

УДК 621.3

МАТЕМАТИЧЕСКИЕ МОДЕЛИ И БЫСТРЫЕ АЛГОРИТМЫ ВЕКТОРНО – РАЗНОСТНОГО КОДИРОВАНИЯ ЦИФРОВЫХ РЕЧЕВЫХ СИГНАЛОВ

К. Ю. Гагарин,

канд. техн. наук, докторант

Санкт-Петербургский государственный политехнический университет

Представлены математические модели и алгоритмы векторно-разностных схем кодирования речевых сигналов на основе многоканальных линейных систем. Предложены алгоритмы вычисления автокорреляционных матриц через быстрые гиперкомплексные преобразования Фурье.

In this paper we present a mathematical models and algorithms of vector-differential coding of speech signals based on multi-channel linear systems theory. Efficient algorithms of autocorrelation arrays computation through hyper complex Fourier transformations are also proposed.

Разностное кодирование относится к методам сжатия цифровых сигналов через кодирование их формы.

Наиболее широкое практическое применение получили скалярно-разностные алгоритмы кодирования, которые строятся на основе скалярно-разностных уравнений

$$d(n) = x(n) - \hat{x}(n), \quad (1)$$

где $x(n)$ – n -й отсчет кодирования сигнала; $\hat{x}(n)$ – предсказанное значение для $x(n)$; $d(n)$ – погрешность предсказания.

На основе разностного уравнения (1) разработаны алгоритмы дельта-модуляции (двухуровневое квантование) и дифференциальной импульсно-кодовой модуляции (многоуровневое квантование), а также алгоритмы с кратко- и долгосрочным предсказанием [1, 2].

В данной статье предложены математические модели векторно-разностного кодирования речевых сигналов с многоканальным линейным предсказанием, которые могут служить одним из способов теоретического обобщения других известных методов сжатия, например, с кратко- и долгосрочным предсказанием или с векторным квантованием [2].

В работах [3, 4] было предложено использовать для кодирования формы речевых сигналов векторно-разностные (ВР) уравнения

$$\bar{d}(m) = X(m) - \bar{X}(m), \quad (2)$$

где $\bar{d}(m)$, $X(m)$ и $\bar{X}(m)$ – векторы соответственно погрешностей предсказания, последовательностей отсчетов сигнала $\{x(n)\}$ и последовательности

предсказанных значений $\{\hat{x}(n)\}$. Схема кодера ВР-кодирования (рис. 1) отличается от скалярно-разностных схем наличием Ф_В-блока формирования векторов и операции сложения-вычитания и квантования выполняются над векторами, т. е. над каждым элементом вектора отдельно и независимо друг от друга. Такой способ кодирования относится к кодированию формы сигналов, образованных одноименными $x_i(k)$ компонентами последовательности векторов $\{X(k)\}$.

Для случая, когда длина вектора определяется величиной периода основного тона речевого сигнала, на рис. 2 приведены графики сигналов $\{x_i(k)\}$ и $\{x(n)\}$. Приведенные примеры форм сигналов $x_i(k)$ и $x(n)$ указывают на высокую степень корреляции между соседними отсчетами внутри и между векторов.

Поэтому как практический, так и теоретический интерес представляют алгоритмы ВР-кодирования, позволяющие учитывать оба вида корреляции отсчетов сигнала.

В работе [5] были предложены математические модели предсказателей в схемах ВР-кодирования в виде многоканальных линейных систем. Рассмотрим более подробно методы ВР-кодирования с учетом особенностей используемых математических моделей.

ВР-кодирование с независимой межканальной обработкой отсчетов сигнала

При независимой межканальной обработке отсчетов сигнала квантование и кодирование осуществляется для каждого i -го сигнала, по которому пе-

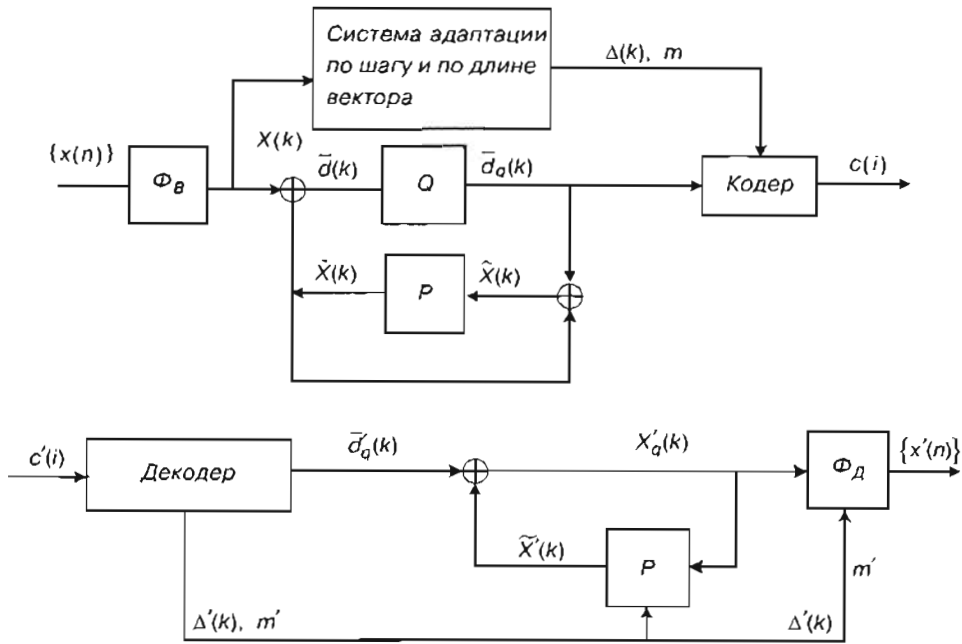


Рис. 1. Схема ВР-кодирования с адаптацией по шагу квантования и по длине вектора

ствляется для каждого i -го сигнала, по которому передается сигнал $x_i(k)$. При этом, согласно векторно-разностному уравнению (2), предсказанные значения $\{\hat{x}_i(k)\}$ формируются независимо друг от друга.

