REVIEW

on a dissertation for acquiring the educational and scientific degree "Doctor"
in a scientific specialty:
01.01.12. Informatics
in a professional direction:
4.6. Informatics and computer science
Author of the dissertation: mag. Eng. Atanas Petrov Uzunov
Topic of the dissertation:
SPEECH DETECTION IN SPEAKER RECOGNITION SYSTEMS
Reviewer: Prof. Dr. Eng. Alexander Bogdanov Bekyarski

1. Relevance of the problem developed in the dissertation in scientific and scientific-applied terms

The topic of the dissertation "SPEECH DETECTION IN DICTOR RECOGNITION SYSTEMS" presented for review naturally reveals the presence of topicality related to the existing and constantly growing requirements for the processing of speech signals through appropriate methods and approaches. With the use of biometric technologies, in order to recognize a speaker, the analysis of the individual characteristics of the speaker increases the accuracy of identification and recognition. This determines the relevance and defines the relevant purpose of this dissertation - development of methods and algorithms for determining robust features designed for algorithms for speech detection and their study in the context of speaker recognition tasks for a voice signal recorded on a telephone channel. The approach and tools proposed in the development of the dissertation are related to speech detection by applying combined methods and algorithms for processing and analysis of speech signals and the use of a priori knowledge. This tendency is the basis of the actuality and perspective expressed in the topic of the dissertation in scientific and scientific-applied terms. In accordance with this assessment of relevance, the relevant focus in the dissertation for proposing, researching and practical use of new methods in the processing of speech signals should be emphasized.

2. Degree of knowledge of the state of the problem and creative interpretation of the literary material

The essence of Chapter 1 of the dissertation directly reveals, justifies and proves the excellent knowledge of the doctoral student of the state of the problem - processing of speech signals transmitted by telephone, in order to determine robust principles by analyzing the individual characteristics of the speaker. This is

achieved through appropriate creative interpretation of the literature, in-depth analysis of the current level of research, development, experiments and practical realizations in the field of speech signal processing. A total of 151 literature sources were used in the process of conducting a thorough review and critical analysis of the existing methods and research in the field of speech signal processing, of which 149 are in Latin and 2 in Cyrillic. All of them are currently relevant scientific publications included in the attached list of references to the dissertation and are from renowned journals, publications from international conferences and electronic sources by leading, currently prominent authors, foreign scientists and specialists in the field of processing speech signals.

3. Correspondence of the chosen research methodology and the set goal and tasks of the dissertation with the achieved contributions

There is a full and well-founded correspondence of the chosen research methodology and the set goal and tasks of the dissertation with the achieved contributions. This statement should include the justification of the goals and objectives in the dissertation, the well-defined significance and necessity in scientific and scientific-applied terms of research in the field of speech signal processing. The correct approach is proposed, according to which the doctoral student has formulated precisely and in a synthesized form both the goal and the tasks in his dissertation, as follows:

Aim of the dissertation: The aim of the dissertation is the formation of robust features designed for algorithms for speech detection and their study in the context of the tasks for speaker recognition in a speech signal recorded on a telephone channel.

In accordance with the defined goal of the dissertation, the following tasks of the dissertation are formulated:

1. Definition of robust features intended for speech detection and based on the properties of the spectral autocorrelation function and the spectrum of group delay and study of their characteristics

2. Development of an approach for determining the boundary points of a voice message including an algorithm for calculating adaptive threshold values and a deterministic finite state machine

3. Development of algorithms for determining boundary points and experimental study of their effectiveness with the proposed in the dissertation features in the verification of speakers with fixed phrases

4. Development of algorithms for detection of speech segments and experimental study of their effectiveness with the features proposed in the dissertation for text-independent identification of speakers.

4. Brief analytical description of the nature and assessment of the reliability of the material on which the contributions of the dissertation are built

The dissertation has a volume of 160 pages. includes an introduction, five chapters, closely related to the formulated main tasks, conclusion (summary of the obtained results), list of the main contributions, list of the dissertation publications, used literature, list of used abbreviations. A total of 151 literary sources were cited, 149 of which were in Latin and 2 in Cyrillic. The dissertation has a total of 48 figures and 27 tables. The numbers of the figures and tables in the abstract correspond to those in the dissertation. The realization of the correctly set goal and respective tasks of the dissertation is the subject of the next five chapters of the dissertation, between which consistency in the exposition, logical connection and methodical presentation of the new important theoretical and practical solutions in the field of dissertation. the topic of the dissertation.

Chapter 2 "Defining Speech Detection Signs Using the Properties of SACF and SGS" discusses some characteristics of the spectral autocorrelation function (SACF) obtained by the FFT spectrum. A method is proposed in which a delta spectral autocorrelation function is obtained by applying a delta filter on the spectral autocorrelation function (SACF). The change of GDS in noise-canceled speech signals is analyzed - based only on the properties of the delta spectral autocorrelation function and by combining it with the modified group delay spectrum. Based on the analysis, five features for speech detection are proposed - MD, log-GDMD, lin-GDMD, BMD and MMD. The first three are intended for detection by time loop analysis, and the last two - for detection by recognition algorithms.

