Speech and Auditory Processing by Humans and Machines

GIAN Course @ IIT Dharwad (ONLINE)
11th - 22nd July, 2022

Prof. Hynek Hermansky, Director CLSP and Professor, Dept of Electrical Engineering, Johns Hopkins University, USA

Prof. S. R. Mahadeva Prasanna, Professor, Dept of Electrical Engineering, IIT Dharwad

....

Overview

Speech is a fundamental mode of human-human communication. Attempts have been made for several decades to use speech as a preferred mode of human-machine communication. In the past one decade there seems to be significant progress in terms of realizing this dream. The same may be attributed to the progress that has been made in the field of artificial intelligence, machine learning, and deep learning. However, whether it is the period before deep learning or currently in the deep learning period, one of the factors that can be attributed to the progress achieved towards human-machine communication are those due to the representations obtained for speech by exploiting the findings from auditory processing. These include perceptual linear prediction (PLP), RASTA, MFCC, TANDEM and so on. The current progress in deep learning may be exploited to discover many more such auditory processing motivated representations. The first step towards the same is to get hold on to the attempts that have been made for speech and auditory processing by humans and machines. The course is proposed to give an in-depth exposure to these attempts along with hands-on tutorials.

Objectives
1. Explaining different speech and perceptual processing methods by humans
2. Presenting different speech and perceptual processing methods by machines
3. Hands on tutorials on different speech and perceptual processing methods by machines
4. Automatic speech recognition using using perceptual processing methods
5. New directions for speech recognition

Modules	July 11 - July 22, 2022
Venue	IIT Dharwad, Dharwad - 580011, Karnataka
You Should Attend If…	- You are a student pursuing BTech/ MS/ MTech/ PhD from ECE/ CSE or equivalent and interested in knowing possibilities of machine learning and deep learning in speech processing. - You are a faculty from academic institutions pursuing research or interested in pursuing research in speech processing. - You are a person from an R & D organization or industry pursuing research and development or interested in speech processing.
Fees	The participation fees for taking the course is as follows: **Participants from abroad : US $500** **Industry/ Research Organizations: INR 5000/-** **Academic Institutions: INR 2500/-** **Students: INR 1000/-**

The Faculty
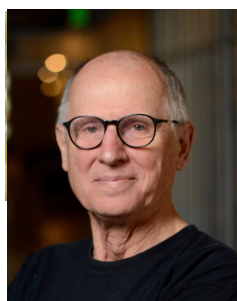
Hynek Hermansky (F'01, SM'92. M'83, SM'78) received the Dr. Eng. Degree from the University of Tokyo, and Dipl. Ing. Degree from Brno University of Technology, Czech Republic. He is the Julian S. Smith Professor of Electrical Engineering and the Director of the Center for Language and Speech Processing at Johns Hopkins University in Baltimore, Maryland. He is also a Research Professor at the Brno University of Technology, Czech Republic. He is a Life Fellow of the Institute of Electrical and Electronic Engineers (IEEE) IEEE, and a Fellow of the International Speech Communication Association (ISCA), was twice an elected Member of the Board of ISCA, a Distinguished Lecturer for ISCA and for IEEE, is the recipient of the 2013 ISCA Medal for Scientific Achievement and the 2020 IEEE James L. Flanagan Speech and Audio Processing Award.

S. R. Mahadeva Prasanna is currently Professor in the Dept. of Electrical Engineering, IIT Dharwad. He was faculty member in the Dept of Electronics and Electrical Engineering, IIT Guwahati from August 2004 to July 2017. He has supervised 21 PhD Theses in the speech processing area, and published about 250 research articles in the national and international journals and conferences. His areas of research interest include speech and handwriting processing.

Course Coordinator : Prof. S. R. M. Prasanna, Dept of Electrical Engineering, IIT Dharwad, Dharwad 580011, Karnataka, prasanna@iitdh.ac.in