

EESIP Seminar

Friday November 9th, 2012 • 11:00-12:00 • Interschool Lab, 7th floor Schapiro CEPSR

Hierarchical Spike Coding of Sound*

Yan Karklin

Laboratory for Computational Vision, New York University

Natural sounds, such as speech and music, exhibit complex statistical regularities at multiple scales. The underlying acoustic events are characterized by precise temporal and frequency relationships, but they can also vary substantially according to the pitch, duration, and other high-level properties of the sound. Learning this structure from data while capturing the inherent variability is an important first step in building auditory processing systems, as well as understanding the mechanisms of auditory perception. Here we develop Hierarchical Spike Coding, a two-layer probabilistic generative model for complex acoustic structure. The first layer consists of a sparse spiking representation that encodes the sound using kernels positioned precisely in time and frequency. Patterns in the positions of first layer spikes are learned from the data: on a coarse scale, statistical regularities are encoded by a second-layer spiking representation, while fine-scale structure is captured by recurrent interactions within the first layer. When fitted to speech data, the second layer acoustic features include harmonic stacks, sweeps, frequency modulations, and precise temporal onsets, which can be composed to represent complex acoustic events. Unlike spectrogram-based methods, the model gives a probability distribution over sound pressure waveforms. This allows us to use the second-layer representation to synthesize sounds directly, and to perform model-based denoising, on which we demonstrate a significant improvement over standard methods.

* Joint work with Chaitanya Ekanadham and Eero Simoncelli.

Yan Karklin received his Ph.D. in Computer Science from Carnegie Mellon University under the supervision of Mike Lewicki. At CMU he was also affiliated with the Center for Neural Basis of Perception. Since 2008 he has been a post-doctoral fellow at New York University and Howard Hughes Medical Institute, working with Eero Simoncelli. His interests lie in computational models of processing in visual cortex, natural image statistics, and hierarchical statistical modeling.



Host: Dan Ellis <dpwe@ee.columbia.edu>