Для определения вектора $\hat{X}_m(k)$ предсказанных значений кодируемого вектора $X(k)$ можно записать векторно-матричное разностное уравнение

$$\hat{X}_m(k) = \sum_{l=1}^p D_m(l) \hat{X}(k-l), \quad (3)$$

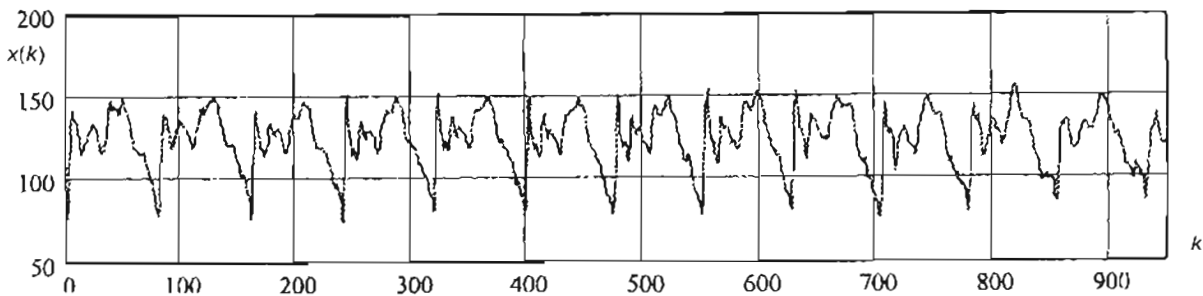
где $D_m(l) = \text{diag}\{\alpha_i(l)\}_{i=0}^{m-1}$ – диагональная матрица коэффициентов предсказания; $\hat{X}(k-l)$ – вектор квантованных значений кодируемого сигнала. Для $p = 1$ выражение может быть записано в виде

$$\hat{X}_m(k) = D_m(1) \hat{X}(k-1),$$

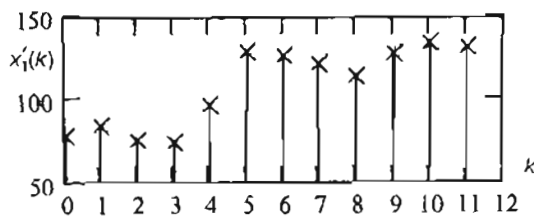
где $D_m(1) = I_m$ – единичная матрица.

Для i -го сигнала $\{x_i(k)\}$ являются справедливыми известные соотношения [6], определяющие каче-

а)



б)



в)

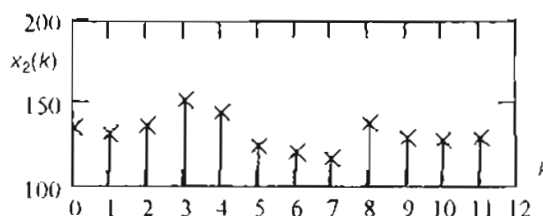


Рис. 2. Графики сигналов: а – $x(n)$, $n = \overline{0,900}$; б – $x_1(n)$, $n = \overline{0,13}$; в – $x_2(k)$, $k = \overline{0,12}$

ство кодирования. Например, отношение сигнал/шум (С/Ш) можно записать через отношение дисперсии погрешности $\sigma_{\hat{d}_i}^2$ предсказания к дисперсии погрешности квантования $\sigma_{e_i}^2$,

$$(C/Ш)_i = \sigma_{\hat{d}_i}^2 / \sigma_{e_i}^2.$$

Для коэффициентов усиления в i -м канале имеем

$$G_{p_i} = \sigma_{\hat{x}_i}^2 / \sigma_{\hat{d}_i}^2.$$

Таким образом, ВР-кодирование с независимой межканальной обработкой может быть осуществлено с использованием в каждом из каналов скалярно-разностных алгоритмов, в том числе алгоритмов с линейным предсказанием.

При этом для определения оптимальных коэффициентов предсказания, представленных матрицей

$$D_m^{(v)}(l) = \text{diag} \left\{ \alpha_v^{(i)}(l) \right\}_{i=1}^m,$$

можно, например, воспользоваться известным [6] соотношением, связывающим коэффициенты предсказания на v -м интервале времени с кратковременной автокорреляционной функцией

$$j = \overline{1, p} \quad R_v^{(i)}(j) = \sum_{l=1}^p \alpha_v^{(i)}(l) R_v^{(i)}(j-l), \quad (4)$$

которое может быть представлено в матричной форме

$$\bar{R}_p = [R_p] \bar{\alpha}_p, \quad (5)$$

где \bar{R}_p – вектор значений $R_v^{(i)}(j)$; $[R_p]$ – автокорреляционная матрица; $\bar{\alpha}_p$ – вектор коэффициентов предсказания $\alpha_v^{(i)}(l)$.

Главной отличительной особенностью ВР-кодирования является обновление на каждом периоде основного тона среднего значения приращения величины отсчетов в каждом канале либо среднего приращения величины отсчетов по всему периоду в целом. Это позволяет значительно расширить возможности ВР-кодирования в части повышения коэффициентов сжатия и помехоустойчивости с сохранением высокого качества восстановления речи.

Экспериментальные исследования ВР-кодирования с дельта-модуляцией показали их значительные преимущества по сравнению с дельта-модуляцией скалярно-разностных алгоритмов.

