Preprocessors for Noisy Speech

Institution: Signition Inc.; P.O. Box 1020, Los Alamos, NM 87544-1020

Principal Investigator: Dr. George Zweig; (505) 662-6855, zweig@earwig.lanl.gov

Objectives: The objective of this project is to develop a preprocessor for speech recognition systems operating in noisy environments. The preprocessor, consisting of a nonlinear inhomogeneous transmission line, will be realized in software, although realization in hardware in FY91 should be possible. More specifically we will:

1) Develop a nonlinear transmission line preprocessor that accurately simulates the mechanics of the mammalian inner ear at all sound pressure levels.

2) Preprocess speech with the nonlinear transmission line and show that there is a substantial improvement in the signal to noise ratio.

3) Assess the desirability and feasibility of implimenting either a digital or analog transmission line on a chip and using it as a preprocessor in the CMU, BBN, or MIT DARPA funded speech recognition systems.

Recent accomplishments: A speech preprocessor based on a nonlinear transmission line model of the cochlea[1] was proposed.[2] Sounds of different intensities, such as voiced and unvoiced speech, are analyzed in radically different ways. The Q's of the preprocessor's nonlinear filters vary with input amplitude, higher Q's (longer integration times) corresponding to quieter sounds. In addition, for quiet sounds, the preprocessor acts like a laser that traps and amplifies low level signals, thereby aiding in their detection and analysis.

Work planed for FY90: Analytical and numerical methods for solving the nonlinear transmission line equations will be developed.

References

1.   G. Zweig, "Cochlear mechanics," accepted for publication in *J. Acoust. Soc. Am.*, 1990.

2.   G. Zweig, "Auditory speech preprocessors," in *Proceedings of the DARPA Speech and Natural Language Workshop*, Morgan Kauffman, 1989.