УДК 621.391.828 В161.691.1

K.I. Prokopenko

SOME METHODS OF SIGNALS COMPRESSION AND ANALYSIS OF THEIR EFFECTIVENESS

Institute of Electronics and Control Systems, NAU, e-mail: prokop@irpen.kiev.ua

A series of widely used methods of speech signals compression is considered. The new methods such as the logarithmic delta-modulation, the algorithms of compression with the application of fast Fourier transformation at a binary partition of a signal fragment, the signal compression in the basis of eigenfunctions of a correlation operator are offered. The comparative analysis of algorithms effectiveness by the relative standard error criterion is carried out.

Introduction

The compression of sound signals is widely used in modern engineering.

It is often applied in data transmission systems with limited channel capacity.

Therefore, the estimation criterion to be put for these systems is the compression coefficient at given limits of the compressed signal from the pre-image.

The compression of sound signals allows:

– to reduce the word length of an entering discrete signal representation;

– to reduce a band of the signal transmission channel.

The so-called delta-modulation methods are widely spread. The impact analysis of the model dynamics allows representing a signal as a sequence of increments with a rather small dynamic range and essentially reducing the word length of data representation

In this paper the comparative analysis of me-thods of sound sequence decomposition by the Fourier function, the basis of the discrete cosine transformation and some other bases are carried out. The algorithm of realization of these methods with the use of the recursave approach is also presented. To estimate of the quality, the criterion of the standard error of difference of a restored signal from a pre-image is applied.

Delta-modulation

Let an entering sequence x_i, $I = 0,1....$ be constructed from r-digit values. The principal stage during the coding is to evaluate the difference between the previous and current values of the signal and to reduce of the word length up to s, $s < r$. As a rule the dynamic range of the entering sequence and restored sequences y_i, $i = 0,1....$ are equal. It is the principal advantage of these methods.

The delta-modulation includes the following operations:

1) $y_1 = x_1$;

2) $h = \frac{2^r}{2^s} K_h$;

3) $b_i = \left[\frac{|x_{i+1} - y_i|}{h}\right] \mathrm{sign}(x_{i+1} - y_i)$;

4) if $\log_2|b_i| > s$,

then $b_i = 2^s \mathrm{sign}(x_{i+1} - y_i)$;

5) $y_{i+1} = y_i + hb_i$

for all i.

Here square brackets mean the operation of taking the whole part. $K_h \in (0,1)$ is a transmission factor of a dynamic range. It is easy to notice, that after realization of the fourth pitch of algorithm, the values will belong to a range $-(2^s)..+(2^s)$, hence they will have word length $s+1$.

Thus compression by $\frac{r}{s+1}$ times is achieved.

These transformations allow to compress the sound sequence by 2–3 times without essential loss of quality.

Logarithmic delta-modulation

In case of logarithmic delta-modulation the difference between next and current r-digit elements is taken by logarithm with base **h**. The obtained outcome has word length **s** and it is kept in output sequence b_i:

1) $y_1 = x_1$;

2) $h = \sqrt[2^s]{2^r K_h}$;

3) $b_i = \begin{cases} [\log_h|x_{i+1} - y_i|]\mathrm{sign}(x_{i+1} - y_i), & x_{i+1} \neq y_i; \\ 0, & \text{else}; \end{cases}$

4) if $\log_2|b_i| > s$,

then $b_i = 2^s \mathrm{sign}(x_{i+1} - y_i)$;

5) $y_{i+1} = y_i + h^{b_i}$

for all i.

These transformations allow to compress the sound sequence by 3–5 times without essential loss of quality (fig. 1).

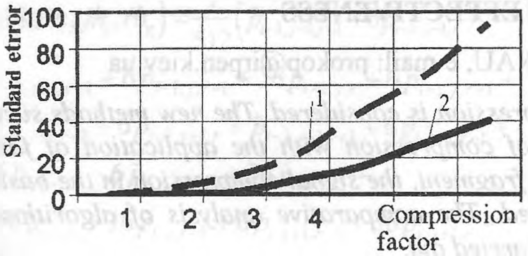

Fig. 1. Dependence of a standard error versus compression factor for usual (*1*) and logarithmic (*2*) delta-modulation

The word length of a pre-image is 16 bit, the coding was carried out by 4-digit values, $Kh = 1$ (fig. 2).

The factor of compression is 4.

