Check for updates

BCI Speller on Smartphone Device with Diminutive-Sized Visual Stimuli

Nuray Serkali, Adai Shomanov, Madina Kudaibergenova, and Min-Ho Lee(✉)

School of Engineering and Digital Sciences, Nazarbayev University, Kabanbay Batyr Avenue 53, Nur-Sultan 010000, Kazakhstan
{nuray.serkali,adai.shomanov,madina.kudaibergenova,minho.lee}@nu.edu.kz

Abstract. In real-world BCI applications, small-sized and low-impact stimuli are more appropriate for smart devices. However, diminishing the stimuli intensity leads to a reduction of P300 amplitude, causing lower system performance. The purpose of this study is to propose a state-of-the-art BCI speller where diminutive (less than 1 mm) visual stimuli were implemented in a smartphone interface. To boost the task-relevant brain components, participants performed a certain mental task according to the given cue signs. Additionally, we applied a data-driven optimization approach to represent the user-specific spatial-temporal features. The results showed 96.8% of spelling accuracy with a maximum ITR of 31.6 [bits/min], which is comparable or even superior to conventional speller systems. Our study demonstrated the feasibility to create more reliable and practical BCI spelling systems in the future.

Keywords: BCI speller · Event related potential (ERP) · Late positive potential (LPP) · Sound imagery · Mental task

1 Introduction

Brain-computer interface (BCI) systems are a non-muscular communication opportunity for people with severe disabilities. BCI helps patients with neuromuscular disorders to project their intention by controlling external devices, such as personal computers, synthesizer for speech, and prostheses [1]. Due to its qualities as optimal price, easy utilization, and no risk, electroencephalography (EEG) is frequently used in researching brain activities [2].

Prior researches have noted that the quality of evoked potentials highly depend on the target stimulus's visual or auditory characteristics [3–5]. Therefore, the main idea of BCI performance improvement was to regularize the shape, color, or intensity of the stimuli (i.e. transferring familiar face pictures to stimulus [3,4], random stimulus [6], and color distinction [5]). These studies basically strengthened the existing visual or auditory stimulus by making them either

Supported by Faculty Development Competitive Research Grant Program (No. 080420FD1909) at Nazarbayev University.

T. Mantoro et al. (Eds.): ICONIP 2021, CCIS 1517, pp. 144–151, 2021.
https://doi.org/10.1007/978-3-030-92310-5_17

more intense or louder, which evoked stronger ERP response, and therefore, caused improvements in the performance.

Moreover, users cannot sit in front of a monitor for long time while doing stimuli tasks due to various factors and task complexities. Thus, in real-world smaller stimuli seem to be more practical. Xu et al. [7] introduced miniature asymmetric visual evoked potentials (aVEPs) to increase speller performance. In the subsequent research, placing unobtrusive visual stimuli outside of the fovea vision, then applying canonical pattern matching method has resulted in a satisfactory performance for classification of ERP components [8].

The aim of this research is to present the idea of diminishing the external properties of visual stimuli in BCI speller systems and maintaining the performance achieved in previous studies. We theorize that performing voluntary mental task could elicit endogenous ERP components, and this complementary synthesis of oddball signals might improve the performance. The mental task for the subjects was to imagine the high-pitch sound, that was played before the experiment, when the target character appeared on screen. Two mental tasks: passive concentration (PC) and active concentration (AC) were designated in order to devaricate the ERP responses.

The hardware setup of this study is a mobile phone screen - speller layout with diminutive visual stimuli (defined as a *dot-speller*), and the subjects were guided to perform the mental task while the target symbol was presented. The paradigms in this experiment were all implemented to present a more user-friendly BCI system which can minimize the unpreventable adverse impact of external stimuli (e.g., noisy visual/auditory stimuli). As a result, a high level cognitive neural activity ERP component that can decode the user's intention, a late positive potential (LPP), was observed.

2 Materials and Methods

2.1 Participants and Data Acquisition

14 healthy subjects (aged 25–33 years, 4 females, 5 BCI naive users) participated in this study. All participants are confirmed to have normal or corrected vision and be free of psychiatric or neurological disorders. During data acquisition, the subjects were sitting on a chair around 80 cm away from the visual stimulus.

