Mismatch compensation in the evaluation of evidence under conditions reflecting those of an actual forensic-voice-comparison case

Ewald Enzinger[1,2]
[1]*School of Elec. Eng. & Telecom., University of New South Wales, Sydney, Australia*
[2]*Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria*
ewald.enzinger@oeaw.ac.at

This presentation demonstrates the application of mismatch compensation techniques in the evaluation of forensic evidence under conditions reflecting those of an actual forensic-voice-comparison case. Several approaches developed in automatic speaker recognition research are considered for use in a forensic-voice-comparison analysis to reduce variability in quantitative measurements made of the acoustic properties of the voices on the suspect and offender recordings caused by mismatched recording conditions. Other aspects of the forensic analysis such as the consideration of the relevant prosecution and defence hypotheses to address in this case, selection of data reflecting the adopted defence hypothesis, statistical modelling, and likelihood ratio calculation are the subject of another proposed presentation.

The offender recording in this case was of a landline telephone call made to a call centre. As well as telephone transmission, it included background noise at the call centre, and it was saved in a compressed format. The suspect recording was of a police interview conducted in a reverberant room with ventilation noise and saved in a compressed format. For this illustration we used recordings from a research database. Procedures are described for simulating the recording conditions of the suspect and offender samples to convert recordings taken from the database. A pair of offender and suspect condition recordings of one speaker was selected as mock offender and suspect samples, respectively, to stand in place of the speakers on the actual casework recording.

We compared the validity and reliability of a forensic-voice-comparison system incorporating feature warping (Pelecanos and Sridharan, 2001) using Gaussian cumulative distribution function matching, probabilistic feature mapping (PFM; Mak et al., 2007), and feature-domain nuisance attribute projection (NAP; Campbell et al., 2008), as well as combinations thereof. While substantial improvements in validity were observed for all techniques, reliability deteriorated. The best performance was obtained by a combination of feature warping and probabilistic feature mapping.

The presentation will include an illustration of how we incorporated the combined feature warping and probabilistic feature mapping compensation method into our forensic-voice-comparison system, the results from testing validity and reliability of this system, and a demonstration of the evaluation of the likelihood ratio for the mock offender and suspect samples.

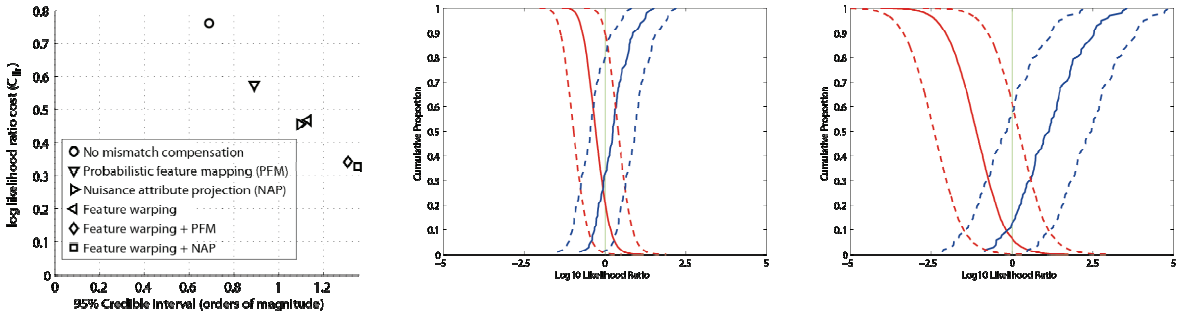

Figure 1 Measures of validity (C_{llr}) and reliability (*log*10 95% credible interval) for systems without and after incorporating mismatch compensation techniques (left); Tippett plots of the system without mismatch compensation (middle) and the system incorporating feature warping and probabilistic feature mapping (right). Solid lines represent likelihood ratios obtained from tests of the system, and dashed lines represent the 95% credible interval.

References

Campbell, W., Sturim, D., Torres-Carrasquillo, P., and Reynolds, D. (2008). A comparison of subspace feature-domain methods for language recognition. Proc. Interspeech, pp. 309–312.

Mak, M.-W., Yiu, K.-K., and Kung, S.-Y. (2007). Probabilistic feature-based transformation for speaker verification over telephone networks. Neurocomputing 71, 1-3, 137–146.

Pelecanos, J. and Sridharan, S. (2001). Feature warping for robust speaker verification. Proceedings of Odyssey 2001: The Speaker Recognition Workshop, pp. 213–218.