ВР-кодирование с межканально-зависимым линейным предсказанием

При межканально-зависимом ВР-кодировании речевых сигналов в качестве математической модели предсказателя может выступать векторно-матричное уравнение

$$\tilde{X}_{n_i}(k) = \sum_{l=1}^p A_m(l) \hat{X}(k-l), \quad (6)$$

где $A_m(l) = [\alpha_{ij}(l)]$, $i, j = \overline{1, m}$ – квадратная матрица коэффициентов предсказания.

Разностному уравнению (6) может быть сопоставлена структурная схема предсказателя (рис. 3), где входами являются компоненты $\hat{x}_i(k-l)$ вектора $\hat{X}(k-l)$, а выходами – компоненты $\tilde{x}_i(k)$ вектора предсказанных значений $\tilde{X}(k)$, определяемые в виде скалярных произведений $(\bar{\alpha}_i(l), \hat{X}(k-l))$ векторов $\bar{\alpha}_i(l)$ и $\hat{X}(k-l)$.

При подстановке (6) в уравнение (2) выражение для вектора погрешности предсказания принимает вид

$$\bar{d}(k) = X(k) - \sum_{l=1}^p A_m(l) \hat{X}(k-l) \quad (7)$$

Оптимальные коэффициенты предсказания, заданные в форме матрицы $A_m(l)$, можно определить из условия минимизации дисперсий $\sigma_{\bar{d}}^2$ элементов вектора $\bar{d}(l)$ посредством решения системы уравнений

$$\frac{\partial \sigma_{\bar{d}}^2}{\partial \alpha_{ij}} = 0,$$

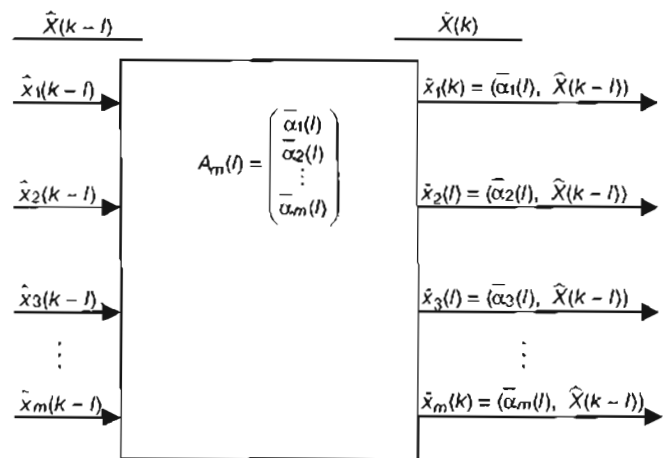
где $\sigma_{\bar{d}}^2 = E(\bar{d}^2(k)) = E((X(k) - \tilde{X}(k))^2)$, $E(\cdot)$ – оператор осреднения.

При этом предполагается, что среднее $\hat{m}_{d_i} = 0$.

Выполнив преобразования для каждого элемента $\tilde{d}_i(k)$, аналогичные скалярно-разностным алгоритмам кодирования, можно получить следующую векторно-матричную форму, связывающую коэффициенты предсказания с блочными корреляционными матрицами

$$R_{xx}(j) = \sum_{l=1}^p A_m(l) R_{xx}(j-l), \quad j = \overline{1, p} \quad (8)$$

От последовательностной векторно-матричной формы (8) можно перейти к блочно-матричной форме, через которую будет задана система линей-



■ Рис. 3. Схема ВР-предсказателя

ных уравнений, решаемых относительно оптимальных коэффициентов предсказания $\alpha_{ij}(l)$:

$$\bar{R}_p = R_p \bar{A}_p, \quad (9)$$

где \bar{R}_p – блочно-матричный вектор-столбец $\bar{R}_p = R'_{xx}(1), R'_{xx}(2), \dots, R'_{xx}(p)$; $\bar{A}(p)$ – блочно-матричный вектор-столбец $\bar{A}' = A'(1), A'(2), \dots, A'(p)$; \hat{R}_p – блочно-теплица корреляционная матрица [7]

$$\hat{R}_p = \begin{pmatrix} R_{xx}(0) & R_{xx}(1) & \dots & R_{xx}(p-1) \\ R_{xx}(1) & R_{xx}(0) & \dots & R_{xx}(p-2) \\ \dots & \dots & \dots & \dots \\ R_{xx}(p-1) & R_{xx}(p-2) & \dots & R_{xx}(0) \end{pmatrix},$$

где $R_{xx}(k) = \frac{1}{N} \sum_{n=0}^{N-k-1} (X(n+k) X'(n))$ – смещенная оценка многоканальной корреляционной матрицы

$$R_{xx} = \begin{pmatrix} r_{11}(k) & r_{12}(k) & \dots & r_{1m}(k) \\ r_{21}(k) & r_{22}(k) & \dots & r_{2m}(k) \\ \dots & \dots & \dots & \dots \\ r_{m1}(k) & r_{m2}(k) & \dots & r_{mn}(k) \end{pmatrix}$$

К блочно-матричной форме (9) можно применить быстрый алгоритм решения заданной через нее системы линейных уравнений. При этом сложность построения быстрых алгоритмов значительно зависит от величины параметра m – количества каналов (или длины вектора). В общем случае матрицы \hat{R}_p не являются теплицевыми, поэтому возможность применения быстрого алгоритма продемонстрируем на разложении Холецкого [6] для значений параметров $m = 2$ и $p = 2$. В этом случае блочно-матричная форма (10) принимает вид

$$\begin{pmatrix} R_{xx}(0) & R_{xx}(1) \\ R_{xx}(1) & R_{xx}(0) \end{pmatrix} \begin{pmatrix} A_2(1) \\ A_2(2) \end{pmatrix} = \begin{pmatrix} R_{xx}(1) \\ R_{xx}(2) \end{pmatrix}. \quad (10)$$

