, дополняющего функции технического устройства.

Программно-аппаратный комплекс выполнен на базе персонального компьютера с использованием оригинальных программных модулей, разработанных и реализованных в средах *Borland Delphi* и *C++ Builder*. Кроме самостоятельно разработанных программ, используется коммерческое программное обеспечение сторонних производителей. Каждый блок, обозначенный на схеме, имеет аппаратную (утолщенный контур на рисунке) либо программную реализацию и выполняет определенные функции. Реализация отдельных блоков рассмотрена в [1, 2].

Функционально блоки установки объединены в три модуля:

- модуль формирования музыкального фрагмента;
- модуль математических преобразований;
- модуль идентификации и распознавания музыкальных объектов.

Рассмотрим работу и функции каждого модуля.

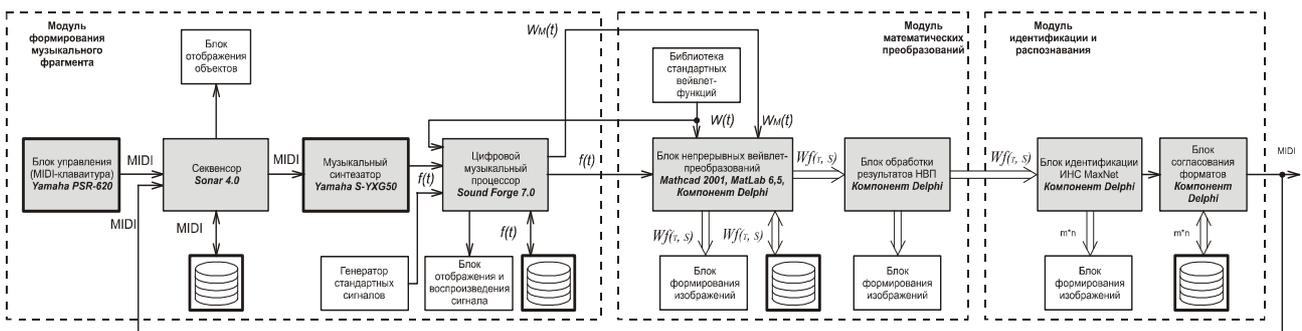


Рис. 3. Структурная схема программно-аппаратного комплекса



Модуль формирования музыкального фрагмента

предназначен для создания экспериментальных сигналов и фрагментов музыкальных произведений для дальнейшей автоматизированной обработки. Модуль позволяет сгенерировать музыкальный фрагмент в требуемом формате. В блоке секвенсора *Sonar 4.0* фирмы *Cakewalk* формируется музыкальный фрагмент в объектном представлении (в виде команд интерфейса *MIDI*). При помощи программы-секвенсора *Sonar* формируются партии всех музыкальных инструментов произведения и объединяются в законченную музыкальную композицию. Все операции совершаются в объектном виде с объектами *MIDI*. Управление секвенсором осуществляется при помощи блока управления. Блок управления *Yamaha PSR-620* представляет собой аппаратно-реализованное устройство, формирующее команды *MIDI*-интерфейса. Блок управления имеет большое количество органов управления, в том числе *MIDI*-клавиатуру, состоящую из 56 клавиш, идентичных клавиатуре фортепиано. Таким образом, сформированное музыкальное произведение в *MIDI*-формате с выхода блока секвенсора поступает на вход музыкального синтезатора.

Музыкальный синтезатор выполняет процедуру преобразования объектного представления музыкальной информации в амплитудно-временное. На выходе блока формируется решетчатая функция $f(t)$, соответствующая АВП сигнала. Частота дискретизации сигнала $f(t)$ $\nu_s=44100$ Гц, количество уровней квантования $N=2^{16}=65536$.

Музыкальный синтезатор может быть реализован как программно, так и аппаратно. В качестве музыкального синтезатора данной установки используется программный синтезатор *Yamaha S-YXG50*. Он способен воспроизводить голоса 625 музыкальных инструментов и формировать многоголосное произведение, с одновременным звучанием до 512 нот.

Для дополнительной обработки дискретизированный сигнал $f(t)$ поступает с выхода блока музыкального синтезатора на вход блока *цифрового музыкального процессора* (ЦМП). В данной установке функции музыкального процессора выполняет программный продукт *Sony Sound Forge 7.0*. И хотя этот ПП достаточно дорогостоящий и обладает сложным интерфейсом, он наиболее функционален из предлагаемых на рынке музыкального ПО. Музыкальный процессор используется для:

- линейной и нелинейной обработки сигнала в цифровом виде;
- визуализации амплитудно-временного представления сигнала и его фрагментов;
- воспроизведения заданного фрагмента сигнала;
- выделения из сигнала $f(t)$ фрагмента для построения вейвлет-функции $fw(t)$;
- сохранения и восстановления сигнала на носителях информации.

Звуковое воспроизведение сформированного и обработанного сигнала осуществляется при помощи цифро-аналогового преобразователя звуковой карты компьютера и динамических акустических систем.

Модуль математической обработки является основой вычислительного эксперимента и предназначен для формирования по АВП сигнала $f(t)$ его амплитудно-

частотно-временного представления (АЧВП) $Wf(\tau, s)$, описывающего мелодическую линию каждого музыкального инструмента.

Модуль математической обработки состоит из двух основных программных блоков:

- блок непрерывных вейвлет-преобразований (НВП);
- блок обработки результатов НВП.

Блок непрерывных вейвлет-преобразований осуществляет непрерывное вейвлет-преобразование дискретизированного сигнала $f(t)$. Основу этого блока представляет программный модуль разработанный в среде *Borland Delphi*. Результатом НВП является двумерный числовой массив, $Wf(\tau, s)$ описывающий амплитудно-частотно-временные характеристики исследуемого сигнала:

$$Wf(\tau, s) = \langle f, w_{s,\tau} \rangle = \int_{-\infty}^{\infty} f(t) w_{s,\tau}(t) dt.$$

Качественно, результаты НВП в значительной степени зависят от выбора функции материнского вейвлета $w(t)$ [3]. Семейство вейвлет-функций $w_{s,\tau}(t)$ получают на основании материнского вейвлета $w(t)$ или $fw(t)$ масштабированием его параметром s и сдвигом на величину τ .

$$w_{s,\tau}(t) = \frac{1}{\sqrt{s}} w\left(\frac{t-\tau}{s}\right).$$

Варьируя параметрами масштабирования s и смещения τ , достигают необходимой локализации как в частотной, так и во временной областях соответственно [4,5].

