Optimizing Stimulus Presentation for a Spatial Auditory P300 Speller

Nicholas Boldt; Shadi Sartipi
Electrical and Computer Engineering Department, University of Rochester

Abstract

Recently, a brain-computer interface (BCI) using virtual sound sources has been proposed for estimating user intention via electroencephalogram (EEG) in an oddball task. However, its performance is still insufficient for practical use.

In this work, we use multiple unique sound images coming from each direction in the form of spoken syllables. We aim to optimize the presentation sequence codebook to make a practical auditory P300 speller. To this aim, we have shown the differences of P300 brain response between playing from one direction versus two different directions. The results show that the brain generates stronger P300 response to the stimuli presentation from two different direction, allowing us to design more efficient auditory BCIs.

Introduction

Brain-computer interfaces provide additional means for humans to interact with their environment. Signals recorded from the brain such as EEG provide meaningful information streams that can be classified into commands that the user intends to execute [1].

The P300 event-related potential is the electrophysiological response to infrequent stimuli occurring during an oddball paradigm. The P300 waveform is characterized by a peak 300 ms after stimulus onset, and is on the scale of millivolts. Hence, it is vital to use proper stimuli to elicit strong P300 responses [2].

Recent research has investigated using sounds played at different directions to elicit P300 responses. Sugi et al. tested the impact of using different stimulus onset asynchronies (SOA), which are the durations between presented sounds, for a spatial auditory P300 speller [3]. They found that SOAs are individual to the user, where most users performed optimally with SOAs around 400 ms.

The P300 response is enhanced by user attention, and as such stimulus order is often randomized [4]. However, the P300 response is also influenced by inter-target times (ITT), where more time between presentations of the attended stimuli enhance the P300 response. Conversely, playing the same stimulus too soon can have negative consequences on the P300.

Hence our goal is to design and test different presentation codebooks that optimize the relationship between SOA and ITTs for many unique auditory stimuli to allow users to perform a speller task. Here, as a first step, we demonstrate the value of utilizing multiple speakers for P300.

Experiment Design

In this study the data corresponds to the evoked P300 waves on an auditory BCI paradigm has been used. A digital electroencephalogram system (Active Two, Bio-Semi, Amsterdam, Netherlands [10]) at 2048 Hz was used for recording the data. The recording device consists of 64 electrodes distributed over the head of the subject based on standard 10-20 system. The distribution of the electrodes was shown in Fig. 1.

The subject participated in two different experiments. First, they were resented with stimuli from two direction simultaneously. On every trial, listeners were cued by a 500 ms audiovisual cue to attend one of the two directions. Second, the stimuli were presented at one of the two directions.

Each trial consisted of playing 26 English letters. For multiple speakers, the stimuli were evenly split, and a participant could attend to their stimulus at a predetermined direction (Fig 2.). The set of possible stimuli at each speaker remains constant, but the order of the stimuli varies within-speaker. Each trial was repeated 10 times for every letter.

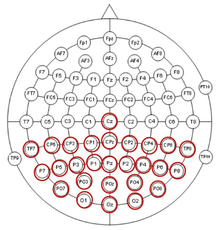

Fig. 1. Electrode placement, The red marked Electrodes used during recording

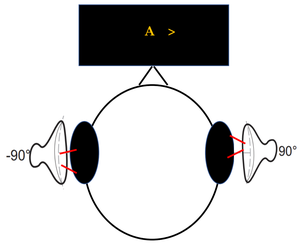

Fig. 2. Flow of the Experiment.

Pre-Processing and Feature Extraction

EEG data preprocessing is conducted as follows: The EEG data were high-pass and low-pass filtered. The low- pass cutoff frequency was 1 Hz and the high-pass cutoff frequency was 30Hz. The data was re-referenced to average.

We consider differences in event related potential (ERP) waveforms to check whether they were evoked by sound from the target direction. Audio P300 response occurs around 200 ms – 350 ms . Thus the first 400 ms is considered.

