

EESIP Seminar

Monday October 15th, 2012 • 14:00-15:00 • Interschool Lab, 7th floor Schapiro CEPSR

Non-negative Hidden Markov Modeling of Audio

Gautham Mysore

Adobe Advanced Technology Labs, San Francisco

Non-negative spectrogram factorization techniques have become quite popular in the last decade as they are effective in modeling the spectral structure of audio. They have been extensively used for applications such as source separation and denoising. These techniques however fail to account for non-stationarity and temporal dynamics, which are two important properties of audio. In this talk, I will introduce the non-negative hidden Markov model (N-HMM) and the non-negative factorial hidden Markov model (N-FHMM) to model single sound sources and sound mixtures respectively. They jointly model the spectral structure and temporal dynamics of sound sources, while accounting for non-stationarity. I will also discuss the application of these models to various applications such as source separation, denoising, and content based audio processing, showing why they yield improved performance when compared to non-negative spectrogram factorization techniques.



Gautham J. Mysore is a senior research scientist in the Advanced Technology Labs at Adobe, San Francisco. His research interests include machine learning and signal processing for various audio applications. He received an M.A. and Ph.D. from the Center for Computer Research in Music and Acoustics (CCRMA) at Stanford University in 2004 and 2010. He also received an M.S. in Electrical Engineering from Stanford University in 2008. In 2010, he was a visiting researcher at the the Gatsby Computational Neuroscience Unit at the University College London. He is currently a member of the IEEE technical committee on Audio and Acoustic Signal Processing.

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

EESIP Seminar Series: <http://labrosa.ee.columbia.edu/eesip>