

Abstract

Mohamed S El-Mahallawy

Text-Independent Egyptian Colloquial Speaker Recognition based on Hidden Markov Model and Sparse Coding

Hidden Markov Model (HMM) is one of the most popular techniques for speech and speaker recognition, while the Sparse Coding (SC) is widely used in face recognition and has not been used widely in speaker recognition. In this paper a comparison between the performances of Sparse Coding and Hidden Markov Model techniques is done on a text-independent speaker recognition task. Speaker recognition is applied on a closed set of 54 speakers speaking short sentences of the Egyptian Colloquial Arabic (ECA). An ergodic HMM is used with Mel Frequency Cepstral Coefficients (MFCCs) features. Sparse Coding is done on the same data the used Sparse Coding classifiers are the non-negative least square (NNLS) and the linear regression classifier (LRC). The result of the comparison is that the Sparse Coding outperforms the Hidden Markov Model, in particular, when the LRC classifier is used.