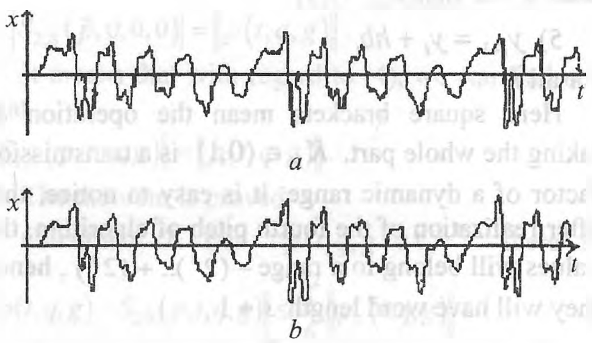

Fig. 2. Result of a sequence processing by the logarithmic delta-modulation method:
a – pre-image; *b* – result of restoration (standard error 10%)

Method of the recursive binary decomposition by orthogonal functions

To reach higher degree of compression the methods of decomposition by the with defined basis functions are used.

It is such methods as Fourier transformation, cosine transformation, transformation with eigenvectors of a correlation matrix of an entering sequence.

The procedure of the recursive binary decomposition of values sequence of a signal $X(t)$ is as follows.

The entering sequence $X(t)$, $t = 0 \ldots N-1$ is divided into fragments of the equal duration. The procedure of binary decomposition for each of these fragments can be conventionally divided into some levels (fig. 3).

At a zero level the set M ($M <= 2^p$) of orthonormalized functions

$$F_0 = [f_{00}, \ldots, f_{0M-1}]$$

is fixed and the standard procedure of decomposition is fulfilled, that is for each function

$$f_{0k}(t) \in F_0$$

the weight factor $h_{0k}, k = 0 \ldots M-1$ of this function in a current fragment is calculated. As the system of functions F_0 is incomplete, the fragment of 2^p elements is divided into the weighted sum M of functions

$$x(t) = \sum_{k=0}^{M-1} h_{0k} f_{0k}(t) + \Delta_0(t),$$

$$t = 0..2^p - 1$$

and the difference $\Delta_0(t)$.

At the first level, the partition of the fragment of the difference $\Delta_0(t)$ on two parts is carried out:

$$\Delta_0^{\,1}(t), t \in (0, 2^{P-1} - 1);$$

$$\Delta_0^{\,2}(t), t \in (2^{P-1}, 2^P - 1).$$

The basis functions F_0 are modified in F_{1k}, $k = 0,1$ in such a manner that

$$f_{1j}(t) = \begin{cases} 2^{\frac{1}{2}} f_{0j}(2t), \ t \in (0, 2^{P-1} - 1); \\ 2^{\frac{1}{2}} f_{0j}(2(t - 2^{P-1})), \ t \in (2^{P-1}, 2^P - 1), \end{cases}$$

$$k = 0,1.$$

Each of the fragments of the zero level difference $\Delta_0^k(t), k = 1,2$ is decomposed by the basis functions F_{1k}, $k = 0,1$ and for each of them the fragments of the difference of the first level $\Delta^k{}_1(t)$, $k = 0,1$ are calculated.

In this paper the application of the criterion of the relative standard error of signal restoration is offered:

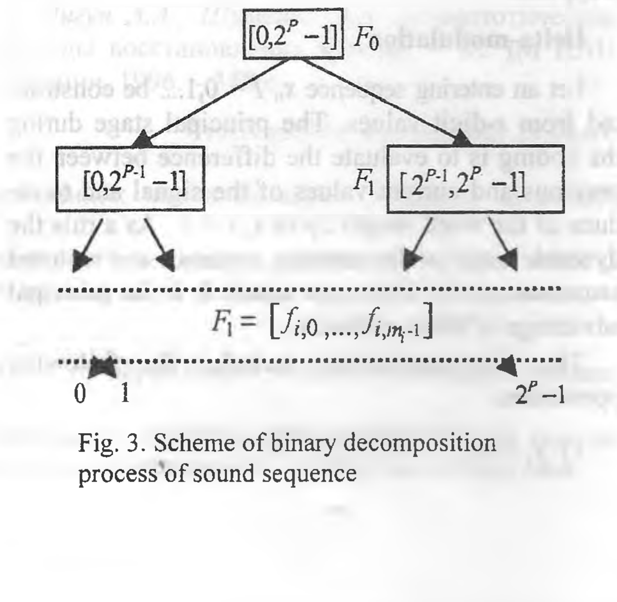

Fig. 3. Scheme of binary decomposition process of sound sequence

$$s = \sqrt{\frac{\sum_{k=0}^{2^p-1}\left(x(k)-x'(k)\right)^2}{\sum_{k=0}^{2^p-1}\left(x(k)\right)^2}}\;100,$$

where $x(k)$ – the original signal; $x'(k)$ – the restored one.

