2021 27th International Conference on Mechatronics and Machine Vision in Practice (M2VIP) | 978-1-6654-3153-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/M2VIP49856.2021.9664989

Cross-Sessions and Cross-Paradigms Analysis for the Problem of Brain-Computer Interface Inefficiency*

Weize Chen, Gan Huang, Zhenxing Hu, Li Zhang, Linling Li, Zhen Liang, Zhiguo Zhang

Abstract— **Brain-Computer Interfaces (BCIs) allow users to make use of brain activity to control external devices directly for rehabilitation and enhancement of human functions. However, the inefficiency problem that a typical BCI system is unable to effectively decode EEG signals in some users, prevents BCI technology from benefitting all users. The proportion of inefficiency varies in the major BCI paradigms, among which Motor Imagery (MI)-based BCI achieves highest (10%-50%). Hence, the question arises as to whether other BCI paradigms, such as P300, could be substitutes for users who cannot be served by certain paradigm. In this work, a cross-paradigms BCI experiment, in which 93 healthy subjects executed BCI tasks including real movement and P300 for two sessions on separated days, was performed to answer the above question. Firstly, the highly correlation between the recognition accuracy in two sessions within subjects for both Sensory Motor Rhythm (SMR) features ($p = 4.47\times10^{-11}$) and P300 features ($p = 2.17\times10^{-3}$) indicated the reproducibility of the subject-level BCI inefficiency in the two paradigms. Further analysis demonstrated no significant correlation between the decoding performance of the SMR and P300 features ($p = 0.604$). The results verified the feasibility of improving BCI decoding performance by replacing certain BCI paradigm with another one when users encounter the problem of BCI inefficiency.**

I. Introduction

Brain-Computer Interfaces (BCIs) create a direct communication access between the human brain and the computer[1], which might serve as an auxiliary control tool for patients with dyskinesia by translating the brain activity into command or control signals[2]. Due to higher security and lower cost compared with invasive BCIs, non-invasive BCIs based on the measurement of scalp electroencephalography (EEG) have been widely used for recording brain signals[3], primarily including Motor Imagery (MI), P300 and Steady-State Visual Evoked Potential (SSVEP). MI-based BCIs require users to imagine limb movement to spontaneously generate Event-Related Desynchronization/Synchronization (ERD/ERS)[4] in the mu (8-13Hz) and beta (18-24Hz) frequency ranges of EEG[5], while P300-based BCIs or SSVEP-based BCIs require users to gaze attentively at visual targets that induce the P300 component[6] or oscillations[7] in the EEG.

However, across the three major BCI paradigms mentioned above, there is an inevitable problem called BCI inefficiency that a typical BCI system cannot work for all users due to its ineffectiveness for some users whose EEG signals are unable to be effectively decoded to attain control[8][9]. In a previous research, about 72.8% of 81 subjects showed a control with the P300 speller in 100% accuracy, while only 11.1% were below 80%[10]. Similarly, about 86.7% of 53 subjects reached an accuracy above 90%, and nobody was below 60% in SSVEP-based BCIs (four classes)[11]. Nonetheless, the problem of inefficiency might be more prominent in MI-based BCIs compared to other BCIs. In a study concerning MI-based BCI (two classes) inefficiency, a high accuracy of above 90% could be achieved by only 6.2% of 99 subjects, while 78.8% were below 80%[12]. Though there were some meaningful studies for the recognition of BCI-inefficient users using physiological or psychological features[13][14], the appropriate solution to the problem of BCI inefficiency is still being exploited.

The previous studies implied that P300-based BCIs or SSVEP-based BCIs might provide users who encounter the problem of inefficiency in certain BCI paradigms, such as MI-based BCIs, with viable alternatives. However, this hypothesis needs to be verified by a broad within-subjects and cross-paradigms research in which each subject is required to perform at least two BCI paradigms. In this work, the problem of BCI inefficiency was investigated in a large-scale BCI dataset (93 subjects) in which subjects executed experiments including real movement and P300 for two sessions on separated days. Cross-sessions correlation analysis was performed to verify the reproducibility of the BCI inefficiency in the two paradigms, and cross-paradigms correlation analysis was performed to explore the relationship of BCI inefficiency between the two types of paradigms.

