FREQUENCY RECOGNITION OF STEADY–STATE VISUALLY EVOKED POTENTIALS USING BINARY SUBBAND CANONICAL CORRELATION ANALYSIS WITH REDUCED DIMENSION OF REFERENCE SIGNALS

Md. Rabiul Islam[*], *Toshihisa Tanaka*[*,†], *Masaki Nakanishi*[‡] *and Md. Khademul Islam Molla*[§]

[*] Tokyo University of Agriculture and Technology, Japan
[†] RIKEN Brain Science Institute, Japan
[‡] University of California, San Diego, USA
[§] University of Rajshahi, Bangladesh

ABSTRACT

This paper presents a frequency recognition method of steady-state visual evoked potentials (SSVEPs) using binary subbands with canonical correlation analysis (CCA). The first subband contains all the target frequencies of SSVEPs. The second one includes the SSVEP signal corresponding to a desired number of higher order stimulus frequencies, which is obtained by filtering out of required range of lower order stimuli. The full dimension of artificial reference signals are used for first subband, whereas a reduced dimension of references is employed for second subband to compute canonical correlation. The weighted sum of the obtained correlation values are used to recognize the frequency of an SSVEP. The experimental results show the superiority of the proposed method compared to the state-of-the-art recognition methods.

Index Terms— Steady-state visual evoked potential (SSVEP), brain–computer interface (BCI), frequency detection, canonical correlation analysis (CCA)

1. INTRODUCTION

Steady–state visual evoked potentials (SSVEPs) are periodic EEG responses at the stimulation frequency and its harmonics elicited by repetitive visual stimulation. In recent years, an SSVEP–based brain–computer interface (BCI) has received increasing attention from researchers due to its advantages of little user training, ease of system configuration, as well as high information transfer rate (ITR) [1]. In SSVEP–based BCIs, users are asked to focus attention on one of the multiple repetitive visual stimulus flickering at different frequencies. The target stimulus can be identified through a frequency analysis such as the discrete Fourier transform (DFT) [2] on which the user is focusing. Moreover, it has been reported that the canonical correlation analysis (CCA) between SSVEPs and sinusoidal templates is very powerful in frequency detection of SSVEPs [3]. However, in the standard CCA–based approach, the canonical correlation value between real world SSVEPs and sine-cosine waves tends to decrease as flickering frequency increases, resulting in decreased accuracy of SSVEP detection in higher frequencies [4].

To tackle this problem, Wang et al. reported that in different frequency regions, the signal-to-noise ratio (SNR) of SSVEPs could have comparable scales that influence to degrade frequency detection accuracy along ascending frequencies [5]. These might be attributed to the fact that the amplitude of background EEG at around lower frequencies (e.g. alpha band) can be more dominant across frequencies leading to degraded amplitude in higher frequencies. The frequency detection based on the DFT or the CCA may fail in the accurate detection of SSVEP frequency due to the attenuated EEG activity in relatively high frequencies. To solve such type of problem, several studies have been exploited in recent years. Tanaka et al. proposed a machine learning-based method, which combined the CCA with linear discriminant analysis (LDA) to improve the detection of SSVEPs in the higher frequency range [6]. Multiway CCA and multi-set CCA has been proposed to optimize the reference signals by combining SSVEP training data in CCA for frequency recognition [7, 8]. Nakanishi et al. also proposed the extended CCA-based method which combined the CCA-based spatial filtering and correlation coefficient between single–trial SSVEPs, training reference signals obtained by averaging training set and reference signals based on sine-cosine waves [9]. Islam et al. generated two levels of data adaptive reference signals to improve significantly at high and lower frequencies derived from the training set of real SSVEP signals rather than using artificial sine-cosine waves with several harmonics [10]. Nevertheless, the training procedure increases the difficulty of system use in real-life applications.

