Proceedings Paper:

Zhang, Y, Xie, SQ orcid.org/0000-0002-8082-9112, Li, Z et al. (3 more authors) (2022) CCA-based Spatio-temporal Filtering for Enhancing SSVEP Detection. In: 2022 IEEE-EMBS International Conference on Wearable and Implantable Body Sensor Networks (BSN). 2022 IEEE-EMBS International Conference on Wearable and Implantable Body Sensor Networks (BSN), 27-30 Sep 2022, Ioannina, Greece. IEEE . ISBN 9781665487917

https://doi.org/10.1109/bsn56160.2022.9928502

CCA-based Spatio-temporal Filtering for Enhancing SSVEP Detection

Yue Zhang, Sheng Quan Xie, Zhenhong Li, Yihui Zhao, Kun Qian, Zhi-Qiang Zhang
School of Electronic and Electrical Engineering
University of Leeds
Leeds, UK
elyzh@leeds.ac.uk, s.q.xie@leeds.ac.uk, z.h.li@leeds.ac.uk,
el14yz@leeds.ac.uk, el14kq@leeds.ac.uk, Z.Zhang3@leeds.ac.uk

***Abstract*—Brain-computer interface (BCI) can provide a direct communication path between the human brain and an external device. The steady-state visual evoked potential (SSVEP)-based BCI has been widely explored in the past decades due to its high signal-to-noise ratio and fast communication rate. Several spatial filtering methods have been developed for frequency detection. However the temporal knowledge contained in the SSVEP signal is not effectively utilized. In this study, we propose a canonical correlation analysis (CCA)-based spatio-temporal filtering method to improve target classification. The training signal and two types of template signals (i.e. individual template and artificial sine-cosine reference) are first augmented via temporal information. Three sets of augmented data are then concatenated by trials. The CCA is performed twice, between the newly obtained training data and each template. The trained four spatial filters can be applied in the following test process. A public benchmark dataset was used to evaluate the performance of the proposed method and the other three comparing methods, such as CCA, MsetCCA, and TRCA. The experimental results indicate that the proposed method yields significantly higher performance. This paper also explored the effects of the number of electrodes and training blocks on classification accuracy. The results further demonstrated the effectiveness of the proposed method in SSVEP detection.**

***Index Terms*—Brain-computer interface (BCI), electroencephalography (EEG), steady-state visual evoked potential (SSVEP), data augmentation**

I. Introduction

Brain-computer interface (BCI) is a human-computer interaction technology that enables people to communicate with an external device directly via brain activities [1]. Among various brain signals, the steady-state visual evoked potential (SSVEP)-based BCI system has been widely explored because it is non-invasive, low cost, and has relatively high information transfer rates (ITR) and signal-to-noise ratio (SNR). In recent decades, SSVEP-based BCI technology has been applied in many applications, such as robotic manipulator grasping [2], speller system [3] and wheelchair control [4].

SSVEPs are periodic neural signals that indicate the electrical responses to visual stimuli at specific frequencies and phases [5]. The evoked responses contain oscillations not only at the stimulus frequency but also its higher harmonics [6]. In the past decades, many target recognition methods were investigated to analyze the features of SSVEPs and detect the subject's intent to operate the peripheral device. The spatial filter-based methods conventionally leverage multi-channel data to achieve target detection, such as minimum energy combination (MEC), common spatial pattern (CSP), multivariate synchronization index (MSI), and canonical correlation analysis (CCA) [7]. Among these methods, CCA gained the most attention due to its ease of use and high efficiency. In recent years, many methods have been proposed to further improve its classification performance. The representatives include extended CCA (eCCA) [8], L1-regularized multiway CCA (L1-MCCA) [9], multiset CCA (MsetCCA) [10], filter bank CCA (FBCCA) [11] and task-related component analysis (TRCA) [12]. Various studies have demonstrated that TRCA is more effective than CCA, MsetCCA, and eCCA in SSVEP-based BCI [12], [13]. Recently, several CCA-based methods employed the concatenation idea in SSVEP classification, showing better performance than TRCA. Wei *et.al* [14] proposed a training data-driven CCA (tdCCA), in which concatenated training data and individual templates were used as the input of CCA to train spatial filters. Similarly, Yuan *et.al* [15] proposed a method in which spatial filters are trained using concatenated individual training signals and sine-cosine reference signals. Although the methods mentioned above have shown effectiveness in SSVEP detection, there is still potential for improvement. The spatial filters designed in most previous methods normally have two functions: to optimise reference signals in CCA such as L1-MCCA and MsetCCA [16] or to optimise a separate correlation analysis procedure (between test data and individual templates) such as TRCA [12] and tdCCA [14]. However, the temporal information in SSVEP signals is not fully utilised and may contribute to improving the classification performance of a SSVEP-based BCI system.