EEG data was recorded via an ActiCap EEG amplifier (Brain Products, Germany) with 32 channels (Fp1-2, F3-4, Fz, F7-8, FC5-6, FC1-2, T7-8, C3-4, Cz, CP1-2, CP5-6, TP9-10, P3-4, P7-8, Pz, PO9-10, O1-2, and Oz). Electrodes with 10–20 system standard were used along with forehead grounded reference on nose (Ag/AgCl electrodes with a maximum impedance of 10 kΩ). The DC artifacts were removed from the data by applying a notch filter with 1 KHz and 60 Hz sampling rate. Finally, Butterworth filter (5^{th} order) with parameters of 0.5 Hz and 30 Hz was used to filter out noise.

The subjects were fully informed of the experiment's objectives. Consent was taken from all subjects in written form. The experiment was revised and

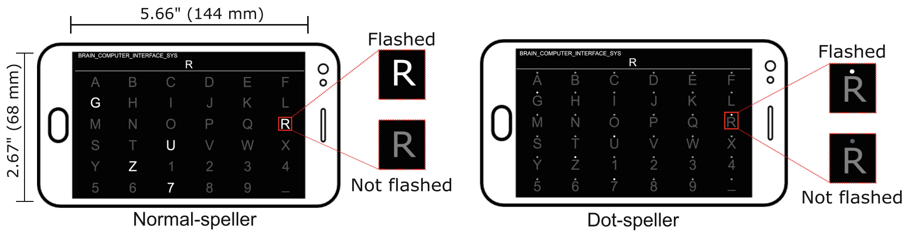

Fig. 1. Illustration of speller systems. Dot- and normal-speller systems were implemented in the smartphone layout

has the approval of the Institutional Review Board at Korea University [1040548-KUIRB-16-159-A-2].

2.2 Dot- and Normal-Speller Experiments

Two groups of mental states were introduced: AC and PC. In the PC condition, subjects were instructed to focus on the given stimuli without any certain mental task. In the AC condition users had an instruction to gaze at the target character and perform a sound imagery task, which was remembered (8 KHz frequency for 2 min) before the experiment. Basically, we validated the spelling performance of the different mental tasks within both dot- and normal-speller layouts.

A normal-speller [9] and a dot-speller were implemented in the smartphone environment. Both of the interface layouts were designed with 36 target visual stimuli ('A'–'Z', '1'–'9', '_'). Individual stimuli had equal positions on the screen by 6 rows and 6 columns. The stimuli were grey colored, and the background was black. The individual letters in the normal-speller were repetitively flashed. In the dot-speller tiny dots (less than 1 mm) were visually positioned on the top of individual letters, and these dot-symbols were flashed instead of the letter itself (see Fig. 1-(d)). Target and non-target trials ratio was 2:10. Note that our experimental approach (e.g., protocol, validation) to the speller experiments was designed based on well-established methods in related studies [3,4,10,11]. A sequence of 12 flashes (i.e. trials) was counted as a single iteration of flashed letters twice within rows and columns. There were overall 10 sequences with around 70 ms stimuli flash, and 150 ms ISI. The newly developed speller layout was presented in a smartphone (Galaxy 9, Samsung) environment with a 1440p OLED/5.8-inch panel using the screen capturing application.

The experimental procedures for spellers were identical. The training phase was conducted offline, and patients were instructed to spell the following phrase, 'BRAIN_COMPUTER_INTERFACE_SYS' (28 characters including spaces '_') according to the task. Therefore, 3360 trials (28 characters $\times$ 10 sequences $\times$ 12 flashes) were collected in each condition. Two classifiers were then constructed based on the training dataset: PC vs. NT and AC vs. NT. In the test phase, subjects had an instruction to spell 'U3RQJSMAUWES2QEF_KOREA UNIVERSITY' (32 characters) according to the given task. After presenting all the letters (i.e., the end of 10 sequences) in every attempt, online feedback was available for users (on top-left corner classifier showed the found target character). A total of 3840 trials (32 characters $\times$ 10 sequences $\times$ 12 flashes) were therefore collected in both conditions, and these test datasets were used to evaluate speller performance.

3 Data Analysis and Performance Evaluations

EEG data were first down-sampled 100 Hz, and then epochs were acquired by extracting individual trials in the interval of [−100 and 1000 ms] referencing the stimulus onset, after which baseline-correction was performed: subtraction of mean amplitudes in [−100 and 0] ms pre-stimulus time interval. Afterwards, ERP responses for NT, PC and AC were investigated in the individual setups (i.e., *normal-speller*, and *dot-speller*). In each session, all trials in the training and test phases were concatenated. The Grand Averaged ERP patterns were then evaluated across all subjects. Decoding accuracy and information transfer rates (ITRs) were calculated along with the individual sequences to evaluate the spelling performances [12]. Note that the training data were used to construct the classifier parameters, and the decoding accuracy was evaluated in the test dataset.

