



INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
SHORT ABSTRACT OF THESIS

Name of the Student : Nagaraj Adiga
Roll Number : 11610235
Programme of Study : Ph.D.
Thesis Title : Glottal Activity Region based Processing for Speech Synthesis
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 : 15/05/2017
Key words for description of Thesis Work : Glottal activity region, Riesz transform, 2-D processing, integrated linear prediction residual, statistical parametric speech synthesis.

SHORT ABSTRACT

Statistical parametric speech synthesis (SPSS) is the mostly preferred synthesizer compared to concatenative synthesis system, due to small footprint and flexibility. However, the naturalness and intelligibility of SPSS are still lagging behind the concatenative synthesis system. In this thesis, glottal activity region based processing for speech synthesis is proposed to improve the quality of speech. Glottal activity regions are perceptually important and constitute the majority of speech sounds. The major contributions of the present thesis are (i) Glottal activity region detection using features like strength of excitation, normalized autocorrelation peak strength, and higher order statistics. (ii) Vocal-tract smoothed spectral envelope computation by applying Riesz transform in the 2-D domain. (iii) Source model is designed with representation for aperiodic and phase components using integrated LP residual. (iv) The combination of suprasegmental, system, and source features for modeling together in SPSS to improve the prosody, naturalness, and intelligibility of SPSS.