Семейства вейвлет-функций, основанные на различных материнских вейвлетах, обладают различными избирательными способностями, как в области частоты, так и времени. Так как операция свертки, лежащая в основе НВП, определяет степень корреляции двух функций (сигнала $f(t)$ и вейвлета $fw(t)$), то идентичные по форме функции сигнала и вейвлета дадут максимально возможные значения амплитуды в массиве результатов НВП. Это свойство использовано для формирования материнских вейвлетов, соответствующих различным музыкальным инструментам. Из аудио-записи звучания одной ноты музыкального инструмента в блоке цифрового музыкального процессора выбирается фрагмент длиной в 4 — 16 периодов квазистационарный по амплитуде и частотному спектру. Количество периодов в звучании музыкального инструмента определяет избирательные способности материнского вейвлета. Кроме того, созданный вейвлет должен удовлетворять двум условиям [6]:

$$\int_{-\infty}^{+\infty} fw(t) dt = 0 \text{ и } fw(t) \rightarrow 0, \text{ при } t \rightarrow \infty.$$

В данной работе в соответствии музыкальным инструментам созданы одноименные материнские вейвлеты, применение которых в НВП позволяет достичь высокой степени идентификации мелодии конкретного музыкального инструмента. Кроме того, использование в качестве базиса НВП вейвлета определенного музыкального инструмента позволяет получать информацию об объектах — нотах, а не о поведении спектра сигнала во времени. В то время как использование в качестве базиса НВП сигнала одной гармоникой дает именно АЧВП сигнала в явном виде. Хотя стандартные функции мате-

ринских вейвлетов не позволяют идентифицировать звучание определенного инструмента, в блоке НВП комплекса реализована возможность их применения из библиотек ПП Mathcad и MatLab.

Результаты НВП весьма специфичны, и, даже при визуальном анализе их графической интерпретации, могут трактоваться по-разному. На рис. 4 приведена графическая интерпретация результатов НВП мелодии В.А.Моцарта. Объектное представление этой же мелодии приведено на рис. 1.б. Ось абсцисс здесь соответствует оси времени, вдоль оси ординат отложены значения частоты, эквивалентные частотам нот фортепиано. Темные точки соответствуют элементам массива $Wf(\tau, s)$ с большими значениями амплитуды.

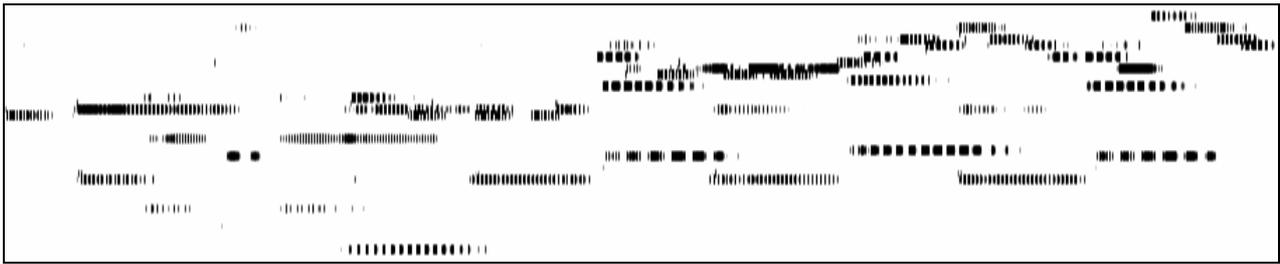


Рис. 4. Графическое представление результатов НВП

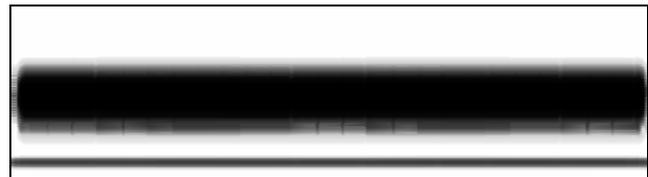
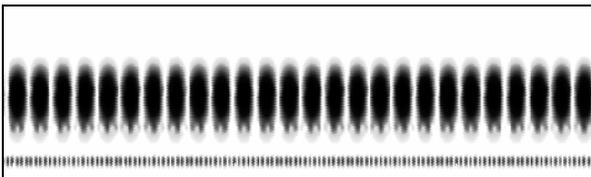


Рис. 5. а) графическое представление результатов НВП $Wf(\tau, s)$; б) идеализированная АЧВХ $Wf^*(\tau, s)$

Результаты НВП содержат большое количество неинформативных участков АЧВП сигнала, которые для повышения качества дальнейшей автоматизированной обработки могут быть устранены. На рис. 5.а приведены результаты НВП сигнала вида $F(t) = \sin(2 \cdot \pi \cdot 25 \cdot t) + \sin(2 \cdot \pi \cdot 120 \cdot t)$. В его спектре присутствуют только две гармонические составляющие. АЧВП сигнала $F(t)$ должно представлять собой две непрерывные во времени линии, соответствующие двум частотам: 25 и 120 Гц. Однако, реальные результаты НВП, представленные на рис. 6.а, содержат два ряда всплесков амплитуды следующих с периодами гармонических составляющих сигнала $F(t)$ (на рисунке показаны черным цветом). Так как сигнал непрерывный во времени, между всплесками амплитуды светлых промежутков не должно быть. Для упрощения графической интерпретации и повышения информативности в модуль математических преобразования включен блок обработки результатов НВП. Блок выполняет сглаживание результатов НВП на основе эвристического алгоритма построения секущей к частотным срезам АЧВП. Так, на каждом частотном срезе АЧВП выполняется сглаживание амплитудных всплесков следующих с периодом, жестко зависящем от текущей частоты среза [2]. На выходе модуля формируется идеализированное АЧВП $Wf^*(\tau, s)$ исследуемого сигнала в виде двумерного массива значений, пригодное для дальнейшей автоматизированной обработки (рис 5.б).

Модуль идентификации и распознавания предназначен для автоматизированной идентификации музыкальных объектов формата MIDI и состоит из двух блоков: блока искусственной нейронной сети (ИНС) и блока согласования форматов. Оба блока реализованы в среде программирования Delphi.

Блок идентификации представляет собой программный модуль, моделирующий работу искусственной нейронной сети. В качестве ИНС используется модифицированная сеть типа «MaxNet», с прямыми связями и варьируемым числом слоев.

Количество слоев ИНС N определяется по формуле:

$$N = \frac{s}{2} - P,$$

где P — уровень полифонической идентификации, зависящий от количества звуков (основных гармоник), присутствующих в сигнале одновременно, которые сеть должна идентифицировать; s — количество масштабов материнского вейвлета, используемого при НВП, определяющих диапазон частот идентифицируемой системой.

