UDC 621.391 (045)

[1]G. F. Konakhovych,
[2]O. Y. Lavrynenko,
[3]V. V. Antonov,
[4]D. I. Bakhtiiarov

A DIGITAL SPEECH SIGNAL COMPRESSION ALGORITHM BASED ON WAVELET TRANSFORM

Educational-Scientific Institute of Air Navigation, National Aviation University, Kyiv, Ukraine
E-mail: [2]oleksandrlavrynenko@gmail.com, ORCID ID: [2]http://orcid.org/0000-0002-7738-161X

***Abstract*—*It is proposed to use developed digital speech signal compression algorithm based on wavelet transform using entropy arithmetic coding, which allows to achieve optimum results in the improvement of data compression while maintaining the intelligibility of speech signals. Substantiated and experimentally proved the feasibility of using the speech compression algorithm. It is implemented the estimation of compression data ratio and bit rate depending on the parameters of the correlation coefficient, the signal / noise ratio, peak signal / noise ratio and root-mean square error, which is the main criterion of performance of speech quality. The results of experimental studies suggest the feasibility of further practical application of the proposed digital speech signal compression algorithm based on wavelet transform into different models of vocoding devices.*

Index Terms—Digital speech signal; compression algorithm; speech coding; wavelet transform; entropy arithmetic coding; compression ratio; bit rate; correlation coefficient.

I. INTRODUCTION

Currently, there are active development and introduction of new means of communication and telecommunications, in particular, modern digital telephone networks, cellular communications and associated customer equipment, and the development of computer telephony and satellite communications. It is caused by the increased needs of businesses and government organizations in the field of quality and quick transmission of voice information.

One of the main problem of the speech signal compression algorithm is the methods of optimal reduction of redundancy of voice data. The solution of the problem will allow to increase the capacity of linear paths and channels in the conditions of specified criterions of communication quality. The reduction of redundancy of voice data, while maintaining the required quality of speech perception allows to transmit the data at lower speeds, thereby increasing the channel capacity.

II. FORMULATION OF THE PROBLEM

A. Goal of the article

1. To develop a digital speech signal compression algorithm based on wavelet transform (WT) using entropy arithmetic coding, which allows to achieve desired results in the enhancement of data compression while maintaining the intelligibility of speech, which in its turn will transmit speech data at low speeds and decreasing the flow of transmitted information, thereby increasing the bandwidth of existing linear paths and channels.

2. To prove experimentally and substantiate the usefulness of using digital speech signal compression algorithm based on WT.

B. The academic novelty of the article

1. The development of the modified digital speech signal compression algorithm based on WT, which increases the efficiency of encoding speech data when transferring data through the communication channels with low bandwidth.

The objectives of the experimental studies.

1) To develop and research the proposed digital speech signal compression algorithm based on WT in the software package Matlab.

2) To calculate the bit rate (BR), the compression ratio (CR), the correlation coefficient (CC), the signal / noise ratio (SNR), peak signal / noise ratio (PSNR) and root-mean square error (RMSE) of experimental example of the speech up and after the application of the compression algorithm.

3) To carry out a comparative evaluation of the CR, CC, SNR and PSNR depending on BR.

4) To carry out a comparative evaluation of the CR, CC, SNR and PSNR depending on CR.

C. The relevance of the research

The developed digital speech signal compression algorithm based on WT providing optimal solutions of the relevant tasks such as digital speech data compression, the maximum speech signal quality at a certain level of compression, the possibility of the release of some bandwidth of the communication channel for the transmission of digital data, security of transmitted speech data, the possibility to implement the algorithm microprocessor with low productivity to reduce the cost of the developed device.

D. Practical importance of the research

Developed and researched digital speech signal compression algorithm based on WT using entropy arithmetic coding provides a reduction in volume of digital speech data with together with speech intelligibility and thereby increases the bandwidth of the communication channel. The results of experimental studies suggest the feasibility of further practical application of the proposed digital speech signal compression algorithm based on wavelet transform into different models of vocoding devices.

III. RESULT

The article presents the developed digital speech signal compression algorithm based on WT using entropy arithmetic coding (Fig. 1).