Further, each of these fragments is divided into two fragments according to the principle of detour of a binary tree in the depth and the operations described above, they are fulfilled recursively, till the variance of the difference calculated at a defined level, reaches the small enough predetermined magnitude.

This process visually represented on fig. 3.

Thus, if at a subsequent stage of expansion the restored signal (on a relative standard error) is, say, 10% less than the pre-image one, the process of decomposition tends to stop.

In fig. 4 the stages of decomposition by the basis functions are schematically represented.

In fig. 5 the results of the comparative analysis of decomposition of a sequence of a speech signal by five various bases are shown:

– Daubechies functions of 4th and 10th orders [4];

– basis of Fourier (FFT), basis of cosine transformation (DCT);

– basis constructed by the eigenvectors of the correlation matrix of the entering sequence (BaseFunc) [5].

Windowsize equals 512 samples.

Fourier transformation (FFT) is calculated by following expression:

$$F_0 = [\sin(w_0 t),\ldots,\ \sin((m_0/2-1)w_0 t),$$
$$\cos(w_0 t),\ldots,\ \cos((m_0/2-1)w_0 t)];$$

$$F_1 = [\sin(w_1 t),\ldots,\ \sin((m_1/2-1)w_1 t),$$
$$\cos(w_1 t),\ldots,\ \cos((m_1/2-1)w_1 t)];$$

$$F_i = [\sin(w_i t),\ldots,\ \sin((m_i/2-1)w_i t),$$
$$\cos(w_i t),\ldots,\ \cos((m_i/2-1)w_i t)];$$

$$t = 0,\ldots, N-1,\quad m_0 = N,$$

$$m_i = \frac{N}{2^i},\ w_i = 2^i w_0.$$

Cosine transformation (DCT) is calculated by formula:

$$F_0 = \left[1/\sqrt{m_0},\sqrt{2/m_0}\cos\frac{(2t+1)k\pi}{2m_0}\right], k = 1,\ldots, m_0-1;$$

$$F_i = \left[1/\sqrt{m_i},\sqrt{\frac{2}{m_i}}\cos\frac{2^{i-1}(2t+1)k\pi}{m_i}\right], k = 1,\ldots, m_i-1;$$