ИНС обрабатывает каждый временной срез АЧВП сигнала, выявляя частотные составляющие с максимальными значениями амплитуды. Количество выявленных максимумов каждого среза косвенно зависит от значения P . С каждой частотной составляющей, отобранной сетью, связано ее значение амплитуды в относительных единицах, которое указывает на громкость звучания ноты определенной частоты.

Разрешение по времени при идентификации музыкальных объектов зависит от параметра τ НВП и частоты осуществления временных срезов на входе сети.

Блок согласования форматов предназначен для формирования музыкального файла в стандарте MIDI из массива данных, полученного на выходе блока идентификации. В соответствие выявленным частотным составляющим выделяются объекты, то есть ноты, а в зависимости от присутствия частоты на смежных временных участках определяется время начала и продолжительность звучания (длительность) каждой ноты.

С выхода блока согласования форматов, по каналу обратной связи, файл в формате MIDI может передаваться на вход секвенсора для возможной обработки, воспроизведения и окончательного применения.

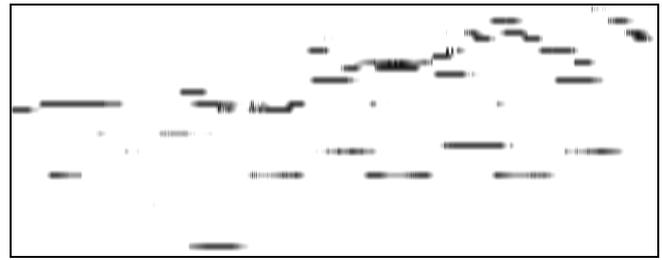
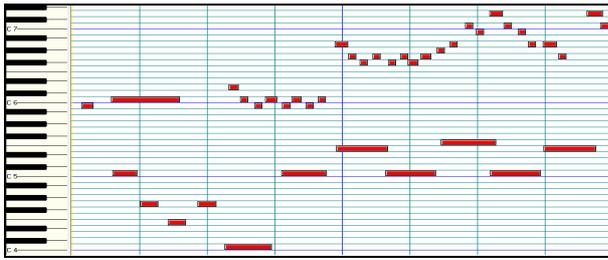


Рис. 6. а) объектное представление полифонической мелодии В.А.Моцарта; б) результат НВП мелодии представленной в виде решетчатой функции.

Результаты экспериментальных исследований

Задачу распознавания нот, как объектов музыкальной информации, можно разбить на две [7]:

1. Определение свойств музыкальных объектов одного музыкального инструмента.
2. Идентификация мелодии одного инструмента на фоне звучания других;

Решение первой задачи заключается в определении высотного положения (тона), времени начала звучания и продолжительности звучания каждой ноты в мелодии. Любой музыкальный инструмент может воспроизвести строго ограниченное число нот. Частоты основных гармоник всех нот, которые могут быть сыграны музыкальными инструментами, дискретизированы и описываются формулой

$$M_i = 27.5 \cdot 2^{\frac{i}{12}},$$

где $M_0=27.5$ Гц — частота ноты «Ля» субконтроктавы, $M_{06}=7040$ Гц — частота ноты «Ля» пятой октавы, i - порядковый номер ноты (на клавиатуре фортепиано), начиная с «Ля» субконтроктавы.

Для исключения неинформативных частот, соответствующих интервалам меньшим музыкального полутона, блок НВП формирует семейство вейвлет функций строго по дискретизированным частотам M_i . На выходе модуля математических преобразований формируется массив $Wf'(\tau, s)$, состоящий из 96 векторов длиной T , каждый из которых описывает поведение одной ноты на всем временном интервале исследования.

Используя частоту дискретизации равной 44100 Гц, модуль НВП формирует сдвиг материнского вейвлета при формировании семейства на величину $\tau = 1/44100$ с. В этом случае время начала звучания ноты $t_n = 44100 \cdot \tau_n$ [с], где τ_n - порядковый номер отсчета

в массиве $Wf'(\tau, s)$, соответствующий начальному моменту звучания ноты. Продолжительность звучания ноты $t_l = 44100 \cdot (\tau_k - \tau_n)$ [с], где τ_k — номер отсчета, соответствующего моменту окончания звучания ноты. А количество отсчетов в массиве $Wf'(\tau, s)$ $T=44100 \cdot t$, где t — время звучания исследуемого музыкального фрагмента.

Ряд экспериментов, проведенных в процессе исследовании свойств НВП и создаваемых вейвлет-функций, показал, что при звучании одного инструмента НВП позволяет достичь высокой степени идентификации нот в мелодии даже при полифоническом звучании, когда в каждый момент времени звучит более одной ноты.

На рис. 6 приведен фрагмент мелодии В.А.Моцарта, сформированный в объектном виде и преобразованный в АВП, а затем обработанный модулем математических преобразований. Визуальный анализ результатов позволяет говорить об идентификации нот, исполняемых с минимально возможным интервалом исполнения по высоте ноты — пол тона (малая секунда).

Второй задачей распознавания музыкальных объектов является задача идентификации звучания одного музыкального инструмента на фоне других.

На рис. 7 приведено: а) объектное представление мелодий двух инструментов — колокольчика и трубы; б) результаты НВП записи звучания обоих инструментов вейвлетом «труба»; в) результаты НВП вейвлетом «колокольчик». Визуальный анализ сформированных изображений позволяет выделить мелодии обоих инструментов. Вейвлет «труба», имея более широкий частотный спектр, позволяет точно идентифицировать мелодию трубы. Однако Вейвлет «колокольчик», выделяя на фоне трубы мелодию колокольчика полностью, идентифицирует и некоторые звуки трубы как звуки колокольчика, давая ошибочные результирующие данные.