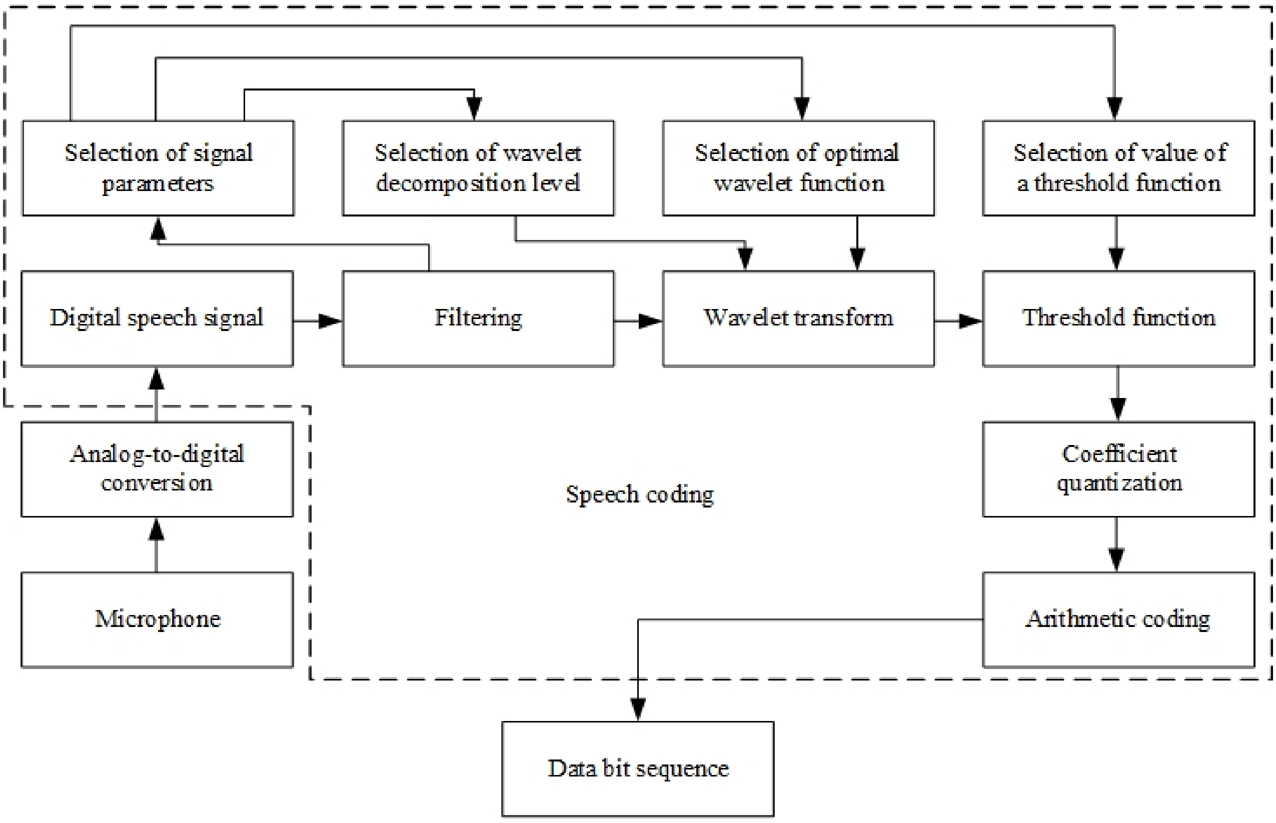

Fig. 1. Digital speech signal compression algorithm based on WT using entropy arithmetic coding

In the research as an input digital speech compression algorithm is used male voice recording with a sampling rate of 8 kHz and the quantization bit depth of 8 bits per sample, which corresponds to the basic digital channels of the telephone network – 64 Kbit / s. The main task of speech signal compression is to reduce the flow of data transmitted over the digital communication channel with a slight deterioration of the restored-term speech at the receiving end [1].

Speech signal compression according to the developed algorithm takes place in several phases. In the first phase for noise reduction and normalization of frequency spectrum of digital speech signal is supplied to 2nd order Butterworth low pass filter, where $n = 5$ with a bandwidth $300 \ldots 3400$ Hz, which is represented as a row vector b and a, having a length $2n+1$ and the polynomial coefficients of numerator and denominator of transmitting function descending powers of z:

$$H(z) = \frac{B(z)}{A(z)} = \frac{b(1) + b(2)z^{-1} + \ldots + b(n+1)z^{-n}}{1 + a(2)z^{-1} + \ldots + a(n+1)z^{-n}}.$$

Butterworth filter cutoff frequency is the frequency at which transmission coefficient module is $\sqrt{1/2}$.

