УДК 004.383.5

ОПТИМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ БИТ В ПАРАУНИТАРНОМ СУБПОЛОСНОМ КОДЕРЕ ИЗОБРАЖЕНИЙ НА ОСНОВЕ АЛГЕБРЫ КВАТЕРНИОНОВ

Н.А. ПЕТРОВСКИЙ

Белорусский государственный университет информатики и радиоэлектроники
П. Бровки, 6, Минск, 220013, Беларусь

Поступила в редакцию 12 ноября 2013

Рассматривается проблема распределения бит в контексте древовидной структуры 2D параунитарного банка фильтров на основе алгебры кватернионов – пакета 4-полосного вейвлет преобразования. Показано, что оптимальное распределение бит будет в случае, если дисперсии ошибок квантователей в каналах кодера равны.

Ключевые слова: параунитарный банк фильтров, многополосное вейвлет-преобразование, кватернион, изображение, дисперсия реконструкции сигнала.

Постановка задачи

В последние несколько декад банки фильтров рассматриваются как наиболее эффективная техника компрессии данных мультимедиа [1]. Банки фильтров применяются в кодировании аудиосигналов, изображений и видеопотока в таких стандартах как JPEG, JPEG2000, JPEG XR, MPEG и H.264/AVC.

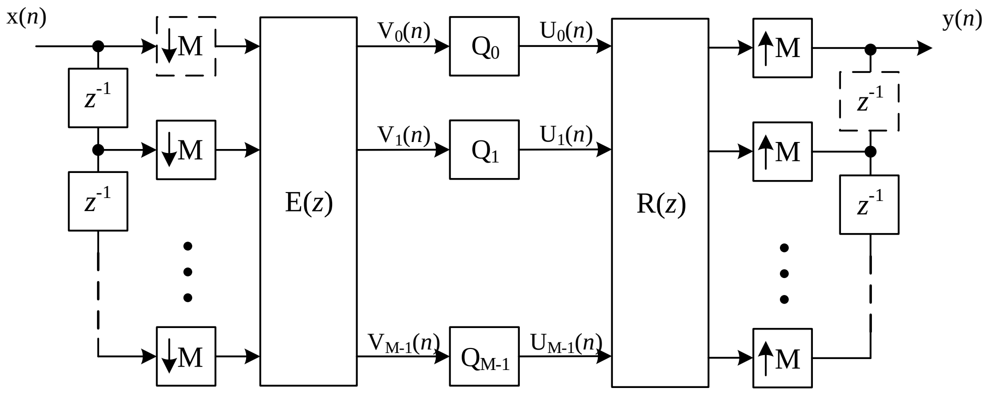


Рис. 1. Схема субполосного кодирования на основе полифазных матриц

На рис. 1 показана схема субполосного кодирования на основе максимально децимированного M-канального банка фильтров, состоящего из систем анализа и синтеза, полифазных матриц $\mathbf{E}(z)$ и $\mathbf{R}(z)$ соответственно. Полифазная структура банка фильтров формулируется следующим образом [1]:

$$\left[\mathbf{H}_0(\mathrm{z}) \quad \mathbf{H}_1(z) \quad \dots \quad \mathbf{H}_{M-1}(z)\right]^{\mathrm{T}} = \mathbf{E}(z^M)\mathbf{e}(z)^{\mathrm{T}},$$

$$\left[\mathbf{F}_0(\mathrm{z}) \quad \mathbf{F}_1(z) \quad \dots \quad \mathbf{F}_{M-1}(z)\right]^{\mathrm{T}} = \mathbf{R}(z)\mathbf{R}(z^M),$$

