

Cognitive Speech Processing

GIAN Course @ IIT Dharwad (ONLINE)

21st Feb. - 4th Mar., 2022

Prof. Hugo Leonardo Rufiner, Director and Professor in FICH-UNL, Santa Fe, Argentina

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

Overview

Speech is one of the most natural ways of communication between humans. Speech signals convey lots of information about the speaker, far from only words, like emotional and health state, identity, age, gender and even height, just to mention some of them. Speech is one of the most studied biosignals and a big effort was put into several years of research to better use it for Human-Computer Interfaces (HCI), although still being a work in progress. Another approach for HCI is to use speech-related biosignals different from the acoustic signal. This has several advantages like complementary discriminative information, acoustic noise robustness, privacy, and helps people who have difficulty with an oral mode of communication, etc. In the last few years, several researchers have tried to better understand the whole process of human speech generation and perception, with an emphasis on cognitive processing of speech. Human oral communication is a complex physiological process that involves the interaction and coordination of several body parts producing lots of signals that can be appropriately measured and analyzed. Examples of these types of signals related to the process of producing or perceiving speech are signals from muscle (EMG), brain activities (EEG, ECoG, fNIRS, MEG, etc), Electroglottogram (EGG) and video recording. Brain signals are especially interesting because they are also used for Brain Computer Interfaces (BCI). In this context, recent investigations proposed new BCI paradigms related to speech communication, like imagined speech. During the imagined speech, the subjects have to imagine pronouncing the word without moving muscles or producing sounds. Some research has been conducted on classifying vowels, syllables and complete words using EEG signals acquired during imagined speech with promising results for this task that can be named as “Speak What You Thought”. Another interesting research line achieved is recognizable speech reconstruction starting from a direct recording of brain signals from listeners, in a way that demonstrated the feasibility of developing systems that can “Hear What You Hear”. Important advances of the last decade in computational intelligence, like the development of deep learning architectures in conjunction with the availability of the great amount of experimental data, are responsible for reaching new levels of performance in such difficult tasks. The course is proposed to give an in-depth exposure to these issues along with hands-on sessions.

Objectives:

The main objectives of the course are the following:

1. Presenting human speech generation and perception, with an emphasis in cognitive processing of speech.
2. Explaining different techniques in speech and related biosignals processing, including computational intelligence and deep learning approaches.
3. Discussing speech related issues in brain computer interfaces.
4. Analyzing and using different biosignals open access databases.
5. Hands on tutorials on different speech related biosignals processing methods.
6. Showing applications across the design, implementation and management of intelligent bioinspired systems and deep neural networks techniques for speech and biosignals processing and recognition.
7. Presenting new directions and challenges for cognitive speech processing.

Course details (lectures)

- Introduction and course overview (1 h)
- Speech processing technologies (1 h)
- Human speech production and perception (3 h)
- Fundamentals of automatic speech recognition (A-ASR) (3 h)
- Basics of brain computer interfaces (BCI) (3 h)
- Neural Networks and Deep Learning (3 h)
- Intelligent and bioinspired acoustic ASR (2 h)
- Speech related biosignals (2 h)
- Speech in BCI (2 h)
- BCI Imagined speech (2 h)
- Biosignals Databases (2 h)
- Applications and new directions in cognitive processing (1 h)

Duration: Feb 21 – Mar 4, 2022 (10 days) : 25 hrs lectures and 20 hrs Tutorials / hands-on

Modules	Feb 21 - Mar 04, 2022 (ONLINE)
Venue	IIT Dharwad, Dharwad - 580011, Karnataka
You Should Attend If...	<ul style="list-style-type: none">• You are a student pursuing BTech/ MS/ MTech/ PhD from ECE/ CSE or equivalent and interested in knowing possibilities of cognitive 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 Participant from Industry/ Research Organizations: Rs. 5000/-

The Faculty



Prof. Hugo Leonardo Rufiner is currently Director and Professor in FICH-UNL, Argentina. He is also associated with CONICET as an Independent Researcher and Full Professor in the Department of Bioengineering, Universidad Nacional de Entre Ríos. His research interests include brain-computer interfaces, Automatic Speech Analysis and Recognition, Time-Frequency analysis and Mathematical Modeling of Biological Systems.



Prof. S. R. M. 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