ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ
INFORMATION TECHNOLOGIES AND TELECOMMUNICATION

UDC 004.415.24 **DOI: 10.18413/2518-1092-2017-2-2-40-48**

Balabanova T.N.
Likhogodina E.S.
Vodounou A.C.
Guryanova O.I.

THE USING OF ORTHOGONAL BASIS FOR THE STEGANOGRAPHIC CODING OF INFORMATION IN MULTIMEDIA

Belgorod State National Research University, 85 Pobedy St., Belgorod, 308015, Russia

e-mail: Sozonova@bsu.edu.ru, elza9313@gmail.ru, aaron.vodounou@gmail.com, Guryanova_o@bsu.edu.ru

Abstract
This article discusses methods for steganographic encoding additional information using three different orthogonal bases. The bases are represented by the functions, which occupies a different bandwidth in the spectrum. There are comparison of the approach to the definition of DCT-coefficients with the approaches used in the methods of the spread spectrum and subband projections. The approaches of the coefficients of the implementation to ensure secrecy by adaptive determination of their value are considered. However, their value is determined based on the energy structure of the segment of the speech signal. Criteria to evaluate steganographic encoding are the secrecy and accuracy of decoding control information. As the control information is a sequence of numbers in binary form. For the proposed principles of adaptation the results of numerical experiments the estimates that determine stealth: mean square error, the distance Itakura-Saito, correlation. In the computational experiments was found the probability of error for bits at different signal-to-noise ratio. The corresponding computational experiments were carried out for all outlined approaches.
Keywords: steganography; orthogonal basis; adaptive threshold of implementation; discrete cosine transform.

УДК 004.415.24

Балабанова Т.Н.
Лихогодина Е.С.
Водуну А.К.
Гурьянова О.И.

ИСПОЛЬЗОВАНИЕ ОРТОГОНАЛЬНОГО БАЗИСА ДЛЯ СТЕГАНОГРАФИЧЕСКОГО КОДИРОВАНИЯ ИНФОРМАЦИИ В МУЛЬТИМЕДИА

Белгородский государственный национальный исследовательский университет, ул. Победы д.85, г. Белгород, 308015, Россия

e-mail: Sozonova@bsu.edu.ru, elza9313@gmail.ru, aaron.vodounou@gmail.com, Guryanova_o@bsu.edu.ru

Аннотация
В данной статье рассматриваются методы стеганографического кодирования дополнительной информации с использованием трех различных ортогональных базисов. Базисы представлены функциями, занимающими разные по ширине полосы частот в спектре. Приведено сравнение подхода к определению ДКП-коэффициентов с подходами, использующимися в методах расширения спектра и субполосных проекций. Рассмотрены подходы выбора коэффициентов внедрения для обеспечения скрытности путем адаптивного определения их величины. При этом их величина определяется, исходя из

энергетической структуры отрезка речевого сигнала. Критериями, оценивающими стеганографическое кодирование, являются скрытность и достоверность декодирования контрольной информации. В качестве контрольной информации используется последовательность чисел в двоичном виде. Для предложенных принципов адаптации в результате вычислительных экспериментов получены оценки, определяющие скрытность: среднеквадратическая ошибка, расстояние Итакуры-Сайто, корреляция. В ходе вычислительных экспериментов была найдена вероятность ошибки на бит при различном отношении сигнал/шум. Соответствующие вычислительные эксперименты были проведены для всех изложенных подходов.
Ключевые слова: стеганография; ортонормальный базис; адаптивный порог внедрения; дискретное косинусное преобразование.

INTRODUCTION

Speech is the most common and natural method of the transmission of information between the people. For the transfer up to the distance spoken language is fixed, and they convert the result of fixation into the code sequence. In the methods of coding, it is possible to isolate a number of the characteristic operations, one of which is the removal of redundancy for decreasing the volume of the transferred code combinations. With the strong decrease of volume (high compression ratio) are possible the changes with which the reproducible speech will be essentially they will differ from the initial. Often this does not influence the transmission of information. In cases when information is important and it is necessary to ensure its authenticity, but the channel capacity of communications does not make it possible to transmit redundant information, in this case for guaranteeing the authenticity it is possible to use methods of cryptography.