In the training phase, k intervals with a step length of 20 ms and an interval length of 100 ms were created. The mean amplitude features [1] in k time intervals were calculated from the ERP trials across all channels. The signed r-squared value [13] was applied to statistically investigate the differences in temporal ERP responses across all channels. A regularized linear discriminant analysis (RLDA) [14] classifier was generated from the selected feature set.

Commonly, for all experiments certain target stimuli were presented 10 times (i.e. 10 sequences), and decoding accuracy was calculated for individual sequence (e.g. the decoding accuracy of last sequence was computed by the taking the average of accumulated epochs of all 10 sequences). The speller layout had 36 classes (the chance level at 2.77%), and the classifier output $f(\mathbf{x}_i)$ was calculated from all the individual letters ($i = 1, ..., 36$). The estimated letter i, which has the highest classification score, was chosen as the desired target symbol. These decision functions were used to provide real-time feedback for the test phase in all individual experiments.

4 Results

4.1 ERP Responses

Typical P300 components [15] were observed in both the normal- and dot-speller experiments as these systems are designed within the oddball paradigm. Mean values for the peak amplitudes in the interval of 300–400 ms (i.e., P300) at the Cz electrode were 1.657 (±1.125)uV, 0.902 (±1.859)uV, 1.837 (±1.122)uV,

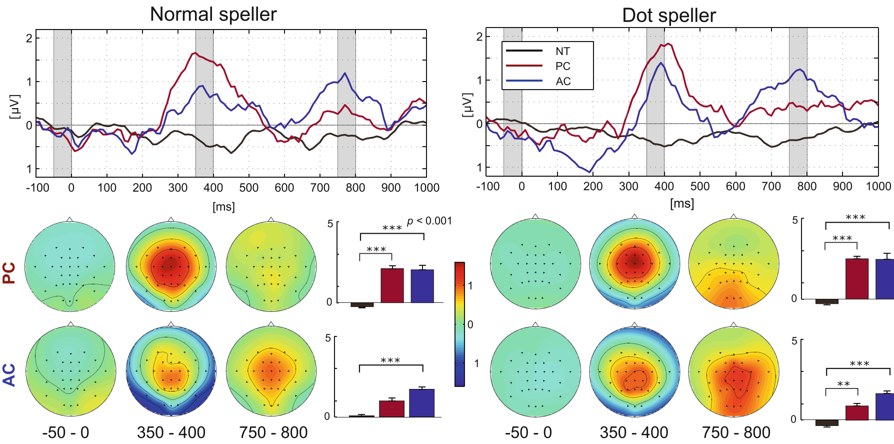

Fig. 2. Average ERP responses at electrode Cz. The scalp plots demonstrate the distribution of signal response for the three different conditions, i.e., *NT*, *PC*, and *AC* trials

and 1.394 (±2.639)uV for normal-passive, normal-active, dot-passive, and dot-active, respectively. The peak amplitude indicates that P300 components were more strongly evoked during passive concentration compared to active task.

Additionally, LPP was observed in the interval of 700–800 ms, and peak amplitudes mean values at the Cz electrode were 0.463 (±1.210) uV, 1.196 (±0.989) uV, 0.476 (±0.996) uV, and 1.244 (±1.105) uV for normal-passive, normal-active, dot-passive, and dot-active, respectively. Contrary to the P300 components, the LPPs were evoked by the active task in both speller systems (see Fig. 2).

4.2 Decoding Accuracy of Normal- and Dot-Speller

Figure 3 indicates the decoding accuracy for target and non-target discrimination in four conditions. Decoding accuracies were calculated from 1 to 10 sequences (x-axis) for individual users as well as the average speller performances. Results demonstrate that active tasks in both speller systems significantly outperformed passive tasks. Precisely, the average accuracies were 53.5%, 83.0%, 62.9%, and 88.8% after sequence four, and 76.3%, 94.1%, 78.1% and 96.8% after the 9th sequence for the normal-passive, normal-active, dot-passive, and dot-active conditions, respectively.

Paired t-tests indicate that active tasks ($p > 0.05$) have similar pattern with the passive tasks ($p > 0.05$) in both speller systems. The active tasks demonstrate higher performance than the passive task in sequences 2–6 and 2–8 for normal-speller and dot-speller respectively. The maximum ITRs were 13.5, 28.6, 18.3, and 31.6 [bits/min] for the normal-passive, normal-active, dot-passive, and dot-active conditions, respectively.

