Abstract

Spoken alphabet recognition as one of the subsets of speech recognition and pattern recognition has many applications. Unfortunately, spoken alphabet recognition might not be a simple task due to highly confusable set of letters as presented in the English alphabets. The highly acoustic similarities that contribute to the confusability may hinder the accuracy of speech recognition systems. One of the confusable set is called the E-set letters which consist of the letters B, C, D, E, G, P, T, V and Z. In this study, we present an investigation of isolated alphabet speech recognition system using the Mel Frequency Cepstral Coefficients (MFCC) and Back-propagation Neural Network (BPNN) for the E-set and for all the 26 English alphabets. Learning rates and momentum rates of the BPNN are adjusted and varied in order to achieve the best recognition rate for the E-set and all the 26 alphabets. By adjusting these parameters, we managed to achieve 62.28% and 70.49% recognition rate for E-set recognition under speaker-independent and speaker-dependent conditions respectively.

References

- M. Karnjanadecha and S. A. Zahorian, "Signal Modeling for Isolated Word
Spoken English Alphabet Recognition with Mel Frequency Cepstral Coefficients and Back Propagation Neural Networks


- K. Daqrouq, "Wavelet Entropy and Neural Network for Text-Independent Speaker Identification," Engineering Applications of Artificial Intelligence, vol. 24, pp. 796-802,
2011.

Index Terms
Computer Science
Artificial Intelligence

Keywords
Mel-frequency Cepstral Coefficients  Mfcc  Error Back-propagation Neural Network  E-set