2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Filter bank approach for enhancement of supervised Canonical Correlation Analysis methods for SSVEP-based BCI spellers

Mario Corral Bolaños, Sheyla Barrado Ballestero and Sadasivan Puthusserypady

Abstract— **Canonical correlation analysis (CCA) is one of the most used algorithms in the steady-state visual evoked potentials (SSVEP)-based brain-computer interface (BCI) systems due to its simplicity, efficiency, and robustness. Researchers have proposed modifications to CCA to improve its speed, allowing high-speed spelling and thus a more natural communication. In this work, we combine two approaches, the filter-bank (FB) approach to extract more information from the harmonics, and a range of different supervised methods which optimize the reference signals to improve the SSVEP detection. The proposed models are tested on the publicly available benchmark dataset for SSVEP-based BCIs and the results show improved performance compared to the state-of-the-art methods and, in particular, the proposed FBMwayCCA approach achieves the best results with an information transfer rate (ITR) of 134.8±8.4 bits/minute. This study indeed suggests the feasibility of combining the fundamental and harmonic SSVEP components with supervised methods in target identification to develop high-speed BCI spellers.**

I. INTRODUCTION

Brain-computer interface (BCI) systems provide an alternative non-muscular communication method by translating brain electrical activities into specific control commands. BCI technologies can be exploited in a wide range of applications [1], [2] among which the BCI-spellers stand out as speech impairment has a critical impact on the quality of life of disabled individuals [3]. Typically, BCI systems rely on the electroencephalogram (EEG) signals features to work. There are three major paradigms in BCI-spellers, namely P300-event-related potential, steady-state visual evoked potentials (SSVEP) and motor imagery (MI) [4]. Of these, the SSVEP-based BCI spellers have attracted much attention due to their high signal-to-noise ratio (SNR) and high communication rate. SSVEP are characterized as voltage oscillations in the visual cortex whose frequency is matching the frequency (and its harmonics) of an external visual stimulus flickering at a specific frequency [5].

Traditionally, the Power Spectral Density Analysis (PSDA) methods are used to identify the flickering stimuli [6]. However, PSDA is sensitive to noise when the signal to be analyzed comes from a single channel. Since then, different multichannel-based methods have been proposed, such as the maximum contrast combination (MCC) [7], the multivariate synchronization index (MSI) [8] or the canonical correlation analysis (CCA), among others. In recent years, deep learning approaches have also been studied [9], [10].

The first two authors contributed equally to this work.

The authors are with the Department of Health Technology, Technical University of Denmark, Kgs. Lyngby 2800, Denmark. {s203065, s203064}@student.dtu.dk, {sapu}@dtu.dk

CCA is initially proposed by Lin *et al.* [11] and became the standard paradigm since its introduction. The CCA-based methods calculate the correlation coefficients between an SSVEP response and reference signals at each stimulus frequency. Many of the improvements of the standard CCA method aim to develop optimal or individual reference signals instead of using artificial sinusoidal signals. Those methods are the so-called supervised methods. Some examples are the individual template based CCA (IT-CCA) [12], the extended CCA (eCCA) [13], the multiway CCA (MwayCCA) [14], the multiset CCA (MsetCCA) [15], the correlated component analysis (CORRCA) [16], or the task-related component analysis (TRCA) [17]. These methods achieve better performance than the standard CCA, although the improvement is dependent on the number of training trials. Recently, the multi-stimulus paradigm has been proposed to overcome this challenge [18]. In parallel, Chen *et al.* [19] introduced a filter bank approach (FBCCA) to take advantage of the harmonics information and enhance the performance without requiring training data.

In this work, we investigate the suitability of combining two types of approaches in the CCA framework: (i) the FBCCA to extract as much information as possible from the harmonics, and (ii) the reference signals optimization to adapt the reference signals to the individual characteristics of the user. In particular, we evaluate the performance of MwayCCA, L1-MwayCCA, MsetCCA and IT-CCA within a FB implementation in order to motivate the development of new paradigms which follow this strategy to achieve the improvement of high-speed BCI spelling systems.

II. MATERIALS AND METHODS

A. Dataset

This study uses the publicly available benchmark dataset for SSVEP [20]. It consists of SSVEP signals from 35 subjects collected using a 64-channels EEG system. For each subject, the experiment included six blocks, each one containing 40 trials corresponding to 40 characters in the speller. Each of these characters is coded within a frequency range of [8−15.8 Hz], with 0.2 Hz separation between them.