To improve the frequency detection accuracies in the higher frequency range, an unsupervised method, which employed the normalized canonical correlation coefficient using background EEG activities, had been proposed [11]. However, such process brings difficulty in practical implementation due to computational cost. Recently, Chen et. al. has implemented the high–speed SSVEP based BCI using filter bank canonical correlation analysis (FBCCA) to improve accuracies of SSVEP detection [12]. The goal of FBCCA is to decompose SSVEPs into subband components so that independent information embedded in the harmonics can be extracted more efficiently than with the standard CCA method. The accuracies of SSVEP frequency detection have increased with respect to the increasing number of subband. The potential problem in this method is that stimulus frequency is designed so that the ranges of fundamental and second harmonics are not overlapped. For example, the range of stimuli between 8 to 16 Hz, we can not generate stimulus after 16 Hz because of that overlapping. Another issue with FBCCA is the subject dependency of its performance. To address these problems, this study proposes a novel method based on a suitable number of lower order stimulus frequency of SSVEP are suppressed to produce the subband signal to improve frequency recognition as well as it reduces number of subband and running time. Finally, BCI experiments are conducted to demonstrate the efficacy of the proposed method that does not require system calibration.

This work was supported in part by JSPS KAKENHI Grant Number 15H04002. Email: tanakat@cc.tuat.ac.jp.

2. METHODS

2.1. Subjects and Experimental Settings

Three males and one female (Sub1, Sub2, Sub3, and Sub4) aged 21-26 (mean 23.3 and S.D. 2.22) took part in our experiment. All subjects gave informed consent, and this study was approved by the research ethics committee of Tokyo University of Agriculture and Technology.

Eight visual flickering stimuli (60×60 mm squares, 5.7 deg view angle and 60 mm distance from neighboring targets) were rendered on 23-inches liquid crystal display (LCD) with a refresh rate of 120 Hz and 1920×1080 screen resolution. Stimulation frequency of each target was 8, 9.2, 10.9, 12, 13.3, 15, 17.1, and 20 Hz and developed under MATLAB using the Psychophysics Toolbox extensions [13]. During the whole experiments, each subject sat on a comfortable chair in front of the display screen at a distance of 60 cm and focused on a flickering target on the screen. They were asked to gaze at one of the visual stimuli indicated in the order of $8, 9.2, \ldots, 20$ Hz for 4 s. There is an interval of 0.5 s between two consecutive trials. Subjects were asked to shift next target at the time.

To evaluate the proposed method for recognizing 8 different stimulus frequencies of SSVEPs, we used the signals observed in electrodes P7, P3, Pz, P4, P8, O1, Oz and O2 for analysis using active electrodes g.LADYbird driven by g.GAMMAbox (Guger Technologies, Austria). The electrodes for GND and reference were AFz and A1, respectively. The signals were amplified by MEG-6116 (Nihon Kohden, Japan), which provides lowpass and highpass analog filters for each channel. In this experiment, we set the cut–off frequencies of the lowpass and highpass filters to 100 Hz and 0.5 Hz, respectively. The amplified signals were sampled by an A/D converter, AIO-163202F-PE (Contec, Japan) with a sampling rate of 1200 Hz. All data were band-pass filtered between 7–50 Hz with an infinite impulse response (IIR) filter. Zero–phase forward and reverse IIR filtering was implemented using the filtfilt() function in Matlab. An offline classifier analyzed the multichannel EEGs with the epoch length of 4 s. We used the following newly proposed methods for determining the command.

2.2. Frequency Recogition Method

An SSVEP can be characterized by sinusoidal-like waveforms at the stimulation frequency and its harmonic frequencies. More specifically, SSVEPs consist of brain responses at identical, harmonic, and sub-harmonic frequencies to the stimulation frequency. In addition to the fundamental frequency component, the harmonic components can provide useful information for frequency detection. Any model which includes both the fundamental and harmonics of stimulus is demanding to increase the frequency recognition in SSVEP based BCI implementation.

