

ROBUST CONTINUOUS 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 recognition (CSR) techniques focussed on applications in Spoken Language Systems (SLS). The effort focusses on developing advanced acoustic modelling, efficient search techniques, rapid enrollment, and adaptation techniques for robust large vocabulary CSR. An additional Lincoln goal is to define and develop application of robust CSR to military and civilian systems, and to expedite effective technology transfer.

BACKGROUND

The Lincoln program began with a focus on improving speaker stress robustness for the fighter aircraft environment. A robust hidden Markov model (HMM) system was developed with very high performance under stress conditions. The robust HMM techniques were then extended to yield state-of-the-art performance on the DARPA Resource Management corpus, using a tied-mixture HMM CSR approach.

Recent work has focussed on the large-vocabulary Wall Street Journal (WSJ) corpus, with vocabularies of 5K, 20K, and up to 64K words. The HMM CSR has been converted to a stack-decoder-based control strategy to operate efficiently with good performance in these tasks.

RECENT ACCOMPLISHMENTS

Recent accomplishments include: (1) development of the stack decoder and demonstration of its effectiveness on vocabularies up to 64K words; (2) development and integration of fast-match and detailed match; (3) further development of acoustic modelling techniques for the large vocabulary task; (4) a full set of evaluation tests in the November 1992 WSJ tests, including (e.g.) a 4.5% error rate on a 5K speaker-dependent test; (5) development of recognition-time speaker adaptation techniques with substantial improvements due to adaptation from both speaker-specific and speaker-independent initial models; (6) participation in and contributions to development of the WSJ corpus,

*THIS WORK WAS SPONSORED BY THE DEFENSE ADVANCED RESEARCH PROJECTS AGENCY. THE VIEWS EXPRESSED ARE THOSE OF THE AUTHOR AND DO NOT REFLECT THE OFFICIAL POLICY OF POSITION OF THE U.S. GOVERNMENT.

including providing baseline language models to all sites; (7) survey and study of opportunities for military and government applications of spoken language technology, and organization of a workshop focussing on technology transfer; and (8) continuing leadership of the DARPA spoken Language Coordinating Committee.

PLANS

Plans for the current program include: (1) development of advanced acoustic modelling techniques; (2) development and improvement of stack-decoder-based HMM for large vocabulary tasks, via development and integration of advanced acoustic models, acoustic fast match, and efficient search techniques; (3) development of technique for integration of stack-based CSR with natural language processors; (4) extension of run-time adaptation techniques to adapt acoustic parameters of the tied-mixture HMM to speaker channel, and environment; and (5) continued investigation of applications opportunities for spoken language systems.