In this study, we proposed a CCA-based spatio-temporal filtering method to enhance the SSVEP recognition performance. The training trial, individual templates and sine-cosine reference signals are all augmented via temporal information. Thus, the augmented data consists of the original data and its multiple time-delayed copies. Correlation analysis is firstly

This work was supported in part by Engineering and Physical Sciences Research Council (EPSRC) (Grant No. EP/S019219/1) and in part by China Scholarship Council (CSC) (Grant No. 201906460007). (Corresponding author: Zhi-Qiang Zhang)

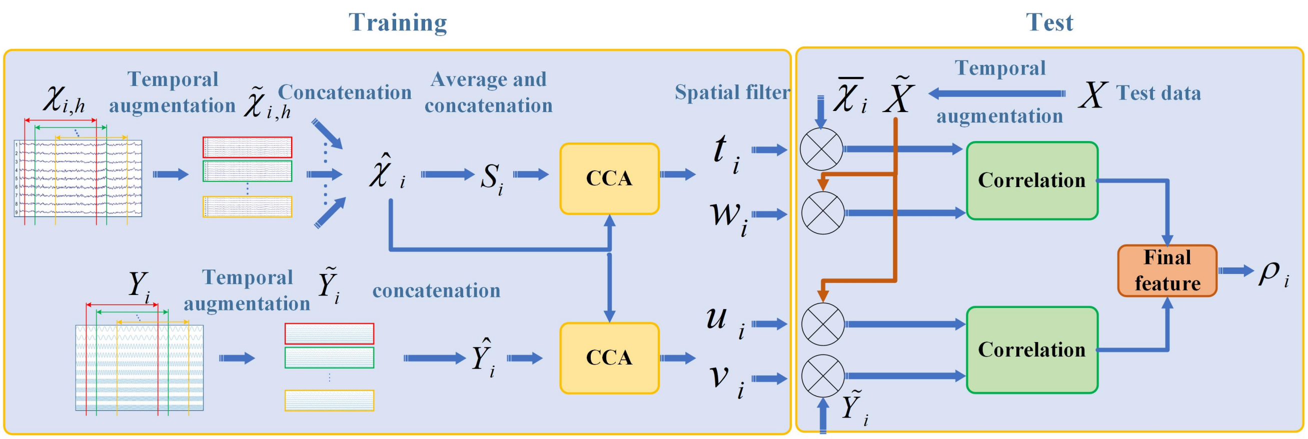


Fig. 1. Diagram of the proposed CCA-based spatio-temporal filtering method for SSVEP classification.

performed between two concatenated matrices constituted by augmented training data and individual templates. Then, CCA is employed again between augmented training data and sine-cosine reference signals. Therefore, four spatial filters are obtained during the training process. The classification performance was evaluated on a 40-target public benchmark dataset. The results demonstrated that the proposed method outperformed CCA, MsetCCA, and TRCA with an average classification accuracy of 90.5% at 1s time window (TW).

The remaining paper is arranged as follows: The dataset and CCA-based spatio-temporal filtering target recognition method are described in Section II. The results and discussion are presented in Section III. Section IV provides the conclusion.

