The Johns Hopkins University

Whiting School of Engineering

Department of Electrical and Computer Engineering

Computational Models of Representation and Plasticity in the Central Auditory System

Dissertation Defense by

Michael Carlin

Graduate Research Assistant

Electrical and Computer Engineering

Abstract

The performance in automated speech processing tasks like speech recognition and speech activity detection rapidly degrades in challenging acoustic conditions. It is therefore necessary to engineer systems that extract meaningful information from sound while exhibiting invariance to background noise, different speakers, and other disruptive channel conditions. In this thesis, we take a biomimetic approach to these problems, and explore computational strategies used by the central auditory system that underlie neural information extraction from sound.

In the first part of this thesis, we explore coding strategies employed by the central auditory system that yield neural responses that exhibit desirable noise robustness. We specifically demonstrate that a coding strategy based on sustained neural firings yields richly structured spectro-temporal receptive fields (STRFs) that reflect the structure and diversity of natural sounds. The emergent receptive fields are comparable to known physiological neuronal properties and can be employed as a signal processing strategy to improve noise invariance in a speech recognition task.

Next, we extend the model of sound encoding based on STRFs to incorporate the cognitive effects of selective attention. We propose a framework for modeling attention-driven plasticity that induces changes to receptive fields driven by task demands. We define a discriminative cost function whose optimization and solution reflects a biologically plausible strategy for STRF adaptation that helps listeners better attend to target sounds. Importantly, the adaptation patterns predicted by the framework have a close correspondence with known neurophysiological data. We next generalize the framework to act on the spectro-temporal dynamics of task-relevant stimuli, and make predictions for tasks that have yet to be experimentally measured. We argue that our generalization represents a form of object-based attention, which helps shed light on the current debate about auditory attentional mechanisms. Finally, we show how attention-modulated STRFs form a high-fidelity representation of the attended target, and we apply our results to obtain improvements in a speech activity detection task.

Overall, the results of this thesis improve our general understanding of central auditory processing, and our computational frameworks can be used to guide further studies in animal models. Furthermore, our models inspire signal processing strategies that are useful for automated speech and sound processing tasks.

Wednesday, February 11, 2015

10 a.m.

Latrobe 107

FOR DISABILITY INFORMATION CONTACT: Janel Johnson (410) 516-7031 janel.johnson@jhu.edu