AF03T006 TITLE: **Application of Cortical Processing Theory to Acoustical Analysis**

TECHNOLOGY AREAS: Human Systems

Objective: Develop a computational model of human auditory processing based upon cortical theory and data. Demonstrate its utility to evaluate and/or improve systems for speech communication and automated recognition of spectro-temporal patterns

Description: Current systems for audio processing and communication, including Automatic Speech Recognition (ASR) systems, are vastly inferior to human auditory performance. They fail to emulate human hearing in the general task of segregating and organizing dynamic acoustic inputs into multiple streams (e.g., multiple talkers). Human hearing is more tolerant of reverberant sound, better able to recognize regularities in timbre and temporal patterns, and less susceptible to interference from non-random background sound.

During the last decade, much has been learned about the physiology, organization, and functions of the auditory cortex. But computational approaches based upon this knowledge have yet to be fully exploited. A basic goal of this research is to formulate a computational model consistent with theory and data regarding cortical processes as well as with cochlear mechanics and early auditory stages. The model should be computationally and experimentally testable. It should provide a quantitative characterization of such "non-intensive" phenomena as the grouping and streaming of sound sequences, and the audibility of spectro-temporal patterns

This computational modeling approach should lead to significant and wide-ranging applications involving the analysis, synthesis, and identification of acoustic signals, medical and industrial diagnostics, speech recognition, and the design and characterization of communication channels. One particular interest is the enhancement of ASR performance in multitalker, noisy, or reverberant environments. Another is the development of improved metrics for speech intelligibility. Currently, the most widely used metrics – the Articulation Index , and the Speech Transmission Index – are based upon restrictive assumptions that require prior specification of the noise conditions affecting intelligibility. These applications will require strong collaboration to link auditory science with speech technology engineering.

Phase I: Design and develop a prototype computational model of human audition to include a representation of higher-order sensory/perceptual mechanisms associated with cortical processing. Demonstrate the model's ability to integrate key neurophysiological and psychoacoustical findings. Identify and define an approach for its application to technical problems in speech processing, auditory displays, or communication systems.

Phase II: Develop and validate the model. Demonstrate applications to technical problems in speech processing, auditory displays, or communication systems. Produce and demonstrate portable technology for these applications.

Dual Use Commercialization Potential: A computational model and the associated technology for Department of Defense applications is equally useful for commercial applications, which may include speech intelligibility assessment, virtual audio displays, ASR systems, medical and industrial diagnostics.

REFERENCES:

1. Hawkins, H. L. & McMullen, T. A. (Eds.) Auditory Computation. [Handbook of Auditory Research, Vol. 6, edited by A.N. Popper & R. R. Fay.] New York: Springer, 1995.

2. Gilkey, R. H. & Anderson, T. R. (Eds.) Binaural and Spatial Hearing in Real and Virtual Environments. Mawah, New Jersey: L. Erlbaum 1997. ISBN 0-8058-1654-2.

3. Greenberg, S. & Slaney, M. (Eds.) Computational Models of Auditory Function. NATO Science Series, IOS Press: Amsterdam, 2001. ISBN 90-5199-457-5

4. DeCharms, R. C., Blake, D. T., & Merzenich, M. M. (1998) Optimizing sound features for cortical neurons. Science, 280, 1439-1443.

5. Deprieux, D. A., et al. (2001) Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. J. Neurophysiol. 85, 1220-1234. [And see references therein.]

KEYWORDS: Computational Audition, Speech Intelligibility, Bioacoustics

TPOC: Dr. Willard Larkin
Phone: (703) 696-7793
Fax:
Email: willard.larkin@afosr.af.mil

Submit proposal (electronically) to Dept. of Defense between March 3 and April 16, 2003.

View DoD instructions on web: http://www.acq.osd.mil/sadbu/sbir/homepg.htm