



Military
University
of Technology

Faculty
of Electronics



Abstract

AUTOMATIC SPEAKER RECOGNITION SYSTEM SUPPORTED BY BEHAVIORAL FEATURES OF SPEECH SIGNAL

Author: Lt Dominik Mały MSc Eng.

PhD student supervisor: Prof. Andrzej P. Dobrowolski, Ph.D., D.Sc., Eng.

Assistant supervisor: Maj. Kamil Kamiński, Ph.D., Eng.

The aim of this dissertation was to develop and implement a set of behavioral features within an existing ASR system to enable an increase in the number of accurate speaker identity identifications, particularly under challenging recording conditions. The dissertation comprises three parts, each illustrating successive stages of the research.

The first part (Chapter 1) introduces the field of voice biometrics, beginning with a description of the human voice generation process, followed by an overview of the structure and classification of automatic speaker recognition systems, and concluding with a literature review. The outcome of this research stage was the determination of the subsequent research direction.

The second part (Chapters 2-4) presents commonly used ASR system architectures, and the algorithms implemented within them. Subsequently, it outlines the successive stages of speech signal analysis from recording to segmentation. The primary focus is placed on aspects related to the origin of features that can be extracted from the speech signal and the methods for their extraction. These issues are thoroughly characterized in the context of the properties of the *ARMiA* system developed at MUT, based on physical features, as well as an original algorithm for automatic speaker recognition based on behavioral features.

The final part of the dissertation (Chapters 5-7) presents conclusions drawn from the research conducted during the design and optimization of the integration of the *ARMiA* system with a proprietary system based on behavioral features. The implementation of system fusion is demonstrated, followed by test results of this solution under conditions of external interference and the use of various acoustic pathways for voice recording. The obtained results confirm the hypothesis proposed in the dissertation, demonstrating that **it is possible to define and extract behavioral features from the speech signal which, when combined with physical features implemented in the existing solution, significantly improve its effectiveness in the presence of external disturbances.**

Key words: speaker recognition, behavioral features, data fusion, genetic algorithm, distinct feature selection

Dominik Mały

Warsaw, 01.04.2025.