An optimized auditory P300 BCI based on spatially distributed sound in different voices

Jing Jin[1*], Ian Daly[2†], Minqiang Huang[1], Yu Zhang[1] and Xingyu Wang [1]
[1] Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai, China
[2]Brain embodiment lab, School of Systems Engineering, University of Reading, Reading, UK
jinjingat@gmail.com, i.daly@reading.ac.uk

Abstract

In this paper, a new paradigm is presented, to improve the performance of audio-based P300 Brain-computer interfaces (BCIs), by using spatially distributed natural sound stimuli. The new paradigm was compared to a conventional paradigm using spatially distributed sound to demonstrate the performance of this new paradigm. The results show that the new paradigm enlarged the N200 and P300 components, and yielded significantly better BCI performance than the conventional paradigm.

1 Introduction

Visual-based P300 brain-computer interfaces (BCIs) requires users' to have control over their gaze direction and, therefore, are not useable by blind patients or other patients who cannot maintain eye gaze [1]. In order to enlarge the group of users who may benefit from use of P300 BCIs, the audio-based P300 BCI was developed [2-3]. It has been shown that an audio paradigm using spatially distributed stimuli is better than using different pitches from one speaker [4]. Höhne et al (2011) presented a paradigm using high, medium, and low unnatural spatially distributed sounds from headphones [5]. However, these two studies did not show that this paradigm was better than a paradigm using same sound played by spatially distributed speakers. In our study, a new audio-based P300 BCI was designed by using three different natural sounds. The three different natural sounds were female, male, and child voices, which were played by six spatially distributed speakers. This paradigm was called the "123 pattern". Another paradigm using unnatural sounds (a "beep"), randomly played by six spatially distributed speakers, was used to evaluate the efficacy of our proposed pattern, this was called the "beep pattern".

2 Methods and Materials

2.1 Subjects

Twelve healthy right-handed subjects (7 male, 5 female, aged 21 to 25 years, mean age 23.2±1.0) participated in this study. All subjects signed a written consent form prior to this experiment and were paid 50 RMB for their participation. The local ethics committee approved the consent form and

* Corresponding author
† Corresponding author

experimental procedure before any subjects participated. All subjects' native language was Mandarin Chinese. Two subjects (labeled S7 and S9) had used a visual-based P300 BCI before this study.

2.2 Stimuli and flash patterns

Six speakers, placed in a semicircle around the participant, were used to present auditory stimuli (see figure 1). The distance between the participant and the speakers was 1.5 m and the distance between two adjacent speakers on the left and right side respectively was 30 degrees. The distance between the last speaker (speaker 3) on the left side and the first speaker (speaker 4) on the right side was 60 degrees (see figure 1). The labels of the six speakers were 1-6 from left to right. Speakers were calibrated to a common stimulus intensity of 45dB. In the experiment, participants were asked to close their eyes and pay attention to the target speaker and count the number of times that the speaker was played.

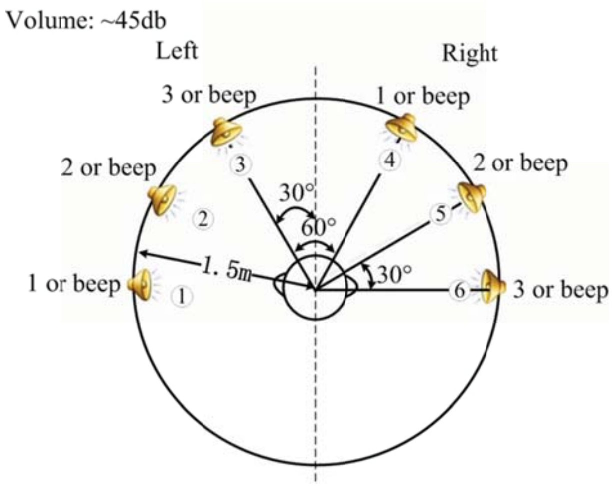

Figure 1: The location of speakers and sounds that were played by the speakers.

2.3 Experimental set up, offline and online protocols

EEG signals were recorded with a g.USBamp and a g.EEGcap (Guger Technologies, Graz, Austria) with a sensitivity of 100μV, band pass filtered between 0.1Hz and 100Hz, and sampled at 256Hz. Data were recorded and analyzed using the BCI platform software package developed through East China University of Science and Technology. We recorded from 24 EEG electrode positions based on the extended International 10-20 system. The electrodes were F3, Fz, F4, T7, C5, C3, C1, Cz, C2, C4, C6, T8, CP5, CP3, CP1, CPz, CP2, CP4, CP6, P3, P1, Pz, P2, and P4. The right mastoid electrode was used as the reference, and the front electrode (FPz) was used as the ground. These electrode positions are more convenient for patients who are lying on a bed and, therefore, not able to easily move their head. Thus, the setup of the BCI was designed to allow rapid adoption by patient groups.

There are two conditions called "123 pattern" and "beep pattern" respectively. The order of the conditions was counterbalanced. In the 123 pattern, auditory presentation of the number "1" was played by speaker 1 in a female voice, auditory presentation of the number "2" was played by speaker 2 in a male voice, and auditory presentation of the number "3" was played by speaker 3 in a child's voice. Speakers 1-3 were located on the left side. On the right side, auditory presentation of "1" was played by speaker 4 in a female voice, auditory presentation of "2" was played by speaker 5 in a male voice, and auditory presentation of "3" was played by speaker 6 in a child's voice. The auditory presentation of numbers was in Chinese. In the beep pattern, only a beep would be played by each of the speakers.

