Abstract

Language Identification is the task of identifying language spoken from unknown user. The main objective is to achieve accurate results in shortest speech segments by using automatic Language Identification system. It works on language classification that involves new language rapid learning identities and reduce the computational complexity. MFCC, GFCC, PLP and the combination of these feature are consider in language identification system. The proposed approach that transforms the spoken words to a represent low dimensional i-vector, on which classification techniques are applied. Feature extraction is done on input audio, Universal background model and i-vector extraction are used in proposed system in order to meet the challenges involved in rapidly making reliable decisions about the spoken language such as Marathi, Hindi and English. For the relevant languages under the different acoustic condition are used to capture robust feature extraction scheme.

References
1. Y.K. Muthusamy, “A Segment Approach to Automatic language identification uses the telephone speech corpus” 1987

Index Terms

Computer Science Pattern Recognition
Keywords

Language Identification, Feature Extraction, Universal background model.