PERCEPTUAL TIME–VARYING MODELLING OF SPEECH SIGNALS FOR ASR COMPRESSION APPLICATION (MonAmOR3)

✱ Author(s) :

Amir Leibman   (Ben–Gurion University, Israel)

Ilan D. Shallom   (Ben–Gurion University, Israel)

✱ Abstract :

Perceptual audio coders and Automatic Speech Recognition (ASR) systems are commonly based on short–time analysis. This paper presents a generalized model for time–varying coefficients based on psychoacoustic properties of the human ear. The proposed model is evaluated in the framework of speaker independent speech recognition using Hidden Markov Models (HMM). The generalized model is compared to the traditional most popular MFCC. The comparison is made with respect to the models baud rate and the total error rate measured in an extensive Speech recognition experiment. The recognition based on the well established speech recognition development environment, the HTK and using the TIDIGIT as the evaluation database. The time varying model achieves better recognition rate in comparison to MFCC, while the proposed model baud rate is about one third of the baud rate that is used in the case of MFCC. In addition, a preliminary evaluation of the model robustness to noise was carried out and is presented.

Menu