

Dehak, N., Kenny, P. and P. Dumouchel. "Continuous Prosodic Features and Formant Modeling with Joint Factor Analysis for Speaker Verification" In *Proceedings of the 8th International Conference of Interspeech 2007*, pp. 1234-1237. Antwerp, Belgium, August 27-31, 2007.

Abstract

In this paper, we introduced the use of formants contours with prosodic contours based on pitch and energy for speaker recognition. These contours are modeled on continuous manners by using the Legendre polynomials on basic unit which represents syllables. The parameters extracted from the Legendre polynomials coefficients plus the syllables duration are modeled with Gaussian Mixture Models (GMM). Factor analysis is used to treat the speaker and channel variability. The results obtained on the core condition of NIST 2006 speaker recognition evaluation show that the use of formant with prosodic information gives an absolute improvement of approximately 3% on equal error rate (EER) compared with the results obtained by prosodic informations alone. However when the formants and the prosodic system scores are fused with a state of the art cepstral joint factor analysis system, we obtain equivalent results to the results obtained when we fused system based on prosodic features alone with the same cepstral joint factor analysis system. This fusion gives a relative improvement of 8.0% (all trials) and 12.0% (English only) on EER compared to cepstral system alone.

Keywords: Speaker recognition, prosodic features, formants, joint factor analysis