

EESIP Seminar

Monday Sep 16th, 2013 • 14:00-15:00 • EE Conference Room, 1300 Mudd

The Blame Game in meeting room ASR: An analysis of feature versus model errors in noisy and mismatched conditions

Dr. Steven Wegmann

ICSI, Berkeley

Given a test waveform, state-of-the-art ASR systems extract a sequence of MFCC features and decode them with a set of trained HMMs. When this test data is clean, and it matches the condition used for training the models, then there are few errors. While it is known that ASR systems are brittle in noisy or mismatched conditions, there has been little work in quantitatively attributing the errors to features or to models. In this talk we will investigate the sources of these errors in three conditions: (a) matched near-field, (b) matched far-field, and a (c) mismatched condition. We undertake a series of diagnostic analyses employing the bootstrap method to probe a meeting room ASR system. Results show that when the conditions are matched (even if they are far-field), the model errors dominate; however, in mismatched conditions features are neither invariant nor separable and this causes as many errors as the model does.

Joint work with Sree Hari Krishnan Parthasarathi, Shuo-Yiin Chang, Jordan Cohen, and Nelson Morgan.



Steven Wegmann has worked at industrial research laboratories on problems in speech processing since 1994, holding positions at Dragon Systems, Lernout & Hauspie, VoiceSignal Technologies, Nuance Communications, and Cisco Systems. He has been a staff researcher at ICSI since 2010 and began leading the Speech Group in 2013. His current research interests are in the areas of automatic speech recognition, diagnostic analysis, and low resource spoken term detection. Earlier in his career, he was a mathematician who specialized in algebraic topology. He obtained his doctorate in mathematics at the University of Warwick while he was a Marshall Scholar.

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

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