



## VOICE-PRINT TRANSFORMATION FOR MIGRATION BETWEEN AUTOMATIC SPEAKER IDENTIFICATION SYSTEMS (EAFS-0475)

**O. Glembek<sup>1</sup>, P. Matejka<sup>1</sup>, L. Burget<sup>1</sup>, P. Schwarz<sup>1</sup>, J. Pesan<sup>1</sup>, O. Plchoť<sup>1</sup>**

<sup>1</sup>Brno University of Technology, Department of Computer Graphics and Multimedia, Brno, Czech Republic

This presentation discusses the scenario of migrating from one forensic automatic speaker identification system (FASIS) to another. In FASIS, an audio recording of a reference speaker is used to train the speaker model. This model is then compared with the model of the tested speaker and a comparison score (in the form of log-likelihood ratio) is computed.

System migration is usually motivated by improving the system recognition accuracy, typically because of technological upgrade, or because of the necessity of processing new kind of data. Unfortunately, such migration usually results in the incompatibility of speaker models and, therefore, in the inability to compare two models. The solution would be to re-train the speaker models and rebuild a model database; however, it may and most likely will happen that the access to the original audio file is unavailable, e.g. due to legal issues. This work introduces a technique of transforming the original speaker models so that –with a slight loss in the accuracy– they are compatible with the new FASIS models. We presents the results on the NIST SRE 2010 evaluation tasks.

Our system is based on the i-vector framework which converts arbitrarily long audio waveform to a fixed-length low-dimensional vector which serves as a speaker model. In this context, the i-vector is sometimes referred to as a voice-print. We use Artificial Neural Networks to restore the original speaker models by mapping them to the new domain. We show that there is approximately 20% relative increase in error rates when substituting the new test speaker models with the restored ones. Normally, the incompatibility of the original speaker models without having the audio files available would make such task impossible.