Military University of Technology

Faculty of Electronics

ABSTRACT

Automatic Speaker Recognition System Based on Cepstral Analysis of Speech Signal and Gaussian Mixture Models

Author: por. mgr inż. Kamil Kamiński

Supervisor: prof. dr hab. inż. Andrzej P. Dobrowolski

The purpose of this dissertation was to develop an automatic speaker recognition system allowing for effective identification of voices in difficult conditions of registration, similar to those of telephone transmission. The dissertation presented the process of designing the system as well as the results of the optimizations and tests carried out with the use of voice databases created at the Military University of Technology as well as commercial databases.

The dissertation consists of four logical parts presenting the subsequent stages of the performed works. The first part (Chapter 1) presents the issue of voice biometry and its determinant manner of human voice organ functioning. Potential application areas of the speaker recognition system were also formulated along with the operational challenges faced by the speaker, which affect the security of the executed authentications.

The next part (chapters 2-6) presents the architecture of the created speaker recognition system based on cepstral analysis of speech signal and Gaussian mixture models. All the stages of speech signal processing from signal acquisition, through its initial processing, generation and selection of individual features, creation of memory-efficient voice models, to the decision on the result of the conducted identification, were discussed in a chronological way. The presented system was subject to numerous modifications and optimizations in the course of the research.

The third part of the work (Chapters 7 to 8) is composed of the results of research carried out during the design and optimisation of the system. It also shows the results of speaker identification achieved by the described system in its finalised form, carried out in various testing variants. The application implementing the developed speaker recognition system in the *Matlab* environment was also presented. This application facilitates adding new voice models and perform speaker identification tests.

The last part of the dissertation is a summary, in which, on the basis of the previously presented research, the thesis of this dissertation was proved, in the following wording: ***it is possible to process the speech signal in a way enabling to create memory-efficient voice models allowing to identify the speaker with high efficiency***. In this part, the predicted further areas of development work of the presented speaker recognition system are also presented.

Key words: speaker recognition, cepstral analysis, Gaussian mixture models, genetic algorithms, selection of distinctive traits

Kamil Kamiński
Warsaw, 21 June 2018