INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
SHORT ABSTRACT OF THESIS

Name of the Student : PARABATTINA BHAGATH

Roll Number : 146101017

Programme of Study : Ph.D.

Thesis Title: STRUCTURAL PROCESSING METHODS FOR SPEECH SIGNAL ANALYSIS

Name of Thesis Supervisor(s) : PROF. P. K. DAS

Thesis Submitted to the Department/ Center : CSE

Date of completion of Thesis Viva-Voce Exam : 09.10.2020

Key words for description of Thesis Work : Phoneme Segmentation, Graph representation of speech, Tree Structures, Geometrical features of speech

SHORT ABSTRACT

Speech signal analysis is a crucial study that helps to develop methods for problems like phoneme segmentation, speech recognition, speaker verification, etc. There are various frameworks and techniques that support these problems. Frameworks like Hidden Markov Modeling and Deep Learning are popular. The frameworks are efficient with large data sets where intensive training is possible. However, this becomes challenging in case of under-resourced language since sufficient data cannot be provided for the intensive training.

To address the needs of these languages, suitable methods are required with the capability to seek for significant clues with less amount of data. Structural processing methods focus on understanding the signals differently compared to signal processing methods. In this approach, a signal is treated as an image rather than a time series with different samples at different time stamps. The need for these methods arises due to the limitations in Hidden Markov Models. HMM contains states in which each state depends on at most two neighboring states. This limit HMM to have a holistic view of the entire signal.

Recent developments in graph signal processing techniques give a way to analyze the signals by using graph data structures. These methods enable to use combination of temporal relations and frequency components while modeling the signals. The thesis addresses the problems of speech characterization and segmentation while considering the above-mentioned issues. Different features like trajectories and Tree structures are proposed and found to be useful for modeling speech signals that can be used further for recognition. Three different features based on trajectories, graph structures and fractals are proposed for segmentation task. The experiments were conducted on Indian accented spoken English vowels, words and TIMIT sentence data. Tree structures and trajectories were found to be useful in characterizing vowels and words, respectively. In the phoneme segmentation experiments, words data were collected from people belonging to different regions of India. The segmentation approaches are ascertained to be appropriate for finding phoneme boundaries of phonetic units in spoken words and sentences. The algorithms and obtained results are discussed in the thesis.

Abstract-TH-2501_146101017