Abstract

Noise robustness is one of the most challenging problem in automatic speech recognition. The goal of robust feature extraction is to improve the performance of speech recognition in adverse conditions. The mel-scaled frequency cepstral coefficients (MFCCs) derived from Fourier transform and filter bank analysis are perhaps the most widely used front-ends in state-of-the-art speech recognition systems. One of the major issues with the MFCCs is that they are very sensitive to additive noise. To improve the robustness of speech front-ends we introduce, in this paper, a new set of MFCC vector which is estimated through three steps. First, the relative higher order autocorrelation coefficients are extracted. Then magnitude spectrum of the resultant speech signal is estimated through the fast Fourier transform (FFT) and it is differentiated with respect to frequency. Finally, the differentiated magnitude spectrum is transformed into MFCC-like coefficients. These are called MFCCs
extracted from Differentiated Relative Higher Order Autocorrelation Sequence Spectrum (DRHOASS). Speech recognition experiments for various tasks indicate that the new feature vector is more robust than traditional mel-scaled frequency cepstral coefficients (MFCCs) in additive noise conditions.

Reference


Index Terms

Computer Science Speech Recognition
Key words

MFCC  Autocorrelation domain  magnitude

spectrum