THEORY AND METHODS OF SIGNAL PROCESSING

UDC 621.391(045)
DOI: 10.18372/1990-5548.56.12928

[1]**O. Yu. Lavrynenko,**
[2]**Y. A. Kocherhin,**
[3]**G. F. Konakhovych**

VOICE CONTROL COMMAND RECOGNITION SYSTEM OF UAV BASED ON STEGANOGRAPHIC-CEPSTRAL ANALYSIS

[1,2,3]Telecommunication Systems and Networks Department, Educational & Research Institute of Air Navigation, Electronics and Telecommunications, National Aviation University, Kyiv, Ukraine
E-mails: [1]oleksandrlavrynenko@gmail.com, [2,3]tks@nau.edu.ua

***Abstract*—*The article suggests the developed voice command recognition system based on steganographic-cepstral analysis and the computational algorithm of mel-frequency cepstral coefficients which is used for creation the basic voice command recognition system on the basis of which can be created more complex voice control solutions. To prove experimentally and substantiate the usefulness of using the developed voice command recognition system and the computational algorithm of mel-frequency cepstral coefficients. The detailed results of preliminary experimental studies of the developed voice command recognition system and the computational algorithm of mel-frequency cepstral coefficients in MATLAB in the case of the identification of control commands "up", "down", "right", "left" spoken by different subjects. The article presents the developed voice command recognition system based on steganographic-cepstral analysis which allows to increase the efficiency of voice command recognition. The conclusion based on the detailed results of preliminary experimental studies shows the usefulness of using the developed voice command recognition system and the computational algorithm of mel-frequency cepstral coefficients and thorough substantiation of the scientific-technical experimental study.*

Index Terms—Classification of the object; semantic identification of the voice commands; Unmanned Aerial Vehicles; voice recognition; mel-frequency cepstral coefficients.

I. INTRODUCTION

In this paper, as proposed and described in detail an algorithm for calculation mel-frequency cepstral coefficients (MFCC), which is used to create a basic voice recognition system, on the basis of which can be built more complex solutions of voice control such as the system of voice radio control functions of Unmanned Aerial Vehicles (UAV) for qualitative decision of tasks in the military intelligence purposes, which is a new approach in this field of application [1].

II. PROBLEM STATEMENT

The main goal of this article is to design the voice command recognition new system based on steganographic-cepstral analysis for controlling the UAV with its subsequent optimization.

The following objectives need to be met in order to fulfill this goal: (a) To do research on the methods of detection individual features of speech signals MFCC, used in voice recognition systems. (b) To do research on the voice recognition system and algorithm for calculation of MFCC in the MATLAB as an example of identification of different subjects spoken commands: “Up”, “Down”, “Right”, “Left”. (c) To carry out a comparative assessment of the calculated values from the selected minimum distance criterion, which is the main indicator of the quality of voice recognition test. (d) To do research on the voice command recognition system based on steganographic-cepstral analysis in software package MATLAB.

III. VOICE CONTROL COMMAND RECOGNITION SYSTEM OF UAV BASED ON CEPSTRAL ANALYSIS

The developed voice command recognition system based on steganographic-cepstral analysis has two operation modes: the learning mode and the recognition mode (testing). These modes include a functional scheme of voice command recognition system based on steganographic-cepstral analysis (Fig. 1), a task which is the primary processing of speech recognition and feature selection, which are mainly used in MFCC. If the system is in the learning mode received in step release values recognition features are saved in the reference database of voice images. When the system is in a state of recognition, the values set of MFCC of subject of controll of voice command sequentially compared to all sets of MFCC values from the database of standard voice samples. The task function of the decision to determine the best result of the comparison to one of the specified criteria, and to give recognition result [1].

This research article considers the approach to solving the problem of recognition of voice commands using MFCC splitting (Fig. 2) for the semantic identification of voice commands. The main objective of this research articles is to select in the speech signal features that are relevant for voice command recognition task, that is, information of a semantic component of the voice control of the subject [2].

