



Interspeech 2018: Perspective Talk - 4

Sriram Ganapathy

Department of Electrical Engineering, Indian Institute of Science, Bangalore, India

Webpage

<http://leap.ee.iisc.ac.in/sriram/>



Title: Speaker and Language Recognition -- From Laboratory Technologies to the Wild

Abstract

Detecting the paralinguistic components of speech like speaker and language is of substantial interest for many commercial, surveillance and security applications. The problem is at least three decades old with some of the early techniques based on simple Gaussian mixture models. A significant advancement in this area came about a decade ago with the advent of joint factor analysis and i-vector models. The last couple of years have seen further breakthroughs with deep embeddings and end-to-end models based on deep learning. With these improvements in modeling speaker and language, the application of the technology has also moved from clean controlled speech data to telephone channel recordings, far-field microphones and more recently to multi-speaker conversations in the wild. In the talk, I will provide a prospective view of the broad research directions in the field of speaker and language recognition. I will also highlight some of the recent advancements from our work on hierarchical end-to-end approaches with relevance modeling.

Biography

Sriram Ganapathy is a faculty member at the Department of Electrical Engineering, Indian Institute of Science, Bangalore. Previously, he was a research staff member at the IBM Watson Research Center. He received his PhD from the Center of Language and Speech Processing, Johns Hopkins University. His research interests include signal processing and machine learning applied to speech recognition, speaker recognition and auditory neuroscience. He is a member of the ISCA and a senior member of the IEEE.