

Kenny, P., Boulianne, G., Ouellet, P. and P. Dumouchel. "Speaker and session variability in GMM-based Speaker Verification" *IEEE Transactions on Audio Speech and Language Processing*, 15 (4) May 2007.

**Abstract**

We present a corpus-based approach to speaker verification in which maximum likelihood II criteria are used to train a large scale generative model of speaker and session variability which we call joint factor analysis. Enrolling a target speaker consists in calculating the posterior distribution of the hidden variables in the factor analysis model and verification tests are conducted using a new type of likelihood II ratio statistic. Using the NIST 1999 and 2000 speaker recognition evaluation data sets, we show that the effectiveness of this approach depends on the availability of a training corpus which is well matched with the evaluation set used for testing. Experiments on the NIST 1999 evaluation set using a mismatched corpus to train factor analysis models did not result in any improvement over standard methods but we found that, even with this type of mismatch, feature warping performs extremely well in conjunction with the factor analysis model and this enabled us to obtain very good results (equal error rates of about 6.2%).

**Keywords:** Speaker verification, Gaussian mixture, factor analysis.