где $\mathbf{e}(z)=\begin{bmatrix}1 & z^{-1} & \dots & z^{-(M-1)}\end{bmatrix}$, $\mathbf{H}_k(z)$ и $\mathbf{F}_k(z)$ – передаточные функции k-го канала банка фильтров анализа и синтеза соответственно. Если матрица $\mathbf{E}(z)$ обратима, то полифазная матрица синтеза $\mathbf{R}(z)$ может быть выбрана как обратная матрица $\mathbf{E}(z)$, тогда перфективная реконструкция достигается. Такой банк фильтров называется перфективно реконструированный банк фильтров или биортогональный банк фильтров. Если $\mathbf{E}^{\mathrm{T}}(z^{-1})\mathbf{E}(z)=\mathbf{I}$ и $\mathbf{R}(z)=\mathbf{E}^{\mathrm{T}}(z^{-1})$, то это специальный класс банка фильтров – параунитарный банк фильтров (ПУБФ). Структура ПУБФ получается путем факторизации его полифазной матрицы $\mathbf{E}(z)$ на элементарные математические преобразования, как правило, это планарные обороты Гивенса. Для ПУБФ анализа с линейной ФЧХ и четного числа каналов M хорошо известна факторизация передаточной полифазной матрицы $\mathbf{E}(z)$ [2]: $\mathbf{E}(z)=\mathbf{G}_{N-1}(z)\mathbf{G}_{N-2}(z)\ldots\mathbf{G}_1(z)\mathbf{E}_0$, где $\mathbf{E}_0=\frac{1}{\sqrt{2}}\mathbf{\Phi}_0\mathbf{W}\mathrm{diag}(\mathbf{I}_{M/2},\mathbf{J}_{M/2})$, $\mathbf{G}_i(z)=\frac{1}{2}\mathbf{\Phi}_i\mathbf{W}\mathbf{\Lambda}(z)\mathbf{W}$, $i=1,\dots,N-1$, $\mathbf{W}=\begin{bmatrix}\mathbf{I}_{M/2} & \mathbf{I}_{M/2}\\ \mathbf{I}_{M/2} & -\mathbf{I}_{M/2}\end{bmatrix}$, $\mathbf{\Lambda}(z)=\mathrm{diag}(\mathbf{I}_{M/2},z^{-1}\mathbf{I}_{M/2})$, $\mathbf{\Phi}_i=\mathrm{diag}(\mathbf{U}_i,\mathbf{V}_i)$, $\mathbf{U}_i$ и $\mathbf{V}_i$ произвольные ортогональные матрицы размера M/2 x M/2. Кроме того, для упрощения обозначенной выше факторизации для $i>0$ матрица $\mathbf{U}_i$ может быть представлена единичной матрицей, тогда $\mathbf{\Phi}_i=\mathrm{diag}(\mathbf{I},\mathbf{V}_i)$. Как видно, ПУБФ может реализовываться на основе решетчатых и лестничных структур, которые, как известно, сохраняют перфективную реконструкцию и при реализации на арифметике с фиксированной запятой.

Положим, что входной сигнал $x(n)$ стационарный, спектральная мощность которого $S_{xx}(e^{j\omega})$, b_k количество бит представления сигнала $V_k(n)$ на выходе k-го канала системы анализа, а $\sigma^2_{V_k}$ есть дисперсия сигнала $V_k(n)$:

$$\sigma^2_{V_k}=\int_{-\pi}^{\pi}S_{xx}(e^{j\omega})\left|H_k(e^{j\omega})\right|^2\tfrac{d\omega}{2\pi}.$$

Динамический диапазон отсчета сигнала $V_k(n)$ на входе k-го квантователя делится на 2^{b_k} интервалов. Вес старшего значащего бита выбирается пропорционально дисперсии $\sigma^2_{V_k}$ таким образом, чтобы вероятность переполнения была одинаковой во всех каналах. Дисперсия ошибки k-го квантователя при условии, что шум квантования аддитивный ко входному сигналу $V_k(n)$ определяется следующим образом [3]:

$$\sigma^2_{q_k}=\varepsilon_k 2^{-2b_k}\sigma^2_{V_k},$$

где ε_k – константа, которая зависит от функции плотности вероятности сигнала $V_k(n)$. Для параунитарной системы, где сохраняется баланс энергий, шум квантования рассматривается на входе системы синтеза $\mathbf{R}(z)$, который усредняется в каналах синтеза и равен средней дисперсии ошибки реконструкции сигнала:

$$\sigma^2_r=\frac{1}{M}\sum_{k=0}^{M-1}\sigma^2_{q_k}=\sum_{k=0}^{M-1}\varepsilon_k 2^{-2b_k}\sigma^2_{V_k}\ /M.$$

