Abstract

The aim of this paper is to extract and select features from speech signal that will make it possible to have acceptable speaker recognition rate in real life. A variety of combinations among formants (F1, F2, F3), Linear Predictive Coefficients (LPC), Mel Frequency Cepstral Coefficients (MFCC) and delta-Mel Frequency Cepstral Coefficients representing features are considered and their effect in speaker recognition is observed. Two similar volume data sets with differed string (words) are considered in the present study. These two data sets are prepared taking into account two differed data sampling rates. The study reveals another interesting fact that the selection of strings in speaker enrollment process is a matter of importance for accurate result. This means that the speaker will be tested for authentication with the same string with which he was enrolled earlier during the time of his first access to the system.

References

- Adjoudj Reda, Boukelif Aoued, "Artificial Neural Network & Mel-Frequency Cepstrum Coefficients-Based Speaker Recognition", 3rd International Conference:
Feature Selection Method for Speaker Recognition using Neural Network

Sciences of Electronic, Technologies of Information and Telecommunications--TUNISIA, March 27-31, 2005


- Talukdar, P. H., Bhattacharjee, U., Goswami, C. K., Barman, J., &quot;Cepstral Measure of Boro Vowels through LPC-Analysis," Journal of the CSI, Vol. 34 No 1, Jan – Mar, 2004


- Braman, J., Kalita, S., Talukdar, P. H., &quot;Features extraction of bodo vowels through lpc-analysis," Proceedings of Frontiers of Research on Speech and Music (FRMS-2004), 2004


Index Terms

Computer Science  Artificial Intelligence

Keywords

Feature Extraction  Feed Forward Neural Network  Speaker Recognition