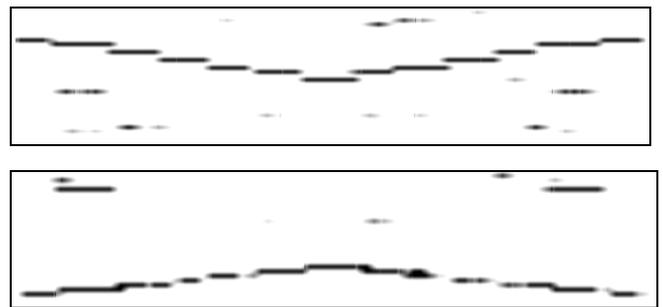
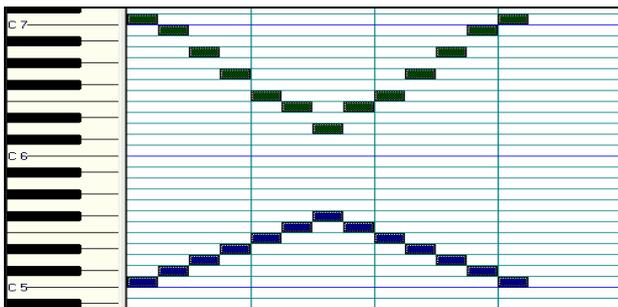


Рис. 7. Объектное представление (а) и результаты НВП музыкального фрагмента исполненного двумя инструментами (б,в).

Заключение

Разработан программно-технический комплекс, позволяющий решать задачу идентификации партии определенного музыкального инструмента с распознаванием в ней объектов-нот: их высотного положения и временных характеристик. Экспериментально доказана возможность идентификации партии полифонического исполнения одним инструментом (когда в каждый момент времени может звучать более одной ноты), а также партии одного инструмента на фоне звучания других.

Для идентификации музыкальных инструментов были сформированы материнские вейвлеты шестнадцати музыкальных инструментов и сложных сигналов: (скрипка, скрипка-пиццикато, фортепиано, орган, труба, колокольчик, рождественский колокольчик (синтезированный искусственно), флейта и др., вейвлеты сигнала синусоидальной, треугольной и прямоугольной форм). Отработана методика формирования материнских вейвлетов, позволяющая получать определенный вейвлет для требуемого музыкального инструмента, т.е. сигнала с заданными характеристиками.

Литература

1. Фадеев А.С., Кочегурова Е.А. К вопросу о преобразовании музыкальных форматов // Современные проблемы информатизации в моделировании и программировании: Сб. трудов XII Междунар. открытой научной конф. – Воронеж: Научная книга, 2006. – С. 255 – 257.
2. Фадеев А.С., Кочегурова Е.А. Подготовка результатов непрерывного вейвлет-преобразования к автоматизированной обработке // Известия Томского политехнического университета, 2006 - т. 309, - № 7. - с. 32-35.
3. Фадеев А.С., Кочегурова Е.А. Выбор вейвлет-функций для анализа музыкальной информации // Современные техника и технологии: Труды XII Международной научно-практической конференции студентов и молодых ученых - Томск, 27-31 марта 2006. - Томск: ТПУ, 2006 - т. 2. - с. 194-196.
4. Чуи К. Введение в вейвлеты. – М.: Мир, 2001. – 412 с.
5. Васильева Л.Г. и др. Преобразования Фурье и вейвлет-преобразования. Их свойства и применение. // Вычислительные методы и программирование. 2002. Т. 3. с 172 - 175.
6. Переберин. А.В. О систематизации вейвлет-преобразований. // Вычислительные методы и программирование. 2001. Т. 2. с 15 - 40.
7. Кочегурова Е.А., Фадеев А.С. Вейвлет анализ в задаче идентификации музыкальной информации // Молодежь и современные информационные технологии: Сб. трудов IV Всеросс. научно-практ. конф. студентов, аспирантов и молодых ученых. – Томск: Изд-во ТПУ, 2006. – С. 149 – 151.

ИНФОРМАЦИОННОЕ ПИСЬМО

Уважаемый коллега!

Приглашаем Вас принять участие в работе 10-ой Международной конференции и выставки «Цифровая обработка сигналов и ее применение – DSPA'2008», которая состоится в Москве 26-28 марта 2008 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 и робототехники
- Цифровая обработка измерительной информации
- Нейрокомпьютерная обработка сигналов и изображений
- Цифровое телевидение и телетрансляция
- Цифровое радиовещание
- ЦОС в системах защиты информации
- Проектирование и техническая реализация систем ЦОС
- ЦОС в открытых системах
- Проблемы подготовки специалистов в области ЦОС

СРОКИ ПРЕДСТАВЛЕНИЯ МАТЕРИАЛОВ

до 20 декабря 2007
Тексты тезисов и докладов

после 20 января 2008
Информация для авторов о приеме докладов на сайте: www.rntores.ru

15 февраля 2008
Рассылка приглашений по почте

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; tor@rgrta.ryazan.ru
[http:// www.dspa.ru](http://www.dspa.ru) и www.rntores.ru (раздел конференции)

ОПТИМИЗИРОВАННЫЕ АЛГОРИТМЫ ПРЕОБРАЗОВАНИЙ ХААРА ДЛЯ ОТЕЧЕСТВЕННОЙ ПЛАТФОРМЫ DSP «ELCORE» И ИХ ПРИМЕНЕНИЕ В ЗАДАЧАХ ГРАФИЧЕСКОЙ ОБРАБОТКИ ДАННЫХ

Милов А.Н.

Введение

ГУП НПЦ «ЭЛВИС» является разработчиком собственной IP-ядерной платформы проектирования систем на кристалле «Мультикор», в состав которой входит библиотека оригинальных ядер цифровых сигнальных процессоров (DSP) серии «ELcore™». Процессоры DSP используются в решении задач таких важных областей, как радиолокация, связь, обработка, кодирование и сжатие сигналов, изображения и звука.

Процессоры серии «Мультикор» являются многоядерными микроконтроллерами с двумя типами ядер[1]: RISC-ядро является MIPS32-совместимым (R3000), выполняет функции управления и является ведущим (master). DSP-ядро «Elcore» является высокопроизводительным 32-битным ведомым (slave) устройством и имеет типичную для многих цифровых процессоров обработки сигналов гарвардскую архитектуру с внутренним параллелизмом по потокам и форматам обрабатываемых данных. Оно предназначено для высокоскоростной обработки информации в форматах с фиксированной и с плавающей точкой. Существенным достоинством данного DSP-ядра для задач обработки компьютерной графики является то, что система его инструкций обеспечивает одновременное выполнение в течение одного командного цикла до двух вычислительных операций и до двух пересылок. Достигается это наличием многопортового регистрового файла и оригинальной архитектуры арифметико-логического устройства (АЛУ).

Все вычислительные операции, поддерживаемые «Elcore», подразделяются на два типа в зависимости от операционных исполняющих устройств:

1. Операции типа OP1 - это операции, исполняемые АЛУ.
2. Операции типа OP2 – это операции, исполняемые при помощи умножителя-сдвигателя (УС).