The rest of this paper is organized as follows. In Section II, the experiment paradigms and the method of signal processing and classification are illustrated in detail. The results and discussions are arranged in Section III and IV respectively. In the end, the conclusion is presented succinctly in Section V.

II. Method

A. Experimental Procedure

Ninety-three healthy subjects (71 females; age 21.1 ± 5.3) participated in the experiment for two sessions with same procedure on separated days. All subjects were naïve BCI users and had no neurological diseases affecting experimental results. Before the experiment, written informed consent from all subjects and ethical approval of the study from the Medical Ethics Committee, Health Science Center, Shenzhen

*Research supported by the National Natural Science Foundation of China (Nos. 81871443), the Science, Technology and Innovation Commission of Shenzhen Municipality Technology Fund (No. JCYJ20190808173819182), and the Shenzhen Peacock Plan (No. KQTD2016053112051497). The authors declare that they have no conflicts of interest.

Weize Chen, Gan Huang, Zhenxing Hu, Li Zhang, Linling Li, Zhen Liang, Zhiguo Zhang are with the School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen, 518060, China; Marshall Laboratory of Biomedical Engineering, Shenzhen, 518060, China (email to huanggan1982@gmail.com; zgzhang@szu.edu.cn)

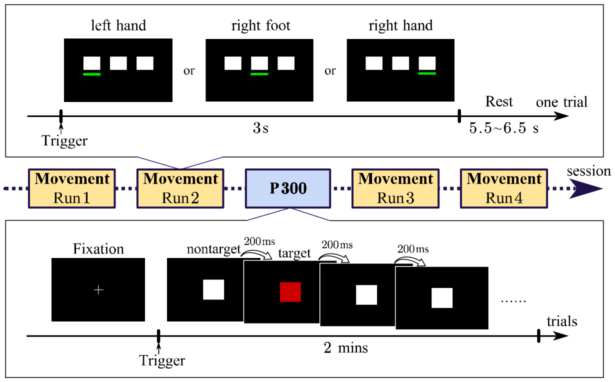


Figure 1. Experimental procedure. For each session, the experiment consisted of four runs of real movement task (60 trials for each run), one run of P300 task (600 trials).

University (No. 2019034) were sought and obtained. In each experiment session, each subject was seated comfortably in a chair in front of a computer screen displaying the visual cue, and instructed to execute the experimental task illustrated in Fig. 1.

In real movement task, the subjects performed real movements by gripping and relaxing repeatedly their left hand (LH) or right hand (RH), or lifting their right foot (RF) as fast as they could in response to the given visual cue for 3 s. No feedback was provided during the online recording. During each session, 240 trials (80 for each movement type) arranged in four runs were presented with the inter-stimulus interval (ISI) 5.5-6.5 s in a random order.

In P300 task, the subjects were asked to gaze attentively at the center square presenting random color stimulus with the red as the target stimulus and white as the nontarget stimulus. Each color square lasted 80 ms and the ISI was set to 200 ms. During each session, 600 trials of the stimuli in which target stimuli came with the possibility of 5%, were arranged in one run lasting 2 mins. Flexible rest time after each run would be allowed as the subjects wanted.

B. Signal Acquisition and Pre-processing

EEG signals were collected by 64 electrodes and referenced to electrode FCz, via BrainAmp (Brain Products GmbH; Germany) with a sampling rate of 1000 Hz. Band-pass filtering between 0.01 and 200 Hz, 50 Hz notch filtering, bad channel interpolation, independent component analysis (ICA) for artifact removal, re-referencing to the average of electrode TP9 and TP10 were applied in sequence to the raw EEG signals.