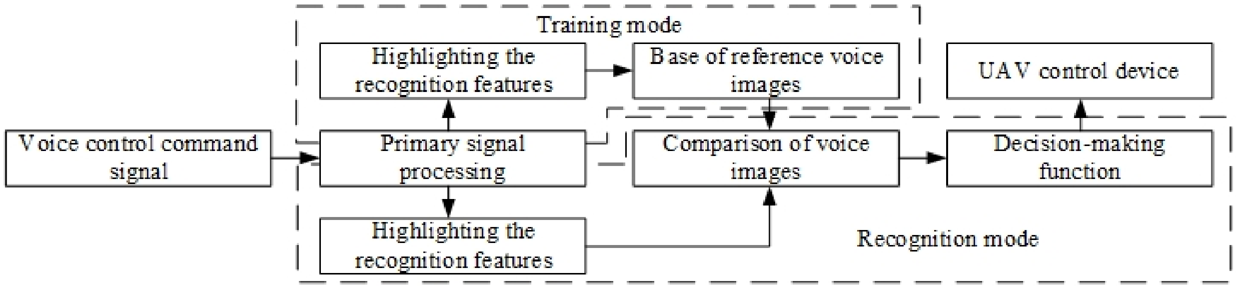

Fig. 1. Voice Control Command Recognition System of UAV based on Steganographic-Cepstral Analysis

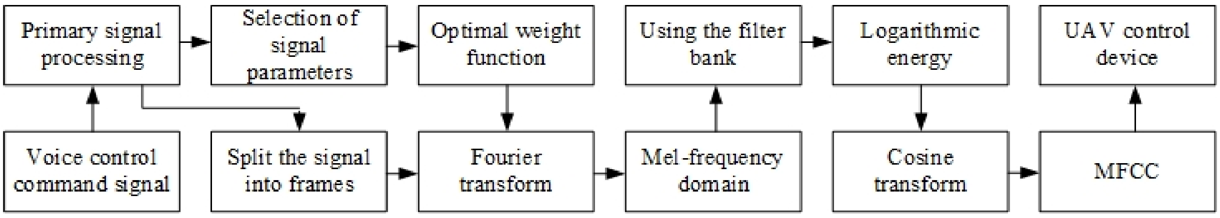

Fig. 2. Algorithm for calculation of MFCC for UAV voice control

Selection of the best parametric features of voice signal is an important task in the development of any UAV voice control system.

Calculation of MFCC includes the following steps. At the input of the algorithm speech signal is supplied frequency band, which is very limited and is in the range from 300 to 4000 Hz. The prevailing approach to speech signal processing procedure is to use short-term analysis. That is, the signal is split into fixed size time window, in which signal parameters do not change. For speech signal the window size is usually chosen in the range of 10-30 ms. For a more accurate signal notation between the windows do overlap, equal to half the length of the window. Then algorithm recognition features selection of (MFCC) is applied to each window. Based on the foregoing, the pre-processing of the speech signal is divided into K frames for N counts, split into $1/2$ of frame length. The input of a discrete Fourier transform (DFT) is fed a sequence readout portion of the speech signal (K frame), investigated in this iteration, $x_0, \dots, x_{N-1}$. Weighting function, and then the DFT is used for a given sequence. In practice, as the weighting function is often used Hamming window which is as follows:

$$w[n] = 0.53836 - 0.46164 \cdot \cos\left(2\pi \frac{n}{N-1}\right),$$
$$n = 0, \dots, N-1$$

where, N is the window length expressed in counting.

Then DFT weighted speech signal can be written as a formula:

$$X[k] = \sum_{n=0}^{N-1} x[n] w[n] e^{\frac{-2\pi j}{N} kn}, \quad k = 0, \dots, N-1.$$

The representation of the speech signal in the frequency domain is divided into bands with the help of the bank (comb) triangular filters. Multiplying the function to filter, we average it at a certain site. Each triangular filter finds a weighted average of the spectral amplitude values corresponding to frequencies in the range between the lower and upper frequency for this filter. If the amplitude corresponds exactly to the average of frequency bands, it is multiplied by a ratio equal to one. When moving the corresponding amplitude value of the frequency from the middle to the lower or upper limit of the coefficient decreases from one to zero.

The result is the weighted average for a given frequency band. Filter's edges are calculated in mel-scale. This scale is the result of research on the human ear's ability to perceive sounds at different frequencies [3].

