

SPOKEN DOCUMENT RETRIEVAL FOR A NATIONAL GALLERY OF THE SPOKEN WORD

John H. L. Hansen

Center for Robust Speech Systems (CRSS)
Department of Electrical Engineering
Erik Jonsson School of Engineering and Computer Science
University of Texas at Dallas, U.S.A.
<http://crss.utdallas.edu/>

RESUMEN

The problem of reliable speech recognition for information retrieval is a challenging problem when data is recorded across different media and equipment. In this talk, we address the problem of audio stream phrase recognition for a new National Gallery of the Spoken Word (NGSW). This will be the first large-scale repository of its kind, consisting of speeches, news broadcasts, and recordings that are of significant historical content. An NSF initiative was recently established to provide better transition of library services to digital format. As part of this Phase-II Digital Libraries Initiative, researchers from Michigan State Univ. (MSU) and Center for Robust Speech Systems (Univ. of Texas, Dallas) have teamed to establish a fully searchable, online WWW database of spoken word collections that span the 20th Century. The database draws primarily from holdings of MSU's Vincent Voice Library which include +60,000 hours of recordings (from T. Edison's first cylinder recordings, to famous speeches such as man's first steps on the moon "One Small Step for Man", to American presidents over the past 100 years). In this partnership, MSU will house the NGSW collection, as well as digitize (with assistance from LDC), catalog, organize, and provide meta-tagging information. MSU is also responsible for a number of engineering challenges such as digital watermarking and effective compression strategies. The team at CRSS-UTD is responsible for developing the robust automatic speech recognition and segmentation for transcript generation, and for the proto-type online audio/metadata/transcript based user search engine (<http://SpeechFind.utdallas.edu>). Additional challenges for SpeechFind include dialect/accent tagging, speaker tracking, cross-language SDR applications, and environmental sniffing concepts.

Biografía: John H. L. Hansen

John Hansen received the PhD and MSEE from Georgia Institute of Technology, and BSEE from Rutgers University, all in Electrical Engineering. He is Department Chair and Professor in Dept. of Electrical Engineering, Erik Jonsson School of Engineering and Computer Science, and Professor in the School of Brain and Behavioral Sciences (Speech & Hearing) at the University of Texas at Dallas (UTD). He also holds the Distinguished Chair in Telecommunications Engineering at UTD. From 1999-2005, he was Professor in the Dept. of Speech, Language and Hearing Sciences (SLHS), and Dept. of Electrical and Computer Engineering (ECE) at the Univ. of Colorado. He also served as Dept. Chair of SLHS. He was 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 from 1988-1999. 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 served as general chairman/organizer of ICSLP-2002, Inter. Conf. on Spoken Language Processing, Denver, CO. He presently serves as IEEE Distinguished Lecturer (2005/06), and previously served as 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 33 PhD and MS thesis students over the past 17 years, and has authored/co-authored 222 journal papers, conference papers, textbooks in the field of speech processing and language technology.