



INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
SHORT ABSTRACT OF THESIS

Name of the Student : Sunil Y.

Roll Number : 08610211

Programme of Study : Ph.D.

Thesis Title: IMPROVING CHILDREN'S SPEECH RECOGNITION UNDER MISMATCHED CONDITION USING ARTIFICIAL BAND WIDTH EXTENSION

Name of Thesis Supervisor(s) : Prof. Rohit Sinha
Prof. S. R. Mahadeva Prasanna

Thesis Submitted to the Department/ Center : Department of Electronics and Electrical Engineering

Date of completion of Thesis Viva-Voce Exam : 14th June 2016

Key words for description of Thesis Work : children's speech, adults' speech, narrowband, wideband, ABWE, mismatched ASR, MFCC based ABWE, Sparse representation based ABWE.

SHORT ABSTRACT

Children's speech production system distinguishes itself from the adults' by shorter vocal tract length and higher pitch value. Due to shorter vocal tract length, formant frequency values shift to higherband (3400-8000 Hz) region. The higher pitch value results in relatively more fluctuations in the spectrum compared to adults. Narrowband (NB, 300-3400 Hz) automatic speech recognition (ASR) performance of children's speech degrades significantly due to loss of information in higher band. This work develops artificial bandwidth extension (ABWE) methods that restore higher band spectral information. The ASR is a connected digit recognition task which has models trained using adults' speech and tested using children's speech, termed as mismatched condition. The ABWE methods using class-specific, age-specific and delta features are developed and used in the children's speech recognition under mismatched condition. All of them show improvement in performance. A computationally efficient architecture for mel frequency cepstral coefficients (MFCC) based ABWE for ASR is developed that avoids vocoder framework for bandwidth extension. In the proposed method, the narrowband MFCC is directly converted into wideband MFCC thus avoiding the synthesis process. Sparse representation based ABWE (SR-ABWE) algorithm is proposed using coupled dictionaries. To further enhance SR-ABWE, least square transformation has been developed to estimate wideband codes from NB interpolated codes. Existing semi-coupled dictionary learning (SCDL) method has been explored for ABWE (SC-ABWE). An improvement in the performance of SC-ABWE is observed in terms of objective quality measures. The significance of SR-ABWE is also demonstrated in children's ASR.