University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology

UVP TECHNICOM

SYSTEM FOR AUTOMATIC TRANSCRIPTION OF AUDIO MEETINGS IN SLOVAK

J. Juhár*, J. Staš, P. Viszlay, M. Lojka, D. Hládek, E. Kiktová, M. Sulír, D. Zlacký, M. Pleva, S. Ondáš, J. Vavrek

Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovakia

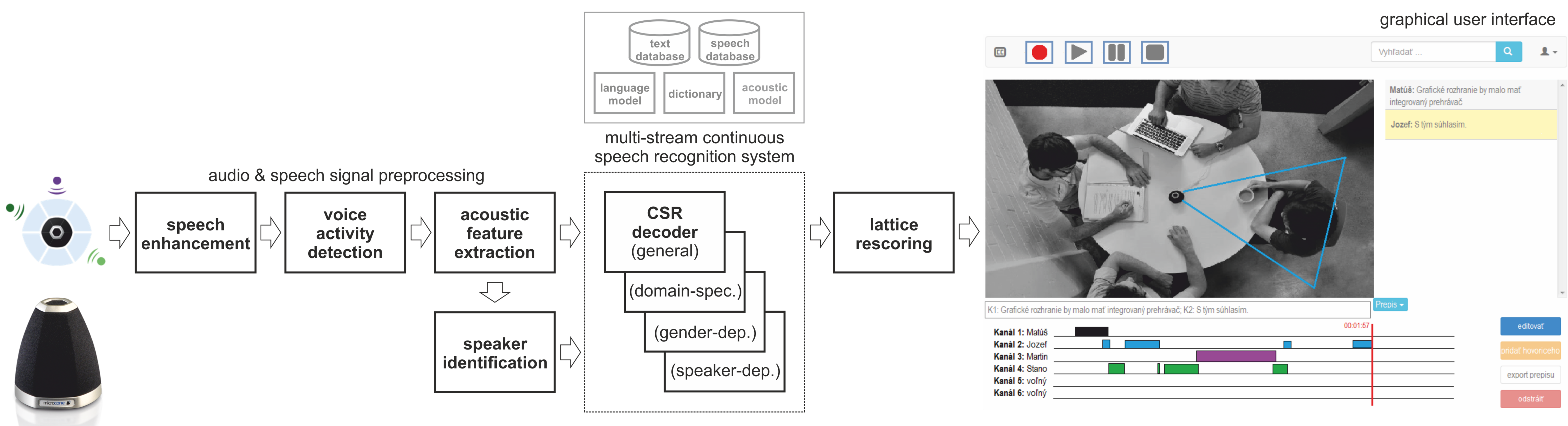
***jozef.juhar@tuke.sk**

Abstract

The main aim of the pilot project is a research and development of the meeting speech recognition system for the Slovak language. The system includes:

- voice recording of meetings in small conference rooms with a limited number of participants using microphone array;
- automatic domain-specific and speaker-dependent speech-to-text transcription in the Slovak language;
- management for storing, browsing, searching and synchronization of transcripts with audio recordings.

Expected outcome is functional prototype, practically applicable software product for automatic transcription of meetings, educational talks and lectures.



System Description

The proposed prototype of the meetings speech recognition system (MSR) in the Slovak language consists of:

- microphone array (Microcone - up to 6 speakers);
- multi-channel speech recorder;
- multi-channel speech enhancement (denoising/dereverberation);
- robust speech/non-speech (voice) activity detection;
- speaker segmentation and verification (including diarization);
- effective acoustic feature extraction and dimension reduction;
- gender-dependent and speaker-dependent acoustic modeling;
- domain-specific and speaker-dependent language modeling;
- knowledge-based lattice rescoring module;
- simple workflow management for storing, browsing, searching and synchronization of audio recordings and transcripts;
- interactive graphical user interface (GUI).

Advanced Algorithms Overview

- speech enhancement using extended spectral subtraction (ESS) and minimum mean square error (MMSE);
- spatiotemporal averaging (SMERSH) and derive glottal closure instants (DYPSA) for multi-microphone speech dereverberation;
- blind audio separation through independent component analysis (ICA);
- voice activity detection based on variance of mel-frequency cepstral (MFC) coefficients and principal component analysis (PCA);
- improving discriminative properties of reduced acoustic features (through PCA) using two-dimensional linear discriminant analysis (2D-LDA);
- unsupervised speaker adaptation through constrained and feature space maximum likelihood linear regression (CMLLR and fMLLR);
- language model adaptation using linear interpolation (LI) of several domain specific models and minimum discrimination information (MDI) approach;
- multi-thread server-based MSR system with two-pass recognition strategy;
- knowledge-based lattice rescoring based on recognized output voting error reduction (ROVER) system.

References

[1] Viszlay, P., Lojka, M., Juhár, J., Class-Dependent Two-Dimensional Linear Discriminant Analysis using Two-Pass Recognition Strategy, In: Proc. of EUSIPCO-2014, Lisbon, Portugal, 2014.
[2] Staš, J., Hládek, D., Juhár J., Recent Advances in the Statistical Modeling of the Slovak Language. In: Proc. of 56th International Symposium ELMAR-2014, Zadar, Croatia, 2014, pp. 39-42.
[3] Staš, J., Juhár, J., Hládek, D., Classification of Heterogeneous Text Data for Robust Domain-Specific Language Modeling, EURASIP Journal on Audio Speech, and Music Processing, 2014(14), 2014.
[4] Lojka, M., Ondáš, S., Pleva, M., Juhár, J., Multi-Thread Parallel Speech Recognition for Mobile Applications, Journal of Electrical and Electronics Engineering, 7(1), 2014, pp. 81-86.
[5] Rusko, M., et al., Recent Advances in the Slovak Dictation System for Judicial Domain, In: Proc. of the 6th Language and Technology Conference, LTC-2014, Poznań, Poland, 2013, pp. 555-560.

Acknowledgements. This poster is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF.

Európska únia
Európsky fond regionálneho rozvoja

„We support research activities in Slovakia/
This project is being co-financed by the European Union“

Operačný program VÝSKUM a VÝVOJ