Итак, положим, что общее количество бит для квантователей всех M канальных сигналов $V_k(n),\ k=0,..,M-1$ фиксировано и равно b. Проблема оптимального распределения бит кватователей может быть сформулирована следующим образом: найти $b_0,\dots,b_{M-1}$ при условии, что общий ресурс битов $b=\sum_{k=0}^{M-1}b_k\big/M=\mathrm{const}$ и минимизации $\min\sum_{k=0}^{M-1}\varepsilon_k 2^{-2b_k}\sigma^2_{V_k}\ /M$.

Таким образом, целью оптимизации является минимизация дисперсии ошибки реконструкции сигнала σ_r^2 при ограничении на общий ресурс битов.

В случае схемы кодирования на основе линейного преобразования, как показано в [3, 4], оптимальное распределение бит для каждого квантователя в каналах будет, когда дисперсии шума квантования канальных сигналов равны. Для параунитарного субполосного кодера важно как распределить общий ресурс бит между канальными квантователями и как выбрать/факторизовать параунитарную матрицу $\mathbf{E}(z)$, чтобы максимизировать коэффициент субполосного кодирования. Решение второй задачи показано в [5]. В данной работе определяется схема оптимального распределения бит в параунитарном субполосном кодере изображений, который реализован на основе алгебры кватернионов.

Параунитарный субполосный кодер

Матрицы $\mathbf{E}_0$ и $\mathbf{R}_i$ в факторизации решетчатой структуры параунитарного банка фильтров анализа, как показано в [5], можно выразить через матрицы умножения кватернионов (левого и правого умножения) следующим образом: $\mathbf{E}_0 = \mathbf{M}^+(Q_0)\mathbf{M}^-(P_0)$, $\mathbf{R}_i = \mathbf{M}^\pm(Q_i),\quad i=1,...,N-1$, где P_0 и все Q_i,– некоторые единичные кватернионы, тогда передаточная функция решетчатой структуры общего ПУБФ анализа будет равна: $\mathbf{E}(z) = \mathbf{M}^\pm(Q_{N-1})\mathbf{\Lambda}(z)\mathbf{M}^\pm(Q_{N-2})\mathbf{\Lambda}(z)\ldots\mathbf{M}^\pm(Q_1)\mathbf{\Lambda}(z)\mathbf{M}^+(Q_0)\mathbf{M}^-(P_0)$. В работе [5] утверждается, что для каждой ортогональной матрицы размерностью 4×4 существует уникальная (с точностью до знака) пара единичных кватернионов P и Q таких, что: $\mathbf{M}^+(P)\cdot\mathbf{M}^-(Q) = \mathbf{M}^-(Q)\cdot\mathbf{M}^+(P)$. Данная факторизация всегда соответствуют ортогональному преобразованию сигналов, даже если ее компоненты квантованы. Это происходит потому, что столбцы каждой матрицы умножения кватернионов составлены из одних и тех же элементов с точностью до знака. Реализация параунитарного банка фильтров анализа с линейной ФЧХ на основе умножителей кватернионов предполагает определение $\mathbf{\Phi_0}$ и $\mathbf{\Phi}_i$ как [6]: $\mathbf{\Phi}_0 = \mathbf{M}^-(P_0)\mathbf{M}^+(Q_0)$, $\mathbf{\Phi}_i = \mathbf{M}^-(P_i),\quad i=1,...,N-1$, где все P_i и Q_0 – некоторые единичные кватернионы.

Для схемы компрессии изображений 1D ПУБФ на кватернионах (N = 3) [6] используется для построения 2D разделяемого ПУБФ (см. рис. 2). Соотношение децимации/интерполяции в 2D схеме ПУБФ, эквивалентной рис. 1, задается двумерной матрицей [1] $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$ и коэффициент децимации определяется как ее определитель, т. е. M = 16.

