Hearing engages in a seemingly effortless way, complex processes and transformations collectively known as auditory scene analysis, through which the auditory system consolidates acoustic information from the environment into perceptual and cognitive experiences. The proposed research seeks to develop a computational model to tackle this challenging phenomenon by closely collaborating with physiological and psychophysical experiments investigating the neural and perceptual mechanisms underlying auditory stream segregation. The intellectual merit of this work is to provide a mathematical framework for the design of new psychoacoustic and physiological experiments of streaming, and for developing effective algorithmic implementations to tackle the “cocktail party problem” in engineering applications.

The model focuses on rigorously testing the hypothesis that streaming of a complex sound from a cluttered acoustic environment can be quantitatively predicted based on its segregation in a higher multidimensional cortical representation that explicitly includes features related to spectral shape and temporal dynamics of the signal. The computational scheme explores techniques of unsupervised learning and the statistical theory of Kalman prediction. The model operates by reconciling acoustic evidence from the input signal (accentuating attributes such as timbre features, e.g. formant bandwidths and overall spectral tilts), and expectations of an internal representation of the different perceptual streams in the environment. This approach yields a robust computational scheme for speaker separation under conditions of speech and music interference; and emulates archetypal streaming percepts typically tested with human subjects.

Project (or PI) Website

http://www.isr.umd.edu/Labs/NSL/nsl.html
Publications