Figure 2 shows diagram of frequency-response characteristic (FRC), phase-frequency response (PFC), impulse response (IR) of synthesized 10th order Butterworth filter with a bandwidth $300 \ldots 3400$ Hz [2].

In the second phase compression algorithm normalized after filtration speech signal after filtering comes on the discrete wavelet transform (DWT) block.

Since the speech signal is a non-stationary random process, then to process it was proposed to use DWT the input of which receives the digital samples of the speech signal, and the output generated wavelet coefficients (WC) [3].

Based on an earlier experimental study as a mother wavelet function is appropriate to use the Daubechies wavelet of order 12.

Calculation of the order N Daubechies scaling filter comes to finding the roots of a polynomial of degree $4N$, whose coefficients

$$a(k)=\frac{\prod\limits_{l=-N+1}^{N}\left(\frac{1}{2}-l\right)}{\prod\limits_{l=-N+1}^{N}(k-l)}, \qquad k=1,\ldots N;$$

for all $k \neq l$ form sets

$$\{a(N)\ \ 0\ \ a(N-1)\ \ 0\ldots 0\ \ a(1)$$

$$1\ \ a(1)\ \ 0\ \ a(2)\ \ 0\ldots 0\ \ a(N)\}.$$

In the third phase WC after DWT come on the thresholding block [4].

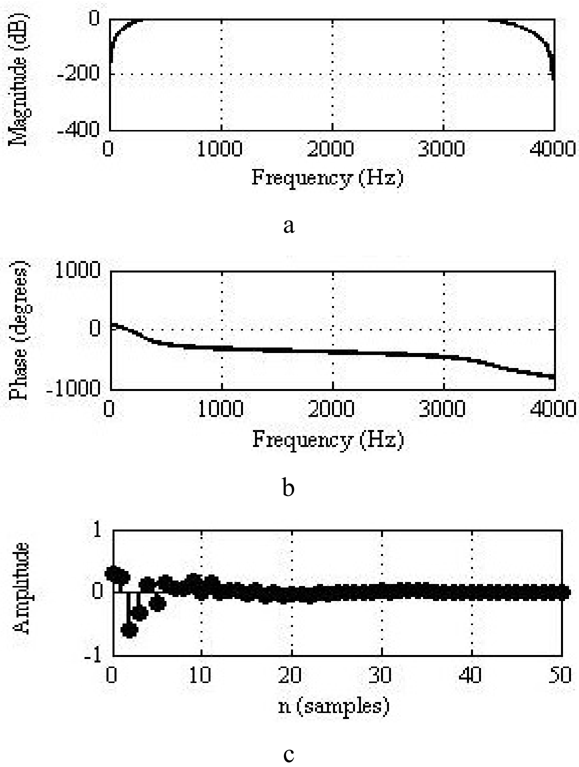

Fig. 2. (a) is the FRC; (b) is the PFC; (c) is the IR of synthesized 10th order Butterworth filter with a bandwidth 300…3400 Hz

