Title: Machine Listening: from sound to meaning

Prof. Kyogu Lee

Music and Audio Research Group

Seoul National University, Republic of Korea

Abstract

Sound is just a series of vibrations of molecules of air (or other medium, such as water) generated by a vibrating object, but the human auditory system has the amazing ability to extract abstract meanings from sound. We can use voice signals to convey very high levels of information, and we can also communicate emotionally through music. However, there is a very large semantic gap between the speech or music signal and the abstract meaning conveyed by it, and it is very difficult to bridge this gap using computer algorithms. In this talk, I will describe how we as humans use our auditory system to perceive sound, and how to build a computational model for machine listening. I will also introduce a variety of machine listening applications, including source separation, speech enhancement, speech synthesis, and music generation.

Bio-sketch

Kyogu Lee received the B.S. degree in Electrical Engineering from Seoul National University, Seoul, Korea, in 1996, the M.M. degree in Music Technology from New York University, New York, in 2002, and the M.S. degree in Electrical Engineering and the Ph.D. degree in Computer-based Music Theory and Acoustics from Stanford University, Stanford, CA, in 2007 and 2008, respectively. He worked as a Senior Researcher in the Media Technology Lab at Gracenote from 2007 to 2009. He is now a professor at the Department of Intelligence and Information at Seoul National University, Seoul, Korea and is leading the Music and Audio Research Group (MARG). His research interests may be summarized as Machine Listening (a.k.a. Computer Audition), where signal processing and machine learning are utilized as two main instruments to better understand human auditory perception and cognition.