Electrical Engineering Colloquium
Dallas Chapter of IEEE Signal Processing Society

Robust Speech Processing & Recognition for In-Vehicle Route Navigation Systems

Prof. John Hansen
University of Colorado at Boulder
IEEE Signal Processing Society Distinguished Lecturer

Tuesday, Feb. 8, 2005
ECSS 2.102 (TI Auditorium), 11:00am

In this talk, we discuss issues for next generation in-vehicle systems. We present recent advances in the formulation and development of speech recognition systems for in-vehicle hands-free route navigation. The system is comprised of a multi-microphone array processing front-end, environmental noise sniffer, robust feature and speech recognition system, and dialog manager and information servers. We also present our new CU-Move speech corpus for in-vehicle interactive speech systems for route planning and navigation collected across the U.S. The CU-Move system is focused on natural conversational interaction between the user and in-vehicle system. In this talk, we consider two main objectives: (i) robust PMVDR feature development and environmental analysis, and (ii) array processing schemes and how they can be used to improve speech quality and speech recognition performance. Traditional MFCC features have been used extensively in speech recognition. PMVDR features are based on a perceptually motivated estimate of the MVDR spectrum and shown to perform well in diverse environments. For robust speech recognition, the concept of environmental sniffing is proposed to seek out knowledge in the acoustic environment, organize this knowledge, and provide decision information to potential speech tasks that will improve performance due to environmental variability. This is a sharp departure from the traditional ROVER paradigm, since we seek to formalize the solution during application of the speech system task. Our results show that an environmental sniffing front-end with a single speech recognizer can outperform a parallel ROVER paradigm for speech recognition in car environments. Finally, we discuss several array processing schemes for speech enhancement and recognition in car environments. Here, we propose a novel combined fixed/adaptive beamforming solution (CFA-BF) that addresses source location calibration and target signal enhancement in real moving car environments. An evaluation using extensive actual car speech data from the CU-Move Corpus shows a 30% reduction in WER for speech recognition.

Biography - John Hansen received his PhD and MSEE from Georgia Institute of Technology, and BSEE from Rutgers University, all in Electrical Engineering. He is Professor in Dept. of Speech, Language and Hearing Sciences (SLHS), and Dept. of Electrical and Computer Engineering (ECE) at the Univ. of Colorado. He also serves as Dept. Chair of SLHS. He is co-founder of the Center for Spoken Language Research, and coordinator of the Robust Speech Processing Group at CSLR. He was a faculty member at Duke Univ., in the Dept. of Electrical Engineering and Dept. of Biomedical Engineering for 11 years. His research interests are focused on speech analysis and modeling, robust automatic recognition, and language technology. He previously served as the Technical Advisor to the U.S. Delegate for NATO, and organized ICSLP-2002, Inter. Conf. on Spoken Language Processing, Denver, Colorado. He presently serves as IEEE Distinguished Lecturer (2005), Associate Editor for IEEE Trans. Speech & Audio Processing (1993-1999), Associate Editor for IEEE Signal Processing Letters (1999-2000), Editorial Board for IEEE Signal Processing Magazine (2000-2002), IEEE Student Branch Coordinator (1988-99). He has supervised 30 PhD and MS thesis students over the past 10 years, and has published 190 journal papers, conference papers, textbooks in the field of speech processing and language technology.