На рис. 3 показаны результаты работы схемы (рис. 2): исходное изображение Lena (512×512) (*а*), декомпозиция изображения полифазной матрицей $\mathbf{E}(z)$ (*б*); реконструированное изображение полифазной матрицей $\mathbf{R}(z)$ (*в*). Эксперимент проводился для ПУБФ с линейной ФЧХ (q_0 = –231/512 + i459/1024 + j0 + k0; p_0 = –7/8 – i3/8+ j0 + k0; p_1 = –3/16+ i15/16+ j0 + k0; p_2 = –9/16 – i13/16+ j0 + k0), которому соответствует 4-х полосное вейвлет-преобразование. Вейвлет-фильтры имеют хорошую частотную избирательность: уровень АЧХ в полосе непропускания для низкочастотных и высокочастотных фильтров составляет около –35 дБ, а для полосовых – около –20 дБ. При этом показатель эффективности кодирования равен 8,1845 дБ. Пиковое значение отношения сигнал-шум здесь составило 53,24 дБ.

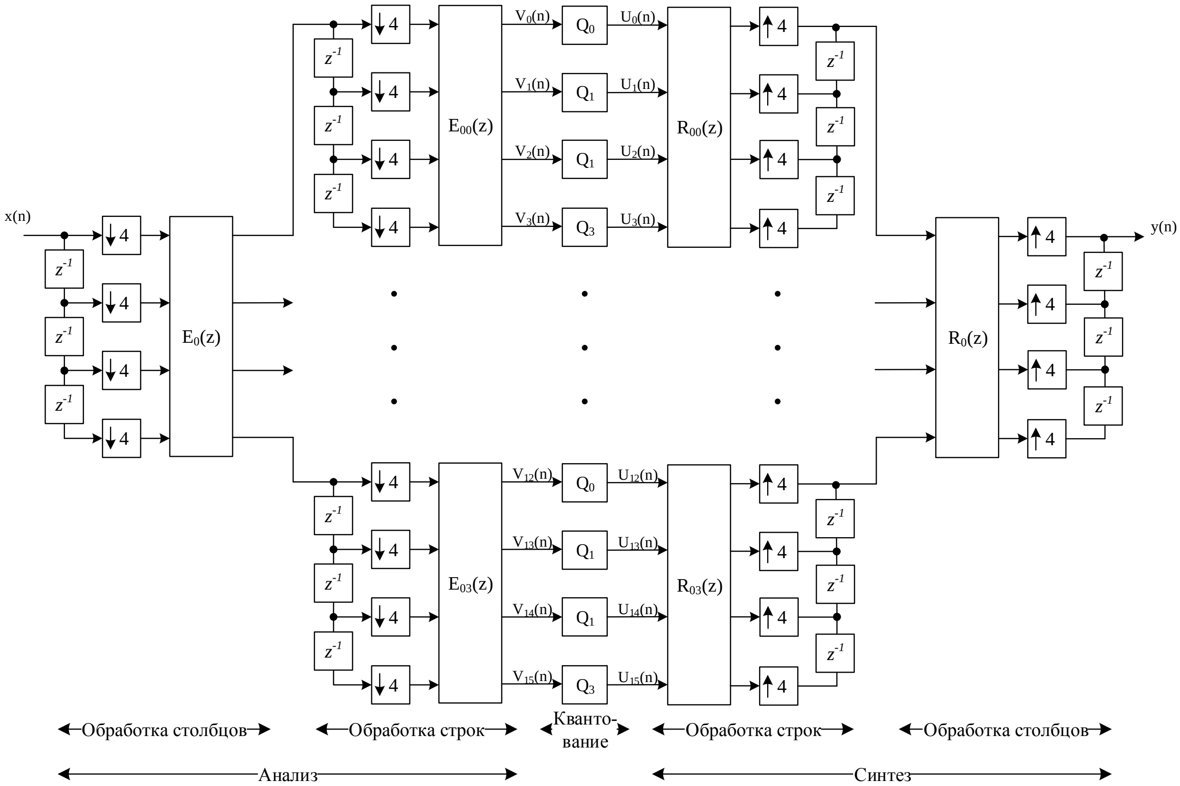
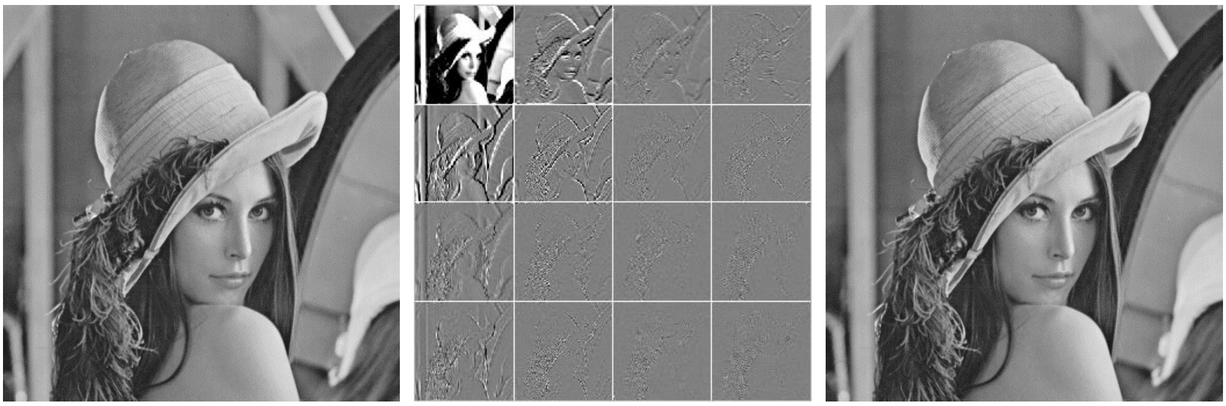