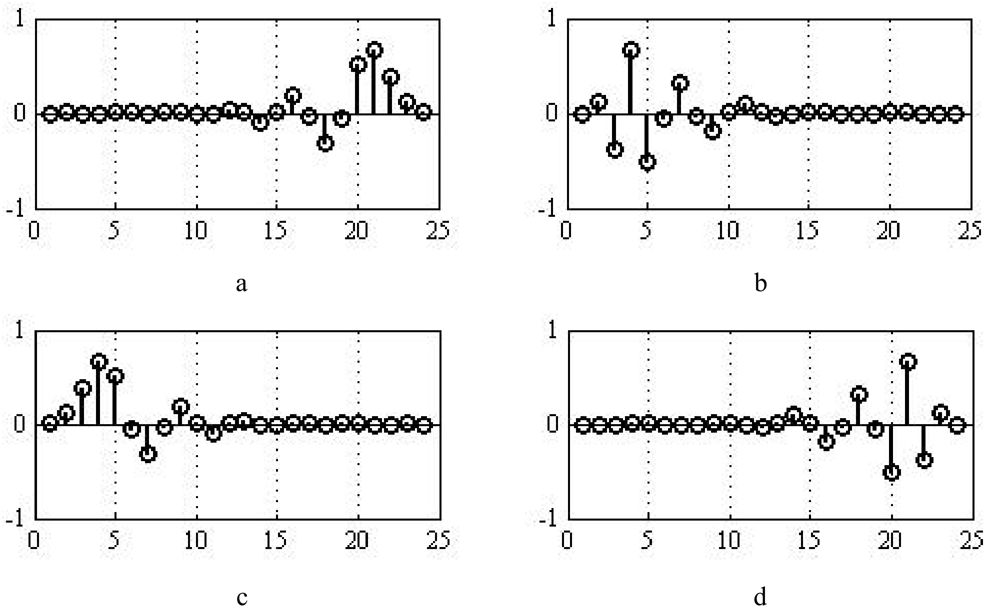

Fig. 3. Orthogonal Daubechies filter of order 12:
(a) is the decomposition low-pass filter; (b) is the decomposition high-pass filter;
(c) is the reconstruction low-pass filter; (d) is the reconstruction high-pass filter

The main property of DWT is that the converted signal is represented by a large number of redundant WC, which, after thresholding are reset WC resets via a given threshold function, WC low or equal to that will be equal to zero, and the rest will remain unchanged.

WC with an absolute value close to zero only contain a small part of the signal energy.

WC resetting results in negligible energy losses.

This property makes DWT attractive for compressing voice data.

The Figure 4 shows a threshold function for processing the WC speech signal, where x – the value of the WC before the threshold, y – the value of WC after threshold, Θ – threshold. Threshold enhancement will increase the degree of redundancy reduction, but at the same time will decrease speech intelligibility. Lowering the threshold to reduce the loss of informational WC, but also reduces the effectiveness of signal compression [5].

$$y = \begin{cases} x, |x| > \Theta \\ 0, |x| \le \Theta \end{cases}$$

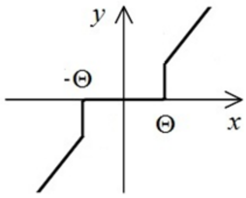

Fig. 4. The threshold function

In the fourth phase of the WC compression algorithm after thresholding is introduced as 8-bit integers. This format is also used for the data transmission. Since detail and approximation of WC are real numbers, then before you performing of speech compression by the arithmetic coding, it is necessary to convert the WC which passed threshold in a numerical range corresponding to the selected format. Otherwise, the WC compression flow will be bigger than speech signal flow. This operation can be performed using the quantization. Thus arises the quantization error, which introduces additional distortion in the transmitted speech signal.

In the fifth phase of the of compression algorithm, the quantized WC encoded by the by arithmetic coding, whereby in the output the compressed bit sequence of speech signal data is generated.

A distinctive feature of the arithmetic coding from well-known coding methods is that neither the encoder or decoder does not store all possible set of code words.

Instead, in the transmission of a particular sequence x the code word $c(x)$ is calculated only for predetermined sequence x. Encoding rules are known to the decoder, and it restores x by $c(x)$, not having a full list of code words.

Consider a discrete binary stationary source without memory with alphabet $X = \{0,1\}$ and with the probability of occurrence units p. The problem consists in the coding sequence $x = \{x1, x2, \ldots, x_n\}$.

The values $p(x)$ and $q(x)$ are calculated using the following recurrence formulas.

In the case when the input character $x_i = 0$:

$$\begin{cases} q(x^i) = q(x^{i-1}), \\ p(x^i) = p(x^{i-1}) \cdot (1-p). \end{cases}$$

In the case when the input character $x_i = 1$:

$$\begin{cases} q(x^i) = q(x^{i-1}) + p(x^{i-1}) \cdot (1-p), \\ p(x^i) = p(x^{i-1}) \cdot p. \end{cases}$$

Let in variable Q is value $q(x^i)$, and in variable P is value $p(x^i)$. Then, the binary arithmetic coding algorithm can be described as follows:

$$Q = 0,\ P = 1.$$

Receive from message source $x = \{x1, x2, \ldots, x_n\}$.
For $i = 1, \ldots, n$ perform the following steps:
If $x_i = 0$, then
$P = P \cdot (1-p)$.
Otherwise
$Q = Q + P \cdot (1-p)$,
$P = P \cdot p$.

