Abstract

This paper considers the problem of acoustic mismatch caused by use of different sensors, in digital gazettes and hand-held devices. In this paper, two complementary features derived from conventional cepstral features are proposed, namely linear/mel spectral subband features (L/M-SSC) and log filter bank energy features (LFBE). The performance of these complementary features is compared with conventional features in acoustic mismatch conditions. To investigate the performance of features alone, all processing and classification steps are kept constant to allow a controlled comparison. A multi-variability speech database (IITG-MV) with acoustic mismatch (different microphones) is used for experimental evaluation. It is observed that all these features show almost equal performance for text independent speaker identification in same acoustic condition. Whereas in mismatch condition, spectral subband centroids (L/M-SSC) features proved to be robust than other features when used alone. Further, use of dynamic features along with channel and noise compensation enhances the percentage identification rate of the system for all cases of acoustic mismatch, with spectral subband centroid features showing comparable performance to that of conventional features.
References


18. Electro Medical and Speech Technology Laboratory, Department of Electronics and
Electrical Engineering, Indian Institute of Technology Guwahati.
http://www.iitg.ernet.in/ece/emstlab/


Index Terms

Computer Science

Signal Processing

Keywords

MFCC, LFCC, Linear/Mel scale spectral subband centroids (L/M-SSC), Log filter bank energy (LFBE)