Abstract

This study is intended to develop a noise robust distributed speech recognizer for real-world applications by employing Cepstral Mean Normalization (CMN) for robust feature extraction. The main focus of the work is to cope with different noisy environments. To realize this objective, Mel-LP based speech analysis has been used in speech coding on the linear frequency scale by applying a first-order all-pass filter instead of a unit delay. Mismatch between training and test phases is reduced through robust feature extraction by applying CMN on Mel-LP cepstral coefficients as an effort to reduce additive noise and channel distortion. The performance of the proposed system has been evaluated on test set A of Aurora-2 database which is a subset of TIDigits database contaminated by additive noises and channel effects. The experiment is conducted on four different noisy environments and the baseline
performance, that is, for Mel-LPC the average word accuracy has found to be 59.05%. By applying the CMN on Mel-LP cepstral coefficients, the performance has been improved to 68.02%. It is found that CMN performs significantly better for different noisy environments.

References

- Itakura, F. and S. Saito, 1968. Analysis synthesis telephony based upon the maximum

Index Terms

Computer Science  Speech Recognition
Keywords
Mel-LPC  bilinear transformation  CMN  Aurora 2 database