Transformation into mel-frequency domain is carried out according to the formula:

$$M = 1127.01048 \ln(1 + F/700).$$

The inverse transformation is expressed in hertz by the formula:

$$F = 700\left(e^{M/1127.01048} - 1\right).$$

The formula for dividing the frequency axis into triangular filters will be as follows:

$$f[m] = \left(\frac{N_f}{Fs}\right) M^{-1}\left(M(F_{\min}) + m\frac{M(F_{\max} - F_{\min})}{N_f + 1}\right),$$

where, N_f is the amount of mel-filters (usually use about 24 filters); Fs is the sampling frequency; $M(F)$ is the frequency transfer function in hertz into frequency in mel-scale, discussed earlier; $M(F_{\max} - F_{\min})$ is the analyzed frequency range in mel-scale, which is divided into N_f evenly distributed overlapping ranges and calculated in the corresponding edges in the linear frequency. We form bank triangular filters according to the following equation:

$$H_m[k] = \begin{cases} 0, & k < f[m-1], \\ \dfrac{(k - f[m-1])}{(f[m] - f[m-1])}, & f[m-1] \le k < f[m], \\ \dfrac{(f[m+1] - k)}{(f[m+1] - f[m])}, & f[m] \le k \le f[m+1], \\ 0, & k > f[m+1], \end{cases}$$

where, $H_m[k]$ are filters weighted coefficients [4] and [5].

The use of the filter is pair-wise multiplication of its values with the values of the spectrum. Because filters have N_f, coefficients will be the same. Filters are applied to the square of the modulus of the DFT coefficients, that is, we need to apply mel-filters not to the values of the spectrum, but to his energy. Do so, you need to calculate the energy for each window, and then take the logarithm of results. It is believed that in this way the sensitivity to noise ratios is decreased. The logarithm of the spectrum energy values are taken and represented as follows:

$$E[m] = \ln\left(\sum_{k=0}^{N-1} |X[k]|^2 H_m[k]\right), \; m = 0,\ldots,N_f - 1.$$

The final step in calculating MFCC to reduce the number of output parameters and de-correlation component is discrete cosine transform (DCT), which is given by the equation:

$$c[n] = \sum_{m=0}^{N_f - 1} E[m]\cos\left(\frac{\pi n\left(m + \frac{1}{2}\right)}{N_f}\right), \; n = 0,\ldots,N_f - 1.$$

This transformation has the property of compactness of energy: greater energy corresponds to a smaller amount of information. The resulting set of values is called MFCC. Coefficient $c[0]$ is not used, as is the energy of the speech signal. Thus, we have a very small set of values, which in recognizing successfully replaces thousands of samples of the speech signal. Typically, it retained only the first few elements (8 to 16), which later produced the identification of voice commands [6]. Figure 3 shows calculated by developed algorithm of MFCC experimental samples of voice commands control subject No.1: “Up”, “Down”, “Right”, “Left”.

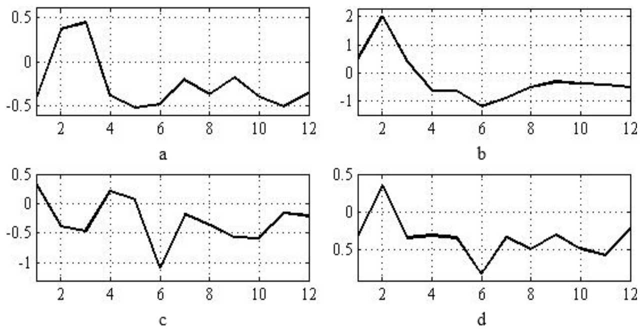

Fig. 3. MFCC voice commands control subject No.1: (*a*) – “Up”; (b) – “Down”; (c) – “Right”; (d) – “Left”

In this system for evaluation of the results of automatic recognition of voice commands a classifier built by the criterion of minimum distance is used. As such indicator figure stands variance of the difference of the expectation value of MFCC saved in base of standard voice samples with expectation value of MFCC at the testing system level [7], [8].

