



**Fall 2009 CIS Colloquium Series**

# **Nonlinear Statistical Modeling of Speech**

**Joseph Picone**

**(Chair, Electrical and Computer Engineering, Temple University)**

*1-2pm, Wednesday, September 2, 2009*

*4<sup>th</sup> Floor Conference Room (Wachman Hall, CC 447)*

**Abstract:** Statistical or machine-learning techniques, such as Hidden Markov models and Gaussian mixture models, have dominated the signal processing and pattern recognition literature for the past 25 years. However, such approaches are prone to overfitting and have problems with generalization. For example, delivering high performance on previously unseen noise conditions remains an elusive goal.

In this presentation, we will review our recent work on applying principles of nonlinear statistical modeling to acoustic modeling in speech recognition. Our goal is to improve recognition performance in noisy environments. We will discuss the use of an extended feature vector containing features based on correlation dimension, correlation entropy and Lyapunov exponents. We will also introduce a new acoustic model based on a probabilistic mixture of autoregressive models.

Experimental results are presented on the Aurora IV large vocabulary speech recognition task in which audio data from a variety of actual noise conditions were digitally added to the standard Wall Street Journal 5K closed-vocabulary task. We will show modest gains in performance can be achieved under matched conditions, but performance degraded under mismatched training conditions.

**Bio:** Joseph Picone received his Ph.D. in Electrical Engineering in 1983 from the Illinois Institute of Technology. He is currently a Professor and Chair of the Department of Electrical and Computer Engineering at Temple University. His primary research interests are currently machine learning approaches to acoustic modeling in speech recognition. For over 25 years he has conducted research on many aspects of digital speech and signal processing. He has also been a long-term advocate of open source technology, delivering one of the first state-of-the-art open source speech recognition systems, and maintaining one of the more comprehensive web sites related to signal processing.

*Refreshments will be served!*