To set the code word as $[-\log P + 1]$ discharges after the decimal point in binary notation of number $(Q + P/2)$.

Now consider decoding algorithm of the arithmetic code. Let the decoder knows alphabet $X = \{0,1\}$, probability of unit p, the length of the messages sequence n and the resulting number from encoder $(Q + P/2)$ approximated to $[-\log P + 1]$ discharges, which we denote for F. The task of the decoder is to calculate the sequence $x = \{x1, x2, \ldots, x_n\}$.

The decoding algorithm of arithmetic code is described as follows:

$Q = 0$, $P = 1$.

For $i = 1, \ldots, n$ perform the following steps:
If $Q + (1-p) \cdot P \ge F$, than
$P = P \cdot (1-p)$.
$x_i = 0$.
Otherwise
$Q = Q + P \cdot (1-p)$.
$P = P \cdot p$.
$x_i = 1$.

In this algorithm, the variable P in step with the number i equal to the probability $p(x^i)$ of the sequence, after the first i characters, and the variable Q is the cumulative probability $q(x^i)$.

In the article the digital speech signal compression algorithm based on WT using entropy arithmetic coding was developed, and modeling in software package Matlab, in particular, it was carried out evaluating the CR data and BR in dependence on the parameters of CC, the SNR, and PSNR and RMSE, which is the main criterion of performance of speech quality.

The first experiment (Table I) was conducted according to the criteria as BR at the output of speech compression algorithm, depending on which values were changed CR, CC, SNR, PSNR, RMSE, thereby giving an objective assessment of speech intelligibility and the level of compression speech data during operation of the algorithm at the given BR.

TABLE I

DEPENDENCE RELATION CR, CC, SNR, PSNR, RMSE ON BR

BR (Kbit/s) output	BR (Kbit/s) input	CR	CC	SNR (dB)	PSNR (dB)	RMSE
8	64	8	0.9614	11.2124	34.0639	0.2750
9	64	7.1	0.9701	12.3023	35.1537	0.2426
10	64	6.4	0.9770	13.4187	36.2744	0.2133
11	64	5.8	0.9822	14.5125	37.3640	0.1881
12	64	5.3	0.9861	15.5819	38.4334	0.1663
13	64	4.9	0.9891	16.6320	39.4835	0.1474

The second experiment (Table II) was conducted according to the criteria as CR of speech data at the output of digital speech signals compression algorithm, depending on which values were changed BR, CC, SNR, PSNR, RMSE, thereby giving an objective assessment of speech intelligibility and speech data transmission speed operation of the compression algorithm at the given BR.

TABLE II

DEPENDENCE RELATION BR, CC, SNR, PSNR, RMSE ON CR

CR	BR (Kbit/s) output	BR (Kbit/s) input	CC	SNR (dB)	PSNR (dB)	RMSE
4	16	64	0.9949	19.9081	42.7596	0.1011
5	12.8	64	0.9885	16.4126	39.2641	0.1511
6	10.6	64	0.9806	14.1443	36.9958	0.1962
7	9.1	64	0.9710	12.4302	35.2817	0.2390
8	8	64	0.9613	11.2008	34.0523	0.2754
9	7.1	64	0.9390	9.2728	32.1243	0.3438

The results of the experiment show that the optimal solution on criteria of the ratio of the speech quality on the level of compression will be proposed speech signal compression algorithm with BR = 9 (9.1) Kbit / s, that allows you to save sufficient CR = 7 (7.1), wherein the main index of criteria of speech quality CC = 0.9701 (0.9710), SNR = 12.3023 (12.4302) dB, PSNR = 35.1537 (35.2817) dB and RMSE = 0.2426 (0.2390) give high results in the conditions of given parameters of BR and CR.

IV. CONCLUSIONS

The article presents a developed digital speech signal compression algorithm based on WT using entropy arithmetic coding was developed, and modeling in software package MATLAB.

The usefulness of using digital speech signal compression algorithm based on WT into different models of vocoding devices experimentally proved and substantiated.

The comparative evaluation of the CC, SNR, PSNR and RMSE depending on BR and CR was carried out.