Раскрывая в уравнении (10) блоки-матрицы, получим

$$\begin{pmatrix} r_{11}(0) & r_{12}(0) & r_{11}(1) & r_{12}(1) \\ r_{21}(0) & r_{22}(0) & r_{21}(1) & r_{22}(1) \\ r_{11}(1) & r_{12}(1) & r_{11}(0) & r_{12}(0) \\ r_{21}(1) & r_{22}(1) & r_{21}(0) & r_{22}(0) \end{pmatrix} \begin{pmatrix} \alpha_{11}(1) & \alpha_{12}(1) \\ \alpha_{21}(1) & \alpha_{22}(1) \\ \alpha_{11}(2) & \alpha_{12}(2) \\ \alpha_{21}(2) & \alpha_{22}(2) \end{pmatrix} = \begin{pmatrix} r_{11}(1) & r_{12}(1) \\ r_{21}(1) & r_{22}(1) \\ r_{11}(2) & r_{12}(2) \\ r_{21}(2) & r_{22}(2) \end{pmatrix}. \quad (11)$$

Таким образом, от блочно-матричной формы (10) удалось перейти к двум обычным векторно-матричным формам представления систем линейных уравнений, для решения которых относительно коэффициентов предсказания $\alpha_{ij}(l)$ можно применить метод Холецкого в обычной векторно-матричной форме.

Для быстрого метода решения систем линейных уравнений с использованием разложения Холецкого

го симметричной матрицы количество умножений пропорционально третьей степени параметра предсказания – p^3 . Поэтому для практических применений параметр p нежелательно выбирать более трех или четырех. Для параметра m желательно иметь примерно такие же ограничения. При этом необходимо учитывать, что параметр m связан с величиной периода основного тона вокализованной речи и может рассматриваться как скользящее окно внутри периода основного тона.

Разностное кодирование с двумерным линейным предсказанием

При двумерном предсказании речевого сигнала $x(i, n)$ уравнения разностного кодирования могут быть записаны в следующем виде:

$$d(i, n) = x(i, n) - \bar{x}(i, n),$$

где $i = \overline{1, m}, n = \overline{0, p-1}$.

В этом случае в качестве математической модели предсказателя можно использовать двумерное разностное уравнение

$$\bar{x}(i, n) = \sum_{l=1}^{p_1-1} \sum_{k=1}^{p_2-1} \alpha(l, k) \bar{x}(i-l, n-k). \quad (12)$$

Несложно заметить, что модель (6) отличается от модели (12) лишь формой представления: модель (6) представлена в векторно-матричной форме, а модель (12) – в последовательностной форме.

Аналогичным образом отличаются корреляционные модели.

Для двумерного сигнала можно записать блочно-теплицевую автокорреляционную матрицу в виде

$$\hat{R}_{p,p_2} = \begin{pmatrix} R_{xx}(0) & R_{xx}(1) & \dots & R_{xx}(p_1-1) \\ R_{xx}(1) & R_{xx}(0) & \dots & R_{xx}(p_1-2) \\ \dots & \dots & \dots & \dots \\ R_{xx}(p_1-1) & R_{xx}(p_1-2) & \dots & \dots \end{pmatrix},$$

где

$$R_{xx}(k) = \begin{pmatrix} r_{xx}(k, 0) & r_{xx}(k, 1) & \dots & r_{xx}(k, p_2-1) \\ r_{xx}(k, 1) & r_{xx}(k, 0) & \dots & r_{xx}(k, p_2-2) \\ \dots & \dots & \dots & \dots \\ r_{xx}(k, p_2-1) & r_{xx}(k, p_2-2) & \dots & r_{xx}(k, 0) \end{pmatrix},$$

$r_{xx}(k, m)$ – коэффициенты двумерной корреляции.

Таким образом, математические методы разностного кодирования с двумерным линейным предсказанием являются эквивалентными математическим моделям многоканального векторно-разностного кодирования.

Быстрые алгоритмы вычисления оценок апериодических автокорреляционных функций

Методы скалярно- и векторно-разностного кодирования с линейным предсказанием базируются на

решении системы линейных уравнений, заданной через автокорреляционную матрицу, относительно коэффициентов предсказания. В свою очередь, автокорреляционная матрица строится на основе аперриодической автокорреляционной функции (АКФ), заданной в последовательностной форме:

$$\hat{r}(k) = \frac{1}{N_1} \sum_{n=0}^{N_1-k-1} x(n) x(n+k), \quad (13)$$

где $k = \overline{0, N_1 - 1}$.

Выражение (13) характеризует смещенную оценку последовательности, которая может рассматриваться как стационарный в широком смысле случайный процесс. Согласно работе [7], использование смещенной оценки АКФ гарантирует положительную полуопределенность автокорреляционной матрицы, необходимую для разрешимости системы линейных уравнений. Заметим, что, пользуясь выражением (13), можно получить лишь первую половину значений АКФ. Вторая половина значений АКФ может быть получена за счет симметрии значений АКФ.

Быстрые алгоритмы вычисления оценки АКФ можно получить либо через векторно-матричные, либо через полиномиальные формы представления периодических АКФ, через которые в том числе можно выразить аперриодические АКФ, заданные выражением (13).