B. Preprocessing

Following [19] preprocessing steps, the signals are band-pass [5−90 Hz] filtered and the initial and last 0.5 seconds of every trial are discarded. Then, the same nine parietal and occipital channels (Pz, PO3, PO5, PO4, PO6, POz, O1, Oz and O2) are used.

C. Existing methods

1) Canonical Correlation Analysis (CCA): A multivariate statistical method used to infer the correlation between two multidimensional variables, $\mathbf{X} \in \mathbb{R}^{I_1 \times J}$ and $\mathbf{Y} \in \mathbb{R}^{I_2 \times J}$. It aims to find two vectors $\mathbf{w} \in \mathbb{R}^{I_1}$ and $\mathbf{v} \in \mathbb{R}^{I_2}$ which maximize the correlation between the linear transformations $\tilde{\mathbf{x}} = \mathbf{w}^T\mathbf{X}$ and $\tilde{\mathbf{y}} = \mathbf{v}^T\mathbf{Y}$, through optimizing the expression:

$$\rho = \max_{w,v} \frac{E[\tilde{\mathbf{x}}\tilde{\mathbf{y}}^T]}{\sqrt{E[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T]E[\tilde{\mathbf{y}}\tilde{\mathbf{y}}^T]}}, \tag{1}$$

where ρ denotes the correlation coefficient. To implement CCA, we designed M matrices corresponding to the M stimulation frequencies ($f_m : m = 1, 2, \ldots, M$) of the SSVEP system. Each matrix (denoted as $\mathbf{Y}_m$), was constructed as a series of sine-cosine waves and work as the reference signal at the m^{th} stimulus frequency and their H harmonics (hence, $I_2 = 2H$). On the other hand, $\mathbf{X}$ represents the multi-channel EEG signal, where I_1 denotes the number of channels and J the number of time points of the signal. Finally, solving the optimization problem between the EEG data and each group of reference signals, the set with the highest correlation is considered the selected frequency.

2) Multiway CCA (MwayCCA): This method aims to optimize the reference signals by maximizing the correlation between the recorded EEG tensor, $\boldsymbol{\chi} \in \mathbb{R}^{I \times J \times K}$ (channel $\times$ time $\times$ trial), and the reference sinusoidal waves. MwayCCA maximizes the correlation between the linear combinations $\tilde{\mathbf{x}} = \boldsymbol{\chi} \times_1 \tilde{\mathbf{w}}_1^T \times_3 \tilde{\mathbf{w}}_3^T$ (where $\times_n$ represents the *mode-n product*) and $\tilde{\mathbf{y}} = \mathbf{v}^T\mathbf{Y}$ through finding the coefficients $\mathbf{w}_1 \in \mathbb{R}^I$, $\mathbf{w}_3 \in \mathbb{R}^K$ and $\mathbf{v} \in \mathbb{R}^{2H}$, leading to an equivalent optimization problem as in Eq. (1). The optimal signal reference for a specific frequency f_m is modeled as:

$$\mathbf{z}_m = \boldsymbol{\chi} \times_1 \tilde{\mathbf{w}}_1^T \times_3 \tilde{\mathbf{w}}_3^T, \quad \mathbf{z}_m \in \mathbb{R}^J. \tag{2}$$

To minimize the possible negative contribution to the reference signal optimization of those trials with artifacts, Zhang *et al.* [14] proposed the implementation of L1-regularization on trial-way, leading to the variation L1-MwayCCA.

3) Multiset CCA (MsetCCA): A generalization of CCA, in which the linear relations between more than two sets of variables are analyzed. This method extracts common features along various trials to develop more natural references than artificial sinusoids. In [15], MsetCCA is implemented to find the spatial filters $\mathbf{w}_{1,m}, \ldots, \mathbf{w}_{K,m}$ that maximizes the overall correlation among the canonical variates $\tilde{\mathbf{z}}_{1,m}, \ldots, \tilde{\mathbf{z}}_{K,m}$, where $\tilde{\mathbf{z}}_{i,m} = \mathbf{w}_{i,m}^T\mathbf{X}_{i,m}$. These canonical variates represent the common features shared among the training data, and the optimized reference signal for a certain stimulus f_m can be calculated as:

$$\mathbf{Y}_m = [\tilde{\mathbf{z}}_{1,m}^T, \ldots, \tilde{\mathbf{z}}_{K,m}^T]^T, \quad \mathbf{Y}_m \in \mathbb{R}^{K \times J}. \tag{3}$$