II. Material and Method

A. Dataset Description

In the benchmark dataset [3], SSVEP data was recorded from thirty-five participates (seventeen females and eighteen males, mean age: twenty-two years). All people were healthy and had normal or corrected to normal vision.

1) Stimulus Design: The stimulation interface includes 5×8 stimulus matrix coded using a joint frequency and phase modulation method. The frequencies range from 8 Hz to 15.8 Hz with an interval of 0.2 Hz. The phase difference between two neighboring stimuli is 0.5π. For each participant, the data contains six blocks of forty trials associated with forty stimuli.

2) EEG Recording: The data were selected from nine electrodes, namely Pz, PO5, PO3, POz, PO4, PO6, O1, Oz, O2. The ground was located between Fz and FPz. The reference was placed on the vertex. The sample rate is 250 Hz.

B. Data Preprocessing

Taking into account the latency delay in human visual system, the signal behind 0.14s is retrieved for analysis. The Chebyshev Type I Infinite Impulse Response (IIR) filter was applied in this work to create band-pass filters. The data were filtered between seven Hz and ninety Hz.

C. Method Description

This study proposed a CCA-based spatio-temporal filtering method to enhance SSVEP detection. The framework of the proposed method was shown in Fig. 1. Suppose $\boldsymbol{\chi}_{i,h} \in \mathbb{R}^{N_c \times N_s}$ is the h-th training trial from i-th frequency where $h = 1, 2, ..., N_t$ and $i = 1, 2, ..., N_f$. Hereafter, N_c, N_s, N_t, and N_f represent the number of channels, the number of samples, the number of training trials, and the number of the frequencies, respectively. For each training trial, an augmented data matrix $\widetilde{\boldsymbol{\chi}} \in \mathbb{R}^{(d+1)N_c \times N_s}$ is defined as:

$$\widetilde{\boldsymbol{\chi}}_{i,h} = [\boldsymbol{\chi}_{i,h}^{\mathrm{T}}, \boldsymbol{\chi}_{i,h,1}^{\mathrm{T}}, ..., \boldsymbol{\chi}_{i,h,d}^{\mathrm{T}}]^{\mathrm{T}} \tag{1}$$

where $\boldsymbol{\chi}_{i,h,d} \in \mathbb{R}^{N_c \times N_s}$ represents training trial $\boldsymbol{\chi}_{i,h}$ delayed by d samples. Thus, the augmented data matrix contains both the original training trial and multiple time-delayed copies [17], [18]. $\hat{\boldsymbol{\chi}}_i = [\widetilde{\boldsymbol{\chi}}_{i,1}, \widetilde{\boldsymbol{\chi}}_{i,2}, ..., \widetilde{\boldsymbol{\chi}}_{i,N_t}] \in \mathbb{R}^{(d+1)N_c \times (N_s \cdot N_t)}$ is the continuous training data constructed by concatenating N_t training trials. The augmented individual template is $\overline{\boldsymbol{\chi}}_i = \frac{1}{N_t} \sum_{h=1}^{N_t} \widetilde{\boldsymbol{\chi}}_{i,h} \in \mathbb{R}^{(d+1)N_c \times N_s}$ which is obtained by averaging all augmented training trials. Similarly, the continuous individual template is concatenated as $S_i = [\overline{\boldsymbol{\chi}}_i, \overline{\boldsymbol{\chi}}_i, ..., \overline{\boldsymbol{\chi}}_i] \in \mathbb{R}^{(d+1)N_c \times (N_s \cdot N_t)}$. SSVEP signals could also be characterized by sine-cosine waves, and the reference signal $\boldsymbol{Y}_i \in \mathbb{R}^{2N_h \times N_s}$ for i-th stimulus can be defined as:

$$\boldsymbol{Y}_i = \begin{bmatrix} \sin(2\pi ft) \\ \cos(2\pi ft) \\ \vdots \\ \sin(2\pi N_h ft) \\ \cos(2\pi N_h ft) \end{bmatrix}, \; t = [1/F_s, 2/F_s, ..., N_s/F_s] \tag{2}$$