Параллельное выполнение двух вычислительных операций DSP-ядром «Elcore» возможно только при условии, что они выполняются на разных операционных устройствах (более подробно см. [2]). DSP «Elcore» позволяет также, оставаясь в рамках одной системы инструкций, наращивать количество вычислительных секций (по SIMD-типу), увеличивая тем самым общую производительность системы.

DSP-ядро «Elcore» имеет два типа встроенной памяти, которые условно обозначаются X и Y. X-память используется как для записи, так и для чтения. Операции обмена с этой памятью могут осуществляться для любого регистра DSP. Y-память является подмножеством X-памяти и может быть использована только для чтения входных данных в регистр R0. Вся встроенная память «Elcore» является

Рассматриваются особенности реализации одномерного и двумерного преобразований Хаара дискретных сигналов на отечественной платформе сигнальных процессоров DSP «Elcore», а также их применение для построения огрубленных версий сеток высот местности и построения списков уменьшенных образов текстур.

двухпортовой. Это означает, что обращения к ней со стороны RISC-ядра или DMA-канала памяти происходят без приостановки конвейера «Elcore».

Преобразование Хаара является простейшим вейвлет-преобразованием, выделяющим высокочастотную (ВЧ) и низкочастотную (НЧ) составляющие сигнала. Рекурсивные формулы одномерного преобразования Хаара для сигнала $x = \{x_j\}_{j \in \mathbb{Z}}$ могут быть записаны следующим образом [3]:

$$x_j^{(i)} = \frac{x_{2j}^{(i+1)} + x_{2j+1}^{(i+1)}}{2}, y_j^{(i)} = \frac{x_{2j}^{(i+1)} - x_{2j+1}^{(i+1)}}{2},$$

где i – уровень разрешения сигнала (для рекурсивной свертки),

j – индекс соответствующего отсчета сигнала x или y ,
 y – ВЧ составляющая сигнала x .

В работе [3] было показано, что двумерное преобразование Хаара можно реализовывать не по общей схеме композиций одномерных преобразований, а по следующим формулам:

$$\begin{aligned} x_{j,k}^{(i)} &= \frac{x_{2j,2k}^{(i+1)} + x_{2j+1,2k}^{(i+1)} + x_{2j,2k+1}^{(i+1)} + x_{2j+1,2k+1}^{(i+1)}}{4}; \\ y_{j,k}^{(i)} &= \frac{x_{2j,2k}^{(i+1)} + x_{2j+1,2k}^{(i+1)} - x_{2j,2k+1}^{(i+1)} - x_{2j+1,2k+1}^{(i+1)}}{4}; \\ y_{j,k}^{\prime(i)} &= \frac{x_{2j,2k}^{(i+1)} - x_{2j+1,2k}^{(i+1)} + x_{2j,2k+1}^{(i+1)} - x_{2j+1,2k+1}^{(i+1)}}{4}; \\ y_{j,k}^{\prime\prime(i)} &= \frac{x_{2j,2k}^{(i+1)} - x_{2j+1,2k}^{(i+1)} - x_{2j,2k+1}^{(i+1)} + x_{2j+1,2k+1}^{(i+1)}}{4}. \end{aligned} \quad (1)$$

В настоящей статье представлен алгоритм реализации этих преобразований на практике и дана оценка производительности для отечественной платформы DSP «Elcore».

Особенности реализации одномерного преобразования Хаара для «Elcore»

Предлагается следующий конвейерный алгоритм обработки данных одномерного преобразования Хаара для чисел в формате IEEE-754 (рис.1Рис.). В строках указана последовательность тактов DSP, в столбцах приведены первые шесть 32-битных регистра из регистрового файла «Elcore».

	R0	R2	R4	R6	R8	R10
1 (AT)+IT						0.5
2 (AT)+IT	R0*R10					0.5
3 (AT)+IT		R0*R10				0.5
4 (AT)+IT	R2±R4		R0*R10			0.5
5 (AT)+IT	(A0)+				R0*R10	0.5
6 (AT)+IT	R0*R10	(A1)+		R6±R8		0.5
7 (AT)+IT		R0*R10	(A0)+			0.5
8 (AT)+IT	R2±R4		R0*R10	(A1)+		0.5
9 (AT)+IT	(A0)+				R0*R10	0.5
10 (AT)+IT	R0*R10	(A1)+		R6±R8		0.5
11 (AT)+IT		R0*R10	(A0)+			0.5
12 (AT)+IT	R2±R4		R0*R10	(A1)+		0.5

Рис. 1. Реализация одномерного преобразования Хаара

Условные обозначения:

	– операция OP1;	0.5	– константное значение в регистре;
	– операция OP2;	R0*R10	– операция умножения регистров в формате с плавающей точкой;

R2±R4 – операция «сложение-вычитание» регистров в формате с плавающей точкой (FAS), результаты помещаются в исходные регистры;

(AT)+IT – загрузка входных данных из Y-памяти и пост-инкремент адресного регистра AT на величину IT (IT=1);

(A0)+,(A1)+ – выгрузка обработанных данных в X-память и пост-инкремент соответствующего адресного регистра (A0, A1).

Алгоритм использует две пары регистров (R2,R4) и (R6,R8) и чередует чтение, обработку и запись данных с одной пары на другую.

Входные данные для обработки загружаются из Y-памяти в регистр R0. Ввиду того, что регистровый файл является многопортовым, параллельно происходит операция деления предыдущего входного отсчета, и результат помещается в регистр R2 (или R6).

Подготовка конвейера осуществляется до четвертого такта включительно. Основная стадия конвейера преобразования занимает восемь тактов (на диаграмме она показана с 5 такта по 12 включительно). На каждой такой стадии из восьми отсчетов входного сигнала образуются по четыре отсчета НЧ и ВЧ составляющих. В течение основной стадии происходит 16 обменов с памятью данных (используется два адресных генератора – AGU и AGU-Y) и 16 арифметических операций с плавающей точкой в формате IEEE-754.

Заметим, что вместо операции FAS могли быть использованы операции FADD (сложение) и FSUB (вычитание), однако в этом случае потребовался бы как минимум еще один регистр из регистрового файла. Кроме того, операция FAS позволила освободить устройство

АЛУ на 5, 9, 13 тактах. Как будет показано ниже, операция FAS наиболее эффективно может быть применена для конвейера двумерного преобразования Хаара.

Итак, производительность приведенного алгоритма ограничена выгрузкой результатов в X-память данных. Поэтому, приведенный алгоритм является полностью оптимизированным для данного устройства.

