PENNSTATE

Technology Available for Licensing

Office of Technology Management - The Pennsylvania State University
113 Technology Center, University Park, PA 16802
814.865.6277 phone; 814.865.3591 fax

Non-Confidential Description - PSU No. 4113
"A Novel System for Noise-Robust Speech Recognition Using Spiking Artificial Neurons"

Keywords/Field of Invention:

Spiking Artificial Neurons, Robust Speech Recognition, Speech Interfaces, Algorithms

Inventors:

Dezhe Jin, Phillip Schafer

Background

Automatic speech recognition (ASR) in noisy background environments is a major problem that currently prevents the widespread use of speech-based human-machine interfaces. Traditional ASR systems use hidden Markov models (HMMs) to encode regularities in the statistics of speech. However, since the statistics of speech can be strongly influenced by the acoustic environment, this approach is not robust in noisy environments, especially when the noise is not seen during the model training.

Invention Description

The subject invention represents a novel speech-encoding scheme based on the spiking responses of artificial neurons responding to brief acoustic features. The spikes are highly resistant to corruption by noise. The artificial neurons are trained with clean speech of a mixture of male and female speakers. The inventors then devised a template-based method for recognizing the spike sequences, which further enhances the system's ability to deal with temporal variability and noise corruption in speech. The system is tested on speech of mixed gender speakers corrupted with babble, car, subway and exhibition hall noise at SNRs of 20 dB through -5dB. The test speech and noise were not used during training of the system. The inventors compared their system with a state-of-the-art HMM-based system on an isolated digit recognition task using the AURORA-2 data set. The system performs similarly to the benchmark in clean and low-noise conditions, but gives a significant increase in performance at 0dB and -5dB, with mean recognition rates of 76% (cf. 60% benchmark) and 48% (cf. 25% benchmark), respectively.

Status of the Invention

The inventors continue to develop this technology under federal research funding to work with other data sets. When scaled to large vocabulary speech recognition tasks, the invention's system can be installed on mobile devices, computers, cars, airplanes and submarines to provide a platform for speech-based interfaces in arbitrary noisy acoustic environments. Potential commercial applications include cell phones, automobiles, military, civilian aviation as well as other communication apparatuses used in loud environments such as first responders.

Contact: **Matthew D. Smith**
Sr. Technology Licensing Officer
The Pennsylvania State University

Phone: (814) 865-6277
Direct: (814) 863-1122
E-mail: mds126@psu.edu