4) Individual Template CCA (IT-CCA): In this case, the reference template is obtained by averaging the training trials of each target as shown below:

$$\mathbf{Y}_m = \frac{1}{K}\sum_{k=1}^{K} \mathbf{X}_k, \quad \mathbf{Y}_m \in \mathbb{R}^{I \times J}. \tag{4}$$

5) Filter Bank CCA (FBCCA): The goal of FB analysis is to decompose SSVEPs into sub-band components so that independent information hidden in the harmonic components can be extracted more efficiently than with standard CCA. We generated n sub-bands applying a high-pass zero-phase Butterworth filter with the cut-offs frequencies for the n^{th} band in the $8n$ Hz. CCA is then applied and for each set of n sub-bands, the obtained correlations are weighted and combined according to:

$$\tilde{\rho}_k = \sum_{n=1}^{N} w(n) \cdot (\rho_k^n)^2, \tag{5a}$$

$$w(n) = n^{-a} + b, \quad n \in [1, N], \tag{5b}$$

where N represents the total number of sub-bands, a, b are two hyperparameters that define the importance of each sub-band, and ρ_k^n represents the obtained correlation for the n^{th} sub-band and the k^{th} stimulus. Finally, the maximum value of $\tilde{\rho}_k$ is used to classify the selected frequency.

D. Filter Bank Enhanced Methods

Our proposal of incorporating the sub-band decomposition to the supervised algorithms represents a natural extension of the FB paradigm and it is applicable to each one of the aforementioned methods.

Consider the EEG tensor $\boldsymbol{\chi}_m \in \mathbb{R}^{I \times J \times K}$ (channel × time × trial) for a particular stimulus frequency, f_m. Firstly, the FB decomposition to every signal is performed, resulting in the four-way tensor, $\boldsymbol{\chi}_m' \in \mathbb{R}^{I \times J \times K \times N}$. Then, the desired optimization method is applied in each of these decomposed sub-bands. Each stimulus frequency thus requires references $\mathbf{Y}_m \in \mathbb{R}^{2H \times J \times N}$, where the third dimension corresponds to each sub-band. When necessary, the artificial sinusoids for each n^{th} sub-band started using frequencies from the n^{th} harmonic, since the n^{th} sub-band mainly aims to extract information from the n^{th} harmonic. Finally, CCA between the sub-band components of new test data and the optimized reference signals is applied and the correlations are combined using Eq. (5a).

E. Experimental Evaluation

We use a leave-one-out cross-validation at block level to evaluate the performance. The nine selected channels are used in the EEG tensor. The number of harmonics (N_h) in the reference signals for the standard CCA and MwayCCA methods is set to 3, while the number of sub-bands (N_{bands}), and the hyperparameters a and b for the FB methods is selected using grid-search and cross-validation in the FBCCA for a fixed time window of 1.5 seconds. The grid-search is performed on the ranges $a \in [1, 2]$, and $b \in [0, 1]$ with steps of 0.25 and N_{bands} from 3 to 6. The values with the highest accuracy are selected. In the case of L1-MwayCCA, the regularization parameter, λ_3, is adjusted at subject level using cross-validation with values between 0.01 and 0.1 with steps of 0.01.

The accuracy and the information transfer rate (ITR in bits/minute, Eq. (6)) are used as performance metrics.

$$ITR = \frac{60}{T}\left(\log_2 M + P\log_2 P + (1-P)\log_2 \frac{1-P}{M-1}\right), \quad (6)$$

where M denotes the number of classes, P the classification accuracy and T the average target selection time (in seconds). In our case, T is the time window employed, which is 0.64s (0.5s of a target cue presented to the subject, and 0.14s to the gaze shifting time) [20]. One-way ANOVA is performed to analyze the significant differences in each time window. Then, paired performance differences are statistically analyzed through *post-hoc* paired sample t-tests. Bonferroni correction is applied to deal with the increased risk of type I errors due to repetitive testing. The results are presented as mean $\pm$ standard error of the mean.

