A Comparison of Regularization Methods in Forward and Backward Models for Auditory Attention Decoding

The decoding of selective auditory attention from noninvasive electroencephalogram (EEG) data is of interest in brain computer interface and auditory perception research. The current state-of-the-art approaches for decoding the attentional selection of listeners are based on linear mappings between features of sound streams and EEG responses (forward model), or vice versa (backward model). It has been shown that when the envelope of attended speech and EEG responses are used to derive such mapping functions, the model estimates can be used to discriminate between attended and unattended talkers. However, the predictive/reconstructive performance of the models is dependent on how the model parameters are estimated. There exist a number of model estimation methods that have been published, along with a variety of datasets. It is currently unclear if any of these methods perform better than others, as they have not yet been compared side by side on a single standardized dataset in a controlled fashion. Here, we present a comparative study of the ability of different estimation methods to classify attended speakers from multi-channel EEG data. The performance of the model estimation methods is evaluated using different performance metrics on a set of labeled EEG data from 18 subjects listening to mixtures of two speech streams. We find that when forward models predict the EEG from the attended audio, regularized models do not improve regression or classification accuracies. When backward models decode the attended speech from the EEG, regularization provides higher regression and classification accuracies.