The performances of Speaker Identification Systems (SIS) are strongly influenced by the quality of the speech signal. Most of these systems are based on Gaussian Mixture Models (GMM) that is trained using a training speech database. The mismatch between the training conditions and the testing conditions has a deep impact on the accuracy of these systems and represents a barrier for their operation in real conditions generally affected by noises disturbances. The Voice Activity Detection (VAD) is a very useful technique for improving the performance of these systems.
systems working in these scenarios. In this paper we have used within the feature extraction process, a robust VAD module, that yield high speech/non-speech discrimination accuracy and improve the performance of the SIS in noisy environments. A set of experiments which we have conducted on our proper database containing 37 Arabic speaker in order to evaluate the performances of our SIS based on gammatone frequency cepstral coefficients (GFCC) front-end combined to VAD algorithm show 7.84% average improvement of Identification Rate (IR) performance of our SIS based on GFCC robust method compared to a baseline MFCC method. 2. 13% average improvement accuracy as a benefit of VAD technique is observed when the Signal per Roise Ratio (SNR) changes from 40 dB to 0dB.

References

Voice Activity Detection for Robust Speaker Identification System

- http://www.speech.kth.se/wavesurfer/

Index Terms

Computer Science  Pattern Recognition

Keywords

Gaussian Mixture Models (gmm)  Mel Frequency Cepstral Coefficients (mfcc)
Gammatone Frequency Cepstral Coefficients (gfcc)

Speaker Identification System (sis)

Voice Activity Detection (vad)