Поскольку конвейер преобразования является довольно коротким, то особо важным моментом является минимизация затрат на выполнение подготовительных действий, связанных с организацией циклов. Для этого в DSP «Elcore» предусмотрена аппаратная поддержка программных циклов типа DO, которая позволяет избежать накладных расходов при организации простых циклов, связанных с инкрементом счетчика циклов и проверкой условия конца цикла. Необходимым требованием при использовании этого аппаратного механизма является то, чтобы число повторов цикла могло быть вычислено до вхождения в цикл. Помимо этого, оно должно находиться в пределах от 1 до 16383.

Основная стадия конвейера преобразования (5-12 такты диаграммы), описанная на языке ассемблера «Elcore», приведена на рис. 2.

```

DO Nd2, LastLabel:
FMPY R0,R10,R8      R2.L,(A0)+      (AT)+IT,R0.L
FMPY R0,R10,R2      FAS R6,R8      R4.L,(A1)+      (AT)+IT,R0.L
FMPY R0,R10,R4      R6.L,(A0)+      (AT)+IT,R0.L
FMPY R0,R10,R6      FAS R2,R6      R8.L,(A1)+      (AT)+IT,R0.L
FMPY R0,R10,R8      R2.L,(A0)+      (AT)+IT,R0.L
FMPY R0,R10,R2      FAS R6,R8      R4.L,(A1)+      (AT)+IT,R0.L
FMPY R0,R10,R4      R6.L,(A0)+      (AT)+IT,R0.L
LastLabel:
FMPY R0,R18,R6      FAS R4,R8      R2.L,(A0)+      (AT)+IT,R0.L

```

Рис. 2. Алгоритм основной стадии конвейера преобразования Хаара.

Реализация двумерного преобразования Хаара для ЦСП «Elcore»

Для того чтобы столь же эффективно реализовать двумерное преобразование Хаара, необходимо представить формулы (1) в другом виде:

$$\begin{aligned}
 x_{j,k}^{(i)} &= \left(\frac{x_{2j,2k}^{(i+1)} + x_{2j+1,2k}^{(i+1)}}{4} \right) + \left(\frac{x_{2j,2k+1}^{(i+1)} + x_{2j+1,2k+1}^{(i+1)}}{4} \right) \\
 y_{j,k}^{(i)} &= \left(\frac{x_{2j,2k}^{(i+1)} + x_{2j+1,2k}^{(i+1)}}{4} \right) - \left(\frac{x_{2j,2k+1}^{(i+1)} + x_{2j+1,2k+1}^{(i+1)}}{4} \right) \\
 y_{j,k}^{\prime(i)} &= \left(\frac{x_{2j,2k}^{(i+1)} - x_{2j+1,2k}^{(i+1)}}{4} \right) + \left(\frac{x_{2j,2k+1}^{(i+1)} - x_{2j+1,2k+1}^{(i+1)}}{4} \right) \\
 y_{j,k}^{\prime\prime(i)} &= \left(\frac{x_{2j,2k}^{(i+1)} - x_{2j+1,2k}^{(i+1)}}{4} \right) - \left(\frac{x_{2j,2k+1}^{(i+1)} - x_{2j+1,2k+1}^{(i+1)}}{4} \right)
 \end{aligned} \quad (2)$$

Заметим, что все комбинации сумм/разностей операндов в (2) можно получить с использованием всего четырех операций FAS. Обозначим:

$$\begin{aligned}
 FAS(n, m)_+ &= n + m, \\
 FAS(n, m)_- &= n - m,
 \end{aligned}$$

$$\begin{pmatrix} N \\ M \end{pmatrix} = FAS(n, m) \Leftrightarrow \begin{cases} N = FAS(n, m)_+ \\ M = FAS(n, m)_- \end{cases}$$

Тогда двумерное преобразование Хаара может быть записано следующим образом:

$$\begin{pmatrix} x_{j,k}^{(i)} \\ y_{j,k}^{(i)} \end{pmatrix} = FAS \left(FAS \left(\frac{x_{2j,2k}^{(i+1)}}{4}, \frac{x_{2j+1,2k}^{(i+1)}}{4} \right)_+, FAS \left(\frac{x_{2j,2k+1}^{(i+1)}}{4}, \frac{x_{2j+1,2k+1}^{(i+1)}}{4} \right)_+ \right)$$

$$\begin{pmatrix} y_{j,k}^{\prime(i)} \\ y_{j,k}^{\prime\prime(i)} \end{pmatrix} = FAS \left(FAS \left(\frac{x_{2j,2k}^{(i+1)}}{4}, \frac{x_{2j+1,2k}^{(i+1)}}{4} \right)_-, FAS \left(\frac{x_{2j,2k+1}^{(i+1)}}{4}, \frac{x_{2j+1,2k+1}^{(i+1)}}{4} \right)_- \right)$$

В соответствии с предложенной формой записи, алгоритм двумерного преобразования Хаара примет вид (рис. 3):

Подготовка конвейера алгоритма занимает 8 тактов, основная стадия конвейера также занимает 8 тактов, в течение которых происходит 16 обращений к памяти и 24 арифметические операции с плавающей точкой. Основная стадия соответствует паре преобразований (1), на входе которых по четыре отсчета двумерного сигнала, на выходе - также по четыре отсчета одного шага двумерного преобразования Хаара.

Заметим, что результаты преобразования могут быть выданы в различные диапазоны X-памяти. Для этого необходимо настроить соответствующие адресные регистры A0, A1, A2, A3.

Основная стадия конвейера предлагаемого алгоритма на языке ассемблера «Elcore» приведена на рис. 4.

	R0	R2	R4	R6	R8	R10	R12	R14	R16	R18	DT
1 (AT)+IT										0.25	N
2 (AT)+DT	R0*R18									0.25	N
3 (AT)-IT		R0*R18								0.25	N
4 (AT)+DT			R0*R18							0.25	N
5 (AT)+IT		R2±R4		R0*R18						0.25	N
6 (AT)+DT			R6±R8	R0*R18						0.25	N
7 (AT)-IT	R2+R6		R2-R6		R0*R18					0.25	N
8 (AT)+DT	(A0)+	R4+R8		R4-R8		R0*R18				0.25	N
9 (AT)+IT		(A1)+			R10±R12		R0*R18			0.25	N
10 (AT)+DT	R0*R18		(A2)+				R14±R16			0.25	N
11 (AT)-IT		R0*R18		(A3)+	R10+R14		R10-R14			0.25	N
12 (AT)+DT			R0*R18		(A0)+	R12+R16		R12-R16		0.25	N
13 (AT)+IT		R2±R4		R0*R18		(A1)+				0.25	N
14 (AT)+DT			R6±R8	R0*R18		(A2)+				0.25	N
15 (AT)-IT	R2+R6		R2-R6		R0*R18		(A3)+			0.25	N
16 (AT)+DT	(A0)+	R4+R8		R4-R8		R0*R18				0.25	N
17 (AT)+IT		(A1)+			R10±R12		R0*R18			0.25	N

Рис. 3. Реализация двумерного преобразования Хаара

Условные обозначения:

(AT) ± IT, (AT) ± DT – загрузка входных данных из Y-памяти и пост-инкремент (декремент) адресного регистра AT на величину регистра смещения IT или DT (IT=1, DT=N),

N – размер матрицы входного двумерного сигнала (предполагается, что матрица имеет NxN элементов).

