# Project Topics

Pranav Sankhe
Indian Institute of Technology, Bombay

April 12, 2019

**Abstract**

In this document I have elaborated over the project topics I would like to work on at CCRMA. The central theme of the projects which are enumerated below is using EEG data to model I have ensured to be very specific and concise about the project description. Having said that, I would be more than happy to discuss them at length we need to.

## An Information Theoretical Perspective towards Audio Stimulus Reconstruction

We wish to model the brain and the EEG mesaurement device as a communication channel where the audio stimulus is the input and the output is the EEG voltage recorded from 128 strategically placed electrodes. We wish to model the entire 'transmission chain' which comprises of stimulus generation, brain processing by the human subject, and the EEG measurements as a nonlinear, time-varying communication channel with memory.
Having done so, we intend to comment on the nature of input, it's degree of reconstrutability, the channel capacity of the modelled channel and neural information transfer from the stimulus to the EEG data.
I have done some preliminary work in this area and have few interesting results which I am interpreting right now. There are still few major issues in modelling the channel and the validation of the results of the same which need to be addressed at this stage.

## Graph Signal Processing to indentify active subnetworks during musical engagement

There have been multiple attempts at analytical modelling of brain to enable us employ our mathematical tools for explaining variety of phenomeon. We choose to look at our brain as a graph. When we are listening to a music piece, there are certain regions of the brain who are involved in the process. We call these regions as subnetworks and hypothesize that they contribute to the overall functioning of the brain given a task. These individual subnetworks are also seen as graphs.

To estimate these underlying subnetworks, we implemente Independent Component Analysis(ICA) in graph paradigm. ICA is applied on the adjacency matrix corresponding to the brain and the independent components obtained are adjacency matrices corresponding to the subnetworks.
The adjacency matrix of the brain as a whole is computed using the EEG data which we have.

What we intend to acheive is that based on the subnetworks we identify, can we relate them with the areas which according to our understanding of brain, are indeed the regions who do the auditory processing. Since we have analytical representation of the regions of the brain which perform auditory processing, interesting interpretation and inferences can be made further.

## Reconstruction of audio waveforms from spike trains of artificial cochlea models

Spiking cochlea models describe the analog processing and spike generation process within the biological cochlea. Reconstructing the audio input from the artificial cochlea spikes is therefore useful for understanding the fidelity of the information preserved in the spikes. This system has been implemented for word recognition. The reconstructed audio under low signal-to-noise (SNR) conditions (SNR $\leq$ –5dB) gave a better classification performance than the original SNR input in this word recognition task [1].

## Connecting Deep Neural Networks to perception of Auditory Signals

What auditory representations contribute to the perception of music? We know from our understanding of human auditory system that it multiple auditory images(representions) are created which are used by our brain to process and makes sense of these signals.
The fact that there exist parallels between the architecture of a CNN and the brain structures from lower or higher cortical areas enables us to use CNNs to understand more about how brain understands audio. The data collected at JHU by Nicholas Huang and Elhilali [2], [3] has been compared with the activations of a CNN trained over audio. Specifically, the follwing audio descriptors have been compared with CNN activations:

- different layers of a CNN with acoustic features extracted directly from the scenes
- perceptual salience obtained from behavioral responses of human listeners
- neural oscillations recorded by electroencephalography (EEG) in response to the same natural scenes.

On the same lines, we follow the same procedure specifically for music. We can compare the CNN activations to musically relevant descriptors like tempo, music entraintment, etc.

## Speaker identity from cortical responses

Human poses the ability to selectively attend to a single speaker in the presence of background noise. The exact mechanism undelying this rather remarkable phenomenon is poorly understood and certainly an interesting research problem in it's own right. [4] N. Mesgarani Et al. have worked towards analyzing encoding of critical features of attended speech from cortical responses. They have designed a simple classifier trained solely on examples of single speakers can decode both attended words and speaker identity. Their ideas of spectrogram reconstruction algorithms can be used for our study.

# References


[1] Zai, Anja T. and Bhargava, Saurabh and Mesgarani, Nima and Liu, Shih-Chii, ”Reconstruction of audio waveforms from spike trains of artificial cochlea models” in Frontiers in Neuroscience, 2015.

[2] Huang, Nicholas and Elhilali, Mounya, ”Auditory salience using natural soundscapes” in The Journal of the Acoustical Society of America, 2017

[3] Snyder, Joel S. and Elhilali, Mounya, ”Recent advances in exploring the neural underpinnings of auditory scene perception” in Annals of the New York Academy of Sciences, 2017

[4] Mesgarani, Nima and Chang, Edward F., ”Selective cortical representation of attended speaker in multi-talker speech perception”, Nature 2012.