A Novel Feature Extraction Technique for Speaker Identification

Abstract

This paper presents a novel feature extraction approach for speaker identification when the speech is corrupted by additive noise. The environmental mismatch between training and
testing data degrades the performance of speaker identification system. The performance degradation is primarily due to presence of background noise when try to match a given speaker to the set of known speakers in a database. Mel frequency cepstral coefficients (MFCCs) are perhaps the most widely used front ends in the state of the art speaker identification systems. One of the major issues with MFCCs is that they are very sensitive to additive noise. To overcome this bottleneck, a temporal filtering procedure on the autocorrelation sequence is proposed to minimize the effect of additive noise. The proposed feature is called Relative Autocorrelation Mel Frequency Cepstral Coefficients (A-MFCC) which is derived based on filtering the temporal trajectories of short time one sided autocorrelation sequence. This filtering process minimizes the effect of additive noise. No prior knowledge of noise characteristics is required. The additive noise can be a colored noise. For speaker identification, Hindi database was constructed from the speech samples of each known speaker. Feature vectors (MFCCs and A-MFCCs) were extracted from the samples by short-term spectral analysis, and processed further by vector quantization for locating the clusters in the feature space. Experimental results indicated that A-MFCCs significantly improved the performance of speaker identification system in noisy environment.

Reference


Index Terms

Computer Science Pattern Recognition

Key words

Speaker identification vector quantization
relative autocorrelation