Abstract

Speech has much capability as an interface between human and computer which comes under the Human Computer interaction (HCI). The major challenge has been the nature of voice is ever varying speech signal.

The paper presents the development of the speech recognition system using Swahili speech database which was collected in three sets: digits, isolated words and sentences from both native and non native speakers of Swahili language.

Different feature extraction techniques deployed in the system are: Linear Prediction Coding (LPC) and Mel-Frequency Coefficients (MFCC). We have used the 12 coefficient features from MFCC and 20 coefficients features from LPC. All these features extracted techniques are applied and tested for the own developed Swahili speech database.
Recognition and verification were done using confusion matrix and Support Vector Machine (SVM) as a classifier for the classification purpose. LDA was tested for the entire dataset for the dimension reduction. LDA gave a good clustering. The performance of the system was checked on basis of their accuracy; Confusion with MFCC 50.9%, confusion with LPC 50.1%, the higher recognition rate in each data set were as follows numeric data: MFCC: 75%, LPC: 72%, isolated word data: MFCC: 65.2% LPC: 66.67%, sentence data MFCC: 63.8%, LPC: 59.6.

References

2. Dat Tat Tran, Fuzzy “Approaches to Speech and Speaker Recognition”, A thesis submitted for the degree of Doctor of Philosophy of the university of Canberra.
12. Volume 248 of the series Advances in Intelligent Systems and Computing pp 21-29

Index Terms

Computer Science
Signal Processing
Keywords

Swahili, Swahili Text corpus, Phonetics, Text Corpus and Speech Corpus, Automatic Speech Recognition