The procedures of the decrease of redundancy, as the methods of cryptography are combined with the use of psychoacoustic models. Naturally, for achievement maximum compression are moved away all frequency-time components, which carry in themselves the redundancy, determined based on psychoacoustic models [1, 2]. This does not make it possible to use excess frequency-time components for coding of additional information. By additional information, we will understand the digital code, which makes it possible to determine the authenticity of speech.

MAIN PART

For the solution of the problem of coding additional information, it is proposed to use the methods, based on the mathematical approach different from that, which was used with the compression. It is worthwhile to note that for guaranteeing the durability of information coding must be accomplished in the space (further the space of coding), and decoding in other space (further the space of decoding).

Ensuring reserve one of requirements imposed to steganographic methods [1-8]. Ensuring reserve is reached when decoding in a component(s) containing the smallest share of energy (fig. 1).

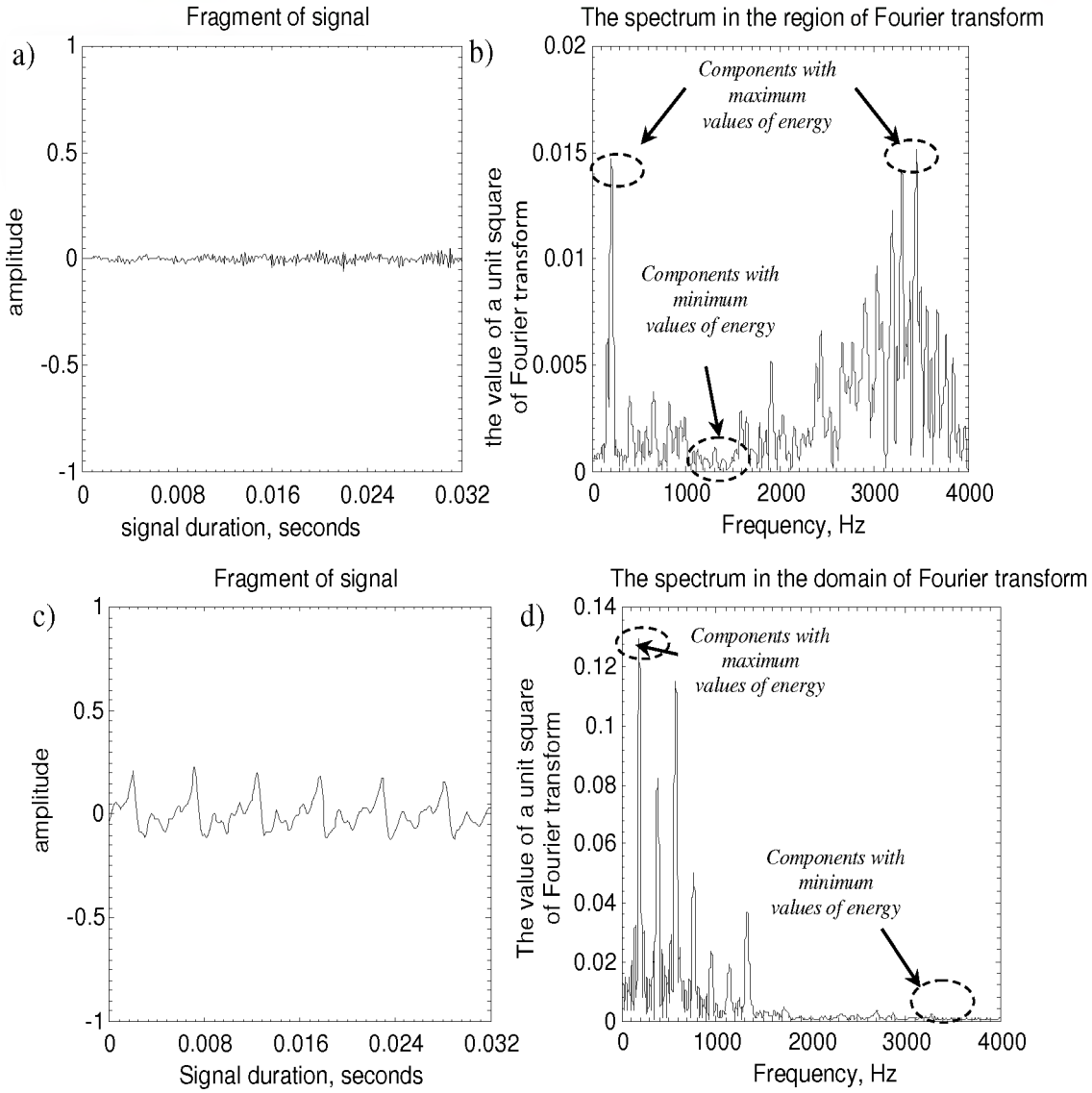

Fig. 1. Audio-signal pieces: a) sound "sh"; b) a range for a sound "sh"; c) a sound "o"; d) a range for a sound "o"

At the same time, not unimportant value plays ensuring probability of an error of decoding of the hidden information, close to zero. Reduction of probability of an error of decoding can be reached thanks to coding of information in a signal