Segment-Based Acoustic Models with Multi-level Search Algorithms for Continuous Speech Recognition

Mari Ostendorf J. Robin Rohlicek
Boston University BBN Inc.

Objective:

The goal of this project is to develop improved acoustic models for speaker-independent recognition of continuous speech, together with efficient search algorithms appropriate for use with these models. The current work on acoustic modelling is focussed on stochastic, segment-based models that capture the time correlation of a sequence of observations (feature vectors) that correspond to a phoneme. Since the use of segment models is computationally complex, we will also investigate multi-level, iterative algorithms to achieve a more efficient search. Furthermore, these algorithms will provide a formalism for incorporating higher-order information. This research is jointly sponsored by DARPA and NSF.

Summary of Accomplishments:

- Investigated the effect on recognition performance of different parameters of the segment model, particularly as a function of training.
- Developed a new approach to modeling time correlation in the context of the segment model.
- Developed new fast approaches to phone recognition, motivated by algorithms from image segmentation, which offers significant computational advantages over dynamic programming.
- Developed a reformulation of hidden Markov models which allows for more general mappings from the acoustic features to state likelihoods.
- Achieved recognition results on the TIMIT database using context-independent models which are comparable to those reported by others using context-dependent models.

Plans:

- Investigate dynamical system models for representing time correlation and context-dependence in the segment model.
- Extend current results to use segmental features and context-dependent models.
- Investigate mechanisms for integrating segment algorithms with the BBN Byblos recognition system.
- Investigate global constraints and features using multi-level algorithms.