The dispersion of the difference of the expectation values of the two samples of voice commands (MFCC), written as follows:

$$D = \frac{\sum_{i=1}^{n}\left(\frac{\sum_{i=1}^{n} x_i}{n} - \frac{\sum_{i=1}^{n} \bar{x}_i}{n}\right)^2}{n},$$

where, x_i is the saved in base of standard voice samples; $\bar{x}_i$ is at the testing system level; n is the amount of MFCC. The decision of semantic identification of the voice commands accepted by the criterion of minimum dispersion, i.e. the least deviation compared MFCC in a certain threshold of recognition which is given by:

$$\begin{array}{l} \text{if } D_{\min} < \Theta \\ \text{"identified!"} \\ \text{else} \\ \text{"not identified!"} \\ \text{end} \end{array}$$

where, $D_{\min}$ is the minimal dispersion; $\Theta = 1 - \Delta$ is the specified threshold of allowable recognition (in practice usually $\Delta = 0.80 \ldots 0.90$ is used). Minimal dispersion, which became a specified threshold of allowable recognition is the best result of the comparison, and therefore, the command is identified (recognized) – "identified!". Otherwise, the voice command fails semantic identification (not recognized) – "not identified!" [9].

IV. Results

All scientific and experimental studies of voice command recognition system outlined below (Table I, II, and III) have been carried out taking into account the minimum distance criterion, which acts as the variance of the difference of the expectations compared MFCC, depending on which the value of the minimum variance $D_{\min}$ were changed, thereby giving an objective assessment of the quality (reliability) of voice recognition commands in the system test mode. The decision of semantic identification of the voice commands adopted by the criterion of minimal dispersion $D_{\min}$ in the specified threshold of allowable recognition $\Theta = 1 - \Delta = 0.15$, where $\Delta = 0.85$.

TABLE I TEST RESULTS OF VOICE COMMAND RECOGNITION SYSTEM OF THE SUBJECT NO.1

Learning	**Testing**			
Subject Control No.1	***Subject Control No.1***			
Voice commands	"Up"	"Down"	"Right"	"Left"
"Up"	0.0311	0.1921	0.2879	0.1479
"Down"	0.5323	0.0648	0.7345	0.4284
"Right"	0.3255	0.5048	0.0123	0.1699
"Left"	0.1737	0.1935	0.1648	0.0112

In the first experiment (Table I) The comparison of MFCC voice commands control of the subject No. 1 were performed: "Up", "Down", "Right", "Left", which are saved in base of standard voice samples with MFCC of voice commands of the same control of the subject No. 1, but already in the system test mode (MFCC of voice command mode test compared with MFCC of voice commands spoken earlier in the system learning mode). The results (Table I) shows that MFCC voice commands control of the subject No. 1 meet the criterion of minimal dispersion $D_{\min}$ in the specified threshold of allowable recognition $\Theta = 0.15$: "Up" is the $D_{\min} = 0.0311$, "Down" – $D_{\min} = 0.0648$, "Right" – $D_{\min} = 0.0123$, "Left" – $D_{\min} = 0.0112$, on this basis, a decision is made about the positive semantic identification of spoken voice commands (voice commands are recognized). In other cases (Table I) clearly shows that $D_{\min}$ values do not match the selected criteria, and therefore MFCC of spoken voice commands do not coincide with MFCC that are saved in base of standard voice samples, voice commands are not recognized.

TABLE II TEST RESULTS OF VOICE COMMAND RECOGNITION SYSTEM OF THE SUBJECT NO.2

Learning	**Testing**			
Subject Control No. 1	***Subject Control No. 2***			
Voice commands	"Up"	"Down"	"Right"	"Left"
"Up"	0.0451	0.3259	0.4090	0.1604
"Down"	0.3770	0.0482	1.1258	0.5460
"Right"	0.2055	0.4988	0.0967	0.1822
"Left"	0.1268	0.3056	0.2764	0.0703

The results (Table II) shows that MFCC of voice commands control of the subject No. 2 meet the criterion of minimal dispersion $D_{\min}$ in the specified threshold of allowable recognition $\Theta = 0.15$: "Up" – $D_{\min} = 0.0451$, "Down" – $D_{\min} = 0.0482$, "Right" – $D_{\min} = 0.0967$, "Left" – $D_{\min} = 0.0703$, hence a decision about the positive result of the spoken voice command recognition is made. In all other cases, the voice commands are not recognized, since the resulting values do not match the specified recognition criteria.