Рис. 2. Схема 2D параунитарного субполосного кодирования



а *б* *в*

Рис. 3. Исходное изображение Lena (512×512) (*а*); декомпозиция изображения полифазной матрицей $\mathbf{E}(z)$ (*б*); реконструированное изображение полифазной матрицей $\mathbf{R}(z)$ (*в*)

Задача оптимального распределения бит по каналам в схеме на рис.2 формируется следующим образом:

найти $b_0, b_1, \ldots, b_{15}$ при условии, что $b=(b_0+b_1+\ldots+b_{15})/16=\mathrm{const}$ и минимизации ошибки реконструкции сигнала $\min\sum_{k=0}^{15}\varepsilon_k 2^{-2b_k}\sigma^2_{V_k}\ /16.$

С другой стороны, задачу оптимизации можно определить как поиск дисперсии сигнала на выходе квантователей при минимизации общей дисперсии ошибки реконструкции сигнала на выходе системы синтеза. Известно, что

$$2^{-2b}=\prod_{k=0}^{15}\left(2^{-2b_k}\right)^{\frac{1}{16}}.$$

После умножений левой и правой части данного выражения на $\prod_{k=0}^{15}\left(2^{-2b_k}\right)^{\frac{1}{16}}$ получается, что $2^{-2b}\prod_{k=0}^{15}\left(\varepsilon_k\sigma_{V_k}^2\right)^{\frac{1}{16}}=\prod_{k=0}^{15}\left(2^{-2b_k}\,\varepsilon_k\sigma_{V_k}^2\right)^{\frac{1}{16}}$.

Здесь левая часть не зависит от числа битов квантователя b_k, т.е. это константа. Пусть $D_k=\varepsilon_k 2^{-2b_k}\sigma_{V_k}^2$ дисперсия сигнала на выходе k-го квантователя, тогда равенство представляется следующим образом: $\mathrm{const}=\prod_{k=0}^{15}D_k^{\frac{1}{16}}$. После возведения левой и правой части в степень 16 получается, что $\mathrm{const}=\prod_{k=0}^{15}D_k$.

Таким образом, задача оптимального распределения бит квантователей в каналах схемы параунитарного субполосного кодирования формулируется следующим образом:
найти $D_0,\ldots,D_{15}$ при условии, что $\mathrm{const}=\prod_{k=0}^{15}D_k^{\frac{1}{16}}$ и минимизации $\min\sum_{k=0}^{15}D_k\;/16$.