component(s) the having overwhelming share of energy of rather synthesizable piece (fig. 1). In this regard, there is a need of the choice between firmness and reserve, for this purpose choose to a component in which coding is carried out. The choice of the fixed threshold or coding in in advance set number components, doesn't provide necessary reserve [9], it is visually illustrated in fig. 1. Apparently from ranges (fig. 1, b and d) sounds "o" and "ш" having different distribution of energy on a frequency axis. The choice, components need to be carried out proceeding from time-and-frequency characteristics of a piece in which reserved coding is carried out, i.e. is adapted to choose to a component for coding [1, 10]. For achievement of high reserve and reduction of probability of a mistake, adaptation under each piece, it is offered to carry out, using the average value having on a component.

Let us consider one of widespread methods of the steganography coding using decomposition of a piece of an audio-signal on DCT coefficients of a look [11, 12]:

$$g_0 = \frac{\sqrt{2}}{N} \sum_{i=1}^{N} x_i \,, \qquad (1)$$

$$g_m = \frac{2}{N}\sum_{i=0}^{N-1} x_i \cdot \cos\left(\frac{(2i+1)\cdot m\pi}{2N}\right),\ m = 1,2,\ldots,(N-1), \quad (2)$$

where x_i – value of signal amplitude; m – number of DCT coefficient; g_m – DCT coefficient.

Results of calculation of DCT coefficients for segments of the audio signals given on fig. 1 are given below.

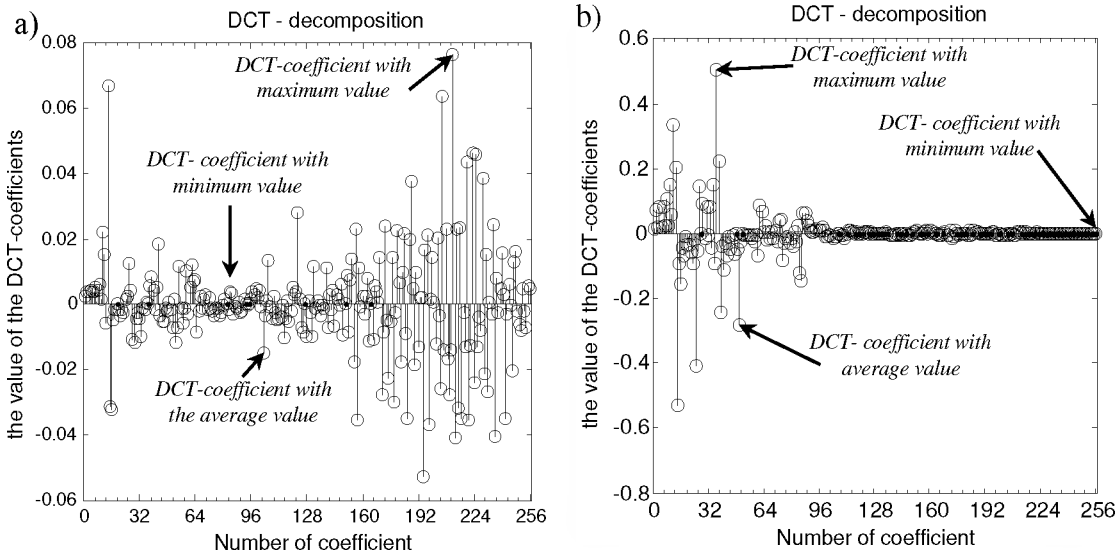

Fig. 2. DCT coefficients: a) sound "sh"; b) sound "o"