where N_h represents the number of harmonics, f is stimulation frequency, and F_s is the sampling rate (i.e. 250 Hz in this study). Similarly, the augmented reference signal $\widetilde{\boldsymbol{Y}}_i$ is constructed by original signal and its time-delay copies. Thus, the concatenated reference signal is represented as $\hat{\boldsymbol{Y}}_i = [\widetilde{\boldsymbol{Y}}_i, \widetilde{\boldsymbol{Y}}_i, ..., \widetilde{\boldsymbol{Y}}_i] \in \mathbb{R}^{2(d+1)N_h \times (N_s \cdot N_t)}$.

The CCA are performed twice for each stimulus in the training stage. Firstly, CCA tries to seeks a pair of spatial filters $\boldsymbol{w}_i \in \mathbb{R}^{(d+1)N_c\times 1}$ and $\boldsymbol{w}_s \in \mathbb{R}^{(d+1)N_c\times 1}$ so that the correlation between two projections $\boldsymbol{w}_x\hat{\boldsymbol{\chi}}_i$ and $\boldsymbol{w}_s\boldsymbol{S}_i$ can be maximized as follows:

$$r_i^1 = \max_{\boldsymbol{w}_i,\boldsymbol{t}_i} \frac{E[\boldsymbol{w}_i^{\mathrm{T}}\hat{\boldsymbol{\chi}}_i\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{t}_i]}{\sqrt{E[\boldsymbol{w}_i^{\mathrm{T}}\hat{\boldsymbol{\chi}}_i\hat{\boldsymbol{\chi}}_i^{\mathrm{T}}\boldsymbol{w}_i]}\sqrt{E[\boldsymbol{t}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{t}_i]}} \tag{3}$$

Correlation analysis is applied again between $\hat{\boldsymbol{\chi}}_i$ and $\hat{\boldsymbol{Y}}_i$:

$$r_i^2 = \max_{\boldsymbol{u}_i,\boldsymbol{v}_i} \frac{E[\boldsymbol{u}_i^{\mathrm{T}}\hat{\boldsymbol{\chi}}_i\boldsymbol{Y}_i^{\mathrm{T}}\boldsymbol{v}_i]}{\sqrt{E[\boldsymbol{u}_i^{\mathrm{T}}\hat{\boldsymbol{\chi}}_i\hat{\boldsymbol{\chi}}_i^{\mathrm{T}}\boldsymbol{u}_i]}\sqrt{E[\boldsymbol{v}_i^{\mathrm{T}}\hat{\boldsymbol{Y}}_i\hat{\boldsymbol{Y}}_i^{\mathrm{T}}\boldsymbol{v}_i]}} \tag{4}$$

The augmentation procedure is also applied to the text data $\boldsymbol{X} \in \mathbb{R}^{N_c\times N_s}$. Suppose the augmented test data is $\widetilde{\boldsymbol{X}} \in \mathbb{R}^{(d+1)N_c\times N_s}$. Once the spatial filter $\boldsymbol{w}_i$, $\boldsymbol{t}_i$, $\boldsymbol{u}_i$, and $\boldsymbol{v}_i$ for i-th stimulus are obtained, the following two Pearson correlation coefficients can be calculated as follows:

$$\rho_i^1 = corr(\widetilde{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{w}_i, \overline{\boldsymbol{\chi}}_i^{\mathrm{T}}\boldsymbol{t}_i),\ i = 1, 2, ..., N_f \tag{5}$$

$$\rho_i^2 = corr(\widetilde{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{u}_i, \widetilde{\boldsymbol{Y}}_i^{\mathrm{T}}\boldsymbol{v}_i),\ i = 1, 2, ..., N_f \tag{6}$$