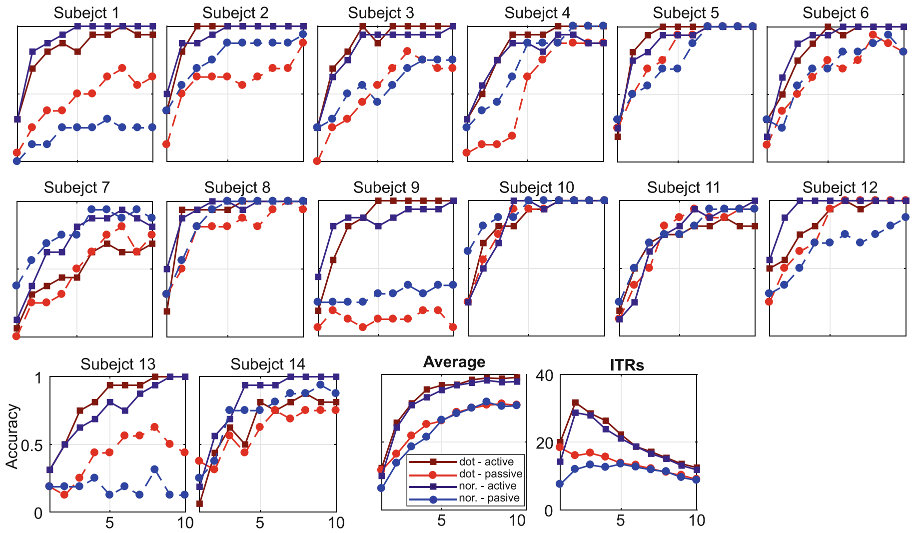

Fig. 3. Decoding accuracy for target and non-target discrimination in the four conditions. The figure indicates the decoding accuracy plots for 14 subjects and the averaged decoding accuracy across all subjects. In the last plot, ITR values for each of the sequences for the four conditions are presented. For each of the subjects, active task experiments produced much higher decoding accuracy than the passive tasks.

5 Discussion

This study aimed to investigate the practical perspective of real-world BCI applications as we studied the performance of a speller system, the application of which combines pixel-level visual stimuli and sound imagery stimuli (dot-speller). The proposed speller systems still rely upon their performance in an oddball paradigm, but with a focus on less obtrusive visual interface. To investigate the efficacy of small visual stimulation, we investigated the differences between a typical speller where the letters themselves were flashed and our novel dot-speller where 0.1 mm dots were flashed instead of the letters on a smartphone. Additionally, we examined the utility of active sound imagery tasks within this setup compared to passive gazing. Our results point to some important implications for future practical BCI interfaces.

Concerning the issue of intentional command, there were two significant intervals for discriminating AC and PC at 300–400 ms (P300) and 700–800 ms (LPP) intervals (see Fig. 2). Interestingly, the PC task showed a stronger P300 component compared to the AC, while the opposite result was found in the LPP. We propose that the LPP stems from the user's active mental task and could be a powerful feature compared to P300 component. While P300 response is the exogenous reaction to the oddball stimulus [9], passive attention can lead to false commands, whereas the active mental would be far more reliable in terms of intention.

Regarding the intention issue, Fig. 3 indicates that the active task significantly outperformed the passive attention task. The average spelling accuracies for the passive task were 76.3% and 78.1% in the normal- and dot-layout systems, respectively. This accuracy is lower than in previous studies where results have shown more than 90% accuracy [3,4,16,17]. This reduced performance is mainly due to our speller system being implemented on a smartphone interface. The indicative stimulus sizes were less than 0.8 cm for the normal-speller letters and only 0.1 mm for the dot-speller layout. As previously was found, smaller-sized and closer-positioned letters can reduce ERP responses, which in turn results in a decreased system performance [18]. Regardless of this shortcoming, the spelling accuracies for the active tasks were 94.1% and 96.8%. This result is comparable to or even outperforms many advanced spelling systems [4,5,16,17,19].