2.2.1. Methods Based on Canonical Correlation Analysis

CCA is a way of determining the linear relationship between two multidimensional variables X and Y using their auto-covariances and cross-covariances. It finds two bases in which the correlation matrix between the variables is diagonal and the correlations on the diagonal are maximized. The dimensionality of these new bases is equal to or less than the smallest dimensionality of the two variables. Given two sets of random variables X and Y, CCA finds w_X and w_Y that maximize the following criteria:

$$\rho(X,Y) = \max_{w_X, w_Y} \frac{w_X^T C_{XY} w_Y}{\sqrt{w_X^T C_{XX} w_X w_Y^T C_{YY} w_Y}} \qquad (1)$$

where ρ is called the canonical correlation, C_{XX} and C_{XX} are the within-sets covariance matrices, and C_{XY} is the between-sets covariance matrix. The maximum value of ρ with respect to w_X and w_X represents the maximum canonical correlation. In the case of the SSVEP frequency recognition $X \in R^Z$ refers to the set of multi-variate (Z-channel) EEG signals representing the SSVEP and Y contains a set of multi-variate reference signals of the same number of data points as X.

For SSVEP frequency recognition, the reference signal $Y = Y_k$ is a pre-constructed signal set for the k^{th} stimulus frequency $f_k (k = 1, \ldots, K)$ are used as the reference signals, and the frequency that gives the maximum canonical correlation is selected. Lin et al. [3] proposed this scheme in an unsupervised way and formed $Y_k \in R^{2H}$ by a series of sine-cosine waves as:

$$Y_k = [\sin(2\pi f_k t), \cos(2\pi f_k t), \ldots, \sin(2\pi H f_k t), \cos(2\pi H f_k t)] \qquad (2)$$

where H is the number of harmonics. The CCA calculates the canonical correlation between multi-channel SSVEPs and the reference signals corresponding to each stimulation frequency.

Chen et al. proposed filter bank canonical correlation analysis (FBCCA) to enhance the performance of the CCA by performing autonomous selection of discriminative fundamental and harmonic of SSVEP X using zero-phase infinite impulse response (IIR) type filter [12]. In [12], three types of filter banks are proposed and the recommended selection is a set of subbands covering multiple harmonic frequency bands. The n^{th} subband started from the frequency at $n \times k$ Hz (where n is the number of subband and k is the starting frequency), and ended at a fixed maximum frequency (88 Hz in the configuration [12]). In the implementation of bandpass filtering, an additional bandwidth of 2 Hz was added to both sides of the passband for each subband. After the filter bank analyis, the standard CCA process was applied to each of the subband components separately and combined the correlation vector $\rho_k = \rho(X, Y_k)$ at stimulus frequency f_k as according to [12]. The frequency of the reference signals with the maximum correlation is considered to be the frequency of SSVEPs.

2.2.2. Proposed Binary Subband CCA

The canonical correlation is computed here between a subset of reference signals and subband components of SSVEPs in addition to the using full dimension of reference signals and original SSVEP respectively to increase the frequency recognition accuracy. The proposed method called binary subband CCA (BsCCA) broadly consists of three steps: i) subband extraction of multichannel SSVEP signals, ii) measure of CCA between subband components and corresponding subsets of artificial reference signals, and iii) frequency recognition based on the measured correlation values. The block diagram (flowchart) of the proposed BsCCA method is illustrated in Fig. 1. It is noted that in FBCCA [12], the filterbank is implemented mainly to capture the relevant harmonics to enhance the frequency recognition accuracy. Unlike the FBCCA [12], only one subband component is extracted from an EEG signal in addition to the original EEG.