The above two correlation coefficients are weighted as the final feature for SSVEP recognition:

$$\rho_i = \sum_{n=1}^{2} sign(\rho_i^n)(\rho_i^n)^2,\ i = 1, 2, ..., N_f \tag{7}$$

The frequency of the test data $\boldsymbol{X}$ can be determined by the following equation:

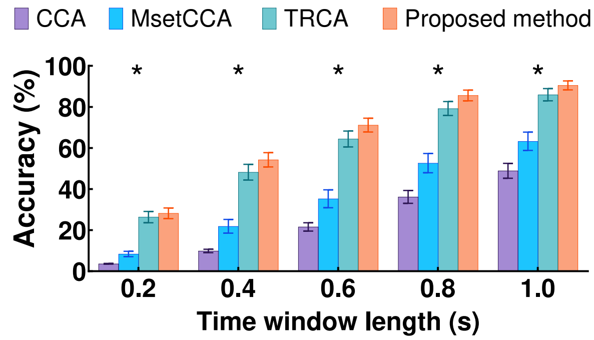
$$f = \underset{f_i}{argmax}\,\rho_i,\ i = 1, 2, ..., N_f \tag{8}$$

III. Results and Discussion

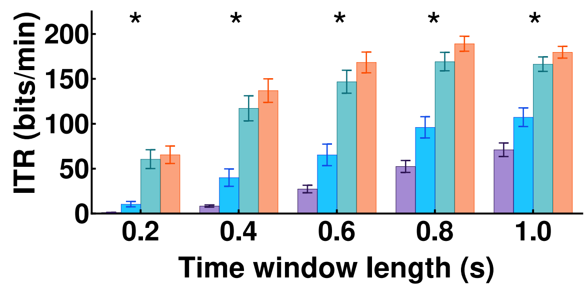
A. Performance Evaluation

Fig. 2 shows performance comparison results for CCA, Msetcca, TRCA, and the proposed method in terms of (a) average classification accuracy and (b) ITRs with the public benchmark dataset. Empirically, $N_h = 5$, $d = 1$ for this dataset. The results illustrate that the proposed method outperformed three comparing methods at all TWs. The highest accuracy for the proposed method is 90.50% with 1 s TW, whereas the highest ITRs is 189.12 bits/min with 0.8 s TW. The one-way repeated measures ANOVA was performed to explore the difference in accuracy and ITRs between the four methods. The results revealed that there are statistically significant differences between these methods at all data lengths.

Fig. 3 shows the probability density of classification accuracy across all subjects for the proposed method and three comparing methods. The violin indicates not only the median values but also the distribution of numeric data. It is apparent that the accuracy provided by all methods show scattered distributions, this is because the dataset collected SSVEP signals from thirty-five subjects. More subjects may result in the violin plot showing more scattered distributions. As shown in Fig. 3, the proposed method always indicates higher median values



(a) Classification accuracy



(b) ITRs

Fig. 2. Performance comparison among various methods using different TWs in terms of (1) averaged classification accuracy and (b) average ITRs across subjects. The error bars indicate standard errors. The asterisks represent significant difference between the four methods provided by one-way repeated-measures ANOVA ($* \, p < 0.0001$).

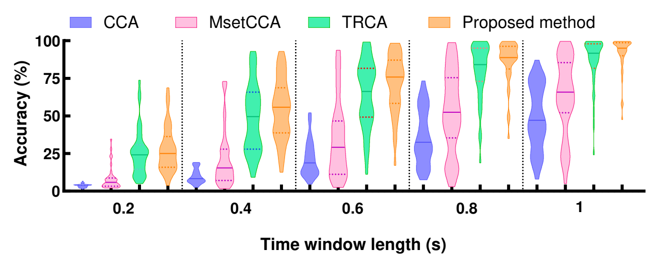


Fig. 3. Violin plots represent the distributions of SSVEP recognition accuracy of all subjects achieved by the four methods with various TWs on the benchmark dataset. Black solid line in each violin indicates median, and two black dotted lines represent interquartile ranges (25% and 75% percentiles).

and a more concentrated distribution. It indicates that the proposed CCA-based spatio-temporal filtering method could attain a more stable performance for SSVEP classification.