For movement experiment, 21 electrodes in motor cortex region were selected for offline analysis (F-5/3/1/z/2/4/6, C-5/3/1/z/2/4/6, P-5/3/1/z/2/4/6). The continues EEG signals were filtered by a 4th order Butterworth digital filter with bandwidth setting 8-30 Hz, and then segmented from 0.5 to 3.0 s with respect to cue onset.

For P300 experiment, 8 electrodes were selected for offline analysis (Fz, Cz, P-3/z/4, Oz, PO-7/8). The continues EEG signals were 0.5-30 Hz band-pass filtered with a 4th order Butterworth digital filter, and then segmented from corresponding stimulus onset to the next 0.6 s.

C. Feature Extraction and Classification

In this study, the Common Spatial Pattern (CSP) algorithm[15] and xDAWN algorithm[16] were applied to perform spatial filtering and feature extraction for both paradigms respectively.

With the purpose of learning spatial filters that maximize the variance of EEG signals from one class while minimizing them from the other class, the CSP method is successfully employed for detecting ERD and ERS[15]. The spatial filters w are optimized by

$$\underset{w}{argmax}\frac{w^T\Sigma_1 w}{w^T\Sigma_2 w} \tag{1}$$

where Σ_1 and Σ_2 are normalized spatial covariance matrix estimated by averaging all training trials respectively from two different classes (e.g., LH vs. RH). By solving the generalized eigenvalue problem,

$$\Sigma_1 w = \lambda\, \Sigma_2 w \tag{2}$$

the m pairs of w corresponding to the largest and lowest eigenvalues are optimal for spatial filtering to discriminate two classes of EEG measurements. In this study, we set $m = 3$ and the logarithm of the EEG signal variance after spatial projection was extracted as Sensory Motor Rhythm (SMR) features and then fed to train a linear discriminant analysis (LDA) classifier for the offline data analysis.

The One-Versus-Rest (OVR) strategy for training classifier was used to extend CSP algorithm to the three-class case, i.e., LH vs. RH vs. RF. Hence, three LDA classifiers, in which the class of highest probability was chosen as the final prediction, were trained for each binary problem (e.g., RH vs. LH plus RF).

Assumed that the P300 component evoked by target stimuli, could be enhanced by spatial filtering from nontarget response and noise, xDAWN aims at learning spatial filters that maximize signal to signal-plus-noise ratio (SSNR)[16]. The spatial filters v are optimized by

$$\underset{v}{argmax}\frac{v^T\,\Sigma_P v}{v^T\Sigma_X v} \tag{3}$$

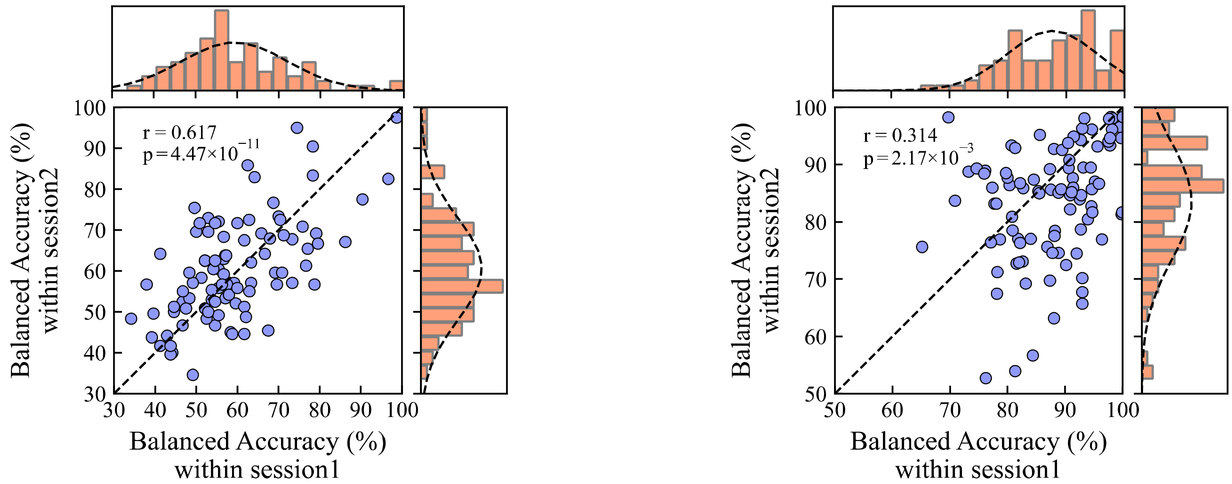
where Σ_P and Σ_X are the averaged spatial covariance matrix of target trials and all trials, respectively. In the same way as CSP does, the spatial filters v could be efficiently estimated by the generalized eigenvalue problem such that