An alternative proposed method [13] choice of coefficients is underwritten method.

Among the calculated DCT coefficients (2), it is possible to select the component defined according to one of rules:

– the DCT coefficient having the minimum value:

$$\left(\|\vec{x}\|^2 - g_k^2\right) = \max_{k\in\{1,2,\ldots,N\}} \varepsilon. \quad (3)$$

– the DCT coefficient the close to mean value:

$$\left(\frac{2}{N}\cdot\|\vec{x}\|^2 - g_k^2\right) = \min_{k\in\{1,2,\ldots,N\}} \varepsilon. \quad (4)$$

– the DCT coefficient having the maximum value:

$$\left(\|\vec{x}\|^2 - g_k^2\right) = \min_{k\in\{1,2,\ldots,N\}} \varepsilon. \quad (5)$$

The operations procedure explained below allows realizing steganographic coding of bit in DCT coefficient:

Input data:

- bit of encoded information of a segment $e_m \in \{-1,1\}$.

- segment duration N.

- values of amplitudes of a segment: $\vec{x} = (x_1, x_2, \ldots, x_i, \ldots, x_N)^T$.

Output data:

- Values of amplitudes of a segment: $\vec{y} = (y_1, y_2, \ldots, y_i, \ldots, y_N)^T$.

1. Let's divide an audio signal into segments $\vec{x}$, the size N of reports.

2. According to conversion (1) we will calculate DCT coefficients for a segment $\vec{x}$, $\vec{g} = (g_0, g_1, \ldots, g_k, \ldots, g_m, \ldots, g_{N-1})^T$ i.e. it is feasible direct DCT conversion.

3. Let's calculate energy of a segment

$$\|\vec{x}\|^2 = \sum_{i=1}^{N} x_i^2. \quad (6)$$

4. It agrees to one of rules (3)-(5) we will define number k of DCT coefficients in which we will realize that coding.

5. We realize coding of bit of information e_m, by means of change of a sign of DCT coefficient:

$$c_k = e_m \cdot abs(g_k). \quad (7)$$

where $abs(\ )$ – the operation discarding a sign y at number; s_k – value of DCT coefficient;

6. We realize the reverse IDCT conversion:

$$y_i = \frac{1}{\sqrt{2}} g_0 + \sum_{m=1}^{k-1} g_m \cdot \cos\left(\frac{(2i-1)m\pi}{2N}\right) + c_k \cdot \cos\left(\frac{(2i-1)k\pi}{2N}\right) + \sum_{m=k+1}^{N-1} g_m \cdot \cos\left(\frac{(2i-1)m\pi}{2N}\right), \quad i = 1,2,\ldots,N. \quad (8)$$

Method of expansion of a range

The essence of a method of expansion of a range consists in addition to a piece of an initial speech signal of the pseudorandom sequence (SSp) according to expression [1, 3, 4, 14]:

$$\vec{y} = \vec{x} + \alpha \cdot e \cdot \vec{u}, \quad (9)$$

where $\vec{x}$ – an initial piece of speech data; $\vec{u}$ – the piece corresponding to the pseudorandom sequence; α – weight coefficient; e – the code display of binary bit of the hidden speech message determined by a formula:

The weight coefficient α defines reserve of system. In works [8, 9] him is offered to be chosen equal:

$$\alpha = \langle \vec{x}, \vec{u} \rangle / \|\vec{u}\|^2. \quad (10)$$

Decoding of bit of control information comes from data by definition of a sign of a scalar product of a piece of data and the pseudorandom sequence:

$$\tilde{e} = sign(\langle \vec{y}, \vec{u} \rangle), \quad (11)$$

where $sign(\ )$ – operation of allocation of a sign.