Так как $D_k=\varepsilon_k 2^{-2b_k}\sigma_{V_k}^2\geq 0$, то значения D_k являются неотрицательными величинами. Известно, что сумма неотрицательных значений минимальна, когда их произведение будет постоянной величиной при условии, если все они равны: $D_0=D_1=\ldots=D_{15}=D$, где D некая наперед заданная величина. Физическая интерпретация данного результата следующая: дисперсия сигналов $D_k=\varepsilon_k 2^{-2b_k}\sigma_{V_k}$ на выходах квантователей всех каналов равна, или другими словами, дисперсия ошибки реконструкции изображения минимальна, когда дисперсии ошибок, внесенные каждым квантователем, равны.

Величина D определяется в процессе задания степени сжатия изображения. Пусть $D_k=D$, тогда

$$2^{-2b_k}=\frac{D}{\varepsilon_k\sigma_{V_k}^2}.$$

Далее следует, что

$$2^{-2b}=\prod_{k=0}^{15}\left(\frac{D}{\varepsilon_k\sigma_{V_k}^2}\right)^{\frac{1}{16}}=\frac{D^{\overbrace{\frac{1}{16}+\cdots+\frac{1}{16}}^{16}}}{\prod_{k=0}^{15}\left(\varepsilon_k\sigma_{V_k}^2\right)^{\frac{1}{16}}}\quad\text{, тогда}$$

$$D=2^{-2b}\prod_{k=0}^{15}\left(\varepsilon_k\sigma_{V_k}^2\right)^{\frac{1}{16}}.$$

Таким образом, оптимальное распределение бит квантователей по каналам параунитарного субполосного кодера будет следующим:

$$b_k=b+\tfrac{1}{2}\log_2\frac{\varepsilon_k\sigma_{V_k}^2}{\prod_{k=0}^{15}\left(\varepsilon_k\sigma_{V_k}^2\right)^{\frac{1}{16}}}.$$

Как следует из приведенных рассуждений, оптимальное распределение бит для каждого квантователя в каналах 2D параунитарного субполосного кодера будет, когда дисперсии шума квантования канальных сигналов равны.

Заключение

Обосновано оптимальное распределение бит в схеме 2D параунитарного субполосного кодера на основе алгебры кватернионов. Так, если число бит представления субполосного сигнала фиксировано, то дисперсии ошибок квантователей в каналах кодера должны быть равны для оптимального распределения бит в параунитарном субполосном кодере. В приложениях кодирования изображений важнее использовать критерий минимизации ошибки реконструкции, чем минимизировать ее дисперсию. Например, можно добавить различные неотрицательные веса для каждого канала и затем пытаться минимизировать взвешенную сумму дисперсий ошибок квантователей. В задаче оптимизации еще можно использовать факт, что средне арифметическое неотрицательных величин больше или равно их средне геометрического. Легко доказать: чтобы минимизировать сумму взвешенных слагаемых, они должны быть равны. Другими словами, дисперсии ошибок квантования в каналах кодера обратно пропорциональны соответствующим весам.

OPTIMAL BIT ALLOCATION IN THE PARAUNITARY SUBBAND IMAGE CODER BASED ON THE QUATERNION ALGEBRA

N.A. PETROVSKY

Abstract

The problem of bit allocation in the context of the tree structure 2D paraunitary filter bank based on the quaternion algebra, i.e. 4-band packet wavelet transform, is considered. It was shown that the bit allocation is optimal if the quantizers error variances in the encoder channels are equaled.

Список литературы

1. *Vaidyanathan P.P.* Prentice-Hall. NJ. 1993.
2. *Chen Y-J., Amaratunga K.S.* IEEE Trans. CAS - II. 2003. Vol. 50, № 12. P. 963–976.
3. *Jayant N.S., Noll P.* Prentice-Hall. NJ. 1993.
4. *Soman A.R., Vaidyanathan P.P.* IEEE Trans. Signal processing. 1993. Vol. 41, № 5. P. 1824–1835.
5. *Parfieniuk M., Petrovsky A.* Signal Processing. 2010. Vol. 90. P. 1755–1767.
6. *Петровский Н.А., Парфенюк М.* // Докл. БГУИР. 2011. № 1 (55). С. 70–74.