$$\Sigma_P\, v = \lambda\, \Sigma_X\, v \tag{4}$$

In this study, the spatial filters v associated with the 4 largest generalized eigenvalue were obtained for spatial filtering. Next, the filtered signals were resampled with a factor 25 and then concatenated together to train an LDA classier.

D. Performance Measure

The performance of both paradigms was validated based on two-fold cross validation method from all trials within sessions (e.g., half of P300 trials in the first session were selected for test, while the other half in the first session were used as training set) according to the well-established approaches.



(a) Movement (b) P300

Figure 2. Session-to-session performance across all subjects by two-fold cross validation within sessions in each BCI paradigm, i.e., (a) Movement and (b) P300.

Due to the significant class imbalance in P300 experiment, we used the balanced classification accuracy as the performance measure for both paradigms[17]. The balanced accuracy is computed as

$$\frac{\sum_i^c \ (m_i / n_i)}{c} \tag{5}$$

where c is the number of classes, n_i is the true number of trials from i^{th} class and m_i is the number of trials that are correctly classified from i^{th} class.

Besides, Pearson's linear correlation coefficient was performed to illustrate the correlation between the decoding performance of SMR and P300 features.

III. Results

A. Statistical Analysis cross Sessions

The session-to-session performance evaluation of each individual paradigm is illustrated in Fig. 2 in the form of scatter plots. The correlation coefficient $r = 0.617$ with $p = 4.47\times10^{-11}$ indicated the repeatability between sessions for SMR features. Besides, the average balanced accuracies of the movement experiment were 59.4% (± 12.8) and 60.5% (± 12.5) for the first and second session respectively, while they were 87.6% (± 7.8) and 83.6% (± 10.2) for P300 experiment. Compared to P300 paradigm, the performance of movement paradigm exhibited lower accuracy and higher variance across subjects.

In this study, we defined the threshold for BCI inefficiency as 60.0% corresponding to movement paradigm, which equivalently represents the subject with balanced accuracy below 60.0% would be considered as individual who encountered the inefficiency problem of movement paradigm. In the same way, the threshold for P300 paradigm was set to 80.0%. As Table I shows, using interval statistics of balanced accuracy results for each paradigm, the proportion of BCI inefficiency in both paradigms were 57.5% and 24.7% respectively. Moreover, the proportion of subjects that achieved balanced accuracy above 80.0% in real movement experiment was only 5.9%, while the proportion achieved 75.3% in P300 experiment. The results indicated the problem of inefficiency is more prominent in movement paradigm compared to P300 paradigm.

TABLE I. Banlanced Accuracy(%) Interval Statistic in Each BCI Paradigm

<table>
<tr><th rowspan="2">Balanced Accuracy (%)</th><th colspan="4">Proportion (%) in Movement</th><th colspan="4">Proportion (%) in P300</th></tr>
<tr><th>Ses1</th><th>Ses2</th><th>Avg</th><th>Sum</th><th>Ses1</th><th>Ses2</th><th>Avg</th><th>Sum</th></tr>
<tr><td>80-100</td><td>4.3</td><td>7.5</td><td>5.9</td><td rowspan="2">43.5</td><td>82.8</td><td>67.7</td><td>75.3</td><td>75.3</td></tr>
<tr><td>60-79</td><td>36.6</td><td>38.7</td><td>37.6</td><td>17.2</td><td>29.0</td><td>23.1</td><td rowspan="4">24.7</td></tr>
<tr><td>40-59</td><td>54.8</td><td>51.6</td><td>53.2</td><td rowspan="3">56.5</td><td>0.0</td><td>3.2</td><td>1.6</td></tr>
<tr><td>20-39</td><td>4.3</td><td>2.2</td><td>3.3</td><td>0.0</td><td>0.0</td><td>0.0</td></tr>
<tr><td>0-19</td><td>0.0</td><td>0.0</td><td>0.0</td><td>0.0</td><td>0.0</td><td>0.0</td></tr>
</table>