TABLE III TEST RESULTS OF VOICE COMMAND RECOGNITION SYSTEM OF THE SUBJECT NO.3

Learning	**Testing**			
Subject Control No.1	***Subject Control No.3***			
Voice commands	"Up"	"Down"	"Right"	"Left"
"Up"	0.0602	0.1657	0.4547	0.1943
"Down"	0.4099	0.0912	1.1869	0.3772
"Right"	0.2149	0.4521	0.0846	0.1784
"Left"	0.1722	0.1946	0.2922	0.0785

The comparison results (Table III): "Up" – $D_{\min} = 0.0602$, "Down" – $D_{\min} = 0.0912$, "Right" –

$D_{\min} = 0.0846$, "Left" – $D_{\min} = 0.0785$, fully comply with the criterion of $D_{\min} < \Theta$, where $\Theta = 0.15$, hence a decision about the positive result of the spoken voice command recognition is made. As for the other obtained resulting values, they do not meet the specified recognition criteria, which means that the voice commands are not recognized.

V. CONCLUSIONS

This paper introduces voice command recognition system based on steganographic-cepstral analysis, which allows to increase the efficiency of voice recognition. The paper details the results of preliminary experimental studies on which conclusions on the desirability of further research and practical application of the developed voice command recognition system based on steganographic-cepstral analysis algorithm for calculation of MFCC, as well as justification of scientific importance of research. The comparative assessment of the calculated values from the selected minimum distance criterion, which is the main indicator of the quality of voice recognition test has been carried out.

In the first experiment (Table I) comparison of MFCC voice commands control of the subject №1 were performed: "Up", "Down", "Right", "Left", which are saved in base of standard voice samples with MFCC of voice commands of the same control of the subject No. 1, but already in the system test mode. The results (Table I) shows that MFCC voice commands control of the subject No. 1 meet the criterion of minimal dispersion $D_{\min}$ in the specified threshold of allowable recognition $\Theta = 0.15$: "Up" – $D_{\min} = 0.0311$, "Down" – $D_{\min} = 0.0648$, "Right" – $D_{\min} = 0.0123$, "Left" – $D_{\min} = 0.0112$, on this basis, a decision is made about the positive semantic identification of spoken voice commands (voice commands are recognized).

In the second experiment (Table II) comparison of MFCC of spoken voice command control subject No. 2 in test mode with MFCC of voice commands control subject No. 1 spoken earlier in the learning mode of the system have been carried out. The results (Table II) shows that MFCC of voice commands control of the subject No. 2 meet the criterion of minimal dispersion $D_{\min}$ in the specified threshold of allowable recognition $\Theta = 0.15$: "Up" – $D_{\min} = 0.0451$, "Down" – $D_{\min} = 0.0482$, "Right" – $D_{\min} = 0.0967$, "Left" – $D_{\min} = 0.0703$, hence a decision about the positive result of the spoken voice command recognition is made.

In the third experiment (Table III) comparison of MFCC of spoken voice command control subject No. 3 in test mode with MFCC of voice commands control subject No. 1 spoken earlier in the learning mode of the system have been carried out. The comparison results "Up" – $D_{\min} = 0.0602$, "Down" – $D_{\min} = 0.0912$, "Right" – $D_{\min} = 0.0846$, "Left" – $D_{\min} = 0.0785$, fully meet the criteria of $D_{\min} < \Theta$, where $\Theta = 0.15$, hence a decision about the positive result of the spoken voice command recognition is made.

A software package was developed that includes tools for compiling the base of reference voice images of subjects of management for training and testing the voice command recognition system for control of an UAV based on the steganographic-cepstral analysis in the MATLAB environment.

REFERENCES

[1] O. Lavrynenko, G. Konakhovych, and D. Bakhtiiarov, "Method of Voice Control Functions of the UAV," *Methods and Systems of Navigation and Motion Control (MSNMC), IEEE 4th International Conference*, 2016, pp. 47–50.

[2] D. Bakhtiiarov, G. Konakhovych, and O. Lavrynenko, "Protected System of Radio Control of Unmanned Aerial Vehicle," *Methods and Systems of Navigation and Motion Control (MSNMC), IEEE 4th International Conference*, 2016, pp. 196–199.

[3] G. F. Konakhovych, O. Y. Lavrynenko, V. V. Antonov, and D. I. Bakhtiiarov, "A digital speech signal compression algorithm based on wavelet transform," *Electronics and control systems*, vol. 48, no. 2, pp. 30–36, 2016.

[4] L. F. Sanchez, H. Abaunza, and P. Castillo, "Safe navigation control for a quadcopter using user's arm commands," *Unmanned Aircraft Systems (ICUAS) 2017 International Conference*, 2017, pp. 981–988.