B. Discussion

The proposed method augments the training data and templates by considering temporal information. Thus, more feature-related knowledge was included in the training process. Besides, in this study, the CCA was employed twice in which the correlation analysis was performed between the training data and each type of template. As a result, both individual-and frequency-related features are extracted in the spatial filters, which results in higher classification performance.

We further explore the influence of the number of channels on classification accuracy. Fig. 4 depicts the averaged accuracy of four methods with different numbers of channels at 0.6s TW. The heat map depicted the comparison results among the four recognition methods. The x-axis indicates the method with corresponding number of channels (from five to nine). The y-axis refers to the subject index, ranging from one to thirty-five. The darker colors indicate that the corresponding method and number of electrodes provide higher accuracy. The proposed method usually shows deeper color than the other three methods with various numbers of electrodes. Besides, when the number of channels increases, the accuracy for four methods generally increases simultaneously.

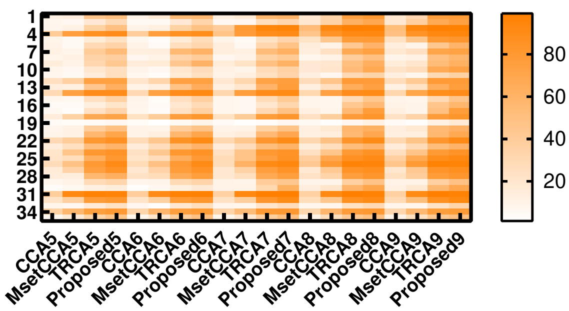


Fig. 4. Heat maps of the SSVEP detection accuracy of four methods with different number of channels at 0.6s TW.

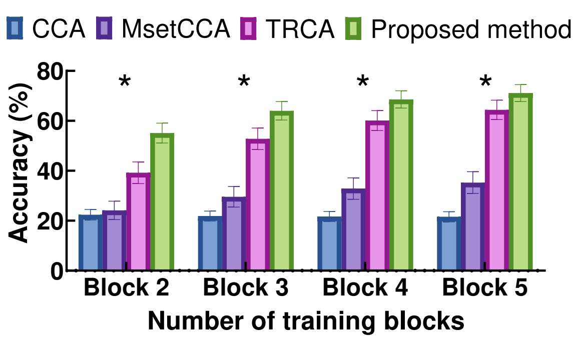


Fig. 5. Barcharts of the four methods' classification accuracy with different numbers of training blocks. The error bars represent standard errors. The asterisks represent significant difference between the four methods provided by one-way repeated-measures ANOVA ($* \, p < 0.0001$).

Fig. 5 depicts how the number of training blocks affects the SSVEP detection performance of the four methods at a 0.6 s TW. It is obvious that the proposed method always provides the best performance. Besides, one-way repeated-measures ANOVA shows that there are significant differences among the four methods with various numbers of training blocks.

IV. Conclusion

In this study, a CCA-based spatio-temporal filtering method was proposed to enhance the classification performance of SSVEP-based BCI systems. Our method achieved data augmentation for the training data, individual template, and sine-cosine reference signals via temporal information. The CCA was employed twice, firstly between augmented training data and individual templates, and secondly between augmented training data and artificial references. Thus, four spatial filters are applied in the test stage. Results based on a public dataset showed that the proposed method could achieve higher classification performance than some popular methods.

References

[1] C. Yang, X. Yan, Y. Wang, Y. Chen, H. Zhang, and X. Gao, "Spatio-temporal equalization multi-window algorithm for asynchronous SSVEP-based BCI," *J. Neural Eng.*, vol. 18, no. 4, p. 0460b7, 2021.

[2] K. Wang, D.-H. Zhai, Y. Xiong, L. Hu, and Y. Xia, "An MVMD-CCA Recognition Algorithm in SSVEP-Based BCI and Its Application in Robot Control," *IEEE Trans. Neural. Netw. Learn. Syst.*, 2021.