Оценку периодической АКФ можно также задать в последовательностной форме

$$\hat{r}(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) x(n+k)_N, \quad (14)$$

где аргумент $n+k$ дискретной функции $x(n+k)$ вычисляется по $\text{mod } N$. Легко показать, что после дополнения в выражении (13) последовательности $\{x(n)\}_{n=0}^{N_1-1}$ нулями с помощью периодической АКФ можно вычислить оценки аперриодической АКФ длины $2N_1 = N$. В векторно-матричной форме оценку периодической АКФ можно записать в виде

$$\hat{r}_N = \frac{1}{N} \hat{S}_N x_N, \quad (15)$$

где \hat{r}_N – вектор значений оценки АКФ; \hat{S}_N – матрица – правый циркулянт с образующим вектором-строкой $\hat{S}_N = x'_N = (x_0, \dots, x_{N-1})$, соответствующим последовательности $\{x(n)\}$. Для выражения (15) известен быстрый алгоритм вычисления АКФ через быстрое преобразование Фурье (БПФ):

$$\hat{r}_N = \frac{1}{N} (F_N^{(\Phi)})^{-1} D_N^* F_N^{(\Phi)} x_N, \quad (16)$$

где $F_N^{(\Phi)}$ – факторизованная форма матрицы преобразования Фурье в поле комплексных чисел, соответствующая тому или иному алгоритму БПФ; D_N^* – диагональная матрица, элементами которой являются компоненты вектора комплексных коэффициентов БПФ от вектора x_N , т. е.

$$D_N^* = \text{diag}\{d_0^*, d_1^*, \dots, d_{N-1}^*\},$$

$$d_N^* = d_0, d_2, \dots, d_{N-1}, d_N = F_N x_N.$$

Если ввести $\bar{P} = |\bar{X}_N|^2 = D_N^* F^{(\Phi)} x_N$ – вектор квадрата модуля коэффициентов БПФ от вектора x_N , то из выражения (16) можно получить известное соотношение для корреляционного метода оценки спектральной плотности мощности сигнала

$$\hat{P}_x(K) = \sum_{n=0}^{N-1} \hat{r}(n) \exp(-i2\pi km / N).$$

Поэтому предлагаемые далее быстрые алгоритмы вычислений оценок АКФ в равной степени могут быть пригодны также для цифрового спектрального анализа.

Алгоритмы вычисления АКФ на основе быстрых гиперкомплексных преобразований Фурье

В работах [7, 8] были предложены алгоритмы быстрых гиперкомплексных преобразований Фурье (ГПФ). Рассмотрим возможности применения данных алгоритмов для быстрого вычисления АКФ.

Матрица гиперкомплексных преобразований Фурье задается в виде

$$F_N^{(1)} = [j_1^{km}],$$

где $k, m = \overline{0, N-1}$, j_1 – мнимая единица, образующая мультипликативную циклическую группу из всех мнимых единиц. Данные ГПФ образованы в расширениях поля рациональных чисел, являющихся также полями.

Матрично-факторизованную форму представления алгоритмов БПФ по основанию два с прореживанием по частоте можно записать

$$F_N^{(\Gamma\Phi)} = \tilde{J}_N \text{diag} \left\{ \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \right\} \text{diag} \left\{ \begin{pmatrix} I_2 & I_2 \\ D_2 & -D_2 \end{pmatrix} \right\} \times \dots \times \begin{pmatrix} I_{N/2} & I_{N/2} \\ D_{N/2} & -D_{N/2} \end{pmatrix},$$

где $D_{N/2} = \text{diag}\{1, j_1, j_2, \dots, j_{N/2-1}\}$, $D_2 = \begin{pmatrix} 1 & 0 \\ 0 & j \end{pmatrix}$, $j = \sqrt{-1}$ – мнимая единица в поле комплексных чисел, j_k – k -я мнимая единица в гиперкомплексной алгебре с $(N/2 - 1)$ мнимыми единицами.

Блок-схема алгоритма вычисления N -точечной периодической АКФ через ГПФ представлена на рис. 4, где через $\hat{R}_{xx}(k)$ обозначена оценка АКФ в области ГПФ. Применяя к ней обратное ГПФ, получаем значения АКФ во временной области $r_{xx}(m)$.

Использование алгоритмов ГБПФ лучше всего продемонстрировать на примере. На рис. 5 и 6 приведены графы прямого и обратного ГБПФ, соот-

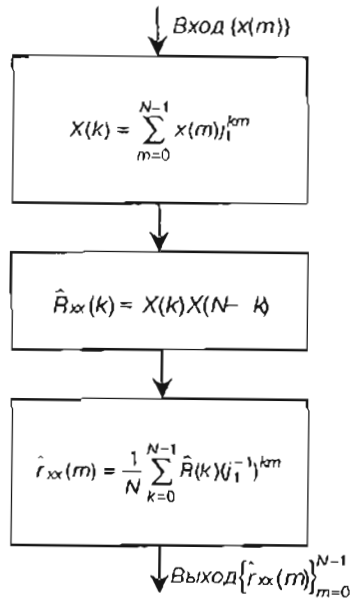


Рис. 4. Блок-схема алгоритма вычисления аperiodической АКФ через ГПФ

ветствующие вычислению аperiodической АКФ через восьмиточечную периодическую АКФ. Значения коэффициентов ГПФ представлены в гиперкомплексной четырехроторной алгебре – алгебре квантернионов $\{1, j, i, k\}$, $i = \sqrt{-1}$, $j = \sqrt{i}$, $k = \sqrt{-i}$.