$$t = 0,\ldots, N-1,\quad i = 1,\ldots,\log_2 N-1.$$

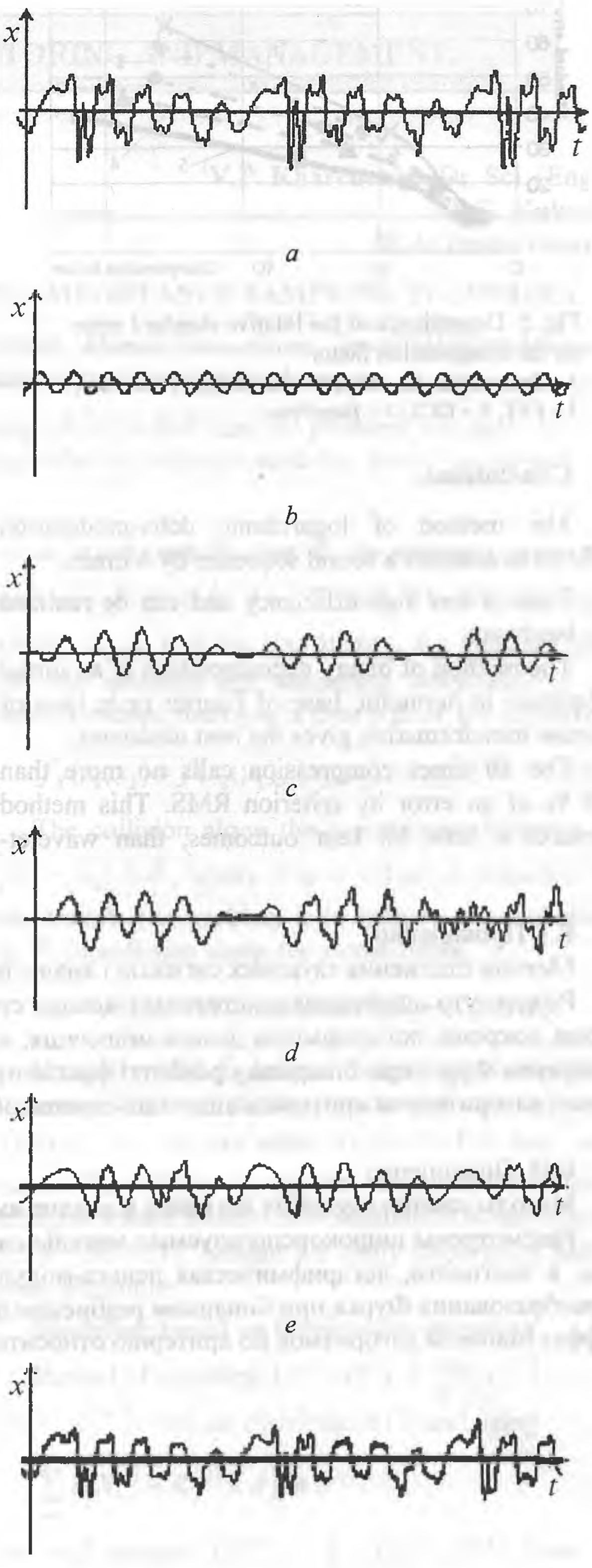

Fig. 4. Stages of speech signal fragment decomposition by the basis functions:
a – original signal; *b* – zero level F_0, standard error 50 %; *c* – first level F_1, standard error 37 %; *d* – second level F_2, standard error 31 %; *e* – third level F_3, standard error 21 %; *f* – fourth level F_4, standard error 11%

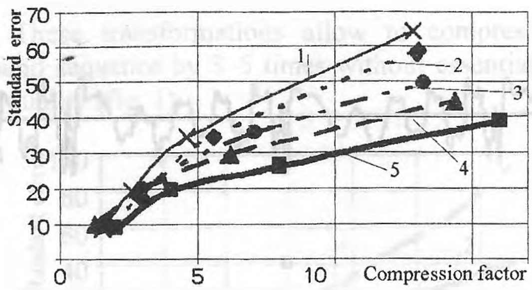

Fig. 5. Dependence of the relative standard error on the compression factor:
1 – Daubechies 4th order, *2* – Daubechies 10th order, *3* – FFT; 4 – DCT; *5* – BaseFunc

Conclutions

The method of logarithmic delta-modulation allows to contract a sound sequence by 4 times.

Thus, it has high efficiency and can be realized by hardware.

The method of binary decomposition of an initial sequence in harmonic base of Fourier or in base of cosine transformation gives the best outcomes.

The 10 times compression calls no more than 40 % of an error by criterion RMS. This method ensures a little bit best outcomes, than wavelet-transformation using Daubechies functions of 4th and 10th orders.

The best outcomes are ensured with the use of transformation with eigenvectors of an entering speech sequence (BaseFunc) correlation matrix. However these algorithms require the large computing expenditures.

References

1. *Ахмед Рао.* Ортогональные преобразования при обработке цифровых сигналов: Пер. с англ. Т.Э. Кренкеля / Под ред. И.Б. Фоменко. – М.: Связь, 1980. – 280 с.

2. *Прокопенко І.Г., Прокопенко К.І.* Програмний комплекс моделювання та обробки сигналів // Вісн. КМУЦА. – 1999. – №2. – С. 121–126.

3. *Digital* cellular telecommunications system (Phase 2) Enchanced Full Rate (EFR) speech transcoding (GSM 06.60 version 4.1.1).

4. *Chui C.K.* An introduction in wavlets // Academic Press. – New York, 1992.

5. *Прокопенко І.Г. Прокопенко К.І.* Представлення мовного сигналу в базисі власних функцій // Матеріали IV Міжнародної наук.-техн. конф. "Авіа-2002", 23-25 квіт. 2002 р. – Т.1. –К.: НАУ, 2002. – С. 11.17–11.20.

Стаття надійшла до редакції 14.10.03.

К.І. Прокопенко

Методи стиснення звукових сигналів і аналіз їх ефективності

Розглянуто широковикористовувані методи стиснення звукових сигналів. Запропоновано нові методи, зокрема, логарифмічна дельта-модуляція, алгоритм стиснення з застосуванням швидкого перетворення Фур'є при бінарному розбитті фрагменту сигналу. Проведено порівняльний аналіз ефективності алгоритмів за критерієм відносної середньоквадратичної похибки.

К.И. Прокопенко

Методы сжатия звуковых сигналов и анализ их эффективности

Рассмотрены широкоиспользуемые методы сжатия звуковых сигналов. Предложены новые методы, в частности, логарифмическая дельта-модуляция и алгоритм сжатия с применением быстрого преобразования Фурье при бинарном разбиении фрагмента сигнала. Проведен сравнительный анализ эффективности алгоритмов по критерию относительной среднеквадратической погрешности.