

Kenny, P., Dehak, N., Dehak, R., Gupta, V. and P. Dumouchel. "The Role of the Speaker Factors in the NIST Extended Data Task" In *Proceedings of IEEE Odyssey 2008 - The Speaker and Language Recognition Workshop*. Stellenbosch, South Africa, January 21-25, 2008.

Abstract

We tested factor analysis models having various numbers of speaker factors on the core condition and the extended data condition of the 2006 NIST speaker recognition evaluation. In order to ensure strict disjointness between training and test sets, the factor analysis models were trained without using any of the data made available for the 2005 evaluation. The factor analysis training set consisted primarily of Switchboard data and so was to some degree mismatched with the 2006 test data (drawn from the Mixer collection). Consequently, our initial results were not as good as those submitted for the 2006 evaluation. However we found that we could compensate for this by a simple modification to our score normalization strategy, namely by using 1000 z-norm utterances in z_t -norm.

Our purpose in varying the number of speaker factors was to evaluate the eigenvoice MAP and classical MAP components of the inter-speaker variability model in factor analysis. We found that on the core condition (i.e. 2–3 minutes of enrollment data), only the eigenvoice MAP component plays a useful role. On the other hand, on the extended data condition (i.e. 15–20 minutes of enrollment data) both the classical MAP component and the eigenvoice component proved to be useful provided that the number of speaker factors was limited. Our best result on the extended data condition (all trials) was an equal error rate of 2.2% and a detection cost of 0.011.