Значения АКФ в области ГПФ:

$$\begin{aligned}
 R(1) = R(7) &= a + jb - kb, \\
 a &= x_0^2 + x_1^2 + x_2^2 + x_3^2, \\
 b &= x_0x_3 + x_1x_2 + x_0x_1 + x_2x_3 = \\
 &= x_0(x_1 - x_3) + x_2(x_1 + x_3); \\
 R(3) = R(5) &= a - jb + kb, \\
 R(2) = R(6) &= (x_0 - x_2)^2 + (x_1 - x_3)^2; \\
 R(4) &= (x_0 + x_2 - x_1 - x_3)^2; \\
 R(0) &= (x_0 + x_1 + x_2 + x_3)^2.
 \end{aligned}$$

Таким образом, алгоритм вычисления АКФ на основе ГБПФ для $N = 8$; количества арифметических операций в поле рациональных чисел; число умножений $M = 10$. число сложений $A = 18$.

Заметим, что алгоритм вычисления аperiodических АКФ через быстрое преобразование Хартли (БПХ) может быть получен посредством подстановки в выражение (16) матрицы

$$F_N = P_N H_N,$$

где $P_N = \frac{1}{N} F_N H_N$ – переходная матрица от матрицы H_N к матрице F_N .

В результате получим

$$\hat{R}_N = \frac{1}{N} H_N^{(\Phi)} [D_N^{(h)}] H_N^{(\Phi)} X_N = \frac{1}{N} \bar{J}_N H_N^{(\Phi)} D_N^{(h)} H_N^{(\Phi)} X_N,$$

где $[D_N^{(h)}] = P_N^{-1} D_N^{(h)} P_N$;

$$D_N^{(h)} = d_0^{(h)} \oplus 1/2 \left[\begin{matrix} \oplus_{k=1}^{N-1} (d_{N-k}^{(h)} + d_k^{(h)}) \\ \oplus_{k=1}^{N-1} (d_{N-k}^{(h)} - d_k^{(h)}) \end{matrix} \right] \bar{I}_{N-1},$$

$\bar{d}_N^{(h)} = H_N X_N$, $d_k^{(h)}$ – k -й компонент вектора $\bar{d}_N^{(h)}$, $\bar{J}_k = 1 \oplus \bar{I}_{N-1}$, \bar{I}_{N-1} – матрица инверсной перестановки

Выводы факторизованных форм $H_N^{(\Phi)}$ матриц H_N приведены в работе [9].

Аппроксимация нормированных аperiodических АКФ функциями среднего значения разностей

Функции среднего значения разностей (ФСР) было предложено [5] использовать для определения величины периода основного тона вокализированной речи. В отличие от модифицированных АКФ они не имеют операций умножения вещественных чисел. В работе [3] был предложен метод прореженных весовых функций, применение которых к ФСР позволяет снизить количество арифметических операций в несколько раз без заметного снижения точности величины периода основного тона.

Покажем, каким образом можно использовать ФСР для аппроксимации нормированных аperiodических АКФ, через которые строятся автокорреляционные матрицы.

В общем виде для оценки ФСР можно записать выражение

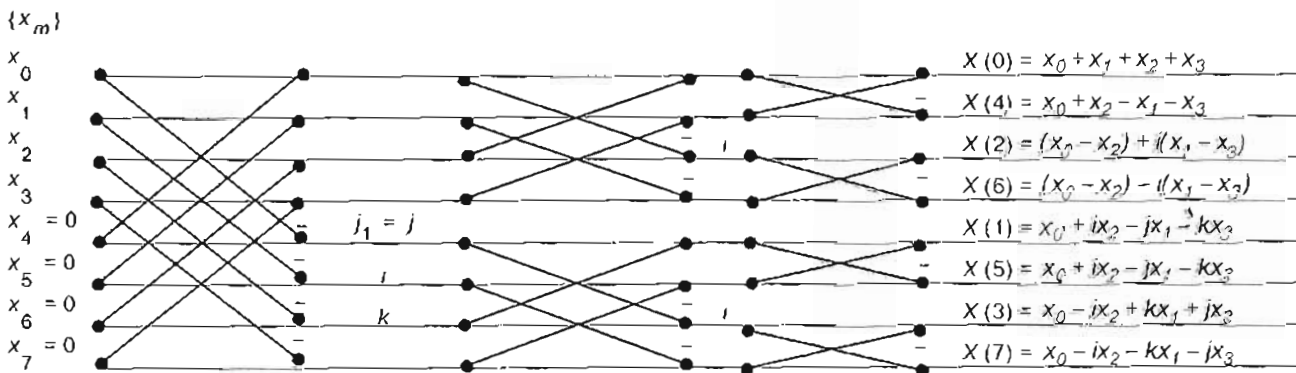


Рис. 5. Граф обратного ГБПФ

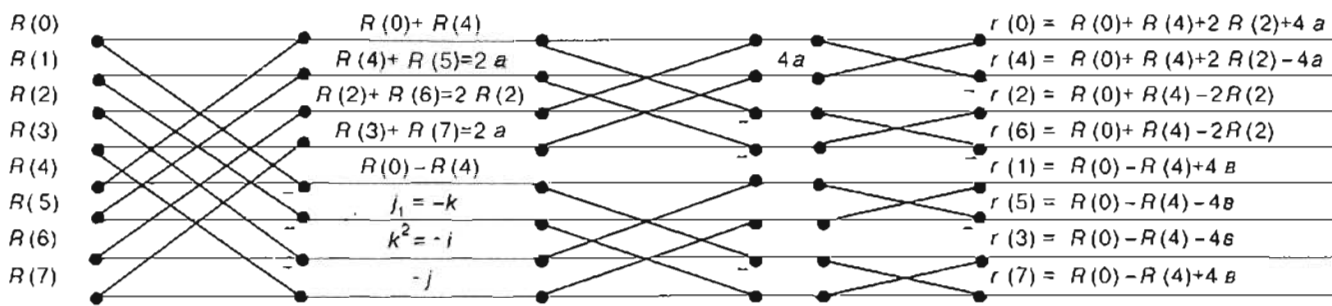


Рис. 6. Граф прямого ГБПФ

$$\hat{\gamma}(k) = \frac{1}{N} \sum_{m=0}^{N-1} |x(m)w_1(m) - x(m \pm k)w_2(m \pm k)|, \quad k = \overline{0, N_i - 1} \quad (17)$$

Применительно к определению величины периода основного тона параметр N_i много больше величины N .