III. RESULTS

A. Overall Performance Comparison

One-way ANOVA showed significant accuracy differences in all the time windows for the compared methods. The lowest p-values are obtained for data lengths $t = 0.75$ s (p-value $< 10^{-6}$) and $t = 1.0$ s (p-value $< 10^{-5}$). Figure 2 depicts the *post-hoc* paired t-test differences in the performance between the different methods and their FB extension. Results showed that the FB implementation outperforms the standard version of the methods, with the only exceptions of data lengths of 0.25 s and for the MsetCCA and ITCCA implementations for short time windows (time ≤ 0.75 and 0.5 s, respectively). In all the cases, the maximum ITR is obtained with the FB approach (See Table I). The maximum ITR is achieved by the FBMwayCCA method (134.8 $\pm$ 8.4), while its regularized version gave very similar results, although statically different (t-test, ITR comparison p-value < 0.05). T-paired tests did not show significant differences between the maximum ITR of the L1-MwayCCA, MwayCCA, FBITCCA, FBMsetCCA and FBCCA. The obtained results with the FBCCA are in accordance with the graphical results presented in [20]. Slightly decreased performance in our model for shorter times than one second might be attributed to the use of a different filter type in the sub-band decomposition, an inferior number of sub-bands, or the number of harmonics in the reference sets.

TABLE I

AVERAGE PERFORMANCE OF EACH METHOD FOR THEIR MAXIMUM ITR

Method	Time (s)	Accuracy	ITR (bits/min)
CCA	1.75	0.74 $\pm$ 0.04	84.4 $\pm$ 6.3
FBCCA	1.5	0.81 $\pm$ 0.03	107.2 $\pm$ 6.0
MwayCCA	1.0	0.71 $\pm$ 0.04	114.2 $\pm$ 9.1
FBMwayCCA	1.0	0.79 $\pm$ 0.04	134.8 $\pm$ 8.4
L1-MwayCCA	1.0	0.71 $\pm$ 0.04	114.5 $\pm$ 9.2
L1-FBMwayCCA	1.0	0.79 $\pm$ 0.04	133.6 $\pm$ 8.5
MsetCCA	1.0	0.65 $\pm$ 0.04	103.2 $\pm$ 9.7
FBMsetCCA	1.25	0.77 $\pm$ 0.04	113.5 $\pm$ 8.5
ITCCA	1.0	0.64 $\pm$ 0.04	99.9 $\pm$ 9.5
FBITCCA	1.25	0.77 $\pm$ 0.04	113.9 $\pm$ 9.4

B. Averaged Accuracy and Computational Time

A large number of decomposed sub-bands implies the involvement of higher frequencies ranges, which potentially can increase the performance of the method. However, the progressive decrease in the SNR of the higher harmonics limits the improvement. Moreover, using more sub-bands implies increasing the computational time.

Figure 1 shows the required time for obtaining a one-stimulus optimized reference (training time), and the time for classifying one SSVEP-stimuli using 1.25 s signal length. This is measured using MATLAB R2020a on an Intel Core i7-10510U 1.80 GHz CPU, 15.8 GB RAM. In every case, increasing the number of sub-bands implies increasing the computational time, but the accuracy reaches its maximum value when using between 3 and 5 sub-bands and then keeps steady or decreases. The standard FBMwayCCA implementation requires the maximum training time, although the computational cost is negligible when compared to the expended time on the training recording process. The testing time using the optimal number of sub-bands is always below 0.05 seconds, confirming the feasibility of the suggested approach.

IV. DISCUSSION AND CONCLUSIONS

This work investigated a new approach that takes advantage of two of the most common CCA-based methodologies. The FB allows better use of the harmonic's hidden information, while the supervised algorithms allow taking advantage of the individual subject characteristics.

All FB methods showed significantly increased performance when compared with their standard version, and better performance when compared with the standard FBCCA method. This approach suggests that it is able to extract more information from the harmonic signals than the ordinary FBCCA approach. The best results are achieved using FBMwayCCA. Previous studies [15], [21], [22] pointed out a better performance of the MsetCCA and IT-CCA methods over the MwayCCA. However, those studies used more trials, so our study suggests increased reliability on MwayCCA if fewer trials are available.