Wilcocson test is used to determine the discriminant features between target and non-target groups.

Results

The results are for considering whole letters together. The set of channels relate to the selected features are "P5, Oz, POz, Cz, CP4, CP2, P2, P4, P6", and " P5, P7, O1, Pz, CPz, P4, P8" for one speaker and two speakers respectively. The ERP for the two of the important channels are shown in Fig. 3. Since the data is imbalance, the AUC is the good metric to show the effectiveness of the method. The ROC for the two experiment is presented in Fig. 4. By considering each letter stimulus individually the ROC increase up to 0.8 for two speaker condition and 0.73 to one speaker condition.

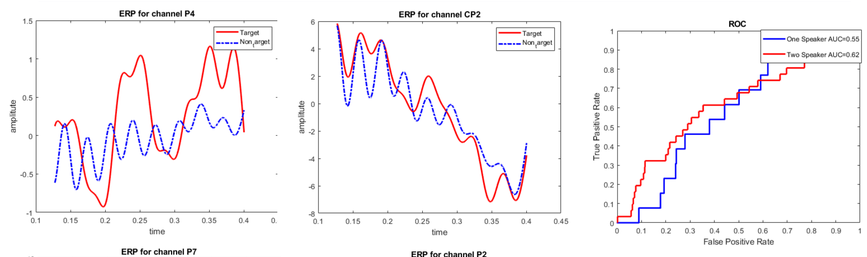

Fig. 3. ERPs for (Right) one speaker experiment (left) two speaker experiment

Fig. 4. ROC gained with SVM classifier

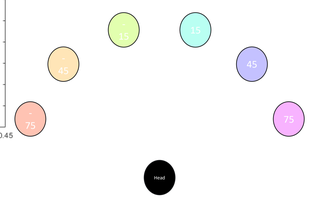

Fig. 5. Six Target Locations {-75,-45,-15,15,45,75}

Conclusions and Future Work

The goal of the experiment is to find a correct SOAs to make the auditory P300 faster to make it feasible to use. The advantage of using two speaker was the enhanced P300 response around 250 ms.

The results are just for one subject. In the future we want to try this in the array of external experiments and play the stimuli with overlaps with correct SOAs. Fig 5 and Fig 6 shows our future plan to continue this research. As shown in Fig 6 the columns would be play randomly during the time.

Deep networks could be effective and by collecting more data trials it would be possible to use deep techniques for high accuracy.

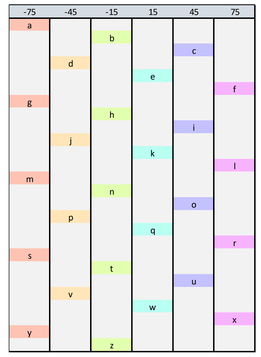

Fig. 6. Pattern of the expected playing order of the stimuli (overlap not depicted)

References

1. B. Edelman, J. Meng, D. Suma, C. Zurn, E. Nagarajan, B. Baxter, C. Cline, and B. He: "Noninvasive Neuroimaging Enhances Continuous Neural Tracking for Robotic Device Control." *Science Robotics*, Vol. 4, No. 31, 2019
2. R. Fazel-Rezai, B. Allison, C. Guger, E. Sellers, S. Kleih, and A. Kübler: "P300 Brain Computer Interface: Current Challenges and Emerging Trends." *Frontiers in Neuroengineering* Vol. 5, 2012.
3. M. Sugi, Y. Hagimoto, I. Nambu, A. Gonzalez, Y. Takei, S. Yano, H. Hokari, and Y. Wada: "Improving the Performance of an Auditory Brain- Computer Interface Using Virtual Sound Sources by Shortening Stimulus Onset Asynchrony." *Frontiers in Neuroscience*, Vol. 12, 2018.
4. J. Polich. "Updating P300: an integrative theory of P3a and P3b." *Clinical Neurophysiology,* Vol. 118, No. 10, 2007.