Multiscale Processing of Complex Acoustic Scenes

Mounya Elhilali
Department of Electrical & Computer Engineering
Johns Hopkins University, Baltimore MD 21218

How our brain is able to effortlessly recognize sound objects in natural soundscapes is one of the major unsolved problems in neuroscience. Though crucial for both engineering and perceptual sciences, the study of the neural underpinnings of sound recognition and auditory scene analysis remains in its infancy. This field is particularly challenged by the lack of integrative theories which incorporate our knowledge of the cortical circuitry in the auditory pathway with the cognitive processes that shape behavior and perception of complex acoustic scenes.

Here, we focus on closing the gap between our knowledge of the perceptual bases of scene analysis with the neural mechanisms along various stages of the auditory pathway; particularly at the cortical level. While most current systems invoke operations akin to processes in the peripheral auditory system (e.g. cochlear spectral analysis, hair cell nonlinearities), they stop shy of incorporating promising capabilities of the central auditory system, most importantly its ability to *adapt* to the demands of an ever-changing acoustic environment and task goals. Recent physiological findings have been amending our views of processing in the auditory cortex; replacing the conventional view of 'static' processing in sensory cortex with a more 'active' and malleable mapping that rapidly adapts to the tasks at hand and listening conditions. This project aims at developing a neuro-computational model of auditory object identification in complex natural scenes. This effort integrates processes at different scales of processing; namely sensory information processing at the peripheral and cortical level at the single neuron and population level, along with cognitive executive functions with a particular focus on goal-directed attention and expectations from top-down prior knowledge. The proposed framework is a 'systems view' to modeling auditory scene analysis. It integrates processes at three scales: a micro-level mapping of complex sounds into a multidimensional cortical feature representation; a meso-level saliency and coherence analysis integrating activity across populations of cortical neurons; and macro-level cognitive feedback of attention and expectations that mediate auditory object formation.