The underlying idea behind this study is that a number of lower frequency stimuli are suppressed to enhance the contribution of the other higher stimulus frequencies in frequency recognition method

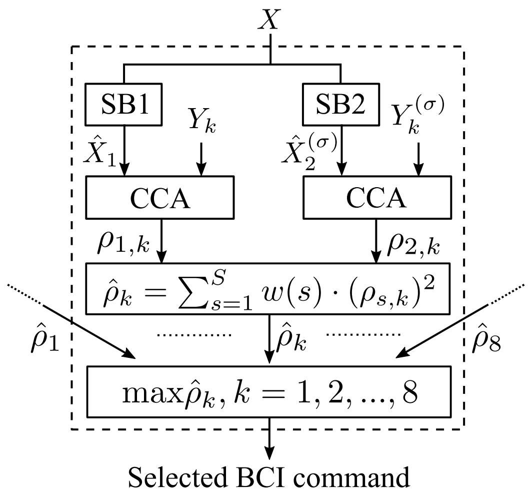

Fig. 1. Flowchart with different components of the proposed method for frequency recognition in an SSVEP based-BCI.

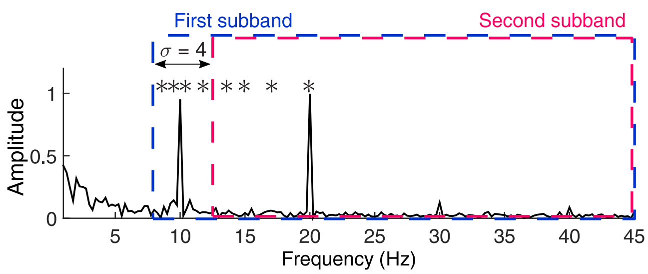

Fig. 2. Subband filtering of EEG signals and two subbands are generated. The first subband ($\hat{X}_1$) contains all the stimulus frequencies, whereas the last subband ($\hat{X}_2^{(\sigma)}$) contains only σ higher order stimulus frequencies. The asterisks ($*$) indicates the command frequencies.

implemented by CCA. To implement this concept and improve the overall recognition accuracy, the EEG signal is decomposed into two subbands. The first subband represented by $\hat{X}_1$ contains all the stimulus frequencies including second harmonics derived from bandpass filtering of original SSVEP between 7.5 Hz and 45 Hz. The second subband contains reduced number of higher order stimuli. It is obtained by filtering-out (suppressing) a desired number of lower order stimulus frequencies. The proposed subband decomposition of eight stimuli SSVEP is illustrated in Fig. 2.

The second subband is obtained by applying zero-phase highpass filter. The lower cut-off frequency of the highpass filter is determined by suppressing a desired number of lower order stimulus frequencies from SSVEP. It is represented as a function of reference stimuli with a suppression parameter σ ($\sigma = 2, \dots, K$) that is the number of lower order stimuli to be suppressed with the maximum number of stimulus frequencies K. The lower cut-off frequency of second subband $\hat{X}_2^{(\sigma)}$ for any value of σ can be expressed as

$$\lambda^{(\sigma)} = f_{\sigma+1} - \delta \tag{3}$$

where f_σ is the σ^{th} stimulus frequency and δ is an offset constant (here $\delta = 0.5$ Hz). The conventional CCA is applied to each subband components individually yielding the correlation values between subband component and the corresponding subset of artificial reference $Y_k^{(\sigma)}$ defined as

$$Y_k^{(\sigma)} = \Phi^{(\sigma)}(Y_k), \tag{4}$$

where $\Phi^{(\sigma)}(\cdot)$ is the function that excludes sinusoidal elements with frequencies less than $\lambda^{(\sigma)}$ and their harmonics. The reason of using this subset is that the proposed subband filtering approach suppresses SSVEP of lower order stimuli keeping the higher orders. It shrinks the classification domain and reduces the size of confusion matrix and hence improves the recognition accuracy. Only the reference signals corresponding to the stimuli contained by the second subband are used in CCA computation. Because of using the constant offset, the adjacent lower order reference stimulus is also kept in the subset of reference signal. For instance, if $\sigma = 2$, only the first reference stimulus is discarded from the set of reference signals.