On the basis of the results of the study (Table I, II) you can demonstrably see, at what speed, with what degree of compression algorithm works with the appropriate quality criteria. The optimal solution in terms of intelligibility at a reasonably good compression will be proposed speech signal compression algorithm with BR = 9 (9.1) Kbit / s, that gives the result CR = 7 (7.1), wherein the main index of criteria of speech quality CC = 0.9701 (0.9710), SNR = 12.3023 (12.4302) dB, PSNR = 35.1537 (35.2817) dB and RMSE = 0.2426 (0.2390) give high results in the conditions of given parameters of BR and CR.

REFERENCES

[1] I. Daubechies, *Ten lectures on wavelets*, CBMS-NSF conference series in applied mathematics, SIAM, 1992.

[2] G. F. Konahovich, D. I. Bakhtiyarov, O. Y. Lavrynenko, "Computer modeling of drone protected control channel," *Science-Based Technologies*, vol. 28, no. 4, pp. 283–290, 2015.

[3] S. A. Mallat, "Theory for multiresolution signal decomposition: the wavelet representation," *IEEE Pattern Anal. and Machine Intell.*, vol. 11, no. 7, pp. 674–693, 1989.

[4] M. J. Shensa, "The discrete wavelet transform: wedding the a trous and Mallat Algorithms," *IEEE Trans. on Signal Processing*, vol. 40, no. 10, pp. 2464–2482, 1992.

[5] G. F. Konahovich, O. I. Davletyants, O. Y. Lavrynenko, and D. I. Bakhtiyarov, "The comparative analysis the Fourier transform, cosine transform and wavelet transform as a spectral analysis of the digital speech signals," *Science-Based Technologies*, vol. 27, no. 3, pp. 210–220, 2015.

Received February 09, 2016

Konakhovych Georgy. Doctor of Engineering. Professor.
Institute of Air Navigation, National Aviation University, Kyiv, Ukraine.
Education: National Aviation University, Kyiv, Ukraine (1968).
Research area: information security systems, telecommunication systems.
Publications: more than 100.

Lavrynenko Oleksandr. Graduate Student.
Institute of Air Navigation, National Aviation University, Kyiv, Ukraine.
Education: National Aviation University, Kyiv, Ukraine (2013).
Research area: speech recognition, digital signal processing.
Publications: more than 10.
E-mail: oleksandrlavrynenko@gmail.com, ORCID ID: http://orcid.org/0000-0002-7738-161X

Antonov Veniamin. Ph.D.
Institute of Air Navigation, National Aviation University, Kyiv, Ukraine.
Education: Orel Higher Military Command School of Communications, Orel, Russia (1992).
Research area: information and communication systems.
Publications: more than 30

Bakhtiiarov Denis. Graduate Student.
Institute of Air Navigation, National Aviation University, Kyiv, Ukraine.
Education: National Aviation University, Kyiv, Ukraine (2012).
Research area: information security systems, telecommunication systems.
Publications: more than 15.

Г. Ф. Конахович, О. Ю. Лавриненко, В. В. Антонов, Д. І. Бахтіяров. Алгоритм стиснення цифрових мовних сигналів на основі вейвлет-перетворення

Запропоновано використовувати розроблений алгоритм стиснення цифрових мовних сигналів на основі вейвлет-перетворення із застосуванням ентропійного арифметичного кодування, який дозволяє домогтися оптимального результату в поліпшенні ступеня стиснення даних при збереженні розбірливості мовних сигналів. Обгрунтовано та експериментально доведено доцільність використання представленого алгоритму стиснення мови. Здійснено оцінювання коефіцієнта стиснення даних і швидкості передачі бітів в залежності від параметрів коефіцієнта кореляції, відношення сигнал / шум, пікового відношення сигнал / шум і середньоквадратичної помилки, які виступають головними показниками критерію якості мови. Отримані результати експериментального дослідження дозволяють зробити висновок про доцільність подальшого практичного застосування запропонованого алгоритму стиснення цифрових мовних сигналів на основі вейвлет-перетворення в різноманітних моделях пристроїв кодування мовних сигналів.

Ключові слова: цифрові мовні сигнали; алгоритми стиснення; кодування мовних сигналів; вейвлет-перетворення; ентропійне арифметичне кодування; коефіцієнт стиснення; швидкість передачі бітів; коефіцієнт кореляції.