Method of subband projections

Also for assessment, the model of a method of subband projections, which is carrying out reserved coding of bits of control information b_m in a piece of speech data $\vec{x}$ is offered [6, 14]:

$$y = \vec{x} + (sign(e_m) \cdot |\alpha| - \alpha) \cdot \vec{q}, \quad (12)$$

The weight coefficient α defines reserve of system. In works [6, 14] him is offered to be chosen equal:

$$\alpha = \langle \vec{x}, \vec{q} \rangle \quad (13)$$

Decoding of control information is carried out by definition of signs of projections α for own vectors $\vec{q}$ of a subband matrix A_r:

$$\tilde{e}_m = sign(\langle \vec{y}, \vec{q} \rangle), \ m \in M, \quad (14)$$

where $\hat{e}_m$ – the symbol decoded by method of subband projections.

Reserve assessment technique

For determination of overall performance of a method, we use indicators the estimating misstatements brought in an audio-signal when coding by the offered approach. For identification of statistics, the following metrics were counted [1, 2, 7, 8, 15]:

Mean square error, *MSE*:

$$MSE = \sum_{i=1}^{N} (x_i - y_i)^2, \quad (15)$$

where x_i - value of amplitude of the initial audio signal; y_i - value of amplitude of the synthesized audio signal.

Correlation ρ:

$$\rho = \frac{\left(\sum_{i=1}^{N}(x_i - \bar{x}) \cdot \sum_{i=1}^{N}(y_i - \bar{y})\right)}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2 \cdot \sum_{i=1}^{N}(y_i - \bar{y})^2}}, \quad (16)$$

where $\bar{x}$ – a constant component of an initial audio-signal; y – a constant component of the synthesized audio-signal.

Changes in a time domain it is also necessary to consider distinctions in frequency area. The measure based on Itakura-Saito's distance is for this purpose used [15, 16]:

$$ISD = \sum_{r=1}^{R} \Delta\omega_r \cdot \left(\frac{\tilde{P}_r}{P_r} + \ln\frac{P_r}{\tilde{P}_r} - 1\right) / \pi, \quad (17)$$

where $\tilde{P}_r$ – value of energy frequency components of an initial piece of data; P_r – value of energy frequency components of the piece of data containing additional information.

The measure makes sense of distance between ranges of two signals and estimates discrepancy between energy of the changed and initial piece of data. At equality of pieces of data the measure addresses in zero.

$$ISD = \sum_{d=1}^{R} \Delta\omega_d \cdot \left(\frac{\vec{y}^T \cdot A_d \cdot \vec{y}}{\vec{x}^T \cdot A_d \cdot \vec{x}} + \ln\frac{\vec{y}^T \cdot A_d \cdot \vec{y}}{\vec{x}^T \cdot A_d \cdot \vec{x}} - 1\right) / \pi, \quad (18)$$

where A_d – a subband matrix [5]; $\Delta\omega_d$ – width of a frequency interval.

As the tool, allowing to make energy calculations, without passing into the frequency area, it is offered to use a mathematical apparatus of subband matrixes [4, 5]:

$$P_r(\vec{x}) = \vec{x}^T A_r \vec{x}, \quad (19)$$

where A_r – the subband matrix determined by elements:

$$A_r = \{a_{i,k}(r)\}, a_{i,k}^r = \sin\left(\frac{2\pi \cdot (\Delta f/2)}{\vartheta_\partial} \cdot (i-k)\right)\cos\left(\frac{2\pi \cdot f_0}{\vartheta_\partial} \cdot (i-k)\right), \quad (20)$$

where i – an element line item in a line of a matrix; k – an element line item in a matrix column; ϑ_∂ – sampling rate; Δf – band width (in case of normalization respectively $\Delta\omega = 2\pi \cdot \Delta f$); f_0 – central frequency (in case of normalization respectively $\omega_0 = 2\pi \cdot f_0$).

