Interspeech 2018: Perspective Talk - 2

Bhuvana Ramabhadran
Google Inc., USA

Webpage
https://scholar.google.com/citations?user=jecEO0EAAAAJ&hl=en

Title: Open Problems in Speech Recognition

Abstract

In this talk, I will focus on the evolution of ideas in speech recognition over the last couple of decades, with emphasis on the key breakthroughs over the last ten years, its impact across spoken language processing in several languages, recent trends and open challenges that remain to be addressed. One such breakthrough is the use of several neural network model variants, which has had an enormous impact on the performance of state-of-the-art large vocabulary speech recognition systems. They have also had impact on keyword search which is the task of localizing an orthographic query in a speech corpus, and is typically performed through analysis of automatic speech recognition (ASR). Using the recently concluded IARPA funded Babel program as an example of a well-benchmarked task that focussed on the rapid development of speech recognition capability for keyword search in a previously unstudied language, I will present the successes and challenges that persist with limited amounts of transcription. Interpreting and understanding the hidden representations of various models remains a challenge today. I will also discuss current research taking advantage of such interpretations to improve robustness to noisy environments, speaker/domain adaptation algorithms, and dialects/accents. I will conclude with relevant metrics to measure speech recognition performance today that include and ignore the bigger picture of end to end user experience.

Biography

Bhuvana Ramabhadran (IEEE Fellow, 2017, ISCA Fellow 2017) currently leads a team of researchers at Google, focussing on multilingual speech recognition and synthesis. Previously, she was a Distinguished Research Staff Member and Manager in IBM Research AI, at the IBM T. J. Watson Research Center, Yorktown Heights, NY, USA, where she led a team of researchers in the Speech Technologies Group and coordinated activities across IBM's world-wide laboratories in the areas of speech recognition, synthesis, and spoken term detection. She was the elected Chair of the IEEE SLTC (2014–2016), Area Chair for ICASSP (2011–2018) and Interspeech (2012–2016), was on the editorial board of the IEEE Transactions on Audio, Speech, and Language Processing (2011–2015), and is currently an ISCA board member. She has served on the editorial board of T-ASLP (2012-2016), technical area chair for ICASSP (2011-2017), Interspeech (2012, 2014-2016), and was one of the lead organizers and technical chair of IEEE ASRU 2011, She has given tutorial and keynote presentations at several international conferences and served as an adjunct professor in Columbia University, where she co-taught a course in speech recognition. She has published over 150 papers and been granted over 40 U.S. patents. She was named a Master Inventor twice by IBM. She is a reviewer for ICASSP, Inerspeech, NAACL, ACL, EMNLP and serves on student dissertation committees and NSF review panels. Her research interests include speech recognition and synthesis algorithms, statistical modeling, signal processing, and machine learning.