TECHNICAL SCIENCES

The efficiency improvement of the permanent voice control over the ATC actions

V.A. Temnikov*, A.V. Peteichuk

Department of information security resources, National Aviation University, Kiev, Ukraine
*Corresponding author. E-mail: temnikov_v@ukr.net

Paper received 05.12.15; Accepted for publication 15.12.15.

Abstract. Security in transport, energy and other industries that use the dispatcher labor significantly depends on their actions correctness. Traffic controller's mistakes can be caused by their being in a nervous and emotional stress that are constantly experienced in relation to high responsibility for their decisions. One of the ways to reduce the accidents number that may occur during the dispatchers personal mistakes proposes to perform an automatic permanent control over the traffic controller's actions during their operations. Current article describes the concepts and main stages of voice control over the air traffic controller (ATC) actions. Proposed control allows preventing staying on the ATC working place unauthorized person or ATC staying in appropriate emotional state (ES). Conclusion about violations is made by means of authentication and monitoring results of ATC. These results are based on analysis of ES informative parameters of the speech signal taken from the microphone output. Presents the study's results of phonemes pitch frequency values of Russian language in a different words and phrases. These results justify the usefulness of monitoring by selected from ATC continuous speech key fragments which are parts of words (phrases) commonly used in the traffic controllers operation. Key speech fragments informative parameters are determined based on separate frames (during authentication by means of short-term analysis) and single phonemes (during ES monitoring) on which key fragments are segmented. Leads the principles of the subsystems construction of recognition and identification the key speech fragments from continuous speech, also developed authentication control system over the ATC actions. The work of the first mentioned subsystems is based on developed approach to the automatic words recognition in continuous speech taken into account the specific of ATC operation. Authentication subsystem is based on trained ANN to recognize the controlled persons. Grounded choices of informative parameters allow increasing the percentage of accurate authentication above 98% and significantly reduce the implementation time, which provided subsystem operation in real time.

Keywords: *Air traffic controller (ATC), authentication, emotional states monitoring, pitch frequency, phoneme*

Introduction

The human-factor aspect has a significant impact on security in various industries. More particularly: the aviation safety significantly reliant on correct action of ATC. Mistakes in their work can be caused by the ATC emotional state (ES) finding that they are constantly experiencing due to the increased responsibility for their decisions.

As one of the ways to solve the problem of impact of human-factor aspect reducing on aviation safety proposed to perform the permanent automatic monitoring of the ATC actions during their duties execution [1]. Under the reducing of impact of human-factor aspect on aviation safety understands the accidents number diminution that can occur due to ATC mistakes.

Said monitoring should be directed on access prevention to information resources that ATC utilize during operations also unauthorized persons as well as persons in inadequate ES.

As a biometric person's feature in which control is performed, authors suggest to use the ATC voice. It makes the possibility to execute the remote control without distracting from the operation (means without direct contact with the body scanning equipment devices).

For permanent monitoring of ATC access to information resources performing being developed the voice control system over the ATC actions (CSAA). It operates in real-time, and will automatically notify the senior manager about detected violations.

The system operation report will serve as the documentary violation evidence.

In a previous author's article [1] were proposed the concept of one of the part of said control, particularly the automatic monitoring ES of ATC during their duties execution.

In current article presents the research results aimed on further ATC access control improvement to information resources.

The control efficiency improvement achieved by ATC authentication function making (injecting). It aims to prevent a finding on a working place an unauthorized dispatcher; utilizing of developed ways to improve performance also the reliability of CSAA work that is being developed.

Authentication and EM monitoring of ATC conducted by means of continuous speech, fixed in audio exchange (dialogue) process with the crew members.

The concept of CSAA block diagram building

1. Operation of ES authentication and monitoring subsystems based on the pattern recognition theory statements. They assign parameterization performing, classification and decision making of the controlled person assignment to desired class [2]. Regarding, the ES identification and monitoring subsystems include in its structure parameterization, classification and decision making modules.

2. The concept is based on the thesis that the ES authentication and monitoring of controlled persons are held by means of key fragments isolated from their continuous speech. These key fragments (or phrases) are frequently used in the ATC operation, including the fixed (standard established) phraseology.

It is caused by means of the following circumstances.