[5] O. Yu. Lavrynenko, G. F. Konakhovych, and D. I. Bakhtiiarov, "Compression algorithm of voice control commands of UAV based on wavelet transform," *Electronics and control systems*, vol. 55, no. 1, pp. 17–22, 2018.

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

[7] G. F. Konahovich, A. I. Davletyants, A. 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.

[8] O. Lavrynenko, G. Konakhovych, R. Odarchenko and D. Bakhtiiarov, "Voice Command Signal Compression Algorithm by Functions of Unmanned

Aerial Vehicles," *Aerospace Technic and Technology*, vol. 130, no. 3, pp. 57–67, 2016.

[9] I. O. Kozlyuk, D. I. Bakhtiiarov, O. Y. Lavrynenko and I. V. Tretiak, "Problems of unauthorized interference to the work of UAV and methods of its solving," *Science-Based Technologies*, vol. 30, no. 2, pp. 206–211, 2016.

Received January 25, 2018

Lavrynenko Oleksandr. Post-graduate student.
Telecommunication Systems and Networks Department, Educational & Research Institute of Air Navigation, Electronics and Telecommunications, National Aviation University, Kyiv, Ukraine.
Education: National Aviation University, Kyiv, Ukraine, (2014).
Research area: speech recognition, digital signal processing, information security.
Publications: more than 15.
E-mail: oleksandrlavrynenko@gmail.com

Kocherhin Yurii. Candidate of Science (Engineering). Associate Professor.
Telecommunication Systems and Networks Department, Educational & Research Institute of Air Navigation, Electronics and Telecommunications, National Aviation University, Kyiv, Ukraine.
Education: National Aviation University, Kyiv, Ukraine, (1989).
Research area: information security systems, telecommunication systems.
Publications: more than 30.
E-mail: tks@nau.edu.ua

Konakhovych Georgy. Doctor of Engineering Science. Professor.
Telecommunication Systems and Networks Department, Educational & Research Institute of Air Navigation, Electronics and Telecommunications, National Aviation University, Kyiv, Ukraine.
Education: National Aviation University, Kyiv, Ukraine, (1968).
Research area: information security systems, telecommunication systems.
Publications: more than 100.
E-mail: tks@nau.edu.ua

О. Ю. Лавриненко, Ю. А. Кочергін, Г. Ф. Конахович. Система розпізнавання голосових команд керування БПЛА на основі стеганографічно-кепстрального аналізу

У даній статті, запропонована розроблена система розпізнавання голосових команд управління на основі стеганографічно-кепстрального аналізу, а також, алгоритм обчислення мел-частотних кепстральних коефіцієнтів, який використовується для створення базової системи розпізнавання голосових команд, на основі якої можуть бути побудовані більш складні рішення голосового управління. Експериментально доведено та обґрунтовано доцільність використання запропонованої системи розпізнавання голосових команд і розробленого алгоритму обчислення мел-частотних кепстральних коефіцієнтів. Детально викладені отримані результати попередніх експериментальних досліджень розробленої системи розпізнавання голосових команд і алгоритму обчислення мел-частотних кепстральних коефіцієнтів в середовищі MATLAB на прикладі ідентифікації вимовлених різними суб'єктами управління команд: «вгору», «вниз», «вправо», «вліво». У даній статті представлена розроблена система розпізнавання голосових команд управління на основі стеганографічно-кепстрального аналізу, яка дозволяє підвищити ефективність розпізнавання голосових команд. Детально викладені отримані результати попередніх експериментальних досліджень, на основі яких зроблені висновки про доцільність подальшого наукового і практичного застосування системи розпізнавання голосових команд управління на основі стеганографічно-кепстрального аналізу і розробленого алгоритму обчислення мел-частотних кепстральних коефіцієнтів, а також, дуже докладне обґрунтування науково-технічної значущості проведеного експериментального дослідження.

Ключові слова: класифікація об'єкта; семантична ідентифікація голосових команд; безпілотний літальний апарат; розпізнавання мови; мел-частотні кепстральних коефіцієнти.