Рассмотрим возможность применения оценки $\hat{\gamma}(k)$ для аппроксимации нормированных смещенных оценок аперiodических АКФ.

Для упрощения выкладок в дальнейшем будем использовать прямоугольные весовые функции с амплитудой, равной единице, так что выражение (17) приобретает вид

$$\hat{\gamma}(k) = \frac{1}{N} \sum_{m=0}^{N-1} |x(m) - x(m+k)|, \quad (18)$$

где $m, k = \overline{0, N-1}, m+k = \overline{(m+k)}_N$.

Функцию (18) по аналогии с АКФ можно рассматривать как периодическую, если параметр $(m+k)$ вычисляется по mod N .

Периодическая функция $\hat{\gamma}(k)$, так же, как и периодическая АКФ, обладает симметрией, т. е. $\hat{\gamma}(k) = \hat{\gamma}(N-k), k = \overline{1, N-1}$

Введем дискретную функцию, образованную через дополнения $\tilde{\gamma}(k) = \gamma_{\max} - \hat{\gamma}(k)$. Полученная функция $\tilde{\gamma}(k)$ будет иметь максимальное значение при $k=0$ (при нулевом сдвиге окна, так же, как АКФ).

Если допустить, что с помощью N -точечной периодической ФСР вычисляется $N/2$ точечная аперiodическая ФСР, то исходная последовательность отсчетов цифрового сигнала $\{x_i\}_{i=0}^{N/2-1}$ должна быть дополнена $N/2$ нулями (так же, как АКФ). Тогда минимальное значение периодической ФСР равно $\tilde{\gamma}_{\min} = \hat{\gamma}(N/2) = 0$. Пронормировав значения $\hat{\gamma}(N/2)$, получим соотношения

$$\tilde{\gamma}_{\min} = \hat{\gamma}(N/2) = \hat{R}(N/2) = 0, \quad \tilde{\gamma}_{\max} = \hat{\gamma}(0) = \hat{R}_{\max} = \hat{R}(0).$$

Поведение функций $\tilde{\gamma}(k)$ и $\hat{R}(k)$ в промежуточных точках между значениями $\tilde{\gamma}_{\max}$ и $\tilde{\gamma}_{\min}$, можно оце-

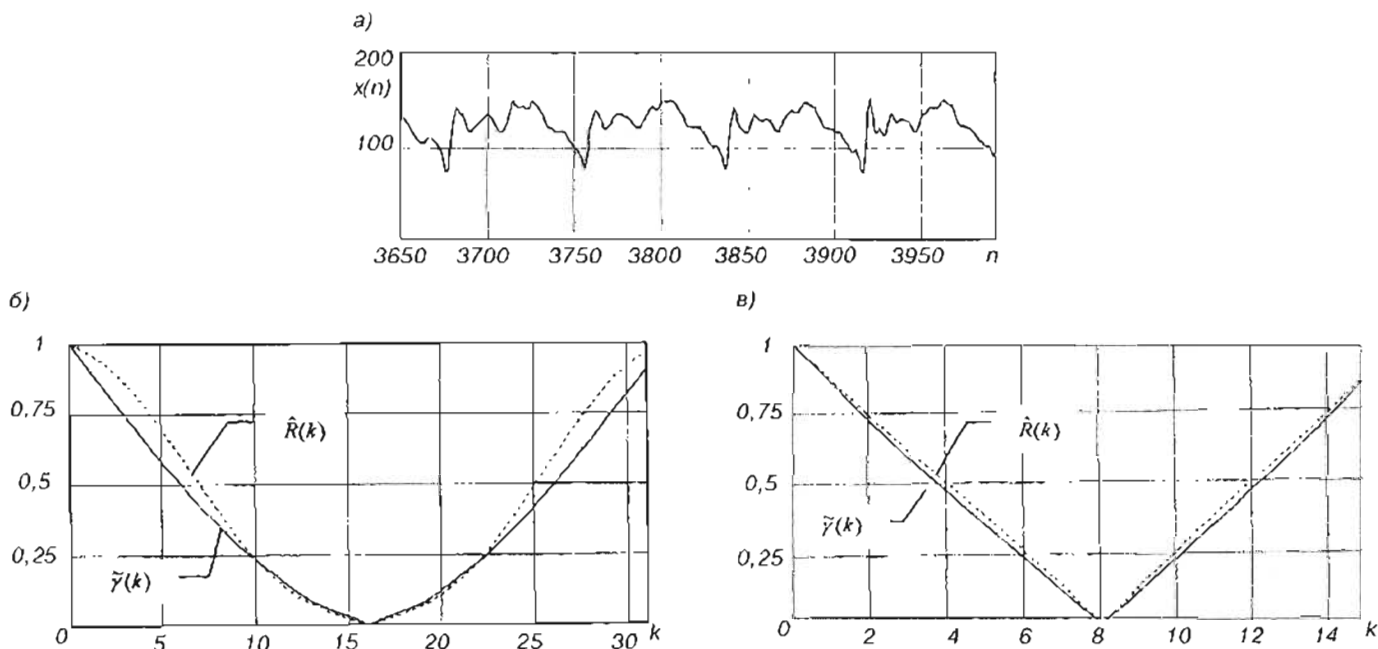


Рис. 7. Графики нормированных АКФ и $\hat{\gamma}(k)$: а - фрагмент локализованного речевого сигнала; б - при $N = 16$; в - при $N = 32$;