Конахович Георгій Филимонович. Доктор технічних наук. Професор.
Інститут аеронавігації, Національний Авіаційний Університет, Київ, Україна.
Освіта: Національний Авіаційний Університет, Київ, Україна (1968).
Напрямок наукової діяльності: системи захисту інформації, телекомунікаційні системи.
Кількість публікацій: більше 100.

Лавриненко Олександр Юрійович. Аспірант.
Інститут аеронавігації, Національний Авіаційний Університет, Київ, Україна.
Освіта: Національний Авіаційний Університет, Київ, Україна (2013).
Напрямок наукової діяльності: розпізнавання мови, цифрова обробка сигналів.
Кількість публікацій: більше 10.
E-mail: oleksandrlavrynenko@gmail.com, ORCID ID: http://orcid.org/0000-0002-7738-161X

Антонов Веніамін Валерійович. Кандидат технічних наук.
Інститут аеронавігації, Національний Авіаційний Університет, Київ, Україна.
Освіта: Орловське вище військове командне училище зв’язку, Орел, Росія (1992).
Напрямок наукової діяльності: інформаційно-комунікаційні системи.
Кількість публікацій: більше 30.

Бахтіяров Денис Ілшатович. Аспірант.
Інститут аеронавігації, Національний Авіаційний Університет, Київ, Україна.
Освіта: Національний Авіаційний Університет, Київ, Україна (2012).
Напрямок наукової діяльності: системи захисту інформації, телекомунікаційні системи.
Кількість публікацій: більше 15.

Г. Ф. Конахович, А. Ю. Лавриненко, В. В. Антонов, Д. И. Бахтияров. Алгоритм сжатия цифровых речевых сигналов на основе вейвлет-преобразования

Предложено использовать разработанный алгоритм сжатия цифровых речевых сигналов на основе вейвлет-преобразования с применением энтропийного арифметического кодирования, который позволяет добиться оптимального результата в улучшении степени сжатия данных при сохранении разборчивости речевых сигналов. Обосновано и экспериментально доказано целесообразность использования представленного алгоритма сжатия речи. Осуществлено оценивание коэффициента сжатия данных и скорости передачи битов в зависимости от параметров коэффициента корреляции, отношения сигнал / шум, пикового отношения сигнал / шум и среднеквадратичной ошибки, которые выступают главными показателями критерия качества речи. Полученные результаты экспериментального исследования позволяют сделать вывод о целесообразности дальнейшего практического применения предложенного алгоритма сжатия цифровых речевых сигналов на основе вейвлет-преобразования в различных моделях устройств кодирования речевых сигналов.

Ключевые слова: цифровые речевые сигналы; алгоритмы сжатия; кодирование речевых сигналов; вейвлет-преобразование; энтропийное арифметическое кодирование; коэффициент сжатия; скорость передачи битов; коэффициент корреляции.

Конахович Георгий Филимонович. Доктор технических наук. Профессор.
Институт аэронавигации, Национальный Авиационный Университет, Киев, Украина.
Образование: Национальный Авиационный Университет, Киев, Украина (1968).
Направление научной деятельности: системы защиты информации, телекоммуникационные системы.
Количество публикаций: более 100.

Лавриненко Александр Юрьевич. Аспирант.
Институт аэронавигации, Национальный Авиационный Университет, Киев, Украина.
Образование: Национальный Авиационный Университет, Киев, Украина (2013).
Направление научной деятельности: распознавания речи, цифровая обработка сигналов.
Количество публикаций: более 10.
E-mail: oleksandrlavrynenko@gmail.com, ORCID ID: http://orcid.org/0000-0002-7738-161X

Антонов Вениамин Валерьевич. Кандидат технических наук.
Институт аэронавигации, Национальный Авиационный Университет, Киев, Украина.
Образование: Орловское высшее военное командное училище связи, Орел, Россия (1992).
Направление научной деятельности: информационно-коммуникационные системы.
Количество публикаций: более 30.

Бахтияров Денис Илшатович. Аспирант.
Институт аэронавигации, Национальный Авиационный Университет, Киев, Украина.
Образование: Национальный Авиационный Университет, Киев, Украина (2012).
Направление научной деятельности: системы защиты информации, телекоммуникационные системы.
Количество публикаций: более 15.