The results evince the relevance of harmonics to obtain more accurate and faster SSVEP based BCI-speller systems. The combination of FBCCA and supervised methods allows

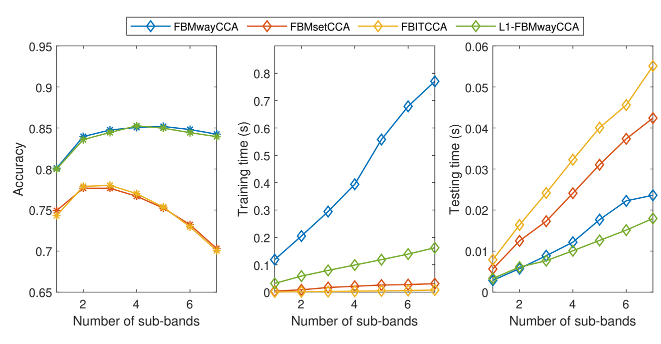

Fig. 1. Averaged accuracy and computational times using a fixed window time of 1.25 seconds.

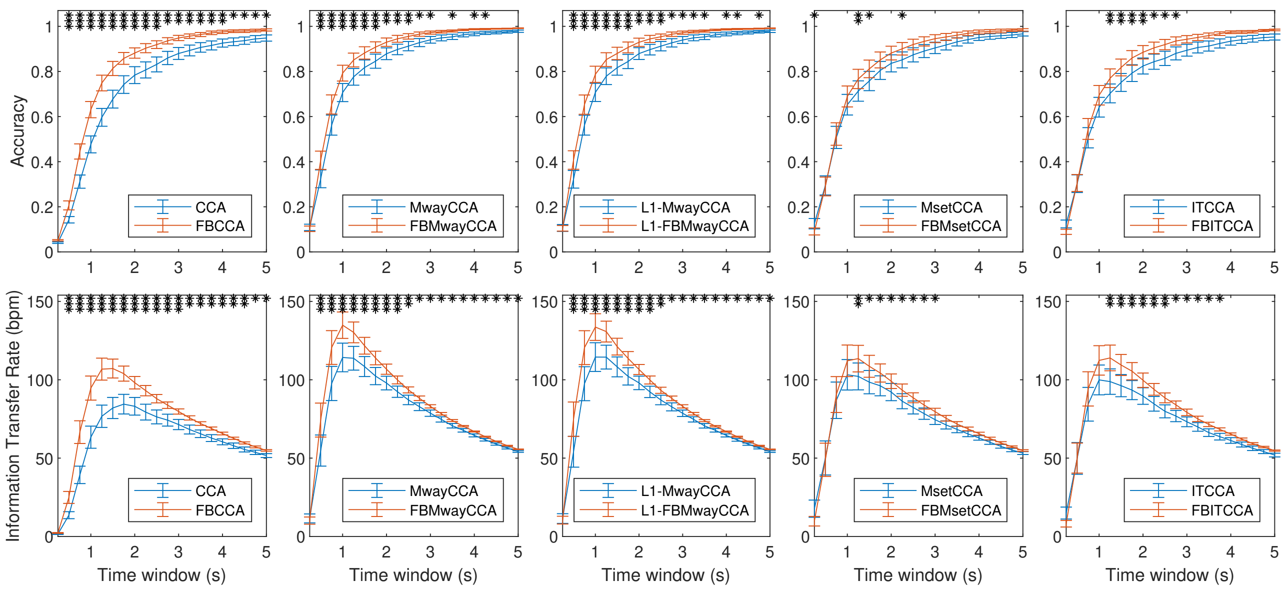

Fig. 2. Accuracy (top) and information transfer rate (below) of the different methods in their standard implementation (in blue) and filter bank extension (in orange), using time windows between 0.25 and 5s with steps of 0.25. Note that $*$ denotes $p < 0.05$, $**$ denotes $p < 0.01$ and $***$ $p < 0.001$.

taking advantage of the individual patient characteristics and harmonics information. Our results clearly motivate the integration of the FB approach in future developments.

REFERENCES

[1] R. Mane, T. Chouhan, and C. Guan, "BCI for stroke rehabilitation: Motor and beyond," J. Neural Eng., vol. 17, no. 4, p. 041001, Aug. 2020, doi: 10.1088/1741-2552/aba162.

[2] 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 Sens. J., vol. 21, no. 2, pp. 1124–1138, 2021, doi: 10.1109/JSEN.2020.3017491.

