A Microphone Array System for Speech Recognition

Harvey F. Silverman
Division of Engineering
Brown University
Providence, RI 02912

The ultimate speech recognizer cannot use an attached or desk-mounted microphone. Array techniques offer the opportunity to free a talker from microphone incumberance. My goal is to develop algorithms and systems for this purpose.

In the past year, we have studied the microphone-array placement problem and come up with some optimal placements for a linear microphone array. In so doing we have developed a new method for general nonlinear optimization which we call the Stochastic Region Contraction method. This allowed us to get optimal solutions to our problem -- globally optimal -- in far less time than simulated annealing would have taken.

We also built a first system for studying linear arrays. The hardware uses one TMS32025 per microphone channel and feeds our parallel processor, Armstrong. Using this facility, we are able to do both time and frequency-domain beam forming, and are able to gather real data from the multiple microphone sources. Currently, we have eight channels and are building another eight.

Current work and work in trhe immediate future includes much on the tracking algorithms. We are currently testing two on both synthetic and our real data. The first applies an interpolative correlation technique, to which a stage of "hyperbolic fit" is added. This fit is accomplished via gradient techniques. The second applies stochastic region contraction to maximize the power over a parameterized spectrum as well as the source location. In addition, we are hypothesizing new architectures and measuring the real effects the real environment.