Robust Speech Recognition Technology Program Summary

Clifford J. Weinstein and Douglas B. Paul, Principal Investigators
Lincoln Laboratory, M.I.T.
Lexington, MA 02173-9108

Program Goals

The major objective of this program is to develop and demonstrate robust, high-performance continuous speech recognizer (CSR) techniques and systems focused on application in spoken language systems (SLS). A key supporting objective is to develop techniques for integration of CSR and natural language processing (NLP) systems in SLS applications. The CSR techniques are based on a continuous-observation Hidden Markov Model (HMM) approach. Efforts are focused on improved HMM training and recognition for high performance and robustness in advanced SLS environments which include variabilities due to spontaneous speech, noise, and task-induced stress. Robustness is also being addressed through a new effort in comparison and development of auditory model front ends for HMM recognizers. The effort in CSR/NLP integration is focused on development of a structured CSR/NLP interface, which will allow effective collaboration with and between other groups developing NLP and/or CSR systems.

Background

The Lincoln program began with a focus on improving speaker stress robustness for the fighter aircraft environment. A robust HMM isolated-word recognition (IWR) system was developed with 99% accuracy under stress conditions, representing more than an order-of-magnitude reduction in error rate relative to a baseline HMM system. A robust CSR system was then developed and integrated into a voice-controlled flight simulator — a simple, but complete SLS involving a stressing, real-time task. The robust HMM recognition system was then adapted and extended to large vocabulary CSR. This effort has included development of a number of new modeling and recognition techniques, including parameter models based on tied Gaussian mixtures, which have resulted in state-of-the-art performance for both speaker-dependent (SD) and speaker-independent (SI) recognition on the DARPA Resource Management (RM) database.

Recent Accomplishments

Recent accomplishments include: (1) development of a new adaptive training strategy, improved semiphone models, and an improved duration model for the tied-mixture HMM recognizer; (2) development and application of stochastic bigram backoff language models to obtain new results on both the resource management (RM) and the Air Travel Information System (ATIS) tasks; (3) modification of the recognizer to work with bigram backoff language models; (4) conversion of a version of the HMM CSR to use a stack decoder controlled search with stochastic language models; and (5) development and implementation of a structured CSR/NLP interface system, with CSR and NLP simulators, and delivery of the interface system software to several other sites in the DARPA SLS community.

Plans

Plans for the current program include: (1) continued development of the tied-mixture HMM CSR system, including adaptive training techniques, mixture weight smoothing, and improved speaker-independent techniques; (2) further development of the stack-decoder-based HMM, for integration with the CSR/NLP interface system and with NLP systems developed at other sites; (3) a comprehensive comparison of auditory model front ends in quiet, noise, and stress conditions, with the goal of developing a more robust front end; and (4) participation in the design of a new CSR corpus with larger vocabulary and higher perplexity than the RM corpus.