[3] Y. Wang, X. Chen, X. Gao, and S. Gao, "A benchmark dataset for SSVEP-based brain–computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 10, pp. 1746–1752, 2016.

[4] Y. Li, J. Pan, F. Wang, and Z. Yu, "A hybrid BCI system combining P300 and SSVEP and its application to wheelchair control," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 11, pp. 3156–3166, 2013.

[5] J. Zhao, W. Zhang, J. H. Wang, W. Li, C. Lei, G. Chen, Z. Liang, and X. Li, "Decision-making selector (DMS) for integrating CCA-based methods to improve performance of SSVEP-based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 5, pp. 1128–1137, 2020.

[6] Y. Chen, C. Yang, X. Ye, X. Chen, Y. Wang, and X. Gao, "Implementing a calibration-free SSVEP-based BCI system with 160 targets," *J. Neural Eng.*, vol. 18, no. 4, p. 046094, 2021.

[7] Y. Zhang, S. Q. Xie, H. Wang, and Z. Zhang, "Data analytics in steady-state visual evoked potential-based brain–computer interface: A review," *IEEE Sensors J.*, vol. 21, no. 2, pp. 1124–1138, 2020.

[8] M. Nakanishi, Y. Wang, Y.-T. Wang, Y. Mitsukura, and T.-P. Jung, "A high-speed brain speller using steady-state visual evoked potentials," *Int. J. Neural Syst.*, vol. 24, no. 06, p. 1450019, 2014.

[9] Y. Zhang, G. Zhou, J. Jin, M. Wang, X. Wang, and A. Cichocki, "L1-regularized multiway canonical correlation analysis for SSVEP-based BCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 6, pp. 887–896, 2013.

[10] Y. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Frequency recognition in SSVEP-based BCI using multiset canonical correlation analysis," *Int. J. Neural Syst.*, vol. 24, no. 04, p. 1450013, 2014.

[11] X. Chen, Y. Wang, S. Gao, T.-P. Jung, and X. Gao, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface," *J. Neural Eng.*, vol. 12, no. 4, p. 046008, 2015.

[12] M. Nakanishi, Y. Wang, X. Chen, Y.-T. Wang, X. Gao, and T.-P. Jung, "Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 1, pp. 104–112, 2017.

[13] Y. Zhang, Z. Li, S. Q. Xie, H. Wang, Z. Yu, and Z.-Q. Zhang, "Multi-objective optimisation-based high-pass spatial filtering for SSVEP-based brain-computer interfaces," *IEEE Trans. Instrum. Meas.*, 2022.

[14] Q. Wei, S. Zhu, Y. Wang, X. Gao, H. Guo, and X. Wu, "A training data-driven canonical correlation analysis algorithm for designing spatial filters to enhance performance of SSVEP-Based BCIs," *Int. J. Neural. Syst.*, vol. 30, no. 05, p. 2050020, 2020.

[15] X. Yuan, Q. Sun, L. Zhang, and H. Wang, "Enhancing detection of SSVEP-based BCIs via a novel CCA-based method," *Biomed. Signal. Proces.*, vol. 74, p. 103482, 2022.

[16] M. Nakanishi, Y. Wang, Y.-T. Wang, and T.-P. Jung, "A comparison study of canonical correlation analysis based methods for detecting steady-state visual evoked potentials," *PloS one*, vol. 10, no. 10, 2015.

[17] F. Qi, Y. Li, and W. Wu, "RSTFC: A novel algorithm for spatio-temporal filtering and classification of single-trial EEG," *IEEE Trans. Neural. Netw. Learn. Syst.*, vol. 26, no. 12, pp. 3070–3082, 2015.

[18] B. Liu, X. Chen, N. Shi, Y. Wang, S. Gao, and X. Gao, "Improving the performance of individually calibrated SSVEP-BCI by task-discriminant component analysis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1998–2007, 2021.