DO Nd2, LastLabel:

```

FMPY R0, R18, R16   FAS R10, R12   R4.L, (A1)+ (AT)+DT, R0.L
FMPY R0, R18, R2    FAS R14, R16   R6.L, (A2)+ (AT)+IT, R0.L
FMPY R0, R18, R4    FAS R10, R14   R8.L, (A3)+ (AT)-DT, R0.L
FMPY R0, R18, R6    FAS R12, R16   R10.L, (A0)+ (AT)+IT, R0.L
FMPY R0, R18, R8    FAS R2, R4     R12.L, (A1)+ (AT)+DT, R0.L
FMPY R0, R18, R10   FAS R6, R8     R14.L, (A2)+ (AT)+IT, R0.L
FMPY R0, R18, R12   FAS R2, R6     R16.L, (A3)+ (AT)-DT, R0.L
LastLabel:
FMPY R0, R18, R14   FAS R4, R8     R2.L, (A0)+ (AT)+IT, R0.L

```

Рис. 4. Основная стадия алгоритма двумерного преобразования Хаара

Как видно из рисунка, на каждом такте происходит по два обмена с памятью данных, одна операция умножения и одна операция «сложение-вычитание». Поскольку выходные данные могут быть записаны только в X-память, предложенный алгоритм также является полностью оптимизированным для данного устройства.

Оценка производительности предложенных алгоритмов

Архитектура операционных блоков DSP-ядер серии «Elcore» включает два режима обработки данных:

- «Один поток инструкций, один поток данных» (Single Instruction Single Data - SISD);
- «Один поток инструкций, множественные потоки данных» (Single Instruction Multiple Data - SIMD).

Отличие SIMD-режима от SISD состоит в том, что за один командный цикл одна арифметическая инструкция выполняется в нескольких одинаковых операционных блоках одновременно (а именно, наборах АЛУ и УС), причем с разными данными. Таким образом, применение режима SIMD в DSP разработки НПЦ «ЭЛВИС» способно ускорить исполнение математической задачи в несколько раз. DSP-ядро «ELcore-24» микросхемы 1892BM2T «MC-24» обладает двумя операционными SIMD-секциями, то есть вычисления в нем могут быть ускорены в два раза.

Преобразование Хаара локально, т.е. затрагивает только ту область сигнала, к которой оно применяется. Поэтому каждый шаг преобразования Хаара любого одномерного или двумерного сигнала можно осуществить, разложив исходный сигнал на несколько составляющих. Например, если необходимо осуществить один шаг двумерного преобразования Хаара для сигнала, заданного двумерной матрицей размером 128x128, то такое преобразование эквивалентно четырем преобразованиям Хаара для составляющих сигнала с размерами 64x64 или двум преобразованиям для составляющих сигнала с размерами 128x64 или 64x128.

Таким образом, производительность приведенных алгоритмов преобразования Хаара может быть увеличена в два раза включением SIMD-режима и разложением исходного сигнала на составляющие.

В таблице 1 представлены характеристики производительности предложенных алгоритмов для различных

DSP-ядер серии «Elcore». Ввиду того, что внутренняя память DSP-ядер реализована непосредственно на кристалле (а, следовательно, существенно ограничена), в таблице также представлены граничные размеры входных сигналов, для которых возможно преобразование без разделения на составляющие.

Для достижения указанной пиковой производительности при обработке непрерывных потоков данных обычно используют блочную обработку. Суть её состоит в разделении сигнала на составляющие блоки такого размера, чтобы одновременно в памяти входных данных DSP могло содержаться два таких блока. Таким образом, в то время, пока DSP производит обработку одного из блоков, DMA-канал производит выгрузку обработанных данных предыдущего и загрузку входных данных следующего блока.

Для эффективного разделения двумерных сигналов на составляющие и последующей их загрузки в Y-память в микроконтроллерах серии «Мультикор» предусмотрен режим двумерной DMA-передачи. Таким образом, нет необходимости в предварительном построении двумерных массивов данных или организации серии DMA-передач.

Применение преобразования Хаара

Преобразования Хаара применяются для построения списков уменьшенных копий текстур с размерами $2^n \times 2^n, n \in N$ (рис. 5) при осуществлении пространственно-аккуратного текстурирования (англ. тер. – mip-mapping) [4].

В работе [3] было показано, что преобразование Хаара может также применяться для сжатия сеток высот поверхностей, параметризуемых на плоскости. На рис. 6 приведены результаты применения одного шага двумерного преобразования Хаара к заданной сетке высот. Слева направо, сверху вниз: монотонные изображения сеток высот размерами 64x64 и 32x32 соответственно, перспективное представление поверхности, заданной сеткой высот, триангуляция исходной сетки высот 64x64, триангуляция сетки высот 32x32, полученной с помощью двумерного преобразования Хаара. Все изображения получены с использованием графической библиотеки с интерфейсом OpenGL для микропроцессоров серии «Мультикор».

Таблица 1.

Характеристики предложенных алгоритмов.

Преобразование	DSP-ядро	Ограничения вх. сигналов, отсчетов	Пиковая производительность, тактов на входной отсчет
Одномерное в формате IEEE-754	Elcore-14(SISD)	12 288	1
	Elcore-24(SIMD)	8 192	0.5
Двумерное в формате IEEE-754	Elcore-14(SISD)	12 100 (110x110)	1
	Elcore-24(SIMD)	8 100 (90x90)	0.5
Двумерное в формате R8G8B8X8 ¹	Elcore-14(SISD)	12 100 (110x110)	2
	Elcore-24(SIMD)	8 100 (90x90)	1

¹ Широко используется для обработки растровых изображений и текстур. Известно так же как 32-битный формат растровых изображений.



