OPINION
on
the dissertation of **Atanas Petrov Ouzounov** on
"SPEECH DETECTION IN SPEAKER RECOGNITION SYSTEMS"
for acquiring PhD degree in Informatics and Computer Sciences
by
Prof. Dr. Gennady Agre

Institute of Information and Communication Technologies at Bulgarian Academy of Sciences

By order № № 55/09.03.2020 of the Director of the Institute of Information and Communication Technologies - Bulgarian Academy of Sciences, I have been approved as a member of the Scientific Jury in connection with the procedure for acquiring educational and scientific degree "doctor" in the doctoral program "Informatics" in professional field 4.6 "Informatics and Computer Sciences" by Atanas Petrov Ouzounov with a dissertation entitled "SPEECH DETECTION IN SPEAKER RECOGNITION SYSTEMS" with the Scientific Supervisor Assoc. Prof. Georgi Gluhchev.

As a member of the Jury I received:

1. A copy of the dissertation for acquiring educational and scientific degree "doctor" in professional field 4.6 "Informatics and Computer Sciences"
2. An abstract of the dissertation (in Bulgarian)
3. An abstract of the dissertation (in English)
4. Professional Curriculum Vitae
5. Copies of the author's publications related to the dissertation.

The dissertation consists of 164 pages, structured in: Introduction, 5 chapters, Conclusion and Bibliography.

1. Actuality of the dissertation topic and relevance of its objectives and tasks

The dissertation is related to the voice biometrics for mobile devices – a fast developing area of the information technologies inspired by the widespread use of mobile phones and voice over Internet applications in recent years. One of the main challenges in this area is the speech detection – recognizing and localizing speech fragments in audio streams received by a device. The rapid development of biometric technologies (including voice biometrics) worldwide, determines the topic of thesis - voice activity detection in speaker recognition systems - as fully up to date, and the objectives and tasks to be solved – as expedient.

2. Knowledge of the state-of-the-art in the field

The PhD candidate showed a good knowledge of the thematic field as well as of topics related to the realization of the dissertation goals. The first chapter analyzes the main approaches for speech detection in the context of systems for speaker detection developed in the last decade. The bibliography contains 151 sources; only 2 of them are in Bulgarian. 41 sources have been published in the last five years.

3. Methodology of the study

The research methodology chosen by the author is derived from the dissertation objectives and tasks – development of robust features for voice activity detection algorithms intended for speaker recognition in audio streams via telephones. It should be noted that each of the algorithms proposed in the dissertation is combined with an experimental study of its effectiveness. In order to fulfill the dissertation goal the author applies the apparatus from the linear algebra (linear transformations, etc.), digital signal processing (correlation analysis, spectral analysis and others) and pattern recognition (neural networks, hidden Markov models, etc.).

4. Short analytical resume of the dissertation

The first chapter is an analytical review of main algorithms for speech detection. Two types of detectors are analyzed – the first tries to find the boundary points of a speech message (Endpoint Detection), and the second - speech fragments (Voice Activity Detection). In both cases the accent of the review is on the systems operating with certified speech corpuses recorded via telephone channel. The main goal of the dissertation, as well as four tasks to be solved in order to satisfy the goal, are formulated in the end of the chapter. The second chapter is devoted to defining speech detection features based on the Spectral Auto Correlation Function (SACF) and on the Group Delay Spectrum (GDS). Two approaches for using SACF are proposed – in the first approach the features are determined only by the SACF properties, while in the second – they are constructed by combination of properties of SACF and of the modified version of GDS. The chapter also considers the main methods for GDS calculation and proposes an analysis of how GDS is changed when an additive noise has been added to the voice signals. The third chapter represents an experimental study of the algorithms for Endpoint Detection in the case when the speaker verification depends on text. An approach for detecting such endpoints is proposed – it is implemented by means of three algorithms, in which the endpoints are formed according to the time contours used. Experimental evaluation of the effectiveness of the proposed algorithms is also presented. The fourth chapter describes an experimental study of Voice Activity Detection algorithms in the case of text-independent speaker identification. Two such algorithms are proposed and analyzed – the first uses a vector representation of the features and a fully connected multilayer neural network as a classifier, and the second uses the threshold-based logic. The fifth chapter presents a telephone speech corpus intended for speaker

recognition that is intensively used for evaluating algorithms described in the dissertation. The corpus includes short phrases, reading text and conversations in Bulgarian as well as just short phrases in English. The conclusion summarizes the content and the results of the dissertation as well as presents the dissertation contributions, the publication activity of the author and the information about citations found.

5. Contributions of the dissertation

The main contributions of the author are stated correctly. However, in my opinion, some of the scientific and applied contribution may be jointed.

6. Assessment of publications related to the dissertation

The author presents 6 publications related to the dissertation – all of them are written without any coauthors. Two of them have been published in the Proceedings of international conferences and the rest four – in the international journals. The publications reflect the content of the dissertation and represent the original achievements of the dissertation' author. According to the "Rules for applying the Law on the Development of the Academic Staff in the Republic of Bulgaria" and the "Regulations for the Specific Conditions for Acquiring Academic Degrees and Positions at IICT-BAS", the PhD candidate must have at least 30 points collected from his publications related to the dissertation that have been published in journals or conference proceedings that are indexed in Web of Science or Scopus. Only four publications (№ 1, 2, 3 and 6) are satisfied these requirements. It should be noted that the last three articles have been published in scientific journals with SJR rank of Scopus. According to the "Regulations for the Specific Conditions for Acquiring Academic Degrees and Positions at IICT-BAS" (which imposes more strict requirements on the quality of publications that the national rules), the mentioned above four publications collect 72 points, which more than two times greater than the requested minimal threshold. In addition, it should be noted that publication № 1, 2 and 6 have 25 independent citations, which also confirms the high quality of these publications and the dissertation.

7. Abstract of the dissertation

The abstract of the dissertation correctly presents the objectives, tasks and the results of the dissertation.

8. Conclusion

From all mentioned above I can to conclude that all requirements of the Law on the Development of the Academic Staff in the Republic of Bulgaria, the Rules for Applying the Low, as well as of the Regulations for the Specific Conditions for Acquiring Academic Degrees

and Positions at IICT-BAS are fully satisfied. I am sure that the dissertation of Atanas Petrov Ouzounov has all the characteristics required by a PhD Thesis. The results presented in the dissertation are significant and contribute to the development of the existing methods for voice biometrics.

All this facts give me the reason for a positive evaluation of the dissertation and I strongly recommend the honored Scientific Jury to award the educational and scientific degree "doctor" in professional field 4.6 "Informatics and Computer Sciences" to Atanas Petrov Ouzounov.

27 May 2020

Sofia