The stimulus on time was 200 ms and the stimulus off time was 440 ms. The inter-stimulus interval was 640 ms. In the experiment, each of the six speakers was played randomly and separately, and each speaker would be played once in one trial. In the offline experiment, there were twelve trials of stimuli in each run, which were needed for one character selection. Fifteen character selections were performed in the offline experiment for each pattern. In the online experiment, there were five trials of stimuli in each run, which were needed for one character selection. Twenty character selections were attempted in the online experiment for each pattern.

In the beginning of each run, there was an auditory cue (in Chinese) from the target speaker to guide the subject to locate the target speaker. The task of the subjects in the experiment was to count the number of times that the target speaker was played.

2.4 Feature extraction procedure

A third order Butterworth band pass filter was used to filter the EEG between 0.1 Hz to 30 Hz. The EEG was down-sampled from 256 Hz to 64 Hz by selecting every fourth sample from the filtered EEG. The first 1000 ms of EEG after presentation of a single stimulus was used to extract the feature.

3 Results

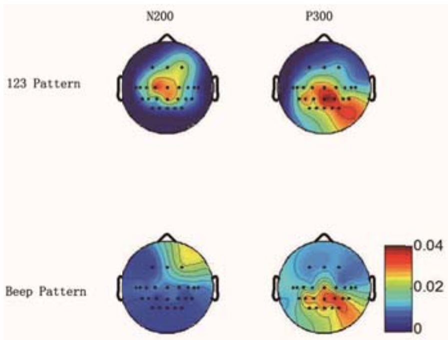

Figure 2: The topographic map of r-squared values for the 123 and beep patterns. The peak point value was used to calculate the r-squared value, which was obtained between 150-300 ms for the N200 and between 300-450 ms for the P300.

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	Avg
123-P(%)	75	90	80	80	90	10	80	80	70	90	60	95	75±22.7
B-P(%)	75	15	95	65	70	10	60	55	90	15	50	30	53±29.2

Table 1: Performance (classification accuracy) from online feedback runs across all 12 participants who participated in online runs. "123-P" denotes the 123 pattern and "B-P" the beep pattern.

Figure 2 shows the topographic map of r-squared values during presentation of the 123 pattern and the beep pattern. It shows that the 123 pattern obtained higher r-squared values than that of the beep pattern in the central region for the N200, and in the central and parietal regions for the P300.

Table 1 shows the online classification accuracy using five trials for constructing the average ERP waveform. Paired samples tests were used to show the difference between the two patterns. This shows that the classification accuracy of the 123 pattern is significantly higher than that of the beep pattern ($t=2.3$, $p<0.05$).

4 Discussion and conclusion

The goal of this study was to prove that an audio-based P300 BCI, using different pitches and spatially distributed speakers, could obtain higher classification accuracy than a comparable BCI using audio stimuli of the same pitch from spatially distributed speakers. Table 1 shows that classification accuracy of the 123 pattern was significantly higher than that of the beep pattern ($t=2.3$, $p<0.05$). The new paradigm used natural sounds, which would lead to high user acceptance and the new paradigm used spatially distributed sounds in difference voices, which allowed users to locate the target easily. However, a longer target to target interval was used in the audio-based BCI compared to the visual-based BCI, which would limit the communication speed of the audio-based BCI.

Neville and Lawson, (1987) explored ERPs from patients in response to spatially distributed tones [6]. Kübler et al., (2009) also tested the audio-based BCI on patients for consciousness assessment and communication [78]. In future, we will also focus on applying this system to patients who cannot control their eye gaze direction and, therefore, validate the performance of this audio-BCI system when it is used by target patient groups.

5 Acknowledgements

This work was supported in part by the Grant National Natural Science Foundation of China, under Grant Numbers 61074113, 61105122, 61203127 and 61305028 and was also supported by the Fundamental Research Funds for the Central Universities (WH1114038, WH1314023).

References

[1] Klobassa D S, Vaughan T M, Brunner P, Schwartz N E, Wolpaw J R, Neuper C, Sellers E W. (2009) Toward a high-throughput auditory P300-based brain-computer interface, *Clin. Neurophysiol*. **120**, 1252-1261

[2] Hill N J, Lal T M, Bierig K, Birbaumer N, Schölkopf B (2004) An auditory paradigm for brain-computer interface NIPS, *books.nips.cc/papers/files/nips17/NIPS2004_0503.ps.gz*

[3] Furdea A, Halder S, Krusienski D J, Bross D, Nijboer F, Birbaumer N, Kübler A 2009 An auditory oddball (P300) spelling system for brain-computer interfaces, *Psychophysiology* **46**, 617-625.

[4] Belitski A, Farquhar, Desain P 2011 P300 audio-visual speller *J. Neural Eng.* **8**, 025022

[5] Höhne J, Schreuder M, Blankertz B, Tangermann M 2011, A novel 9-class auditory ERP paradigm driving a predictive text entry system *Front Neurosci* **22**,5:99

[6] Beville H J, Lawson D 1987 Attention to central and peripheral visual space in a movement detection task III. Sepatate effects of auditory deprivation and acquisition of a visual language *Brain Res* ***405***, 284-294

[7] Kübler A, Furdea A, Halder S, Hammer E M, Nijboer F, Kotchoubey B 2009 A brain-computer interface controlled auditory event-related potential (p300) spelling system for locked-in patients *Ann N Y Acad Sci* **1157**, 90-100.