Public Abstract
First Name:Rusheng
Middle Name:
Last Name:Hu
Adviser's First Name:Yunxin
Adviser's Last Name:Zhao
Co-Adviser's First Name:
Co-Adviser's Last Name:
Graduation Term:FS 2006
Department:Computer Science
Degree:PhD
Title:Statistical Optimization of Acoustic Models for Large Vocabulary Speech Recognition

This dissertation investigates statistical optimization of acoustic models in large vocabulary speech recognition. Two new optimization methods are proposed for phonetic decision tree (PDT) search and Hidden Markov Modeling (HMM)—the knowledge-based adaptive PDT algorithm and the HMM gradient boosting algorithm.

In speech recognition, phonetic decision trees are used to map logical context-dependent phones to acoustic features. The knowledge-based PDT algorithm assumes that this mapping logic can be modeled through implicit usage of the linguistic information extracted from a large data set. A computational efficient algorithm is developed and evaluated on the Telehealth conversational speech recognition task.

Another important part of acoustic modeling is the modeling of speech dynamics, which is typically modeled by Hidden Markov Models. In conventional model estimation approaches, there are difficulties in determining the optimal model structure, which is also known as the model selection problem. The proposed HMM gradient boosting method is based on a function approximation scheme from the perspective of optimization in function space rather than the parameter space, and hence provides a new scheme which can jointly optimize model structure and parameters. This approach is evaluated on the World Street Journal (WSJ) task.

Experimental results have confirmed that both algorithms proposed in here outperformed existing approaches.