6 Conclusion

In this article, we proposed and tested a novel concept of lowering the impact of external stimuli (visual) while maintaining high classification accuracy by generating endogenous ERP components through a mental task (sound imagery) the user executes. The experiment was run and compared between four possible combinations of stimuli, which allowed us to study the impact of both external (normal- and dot-speller) and mental (passively attending and active concentrating) stimuli. As a result, executing a mental task improved the performance significantly, and the dot-speller showed higher accuracy than the traditional speller. The experiments were taken by healthy (normal or corrected vision) individuals, so further study should be conducted to find out the proposed feature's applicability for users with eye impairments. The feature proposed in this article demonstrates superior potential to create more reliable, user-friendly BCI spelling systems in the future.

References

1. Lee, M.H., et al.: EEG dataset and OpenBMI toolbox for three BCI paradigms: an investigation into BCI illiteracy. GigaScience **8**(5), giz002 (2019)
2. Nicolas-Alonso, L.F., Gomez-Gil, J.: Brain-computer interfaces, a review. Sensors **12**(2), 1211–1279 (2012)
3. Yeom, S.K., Fazli, S., Müller, K.R., Lee, S.W.: An efficient ERP-based brain-computer interface using random set presentation and face familiarity. PLoS ONE **9**(11), e111157 (2014)
4. Li, Q., Liu, S., Li, J., Bai, O.: Use of a green familiar faces paradigm improves p300-speller brain-computer interface performance. PLoS ONE **10**(6), 1–15 (2015). https://doi.org/10.1371/journal.pone.0130325
5. Li, Q., Lu, Z., Gao, N., Yang, J.: Optimizing the performance of the visual P300-speller through active mental tasks based on color distinction and modulation of task difficulty. Front. Hum. Neurosci. **13**, 130 (2019)
6. Kubová, Z., Kremlacek, J., Szanyi, J., Chlubnová, J., Kuba, M.: Visual event-related potentials to moving stimuli: normative data. Physiol. Res. **51**(2), 199–204 (2002)

7. Xu, M., Xiao, X., Wang, Y., Qi, H., Jung, T.P., Ming, D.: A brain-computer interface based on miniature-event-related potentials induced by very small lateral visual stimuli. IEEE Trans. Biomed. Eng. **65**(5), 1166–1175 (2018)
8. Xiao, X., Xu, M., Jin, J., Wang, Y., Jung, T.P., Ming, D.: Discriminative canonical pattern matching for single-trial classification of ERP components. IEEE Trans. Biomed. Eng. **67**(8), 2266–2275 (2019)
9. Farwell, L.A., Donchin, E.: Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroencephalogr. clin. Neurophysiol. **70**(6), 510–523 (1988)
10. Lee, M.H., Williamson, J., Won, D.O., Fazli, S., Lee, S.W.: A high performance spelling system based on EEG-EOG signals with visual feedback. IEEE Trans. Neural Syst. Rehabil. Eng. **26**(7), 1443–1459 (2018)
11. Lee, M.H., et al.: OpenBMI: a real-time data analysis toolbox for brain-machine interfaces. In: 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 001884–001887. IEEE (2016)
12. McFarland, D.J., Sarnacki, W.A., Wolpaw, J.R.: Brain-computer interface (BCI) operation: optimizing information transfer rates. Biol. Psychol. **63**(3), 237–251 (2003)
13. Blankertz, B., Lemm, S., Treder, M., Haufe, S., Müller, K.R.: Single-trial analysis and classification of ERP components: a tutorial. NeuroImage **56**(2), 814–825 (2011)
14. Friedman, J.H.: Regularized discriminant analysis. J. Am. Stat. Assoc. **84**(405), 165–175 (1989)
15. Bennington, J.Y., Polich, J.: Comparison of P300 from passive and active tasks for auditory and visual stimuli. Int. J. Psychophysiol. **34**(2), 171–177 (1999)
16. Jin, J., et al.: The changing face of P300 BCIs: a comparison of stimulus changes in a P300 BCI involving faces, emotion, and movement. PLoS ONE **7**(11), e49688 (2012)
17. Lu, Z., Li, Q., Gao, N., Yang, J.: The self-face paradigm improves the performance of the P300-speller system. Front. Comput. Neurosci. **13**, 93 (2019)
18. Cao, T., Wan, F., Wong, C.M., da Cruz, J.N., Hu, Y.: Objective evaluation of fatigue by EEG spectral analysis in steady-state visual evoked potential-based brain-computer interfaces. Biomed. Eng. Online **13**(1), 28 (2014)
19. Gu, Z., Chen, Z., Zhang, J., Zhang, X., Yu, Z.L.: An online interactive paradigm for P300 brain-computer interface speller. IEEE Trans. Neural Syst. Rehabil. Eng. **27**(2), 152–161 (2019)