Mean squared error (MSE) measures the relative difference between the energy of segments signals in the time domain. This measure allows identifying the differences in the envelopes of the amplitudes of the segments of speech signals. The fewer changes can be made when introduced additional information, the closer the value for this score to zero [15]:

$$\sigma = \sum_{n=1}^{N}(x_n - y_n)^2 \Big/ \sum_{n=1}^{N} x_n^{\,2}. \quad (21)$$

where x_n – value of amplitude of the initial segment of data; $\tilde{x}_n$ – value of amplitude of the segment of data containing additional information, N – the number of counting of the compared segments of signals.

Reliability assessment technique

Assessment of reliability of decodable information, we will carry out proceeding from probability of an error (BER) [1, 7]:

$$BER = \frac{1}{M}\sum_{m=1}^{M}\left(\left((sign(e_m)+1)/2\right) \oplus \left((sign(\tilde{e}_m)+1)/2\right)\right). \quad (22)$$

where M - the number of encoded bits; $\oplus$ - operation "the amount on the module two"; $sign(\ )$ – operation of separation of a sign; $\tilde{e}_m$ – decodable bit.

Results of simulation

For check of operability of a method based on DCT-conversion, audio signal fragments with sampling rate 8 kHz and digit capacity of 16 bits were used [10]. The general duration of speech material made 23 minutes, lasting 0,032sec, (the segments, which are not containing energy - pauses, were excluded from material). As noise 10^9 not repeating PSP, segments were taken. As a result of simulation it was implemented 10^9 bit, results of simulation are provided to tab. 1.

Table 1

Reliability assessment

№	*BER*	*Noise to signal*			
		0.001	0.01	0.1	1
1	Maximum DCT, (3, 8)	≈ 0	≈ 0	≈ 0	$2.1646*10^{-5}$
2	Average DCT (4, 8)	≈ 0	≈ 0	$2,4071*10^{-5}$	0,0396
3	Minimum DCT, (5, 8)	0,1181	0,1230	0,1247	0,1252
4	SSp, (8, 9, 10)	0,1285	0,1290	0,1439	0,2133
5	SubBand, (8, 10, 12)	0,0219	0,0675	0,1803	0,3345

Results of reserve of the introduced information are given in tab. 2 for the parameters of modelling specified above.

Table 2

Reserve assessment

№	Choice of coefficients principle	*MSE*	ρ	σ	*ISD*
1	Maximum, (3)	2.427 E-0	0.8472	5.41 E-01	3.712
2	Average (4)	2.875 E-3	0.9960	4.28 E-04	1.054
3	Minimum, (5)	8.341 E-8	0.9999	2.31 E-16	0.023
4	SSp	1.102 E-3	0.9923	0.14 E-03	0.031
5	SubBand	3.256 E-3	0.9931	1.21 E-16	0.003

CONCLUSIONS

The given algorithm is optimum from a position of the accounting of frequency properties of the audio-signal containing digital submission of the speech as solving the rule considers uneven distribution of energy on a frequency strip and perception of a sound by the person. Use of DCT coefficient with average value of energy, for reserved coding of information, will allow to reduce by two orders changes of energy in we synthesize an audio-signal piece.

References

1. Fridrich, J. 2012. Steganography in digital media: Principles, algorithms, and applications. Steganography in Digital Media, P. 1-441.
2. Furui, Sadaoki. 2000. Digital speech processing, synthesis, and recognition. 2nd ed., rev. and expanded
3. Cox I. J., Kilian J., Leighton F. T., Shamoon T. Secure spread spectrum watermarking for multimedia // IEEE transactions on image processing. – 1997. – V. 6, № 12. – P. 1673-1687.
4. Nedeljko Cvejic, Tapio Seppanen. 2004. Spread spectrum audio watermarking using frequency hopping and attack characterization. Signal Processing 84. P. 207 – 213.
5. Lykholob, P.G. Research of sensitivity of some measures of quality assessment of hidden information in the audio content [Текст] // Medvedeva, A.A., Likhogodina, E.S., Mishina, O.O. RESEARCH RESULT. Information technologies. №4. v.1. 2016. pp.21-25 URL: http://rr.bsu.edu.ru/media/information/2016/4/3_it.pdf DOI: 10.18413/2518-1092-2016-1-4-21-24
6. Zhilyakov E.G., Pashintsev V.P., Belov S.P., Likholob P.G. About the secretive method of encoding control information in the speech data// Infocommunicatsionnye technologii. – Samara, 2015. – V. 13, № 3. – P. 325-333.
7. Fridrich, J. 2012. Steganography in digital media: Principles, algorithms, and applications. Steganography in Digital Media, P. 1-441.
8. Furui, Sadaoki. 2000. Digital speech processing, synthesis, and recognition. 2nd ed., rev. and expanded
9. GOST 16600-72. The transmission of speech by radio communication paths. The requirements for intelligibility of speech and methods of articulation measurements [Sound recording] / GOST 16600-72; isp.: D.I. Biblev. – Belgorod: NIU