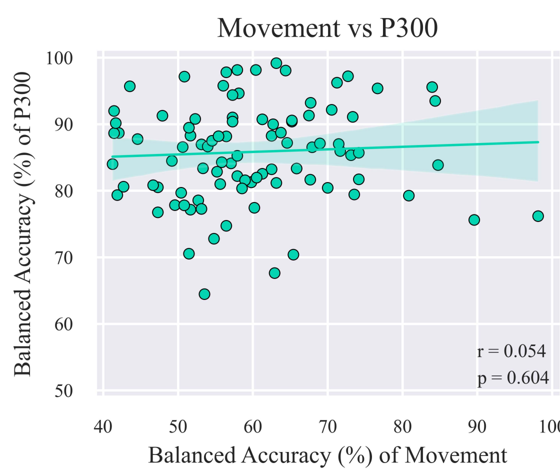


Figure 3. Paradigm-to-paradigm performance comparison with correlation coefficient r and p value of t-test across all subjects.

TABLE II. INEFFICIENCY CATEGORY STATISTIC OVER PARADIGMS

Inefficiency		Proportion (%)		
Movement	*P300*	*Ses1*	*Ses2*	*Avg*
×	×	36.6	31.2	**33.9**
×	√	4.3	15.1	**9.7**
√	×	46.2	36.6	**41.4**
√	√	12.9	17.1	**15.0**

B. Statistical Analysis cross Paradigms

As Fig. 3 shows, the paradigm-to-paradigm performance comparison and the correlation coefficient between both paradigms were calculated based on the average balanced accuracies in the first and second session. The correlation coefficient r was 0.054 and failed the t-test ($p = 0.604$) with the hypothesis of the existence of inefficiency correlation, which indicated there is no significant correlation between the decoding performance of SMR and P300 features within subjects.

Based on the statistics results on inefficiency category over paradigms in Table II, there were almost 41.4% subjects failed the movement task but performed the P300 task successfully. Meanwhile, about 9.7% subjects presented the opposing situation. The results presented that P300 paradigm could work for certain subjects whose EEG could not be recognized by movement paradigm, and movement paradigm could also provide an alternative for a part of subjects who were unable to be served by P300 paradigm.

IV. DISCUSSION

The proportion of the subjects considered as individuals encountering BCI inefficiency problem highly depends on the definition of threshold, while there are no acknowledged guidelines to determine it. For the purpose of practicability in reality, which means users could perform BCI tasks successfully with an acceptable tolerance, the thresholds for BCI inefficiency were defined as 60.0% and 80.0% corresponding to movement paradigm (three classes) and P300 paradigm (two classes) respectively in this work. Though the customized thresholds might lead to higher proportion of BCI inefficiency than the results proposed in the literatures[18], the results indicated that BCI paradigm based on endogenous potentials (i.e., movement) indeed presents more serious problem of inefficiency compared to BCI paradigm based on exogenous potentials (i.e., P300).

For the exploration of BCI inefficiency, in this work real executed movement is used instead of imagery movement to ensure the consistency of the movement and the participant's participation and eliminate the effect of train varied among participants. In spite of this, without the involvement of motor imagery, the evidence illustrated in this work could not directly demonstrate the in-correlation in inefficiency problem between MI-based BCIs and P300-based BCIs. However, it was previously reported that MI and real movement can produce similar SMR activity[4].