Based on the research conducted in Chapter 2 and the corresponding analysis in Chapter 3 "Algorithms for determining boundary points in text-dependent verification of speakers. Experimental Research "a comparative experimental analysis of the effectiveness of the proposed in ch. 2 signs intended for speech detection by analysis of time contours. The following are selected as reference traits: a trait obtained by a combination of signal energy and spectral entropy (Energy-Entropy (EE) parameter); Spectral Entropy with Normalized frame Spectrum (SENS parameter); Modified Teager's Energy (MTE parameter) and Long-term Spectral Divergence (LTSD parameter). Methods and algorithms for

determining the boundary points of short phrases have been proposed, developed and researched.

Based on the obtained experimental results, the following three conclusions were made: First - the detectors based on the log-GDMD feature in all tests are superior to those based on LTSD; Second - the detection accuracy when using a finite state machine with adaptive thresholds is always superior, with the same feature, that obtained by the hangover algorithm and Third - in terms of error in 24 verifications in most cases the detector with automatic state machine and adaptive thresholds with the same feature, the one with the hangover algorithm, but the difference between them is not statistically significant. The accuracy of detection was experimentally investigated by analyzing the differences between the manually determined boundary points and those obtained by the proposed algorithms. It has been established how different detection algorithms affect the accuracy of recognition in text-dependent speaker verification.

The methods and algorithms proposed and developed in Chapter 3 are the main ones, through which the achievement of the goal in the dissertation is theoretically substantiated and argued - speech detection algorithms and their research in order to identify and recognize speakers.

In Chapter 4 "Speech detection algorithms for text-independent speaker identification. Experimental Research "a comparative experimental analysis of the effectiveness of the proposed in ch. 2 signs designed for speech detection. Two algorithms for speech detection - VAD-1 and VAD-2 - were experimentally studied. Two speech detectors were studied, which are in fact binary classifiers. The announcer identification algorithm, regardless of the text as a classifier, uses a multilayer perceptron with a single hidden layer and a training algorithm for error backpropagation. The experimentally determined results of the algorithms for speech detection and speaker identification are presented in tabular and graphical form. Conclusions are made about their effectiveness. The influence of speech detection algorithms on the accuracy of recognition in a system for text - independent speaker identification has been studied. The experiments were performed with spoken data in Bulgarian, recorded on a telephone channel.

Chapter 5 "BG-SRDat - corpus with voice data recorded on a telephone channel and intended for speaker recognition" describes the corpus BG-SRDat (Bulgarian language Speaker Recognition DATa) q containing speech recorded on a telephone channel (fixed and mobile telephones and via VoIP) and including phrases and conversations in Bulgarian and only phrases in English. It is used for research in the field of speaker verification through fixed phrases (in Bulgarian and English), text-independent speaker identification and speech detection. The

main trend in the future development of the corpus will be its gradual transformation into a corpus containing voice data received only from mobile devices.

5. Scientific and/or scientific-applied contributions of the dissertation work

The main results achieved in this dissertation are summarized and presented in a synthesized form in the form of the following scientific-applied and applied contributions:

Scientific contributions:

1. A method is proposed in which by applying a delta filter on the spectral autocorrelation function the so-called delta spectral autocorrelation function (Chapter 2, item 2.1.3)

2. A theoretical analysis of the change of the group delay spectrum in speech signals noisy with additive noise has been performed (Chapter 2, item 2.2.3).

3. The following approaches are proposed: for the calculation of signs for speech detection based on the properties of the delta spectral autocorrelation function. Through this approach, three features are defined (Chapter 2, item 2.1.4); to calculate signs for speech detection based on a combination of the modified group delay spectrum and the delta spectral autocorrelation function (Chapter 2, item 2.2.4); for determining the boundary points of a voice message, including an algorithm for calculating adaptive thresholds and a deterministic finite state machine (Chapter 3, items 3.4.2-3).

Scientific and applied contributions:

1. A comparative experimental analysis of: proposed in ch. 2 signs in relation to selected reference ones, where it is proposed to use the Euclidean distance between Z-normalized time contours, calculated for each sign, respectively from pure and noisy signal (Chapter 3, item 3.3); of the effectiveness of the proposed in ch. 2 features in relation to selected reference ones when using them in the developed algorithms for determination of boundary points (Chapter 3, item 3.6.3 and Chapter 3, item 3.6.4); of the effectiveness of the proposed in ch. 2 parameters compared to selected reference ones when using them in VAD-1 and VAD-2 (Chapter 4, item 4.6.5 and item 4.8.4, Chapter 4, item 4.7 and item 4.9).