Рис. 5. Список уменьшенных копий текстур

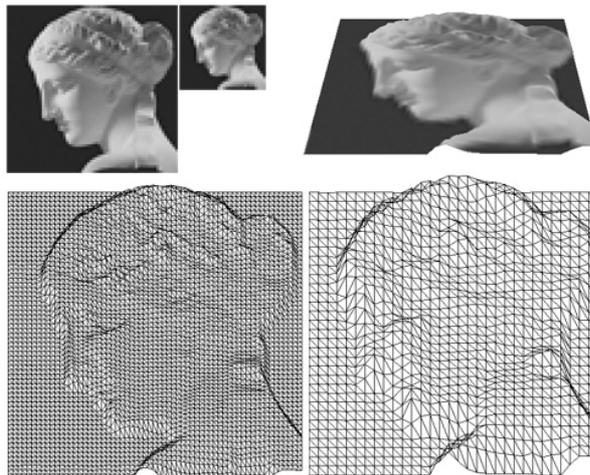
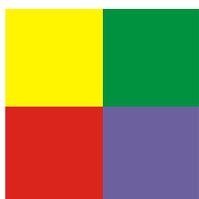


Рис. 6. Один шаг двумерного преобразования Хаара для заданной сетки высот

Литература

1. «Микросхема интегральная 1892ВМЗТ. Руководство пользователя», ГУП НПЦ «ЭЛВИС», Москва, 2005 г.
2. «DSP-ядро ELcore-x4. Система инструкций», ГУП НПЦ «ЭЛВИС», Москва, 2006 г.
3. Переберин А.В., диссертация кандидата физико-математических наук «Многомасштабные методы синтеза и анализа изображений», ИГМ им. Келдыша, Москва, 2002 г.
4. Lance Williams, «Pyramidal parametrics», Computer graphics, volume 17, number 3, 1983.



MATROX
IMAGING

Компания Matrox Imaging более 30 лет успешно поставляет свои ПК-ориентированные программные и аппаратные технологии разработчикам *систем машинного зрения, анализа объектов и изображений, цифрового видео наблюдения, систем визуализации в медицине и биологии, систем управления технологическими процессами.*

Линейка продуктов включает в себя: *фрейм-грабберы, процессоры обработки изображения, видео смарт-камеры, готовые системы машинного зрения, программное обеспечение для анализа изображения (библиотеки MIL). Штаб квартира в Монреале (Канада), офисы и представительства в 20 странах мира.*

Авторизованный представитель компании **Matrox Imaging** в России – **System Video Graphics Alliance (SVGA).**

КРАТКИЕ СООБЩЕНИЯ

УДК 621.372

ЭФФЕКТИВНЫЕ ОКОННЫЕ ФУНКЦИИ НА ОСНОВЕ ФУНКЦИЙ ГАУССА-ЭРМИТА

Вашкевич М.И.

Рассматривается метод расчета эффективных оконных функций, используемых при гармоническом анализе сигналов с применением ДПФ.

Как известно, задача минимизации мощности спектра оконных функций на заданном частотном интервале приводит к использованию вытянутых сфероидальных волновых функций (ВСВФ). Особенность их в том, что они являются собственными (инвариантными) функциями преобразования Фурье на ограниченном интервале. Это означает, что преобразование Фурье переводит каждую ВСВФ в такую же функцию с точностью до константы. Значит, задав во временной области достаточно «узкую» функцию, можно получить функцию ограниченную и в частотной области. Одной из наиболее известных и удачных аппроксимаций ВСВФ являются оконные функции Кайзера. Кроме ВСВФ известен ряд функций, инвариантных к преобразованию Фурье в бесконечных пределах. Функции Гаусса-Эрмита, как один из примеров таких функций, имеют вид:

$$\varphi_n = \omega_n(x) \cdot e^{-x^2/2},$$

где $\omega_n(x)$ - многочлен Эрмита n -го порядка.

Из сказанного выше можно сделать вывод, что функции Гаусса-Эрмита (ФГЭ) могут являться неплохим «материалом» для построения оконных функций. Тем более, что в [1] было показано, что энергия ФГЭ сконцентрирована на ограниченном отрезке, как во временной, так и в частотной области.

Окно, которое исследовалось в данной работе, имеет вид:

$$\omega(k) = e^{\frac{(k/N)^2}{2\sigma^2}} \times \frac{H_4\left(\frac{k \cdot \alpha}{N}\right)}{H_4(0)}, \quad k \in \left[-\frac{N}{2}; \frac{N}{2}\right],$$

N – половина ширины окна,

$H_4(x)$ - многочлен Эрмита 4-го порядка.

В этой формуле между α и σ существует определенная эмпирически установленная зависимость: $\sigma(\alpha) = 0.96275 - 0.8375 \cdot \alpha$. Параметр α может принимать значения из интервала $[0; 0.5]$. По сути α - это константа, аналогичная константе β в окне Кайзера. При увеличении α происходит расширение главного лепестка, но при этом падает уровень боковых лепестков.

Сравнение данного окна производилось с окном Кайзера в среде MatLab 6.5 (при помощи утилиты wvtool). Параметр β в окне Кайзера подбирался таким образом, чтобы ширина главного лепестка была одинаковой у обоих окон. Во всех случаях длина окна $N=65$. В таблице 1 приведены наиболее интересные из полученных результатов.

Таблица 1.

Сравнение экспериментального окна с окном Кайзера

Экспериментальное окно			Окно Кайзера			разность уровней боковых лепестков (дБ)
α	$K_y, \%$	УБЛ (дБ)	β	$K_y, \%$	УБЛ (дБ)	
0,30	0,67	25,3	2,91	0,70	23,6	1,7
0,35	0,33	29,0	3,34	0,35	26,3	2,7
0,40	0,11	34,3	3,95	0,13	30,2	4,1
0,45	0,03	39,6	4,72	0,03	35,4	4,2
0,50	0,01	44,6	5,76	0,01	42,3	2,3

Здесь: K_y - коэффициент утечки, который показывает, какая доля общей мощности сосредоточена в боковых лепестках его спектра. УБЛ - уровень боковых лепестков (уровень максимального из боковых лепестков относительно главного лепестка).

Как видно из полученных результатов, предлагаемое окно имеет несколько лучшие показатели по сравнению с окном Кайзера (при $\beta = [3; 6]$), и поэтому может успешно применяться при гармоническом анализе сигналов с применением ДПФ.

Литература

1. Хонина С.Н., Котляр В.В. Влияние дифракции на изображения, согласованные с вытянутыми сфероидальными волновыми функциями // Компьютерная оптика, ИСОИ РАН, Самара-Москва, 2001, №21, с. 64-69.