

# СЕКЦИЯ 1 ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ТЕХНИЧЕСКИХ СИСТЕМАХ

УДК 621.391

С.Л. Бабаринов

babarinov@bsu.edu.ru

ФГАОУ ВПО «Белгородский государственный национальный исследовательский университет»

## ОБ ОДНОМ МЕТОДЕ ИЗМЕНЕНИЯ ТЕМБРАЛЬНЫХ ХАРАКТЕРИСТИК РЕЧИ

Сущность представляемого в докладе метода заключается в использовании оптимальной субполосной коррекции частотных параметров речи, позволяющей целенаправленно воздействовать на распределение энергий по частотным интервалам, таким образом компенсировать искажения вносимые помехами и шумами, а также изменять частотный состав речевых сигналов с целью улучшения тембральных характеристик. Основная задача заключается в переносе энергии из одного частотного интервала в другой.

Речевой сигнал разбивается на отрезки величиной  $N$  для дальнейшей обработки:

$$\vec{x}_N = (x_1, \dots, x_N)^T \quad (1)$$

Частотный диапазон можно представить в виде  $R$  интервалов:

$$R = (N - 2) / 4; \quad (2)$$

При этом их ширина определяется как:

$$\Delta v = v_r - v_{r-1} \quad (3)$$

Для нулевого интервала:  $v_1 = 2\pi / N, v_0 = 0$ .

Для первого и последующих интервалов:  $2\Delta v = 4\pi / N$ .

А центральная частота определяется как:

$$\omega_r = r * 4\pi / N; r = 0 \dots R; \quad (3)$$

Важным условием является сохранение доли энергии сосредоточенной в заданном частотном интервале, для этого определяется доля энергии сигнала, приходящаяся на каждый интервал [1]:

$$\sum_{k=1}^N x_k^2 = \int_{-\pi}^{\pi} |X(\omega)|^2 d\omega / 2\pi, \quad (5)$$

$$P_r = \int_{\omega \in \Omega_r} |X(\omega)|^2 d\omega / 2\pi, \quad (6)$$

где  $X(\omega) = \sum_{k=1}^N x_k \exp(-j(k-1)\omega)$ .

Подставив выражение (5) в (6) получаем представление для сосредоточенной в частотном интервале доли энергии сигнала:

$$P_r = \vec{x}^T A_r \vec{x} \quad (7)$$

где  $A_r = \{a_{ik}^r\}, i, k = 1, \dots, N; r = 1, \dots, R$  представляет собой матрицу с элементами вида [1]:

$$a_{ik}^r = \frac{\sin[v_r \cdot (i-k)] - \sin[v_{r-1} \cdot (i-k)]}{\pi(i-k)} = 2 \frac{\sin(\Delta v \cdot (i-k))}{\pi(i-k)} \cdot \cos(\omega_r \cdot (i-k)) \quad (8)$$

Таким образом, определив  $a_{ik}^0$  можно вычислить значение  $a_{ik}^r$  в любом частотном интервале:

$$a_{ik}^r = 2 \cdot a_{ik}^0 \cos(\omega_r \cdot (i-k)) \quad (9)$$

Применяя фильтрацию, находим долю энергии сосредоточенную в требуемом интервале:

$$\vec{y}_r = A_r \vec{x} \quad (10)$$

Определим синусные и косинусные компоненты огибающей сигнала:

$$C_r = \text{diag}(\cos \omega_0 \dots \cos \omega_r); \quad (11)$$

$$S_r = \text{diag}(\sin \omega_0 \dots \sin \omega_r); \quad (12)$$

Осуществить перенос энергии сигнала на определенное количество интервалов смещением  $\rho$ :

$$C_r = \text{diag}(\cos(\omega_0 + \rho) \dots \cos(\omega_r + \rho)); \quad (13)$$

$$S_r = \text{diag}(\sin(\omega_0 + \rho) \dots \sin(\omega_r + \rho)); \quad (14)$$

где  $\rho = a * (2\pi / N)$ , где  $a = 1 \dots R$ ;

Для каждого частотного интервала:

$$Ca_r = A_0 * C_r \quad (15)$$

$$Sa_r = A_0 * S_r \quad (16)$$

Определим для каждого отрезка матрицы:

$$Zc_r = Ca_r * \vec{x} \quad (17)$$

$$Zs_r = Sa_r * \vec{x} \quad (18)$$

Осуществим умножение матриц (13, 14) на матрицы (18, 19) и получим для каждого отрезка:

$$Xcv_r = Cv_r * Zc_r \quad (20)$$

$$Xsv_r = Sv_r * Zs_r \quad (21)$$

Синтез измененного речевого сигнала:

$$\vec{X} = y_r + Xsv_r + Xcv_r \quad (21)$$

В ходе экспериментальных исследований разработанного метода было произведено смещение спектра исходного речевого сигнала на значение  $2\pi / N$  и  $-2\pi / N$ , что привело к изменению тембральных характеристик речи, рис. 1.

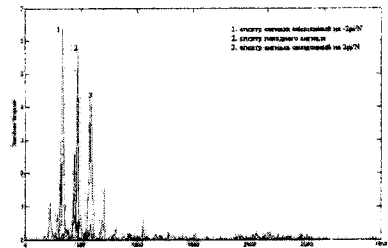


Рис. 1. Результаты экспериментов

### Список литературы

1. Жильяков Е.Г. Вариационные методы анализа и построения функций по эмпирическим данным на основе частотных представлений: моногр. / Е.Г. Жильяков. - Белгород: Изд-во БелГУ, 2007. - 160с

<sup>1</sup> Работа выполнена в рамках проекта № 8.2251.2011