Лавриненко Олександр Юрійович. Аспірант.
Кафедра телекомунікаційних систем, Навчально-науковий інститут аеронавігації, електроніки та телекомунікацій, Національний авіаційний університет, Київ, Україна.
Освіта: Національний авіаційний університет, Київ, Україна, (2014).
Напрямок наукової діяльності: розпізнавання мови, цифрова обробка сигналів, захист інформації.
Кількість публікацій: більше 15.
E-mail: oleksandrlavrynenko@gmail.com

Кочергін Юрій Анатолійович. Кандидат технічних наук. Доцент.
Кафедра телекомунікаційних систем, Навчально-науковий інститут аеронавігації, електроніки та телекомунікацій, Національний авіаційний університет, Київ, Україна.
Освіта: Національний авіаційний університет, Київ, Україна (1989).
Напрямок наукової діяльності: системи захисту інформації, телекомунікаційні системи.
Кількість публікацій: більше 30.
E-mail: tks@nau.edu.ua

Конахович Георгій Филимонович. Доктор технічних наук. Професор.
Кафедра телекомунікаційних систем, Навчально-науковий інститут аеронавігації, електроніки та телекомунікацій, Національний авіаційний університет, Київ, Україна.
Освіта: Національний авіаційний університет, Київ, Україна (1968).
Напрямок наукової діяльності: системи захисту інформації, телекомунікаційні системи.
Кількість публікацій: більше 100.
E-mail: tks@nau.edu.ua

А. Ю. Лавриненко, Ю. А. Кочергин, Г. Ф. Конахович. Система распознавания голосовых команд управления БПЛА на основе стеганографически-кепстральном анализе

В данной статье, предложена разработанная система распознавания голосовых команд управления на основе стеганографически-кепстрального анализа, а также, алгоритм вычисления мел-частотных кепстральных коэффициентов, который используется для создания базовой системы распознавания голосовых команд, на основе которой могут быть построены более сложные решения голосового управления. Экспериментально доказать и обосновать целесообразность использования предложенной системы распознавания голосовых команд и разработанного алгоритма вычисления мел-частотных кепстральных коэффициентов. Подробно изложены полученные результаты предварительных экспериментальных исследований разработанной системы распознавания голосовых команд и алгоритма вычисления мел-частотных кепстральных коэффициентов в среде MATLAB на примере идентификации произнесенных разными субъектами управления команд: «вверх», «вниз», «вправо», «влево». В данной статье представлена разработанная система распознавания голосовых команд управления на основе стеганографически-кепстрального анализа, которая позволяет повысить эффективность распознавания голосовых команд. Подробно изложены полученные результаты предварительных экспериментальных исследований, на основе которых сделаны выводы о целесообразности дальнейшего научного и практического применения системы распознавания голосовых команд управления на основе стеганографически-кепстрального анализа и разработанного алгоритма вычисления мел-частотных кепстральных коэффициентов, а также, основательное обоснование научно-технической значимости проведенного экспериментального исследования.
Ключевые слова: классификация объекта; семантическая идентификация голосовых команд; беспилотный летательный аппарат; распознавание речи; мел-частотные кепстральные коэффициенты.

Лавриненко Александр Юрьевич. Аспирант.
Кафедра телекоммуникационных систем, Учебно-научный институт аэронавигации, электроники и телекоммуникаций, Национальный авиационный университет, Киев, Украина.
Образование: Национальный авиационный университет, Киев, Украина (2014).
Направление научной деятельности: распознавания речи, цифровая обработка сигналов, защита информации.
Количество публикаций: более 15.
E-mail: oleksandrlavrynenko@gmail.com

Кочергин Юрий Анатольевич. Кандидат технических наук. Доцент.
Кафедра телекоммуникационных систем, Учебно-научный институт аэронавигации, электроники и телекоммуникаций, Национальный авиационный университет, Киев, Украина.
Образование: Национальный авиационный университет, Киев, Украина (1989).
Направление научной деятельности: системы защиты информации, телекоммуникационные системы.
Количество публикаций: более 30.
E-mail: tks@nau.edu.ua

Конахович Георгий Филимонович. Доктор технических наук. Профессор.
Кафедра телекоммуникационных систем, Учебно-научный институт аэронавигации, электроники и телекоммуникаций, Национальный авиационный университет, Киев, Украина.
Образование: Национальный авиационный университет, Киев, Украина (1968).
Направление научной деятельности: системы защиты информации, телекоммуникационные системы.
Количество публикаций: более 100.
E-mail: tks@nau.edu.ua