The main informative speech signal parameters are: voice pitch frequency (further just pitch) and related to it parameters such as melodic contour angularity, dispersion, central tendency. These parameters frequently used for ES determining. [3-7].

Generally admitted that the voice pitch increasing indicates about person excitation; therefore the reduction - moving a person in constrained, depression state [8].

Meanwhile, the authors during the research received the results, which indicate that the pitch value can considerably vary. Not only in the pronunciation of different words by the same person, but to be significantly different for the same phonemes entering them into the various speech fragments.

Below are results of voice pitch measurements of various key words and phrases made up to the fixed ATC phraseology, as well as the phonemes included to these keywords.

In tab. 1 as an example shown the average value of pitch for several phrases that made up to the fixed phraseology.

Table 1. Pitch for different phrases

Key phrase	Pitch, Hz
«Вас понял» (Vas ponyal)	198
«Так точно» (Tak tochno)	186
«Внимание, всем воздушным судам» (Vnimanie vsem vozdushnym sudam)	221

As seen from the tab. 1, pitch is tending to change from one keyword phrase to the next for the same speaker. Data are presented for key phrases pronounced by the same speaker in the same ES (norm state).

The pitch variance for different keywords may vary up to 40 Hz.

As an evidence of the fact that variability values of the features that characterize the same phonemes in keywords pronounced by the same speaker, tab 2 shows the pitch variation for phonemes "E", are located in different places of «Ответьте диспетчеру» phrase. We note that the human ear perceives the specified command as «атвЕттЕ диспЕчЕру».

Disregarding the mentioned results of research can lead to unreliable operation of CSAA.

In accordance to it, the part of CSAA building concept is ATC control performing by means of informative parameters that are described a particular - a key – a speech fragment.

Consequently into the CSAA structure included subsystem of speech fragments recognition and selection from ATC continuous speech key phrases.

Table 2. Pitch values for phonemes «E», are located in different parts of key phrases

Phoneme	F0, Hz
E1	133,7
E2	144,8
E3	172
E4	116

Built algorithm that is based on the developed block diagram building CSAA concept includes the following main steps:

1. The speech signal sampling (discretisation) from the microphone output by means of which the communication (dialogue) between the ATC and crew members implements.
2. The speech signal pre-processing. Is meant that the signal segmentation to voiced speech fragments, and noise reduction. The noise presence in the analyzed speech fragments can reduce the reliability the CSAA operation.
3. The ATC continuous speech segmentation to the speech fragments. Their recognition and key continuous speech pronouncing selection by parameters of which perform the ATC ES authentication and monitoring.
4. The ATC authentication by informative speech fragments parameters utilization that are intended into authentication process.
5. The ATC ES monitoring by means of speech fragments informative parameters system. It intended for monitoring, in the case of the ATC passage authentication procedures (ATC ES monitoring subsystem operation is described be the authors in [1]).

Below are shown the principles of the basic subsystems CSAA performed the processing of speech fragments in different algorithm stages.

Speech fragments recognition subsystem and key fragments selection from ATC continuous speech.

Subsystem of identification and key speech fragments selection from continuous speech based on developed approach to the automatic words recognition from ATC continuous speech taken into account the specific of their operation.

At the core of the developed approach lies the segmentation into phonemes and pauses with further recognition. The approach, based on which, proposed to perform the key fragments recognition and selection developed with taking into account about features of the ATC duties.

After segmenting the speech fragment represents as a sequence of phonemes and pause, which are shown in the form of cells. Each cell corresponds to a single phoneme or pause and defines the boundaries of beginning and end. At this point it is very important to choose the correct segmentation method that will give reliable results.

For the method developing aimed on accurate segmentation currently utilizing learned framework based on Markov models. After perform the classes cells assignment with a consistent comparison of the speech fragments with a keywords dictionary.

At the beginning the assignment is performed for the deaf consonant phonemes because are the easiest for recognizing. It is a unique phonemes class without pitch. After that are allocated voiced consonants and vowels. Voiced consonants are characterized by the noise presence unlike of the vowel. Voiced consonants noise level will be much bigger than the vowels. Even at this stage, the obtained pattern can be compared to fill in the keywords dictionary (a cells sequence, which had been assigned the class). It should be noted that in the database (dictionary) key fragments are included in two forms: a sequence of phoneme classes (in the form of cells) and key parts of speech.