[3] S. H. Felgoise, V. Zaccheo, J. Duff, and Z. Simmons, "Verbal communication impacts quality of life in patients with amyotrophic lateral sclerosis," Amyotroph. Lateral Scler. Front. Degener., vol. 17, no. 3–4, pp. 179–183, May 2016, doi: 10.3109/21678421.2015.1125499.

[4] A. Rezeika, M. Benda, P. Stawicki, F. Gembler, A. Saboor, and I. Volosyak, "Brain–computer interface spellers: A review," Brain Sci., vol. 8, no. 4, 2018, doi: 10.3390/brainsci8040057.

[5] A. M. Norcia, L. G. Appelbaum, J. M. Ales, B. R. Cottereau, and B. Rossion, "The steady-state visual evoked potential in vision research: A review," J. Vis., vol. 15, no. 6, p. 4, 2015, doi: 10.1167/15.6.4.

[6] M. Cheng, X. Gao, S. Gao, and D. Xu, "Design and implementation of a brain-computer interface with high transfer rates," IEEE Trans. Biomed. Eng., vol. 49, no. 10, pp. 1181–1186, Oct. 2002, doi: 10.1109/TBME.2002.803536.

[7] G. R. Kiran Kumar and M. Ramasubba Reddy, "Constructing an exactly periodic subspace for enhancing SSVEP based BCI," Adv. Eng. Informatics, vol. 44, p. 101046, Apr. 2020, doi: 10.1016/j.aei.2020.101046.

[8] Y. Zhang, P. Xu, K. Cheng, and D. Yao, "Multivariate synchronization index for frequency recognition of SSVEP-based brain-computer interface," J. Neurosci. Methods, vol. 221, pp. 32–40, 2014, doi: 10.1016/j.jneumeth.2013.07.018.

[9] H. Vu, B. Koo, and S. Choi, "Frequency detection for SSVEP-based BCI using deep canonical correlation analysis," in 2016 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2016 - Conference Proceedings, Feb. 2017, pp. 1983–1987, doi: 10.1109/SMC.2016.7844531.

[10] J. J. Podmore, T. P. Breckon, N. K. N. Aznan, and J. D. Connolly, "On the Relative Contribution of Deep Convolutional Neural Networks for SSVEP-Based Bio-Signal Decoding in BCI Speller Applications," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 27, no. 4, pp. 611–618, Apr. 2019, doi: 10.1109/TNSRE.2019.2904791.

[11] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs," IEEE Trans. Biomed. Eng., vol. 53, no. 12, pp. 2610–2614, Dec. 2006, doi: 10.1109/TBME.2006.886577.

[12] G. Bin, X. Gao, Y. Wang, Y. Li, B. Hong, and S. Gao, "A high-speed BCI based on code modulation VEP," in Journal of Neural Engineering, Apr. 2011, vol. 8, no. 2, p. 025015, doi: 10.1088/1741-2560/8/2/025015.

[13] M. Nakanishi, Y. Wang, Y. Te 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. 6, pp. 1–18, 2014, doi: 10.1142/S0129065714500191.

[14] 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, doi: 10.1109/TNSRE.2013.2279680.

[15] 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. 4, 2014, doi: 10.1142/S0129065714500130.

[16] Y. Zhang et al., "Correlated component analysis for enhancing the performance of SSVEP-based brain-computer interface," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 26, no. 5, pp. 948–956, May 2018, doi: 10.1109/TNSRE.2018.2826541.

[17] M. Nakanishi, Y. Wang, X. Chen, Y. Te 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, Jan. 2018, doi: 10.1109/TBME.2017.2694818.

[18] C. M. Wong et al., "Learning across multi-stimulus enhances target recognition methods in SSVEP-based BCIs," J. Neural Eng., vol. 17, no. 1, 2020, doi: 10.1088/1741-2552/ab2373.

[19] 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, Aug. 2015, doi: 10.1088/1741-2560/12/4/046008.

[20] 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, 2017, doi: 10.1109/TNSRE.2016.2627556.

[21] M. Nakanishi, Y. Wang, Y. Te 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, p. e0140703, Oct. 2015, doi: 10.1371/journal.pone.0140703.

[22] B. Liu, X. Huang, Y. Wang, X. Chen, and X. Gao, "BETA: A Large Benchmark Database Toward SSVEP-BCI Application," Front. Neurosci., vol. 14, p. 627, Jun. 2020, doi: 10.3389/fnins.2020.00627.