

A Robust Preprocessor for Speech-Recognition Systems

George Zweig, Principal Investigator

Hearing Research Laboratory

Signition, Inc.

P.O. Box 1020

Los Alamos, NM 87544-1020

PROJECT GOALS

Signition is designing a preprocessor chip for speech-recognition systems operating in noisy environments. Current preprocessors, such as those based on moving Fourier transforms or linear predictive coding, are linear and not effective when speech is embedded in high levels of noise.

Specifically we are:

- Developing a nonlinear transmission line speech preprocessor that accurately simulates the mechanics of the inner ear over the entire range of sound pressure levels found in speech.

- Preprocessing speech with the nonlinear transmission line and demonstrating that there is a substantial improvement in the signal-to-noise ratio at the output taps of the line as compared with the output of conventional linear speech preprocessors.

- Assessing the feasibility of creating a digital transmission line on a chip to use as a preprocessor in the CMU, BBN, or MIT DARPA funded speech recognition systems.

RECENT RESULTS

- Deduced the mechanics of hearing at low sound pressure levels from measurements of basilar membrane motion. The mechanical properties of individual sections of the cochlea have been determined by examining their collective response to tones of different frequency. The conclusion is that each section acts like a harmonic oscillator with *negative* damping. The oscillator is controlled by negative feedback that drives the oscillator with a force proportional to the displacement of the oscillator at an earlier time. The time delay is approximately 1.75 times the oscillator's period. Thus, the inner ear is "active," creating sound that interferes "intelligently" with the incoming sound.

- Showed that the inner ear acts as a "subthreshold

laser" that traps and amplifies quiet sounds, thereby aiding in their detection and analysis. This system was modeled as nonlinear transmission line with active circuit elements and negative feedback.

- Developed a fourth-order Runge-Kutta method for solving the nonlinear transmission-line equation.

- Solved the nonlinear transmission-line equation for a variety of simple nonlinearities using clicks and pure tones as inputs.

- Determined the form of the nonlinearity by comparing solutions of the transmission-line equation with Rhode's corresponding data.

- Showed that two frequencies interact with one another by affecting each other's region of amplification. Such two-tone suppression, the auditory analogue of visual lateral inhibition, leads to formant sharpening.

- Developed a transmission-line parallel-processing algorithm for the Connection Machine.

- Studied the feasibility of creating digital VLSI chip preprocessors that emulate the nonlinear transmission line. Preliminary results are encouraging.

PLANS FOR THE COMING YEAR

- Solve the nonlinear transmission line equations for noisy speech and provide a preliminary qualitative assessment of the signal processing capabilities of the nonlinear transmission line.