BelGU, 2016. – 1380 sec. – Access mode: https://www.researchgate.net/publication/312167036_Recording_Gost_16600-72 DOI: 10.13140/RG.2.2.33677.74720

10. Kisilenko A.V., Likhogodina E.S., Likholob P.G. About choice of the place for hiding information [Text] / Kisilenko A.V., Likhogodina E.S., Likholob P.G. // Sovremennoe obschestvo, obrazovanie i nauka. Sbornik nauchnyh trudov po matherialam Mezhdunarodnoi nauchno-practicheskoi konferencii: v 9 chastyah. – Tambov: OOO "Konsaltingovaya kompaniya Yukom ", 2014. – P. 76-78

11. Signal processing with lapped transforms. / Malvar H. S. – Boston: Artech House, 1992.

12. Ahmed N., Natarajan T., Rao K. R. Discrete cosine transform // IEEE transactions on Computers. – 1974. – V. 100, № 1. – P. 90-93.

13. On uniqueness of determination of identity-relevant frequency bands in the sounds of Russian speech affected by noise [Текст] / Zhilyakov E.G., Likholob P.G., Kurlov A.V., Medvedeva A.A. // // Nauchnye vedomosti Belgorodskogo gosudarstvennogo universiteta. Seriya: Economica. Informatica. 2016. V. 37. № 2 (223). P. 167-173

14. Evgeny G. Zhilyakov, Sergey P. Belov, Likholob P. G., Pashintsev V. P. On the Steganography in Voice Data // Asian Journal of Information Technology. – 2016. – V. 15, № 12. – P. 1949-1952.

15. Zhilyakov E.G., Likholob P.G., Medvedeva A.A., Prochorenko E.I. Research of the sensitivity of certain quality measures to hide information in the speech data // Nauchnye vedomosti Belgorodskogo gosudarstvennogo universiteta. Seriya: Economica. Informatica. – Belgorod, 2016. – V. 9, № 230. – P. 174-179.

16. Zhilyakov E.G. Optimal subband methods of analysis and synthesis of signals of finite duration. Automation and mechanics. – M.: Akademicheskiy nauchno-izdatelskiy, proizvodstvenno-poligraficheskiy i knigoraspredelitelskiy tsentr Rossiyskoi akademii nauk “Izdatelstvo “Nauka” № 4, 2015г. P. 51-66

Balabanova Tatyana Nikolaevna, Candidate of Technical Sciences, associate Professor, Department of information and telecommunication systems and technologies
Likhogodina Elizaveta Sergeevna, student, Department of information and telecommunication systems and technologies
Vodounou Aaron Candide, student, Department of information and telecommunication systems and technologies
Guryanova Oksana Igorevna, student, Department of mathematical and software information systems

Балабанова Татьяна Николаевна, доцент кафедры информационно-телекоммуникационных систем и технологий, кандидат технических наук
Федеральное государственное автономное образовательное учреждение высшего образования

Лихогодина Елизавета Сергеевна, студент кафедры информационно-телекоммуникационных систем и технологий
Водуну Аарон Кандид, студент кафедры информационно-телекоммуникационных систем и технологий
Гурьянова Оксана Игоревна, студент кафедры информационно-телекоммуникационных систем и технологий