Unsupervised decoding of matched versus mismatched EEG responses to speech

Nicolas Heintz*,† Tom Francart† Alexander Bertrand*
KU Leuven
*Dept. of Electrical Engineering (ESAT), STADIUS
Kasteelpark Arenberg 10, B-3001 Leuven, and
†Dept. of Neurosciences, ExpORL
Herestraat 49 bus 721, B-3000 Leuven
nicolas.heintz@kuleuven.be

Abstract

In recent years, there has been a rapid development of new algorithms to model how the human brain responds to attended speech based on electroencephalography (EEG) signals. Such models are not only interesting to study the brain, but can also be used to create objective markers of speech intelligibility, which is in high demand in audiology, or to create novel brain-computer interfaces such as neuro-steered hearing aids [1, 2]. We designed an unsupervised algorithm that adaptively trains such models, and is ideally suited for online, wearable devices.

The exact interaction between the attended speech and the recorded EEG signals is different for each person. Subject-specific models are therefore required for optimal performance, yet they require long and impractical training sessions for each subject. Subject-independent models avoid such individualised training sessions, yet perform significantly worse [2]. Moreover, if these models are used in real-life applications such as neuro-steered hearing aids, they should ideally also adapt to changes in, e.g., the neural response to speech, the electrode characteristics, the listening environment, etc. This is possible by updating the model on the fly, using the newly recorded incoming EEG and speech signals [2]. Crucially, the model should only be updated when a person actually hears and attends to the speech, to avoid updating the model with irrelevant, noisy signals.

These models thus require a classifier that can detect when the person is actively listening to that speech, and when this is not the case. We model this case by detecting whether a given EEG segment is a response to a given speech segment, i.e., whether both are 'matched' or not [1]. However, such a classifier also needs to model the interaction between EEG signals and attended speech, and should thus also be updated adaptively, creating a chicken-and-egg problem. Inspired by [2], we designed a new, unsupervised algorithm that first correlates the recorded EEG to the speech with a canonical correlation analysis (CCA)-based model, then uses the correlations to perform the match-vs-mismatch classification and finally only updates the CCA-based model when this is the case. This iterative model quickly converges to an equilibrium operating point in which it only performs about 1% worse than the corresponding supervised model that has access to the match-vs-mismatch labels.

References

[1] A. de Cheveigné, M. Slaney, S. A. Fuglsang, and J. Hjortkjaer, "Auditory Stimulus-response Modeling with a Match-Mismatch Task," *Journal of Neural Engineering*, 2021.

[2] S. Geirnaert, T. Francart, and A. Bertrand, "Unsupervised Self-Adaptive Auditory Attention Decoding," *IEEE Journal of Biomedical and Health Informatics*, pp. 1–1, 2021.