Speaker Recognition System for Telephone Lines

Shaikha Ali AL-Hajri

Abstract

Speaker recognition is a biometric modality that uses an individual's voice for the recognition purposes. The whole process relies on features influenced by both the physical structure of a person's vocal tract and the behavioural characteristics of individual speakers. A speaker recognition system typically consists of two main parts: feature extraction part and pattern matching part.

There are many applications that can make use of speaker recognition systems. Remote authentication using phone lines can be one of the popular applications for speaker recognition. The most crucial obstacle affecting the remote authentication through phone lines is the voice data clarity. Phone lines introduce random noise to speech data and hence making it difficult to correctly identify the speakers.

In this project, several experiments were conducted for the purpose of speaker identification; particularly for noisy data. The system was tested mainly with two publicly available speech databases: TIMIT and NTIMIT. TIMIT data was used twice; as a clear data and as a noisy data after adding fixed values of SNR. The techniques this project focuses on when it comes to feature extraction are MFCC and DCT-Cepstrum Histogram. In terms of pattern matching techniques, this project is concerned with GMM.

The speaker selection in this project was done in two manners: alphabetically and randomly. Moreover, detailed testing was done to obtain the optimal set of parameters in order to get the highest identification efficiency. That was done by fine-tuning three main parameters; number of gaussians, number of Iterations and number of the cepstral coefficients.

This thesis introduces the use of historical speech files in the speaker identification testing phase. Usage of one historical speech file achieved an identification efficiency of 85% with a population size of 400 speakers from the NTIMIT. Increasing the number of re-used speech files to two increased the efficiency for the same population size to 96%.