

Ensemble Methods in Large Vocabulary Continuous Speech Recognition

Xin Chen

Dr. Yunxin Zhao, Thesis supervisor

Abstract

Combining a group of classifiers and therefore improving the overall classification performance is a young and promising direction in Large Vocabulary Continuous Speech Recognition (LVCSR). Previous works on acoustic modeling of speech signals such as Random Forests (RFs) or Phonetic Decision Trees (PDTs) has produced significant improvements in word recognition accuracy. In this thesis, several new ensemble approaches are proposed for LVCSR and experimental evaluations have shown absolute accuracy gains up to 2.3% over the conventional PDT-based acoustic models in our telehealth conversational speech recognition task.

The word accuracy performance improvement achieved in this thesis work is significant and the techniques have been integrated in the telemedicine automatic captioning system developed by the SLIPL group of the University of Missouri – Columbia.