



INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
SHORT ABSTRACT OF THESIS

Name of the Student : BANRISKHEM K. KHONGLAH
Roll Number : 126102033
Programme of Study : Ph.D.
Thesis Title: Speech Knowledge based Broadcast Audio Classification and Phone Recognition
Name of Thesis Supervisor(s) : Prof. S. R. M. Prasanna
Thesis Submitted to the Department/ Center : EEE
Date of completion of Thesis Viva-Voce Exam : 21-OCT-2017
Key words for description of Thesis Work : speech-specific, classification, speech enhancement, phone recognition

SHORT ABSTRACT

In this work an alternate approach for automatic transcription of the anchor speakers' speech in broadcast audio for Indian News Channels is proposed. Some preprocessing methods are required in order to take care of the presence of scenarios like speech with background music and pure music which correspond to the news headlines and voice-over, in addition to the clean speech. The first preprocessing task is the speech versus music classification module which involves the use of speech specific features for classification. The speech output may have some of the segments containing background music, due to the speech specific nature of the features. These segments are then passed through clean speech versus speech with background music classification module which involves the use of the features based on the average and relative characteristics of the vocal tract system. The speech with background music is enhanced using the temporal, spectral processing and perceptual methods where the source information is mostly exploited to obtain the enhanced speech. Finally the clean and enhanced speech segments are passed through the phone recognition system and the final output will be the transcription of clean and enhanced speech, with improved accuracy compared to directly passing the broadcast audio through the phone recognition system.