



Daniel Povey
Johns Hopkins University

Friday, October 12
2:30PM, Doherty Hall 2315

Subspace Gaussian Mixture Models for Speech Recognition

In speech recognition, the Subspace Gaussian Mixture model is a modeling framework based on the conventional HMM-GMM framework which reliably gives better results, especially when the amount of training data is limited. I will describe this framework, and talk about how these models are trained and how they are efficiently evaluated. The talk will also address some of the practical issues encountered in training them, and their advantages and disadvantages compared to the conventional Gaussian Mixture Model.

Bio: Daniel Povey completed his PhD at Cambridge University in 2003, and after spending just under ten years working for industry research labs (IBM Research and then Microsoft Research), joined Johns Hopkins University in 2012. His thesis work introduced several practical innovations for discriminative training of models for speech recognition, and made those techniques widely popular. At IBM Research he introduced feature-space discriminative training, which has become a common feature of state-of-the-art systems. He also devised the Subspace Gaussian Mixture Model-- a modeling technique which enhances the Gaussian Mixture Model framework by using subspace ideas similar to those used in speaker identification. At Microsoft Research and then at Johns Hopkins University, he has been creating a speech recognition toolkit "Kaldi", which aims to make state-of-the-art speech recognition techniques widely accessible.