2. The following algorithms have been developed: three algorithms for determining boundary points, based on the proposed approach and formed according to the used time contours (Chapter 3, item 3.5); two algorithms for

detection of speech segments using a classifier based on a multilayer neural network scalar, vector features and threshold logic (Chapter 4, item 4.6 and item 4.8).

Applied contributions:

1. Development, with the emphasized participation of the doctoral student, as a chief scientific consultant in the company Fadata, Ltd., of the project "Speaker Recognition API (SR-API)" project, which is used (in Chapter 3) in the program implementation of the algorithm for determination of boundary points in two applications for verification of speakers - DTW and HMM.

6. Assessment of the degree of personal participation of the doctoral student in the contributions

The personal participation of the doctoral student in the presented scientific-applied and applied contributions to the dissertation is indisputable and confirmed, both by the exposition in the dissertation and the abstract, and by the publishing activity (6 scientific publications), where the doctoral student has no co-authors.

7. Evaluation of the publications on the dissertation

The main achievements in the dissertation are popularized in 6 scientific publications, one of which is referred to in the Web of Science, three are referred to in Scopus. All publications are of theoretical and applied significance, related to the dissertation and the professional field 4.6. Informatics and computer science. The doctoral student presented a report on observed citations - a total of 25 copies. I accept the publishing activity as completely sufficient in volume, at a high scientific level and sufficiently popularized nationally and internationally.

8. Use of the results of the dissertation work in the scientific and social practice

The declared participation of the doctoral student in the presented "Information for participation in contracts related to the topic of the dissertation" should be accepted as real proof for the use of the results of the dissertation work in the scientific and social practice: the doctoral student is a participant in two contracts with the Research Fund. Biometric parameters for identification "and" Methods and algorithms for analysis of combined biometric information "; Contract from the VI Framework Program - №507634 (2004-2007), BIOSECURE - Network of Excellence - project from the 6th Framework Program for identification by biometric indicators; two internal contracts of IIT-BAS "Identification of biometric parameters" and "Recognition of biometric parameters". The PhD

student was the lead scientific consultant in a contract for the development of an application interface for speaker recognition - Speaker Recognition API (SR-API). The doctoral student's participation in these contracts, closely related to the topic of the dissertation, as well as their successful implementation, clearly show the doctoral student's desire to actually apply the algorithms and software applications developed in the dissertation for speech detection in recognition systems.

9. Assessment of the compliance of the abstract with the requirements for its preparation, as well as of the adequacy of the reflection of the main positions and the contributions of the dissertation

The precise review of the abstract to the dissertation work shows the full compliance of the abstract with the requirements for its preparation, as well as the adequacy of reflecting the main points and contributions of the dissertation. The abstract to the dissertation can be reasonably evaluated and characterized as a synthesized version of the dissertation, accurately reflecting its main content, theoretical and practical results and defined scientific and applied contributions contained in the full text of the dissertation.

10. Opinions, recommendations and notes

All theoretical and practical statements in the dissertation are presented in a reasoned, correct and methodologically appropriate sequence. Therefore, when reviewing this dissertation, only minor editorial inaccuracies and omissions can be pointed out, some of which are the following:

1. The conclusions to the first chapter (1.4. Conclusion) should critically emphasize on the basis of the extensive review of the existing methods the need for new research in the field of methods and algorithms for speech detection in order to argue more clearly. the set goal and tasks of the dissertation.

2. The presented several scientific and applied contributions (№ 1, № 3, № 4, № 6 and № 7) related to the conduct of a "comparative experimental analysis" can be combined (without diminishing their significance), as they all refer to "those proposed in ch. 2 signs in relation to selected reference ones ", but by applying different comparison criteria.

3. Scientific and applied scientific contributions should be presented in a comparative aspect with existing similar methods, approaches and algorithms, in order to emphasize even more clearly the specific contribution and positive effect of the new or modified methods, approaches and algorithms developed in the dissertation.

4. I believe that the methods, approaches and algorithms developed in the dissertation have value and practical application in speech detection for speaker recognition not only from speech signals recorded on a telephone channel, which is useful and should be reflected in the dissertation contributions.

I assume that these remarks and recommendations will be valuable and useful in the future research and development work of the doctoral student.

11. Conclusion

I believe that the overall positive assessment of the theoretical and practical results achieved in the development of this dissertation, defined as scientific and applied contributions and reflected in a sufficient number of scientific publications and in relevant scientific journals and conferences, are quite sufficient grounds for clear positive conclusion regarding the qualification of the doctoral student and his / her confirmed in the dissertation work qualities of a scientist in the chosen scientific field. Therefore, I propose to the esteemed scientific jury to award the educational and scientific degree "Doctor" to Mag. Eng. Atanas Petrov Uzunov in professional field 4.6. Informatics and computer science.

Date: 11.06.2020.   JURY MEMBER:

Prof. Dr. Alexander Bogdanov Bekyarski