The next stage is to divide the class into subclasses of cells and recognition of phonemes assigned to a particular subclass.

The vowels class dividing into subclasses. Vowel phonemes recognition is based on results of comparative analysis of the different formant frequencies values of phonemes and their correlation (average values of the first two formant frequencies F1 and F2 and their relation to the speakers of both sexes). These parameters are basic parameters for vowel phonemes recognition.

Vowel phonemes class is divided into two subclasses, the А-У-Е (ratio formant frequencies F1 and F2 lie in the 1.3 to 4.3 range) and О-И-Ы (the ratio of the frequencies F1 and F2 lie in the 5.6 to 8.5 range).

Further execute the assignment of phoneme probability of belonging to one of the subclasses.

Accuracy of phonemes classifying to one of the subclasses is checked by comparing the current pattern with the existing in dictionary. Similar procedures are performed in relation to the vowel phonemes.

Applying the mentioned approach can improve the subsystem performance in comparison to its construction based on artificial neural networks.

ATC authentication subsystem

ATC authentication is performed using the short-term analysis method (prescribing the key speech fragment to the short duration frames [9]).

The authentication subsystem operations based on aggregate classification of informative parameters are calculated by linear cepstrum coefficients prediction. The cepstrum coefficients system construction based on coefficients linear prediction determined by the fact that, in accordance with [8] coefficients linear prediction relatively weak depend on the person ES.

Authentication subsystem is based on trained ANN [10] for the controlled persons recognizing.

Research has shown that as the ANN advisable to use multilayer perceptrons with one hidden layer.

Specific values of ANN parameters and informative speech fragments were determined in the process of modules classification testing on the criteria of the maximum percentage of accurate authentication.

ANN applying with multiple outputs allows increasing the access control reliability - in case of informed choice of the accurate authentication parameters can reached above 98%.

Conclusion

1. Proposed to perform the classification of controlled persons by means of key speech fragments are allocated from ATC continuous speech. Said key speech fragments are parts of words and phrases frequently used during working process, also included to the fixed by standards professional phraseology.

2. Applying of developed approaches performance improvement and reliability of subsystem authentication, including, and grounded choice of its parameters, allowed to increase the percentage of accurate authentication above the 98%. Also significantly reduce the implementation time, which provided subsystem operation in real time.

3. Applying the developed system of ATC control over their actions will allow significantly reduce the impact of human factors aspect on the aviation safety. It performed by means of preventing of possible illegal actions by unauthorized persons, also reducing the errors likelihood that can allow the controllers that are in an improper ES.

REFERENCES

1. Temnikov, V, Peteichuk, A. The concept of construction an automatic system for ATC emotional condition monitoring // Science and Education a New Dimension. Natural and Technical Sciences, III(6), Issue:54, 2015 – P. 52-54.
2. Ramiashvili, G.S. Automatic speaker recognition by voice // M.: Radio and Communication, 1981. – 224 p.
3. Frolov, M.V. Monitoring of functional state of human operator – M.: Nauka, 1987. – 197 p.
4. Taubkin, V.L. Recognition of the emotional state of a human operator using the parameters of the speech signal // Thesis. ...PhD. – M., 1977. – 160 p.
5. Williams, C.E., &Stevens, K.N. (1969). On determining the emotional state of pilots during flight: An exploratory study. Aerospace Medicine, 40, 1369-1372.
6. Kuroda, I., Fujiwara, O., Okamura, N., & Utsuki, N. (1976). Method for determining pilot stress through analysis of voice communication. Aviation, Space, and Environmental Medicine, 47, 528-533.
7. Congleton, J.J., Jones, W.A., Shiflett, S.G., Mesweeny, K.P., Huchingson, R.D. 1997). An evaluation of voice stress analysis techniques in a simulated AWACS environment. International Journal of Speech Technology, 2, 61-69. doi: 10.1007/BF02539823.
8. Tkachenya, A.V. The problem of the emotions stability informative features in speech recognition problem // Journal BGU, Part1. – 2014. – №3. – P. 56-61
9. Rabiner, L., Gould B. Theory and Application of Digital Signal Processing. M.: Mir, 1978. – 848p.
10. Haykin, S.Neural networks. // 2nd ed. Tr. From Eng. – M.: Publishing house "Williams", 2006. – 1104 p.