V. CONCLUSION

To investigate the problem of BCI inefficiency cross sessions and paradigms, we performed statistics analysis on a large-scale dataset with 93 subjects during two sessions of the same experiment procedure of movement and P300 task. Cross-sessions analysis indicated the reproducibility of the subject-level BCI inefficiency in both paradigms. In addition, the results not only showed that P300 features possess higher recognition accuracies than SMR features, but also demonstrated that the problem of inefficiency is greater in movement paradigm (56.5%) compared to P300 paradigm (24.7%) with inefficiency thresholds defined as 60.0% and 80.0% respectively. Cross-paradigms analysis indicated no significant correlation ($r = 0.054$, $p = 0.604$) between decoding performance of the two paradigms. The results also presented that P300 paradigm could be a substitute for a part of subjects (41.4%) whose EEG signals cannot be decoded effectively by movement paradigm, and vice versa (9.1%).

REFERENCES

[1] J. R. Wolpaw *et al.*, "Brain-computer interface technology: a review of the first international meeting," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 164-173, 2000.

[2] J. R. Wolpaw *et al.*, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, pp. 767-791, 2002.

[3] H. Cecotti, "Spelling with non-invasive Brain-Computer Interfaces - Current and future trends," *J. Physiol. Paris*, vol. 105, pp. 106-114, 2011.

[4] G. Pfurtscheller and F.H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clin. Neurophysiol.*, vol. 110, pp. 1842-1857, 1999.

[5] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," in *Proc. of IEEE*, vol. 89, pp. 1123–1134, 2001.

[6] D. J. Krusienski *et al.*, "A comparison of classification techniques for the P300 Speller," *J. Neural Eng.*, vol. 3, pp. 299-305, 2006

[7] M. Cheng *et al.*, "Design and implementation of a brain-computer interface with high transfer rates," *IEEE Trans. Biomed. Eng.*, vol. 49, pp. 1181-1186, 2002.

[8] K. K. Ang *et al.*, "A large clinical study on the ability of stroke patients to use an EEG-based motor imagery brain-computer interface," *Clin. EEG Neurosci.*, vol. 42, pp. 253-258, 2011.

[9] B. Blankertz *et al.*, "The Berlin Brain--Computer Interface: accurate performance from first-session in BCI-nave subjects," *IEEE Trans. Biomed. Eng.*, vol. 55, pp. 2452-2462, 2008.

[10] C. Guger *et al.*, "How many people are able to control a P300-based brain-computer interface (BCI)?" *Neurosci. Lett.*, vol. 462, pp. 94-98, 2009.

[11] C. Guger *et al.*, "How many people could use an SSVEP BCI?" *Front. Neurosci.*, 6:169, 2012.

[12] C. Guger *et al.*, "How many people are able to operate an EEG-based brain-computer interface (BCI)?" *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, pp. 145-147, 2003.

[13] X. Shu *et al.*, "Fast Recognition of BCI-Inefficient Users Using Physiological Features from EEG Signals: A Screening Study of Stroke Patients," *Front. Neurosci.*, vol. 12, no. 93, 2018.

[14] E. M. Hammer *et al.*, "Psychological predictors of SMR-BCI performance," *Biol. Psychol.*, vol. 89, pp. 80-86, 2012.

[15] B. Blankertz *et al.*, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Proc. Mag.*, vol. 25, pp. 41-56, 2008.

[16] B. Rivet *et al.*, "xDAWN algorithm to enhance evoked potentials: application to brain-computer interface," *IEEE Trans. Biomed. Eng.*, vol. 56, pp. 2035-2043, 2009.

[17] H. He and D. Wu, "Transfer learning for brain-computer interfaces: A Euclidean space data alignment approach," *IEEE Trans. Biomed. Eng.*, vol. 67, pp. 399–410, 2020.

[18] T. Liu *et al.*, "Reduce brain computer interface inefficiency by combining sensory motor rhythm and movement-related cortical potential features," *J. Neural Eng.*, vol. 17, no. 3, 2020.