■ **Оценки количества арифметических операций быстрых алгоритмов**

N- длина дополненных N/2 нулями последовательностей $\{x_i\}_{i=0}^{N-1}$	Количество нетривиальных умножений			Количество сложений-вычитаний		
	АКФ через БПХ	ФСР	ГБПФ	АКФ через БПХ	ФСР	ГБПФ
4	4	-		12	4	
8	12	-	10	44	16	18
16	48	-	24	128	64	39
32	152	-	68	348	256	110
64	427	-	220	884	1024	385
128	1096	-	780	2146	4096	1441
256	2696	-	1924	5042	8384	4608

нить с помощью функции $\Delta(k)$, отражающей погрешность отклонения

$$\Delta(k) = \frac{\hat{R}(k)}{\hat{R}_{\max}} - \frac{\tilde{\gamma}(k)}{\tilde{\gamma}_{\max}}$$

где

$$\hat{R}_{\max} = R(0) = \frac{1}{N} \sum_{i=0}^{N-1} x_i^2,$$

$$\tilde{\gamma}_{\max} = \tilde{\gamma}(0) = \frac{1}{N} \sum_{i=0}^{N-1} |x_i|.$$

Для практического применения интерес представляет возможность использования дополненных нормированных ФСР для вычисления N/2 оптимальных коэффициентов предсказания из системы линейных уравнений, заданных через теплицеву матрицу, образующей строкой которой может быть использованы N/2 значений ФСР вместо N/2 значений АКФ.

В таблице представлены сравнительные оценки количества арифметических операций для вычисления аперiodических АКФ через БПХ, ГБПФ и через N-точечные периодические ФСР. По таблице можно заключить, что для длин периодических АКФ и ФСР от N = 4 до N = 64 в алгоритмах вычисления АКФ через БПХ общее количество арифметических операций сложений и умножений больше количества сложений-вычитаний, необходимых для вычислений ФСР, которое может быть отражено выражением $A_{\text{ФСР}} = N^2 / 4$.

Литература

1. Брауде-Золотарев Ю. Сжатие речи. – Компьютерра. – 1998. – № 15 (293). – С. 28–31.
2. Spanias A. Speech codig: A Tutorial Review ASU — T.R.C. Technical report. T.R.C. – С.Р. – ASP – 9201. – July 1992. – P. 21–23.
3. Гагарин К. Ю. Быстрые алгоритмы векторно-разностного кодирования речевых сигналов. Сб. научн. тр. СПбГТУ, сер. ВТАРЭ. – № 457. – СПб., 1995. – С. 84–89.
4. Гагарин Ю. И., Раххал М. Математические модели предсказателей в схемах векторно-разностного кодирования речевых сигналов на основе многоканальных линейных систем. Сб. научн. тр. – сер. ВТАРЭ. – № 487. – СПбГПУ, 2002. – С. 80–81.
5. Рабинер Л. В., Шафер Р. В. Цифровая обработка речевых сигналов. – М.: Радио и связь, 1981. – 495 с.

В заключение следует отметить, что опубликованные в данной работе результаты были получены при финансовой поддержке Минобразования РФ НИР-грант ТО2-03.2-2731

На рис. 7 приведены графики функций $\hat{R}(k)$ и $\tilde{\gamma}(k)$, из которых можно заключить, что характер изменения значений $\hat{R}(k)$ и $\tilde{\gamma}(k)$ является одинаковым, хотя в отдельных случаях наблюдается отклонение значений $\hat{R}(k)$ и $\tilde{\gamma}(k)$ до 30 %. [10]

Выводы

1. Полученные математические модели ВР-кодирования речевых сигналов во временной области позволили использовать для нахождения оптимальных коэффициентов многоканального (двумерного) линейного предсказания быстрые алгоритмы решения систем линейных уравнений одномерного предсказания.

2. Для вычисления оценок АКФ предложены быстрые алгоритмы через гиперкомплексные преобразования Фурье и через дополнения функций среднего значения разностей. Приведенные сравнительные данные по количеству арифметических операций позволяют заключить, что для длин АКФ, не превышающих 128 точек, предлагаемые алгоритмы являются наиболее предпочтительными. Кроме того, все арифметические операции в быстрых алгоритмах выполняются в поле рациональных чисел.

6. Марпл С. Л. Цифровой спектральный анализ и его приложения. – М., Мир, 1990. – 584 с.
7. Гагарин Ю. И., Гагарин К. Ю. Гиперкомплексные быстрые преобразования Фурье в расширениях поля рациональных чисел // Труды СПбГТУ, серия ВТАРЭ. – 1998. – № 472. – С. 77–80.
8. Гагарин К. Ю. Быстрые гиперкомплексные преобразования Фурье в расширениях поля рациональных чисел с длиной преобразования, факторизованной взаимно простыми множителями // Труды СПбГТУ, серия ВТАРЭ. – 2000. – № 480. – С. 89–92.
9. Гагарин Ю. И. Математические модели и алгоритмы быстрых ортогональных преобразований. – СПбГТУ, 1999 – 400 с.
10. Раххал М. Методы и алгоритмы векторно-разностного кодирования цифровых аудиосигналов. – СПбГПУ, кандидатская диссертация, 2004. – 126 с.