For the k^{th} stimulus frequency, we obtain two canonical correlations in two subbands:

$$\rho_{1,k} = \rho(\hat{X}_1, Y_k), \;\; \rho_{2,k} = \rho(\hat{X}_2^{(\sigma)}, Y_k^{(\sigma)}) \tag{5}$$

The feature of target identification could be defined as:

$$\hat{\rho}_k = \sum_{s=1}^{S} w(s)(\rho_{s,k})^2 \tag{6}$$

where S is the number of subbands. Note that in this binary case, $S = 2$. The weights for the subband components were defined as $w(s) = s^{-a} + b, \; s \in [1 \; S]$; where a and b are constants that maximize the classification performance defined as according to [1]. In practice, a and b can be determined using a grid search method using an offline analysis. The weighted correlation value $\hat{\rho}_k^{(\sigma)}$ corresponding to all stimulus frequencies $k = 1, 2, \dots, K$ are used for determining the frequency of the SSVEPs. The frequency of the reference signals with the maximum $\hat{\rho}_k$ is considered to be the frequency of the SSVEPs. Fig. 1 illustrates the procedure of target identification method using subband CCA.

3. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed BsCCA, eight different stimulus frequencies are used for SSVEP frequency recognition. Each frequency has 20 trials and each trial contained the data for each of 8 different stimulus frequencies. The method does not require training data and its average recognition accuracy is evaluated on the direct validation of 20 runs.

The performance of the proposed BsCCA varies with different values of suppression parameter σ as illustrated in Fig. 3(L). The maximum accuracy is achieved with $\sigma = 4$, i.e., suppression of four lower order stimuli with keeping the higher order four stimuli. Only two subbands (first one includes all stimuli and the number of stimuli in second subband is defined as a function of σ) are used in frequency recognition and hence the computational cost is reduced. The parameters a and b used to calculate the weighting factor in Eq. (6) are set to 1.25 and .25 respectively according to [12].

The underlying assumption of using such suppression method is that the recognition performance of higher order stimuli is decreased with respect to the lower order stimuli using the standard CCA with artificial reference signals proposed by Lin et al. [3] as shown in Fig. 3(R). The recognition accuracy up to 12 Hz is reasonable, whereas it is decreasing for the higher order stimulus frequencies. The performance of the proposed method (with $\sigma = 4$)

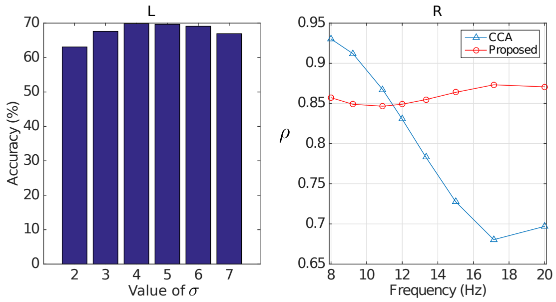

Fig. 3. Average (across all subjects) frequency recognition accuracy as a function of σ (L). Canonical correlation coefficients of SSVEPs at different stimulation frequencies (R).

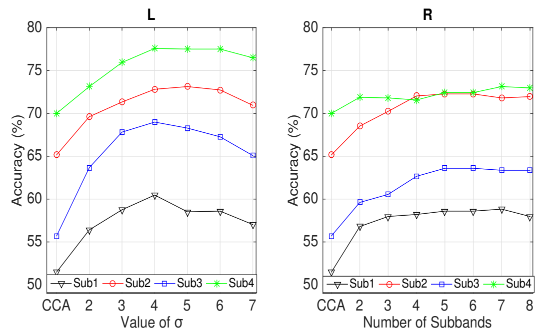

Fig. 4. Frequency recognition accuracy obtained by the standard CCA [3] and the proposed BsCCA with $\sigma = 2, \ldots, 7$ (L); recognition accuracy obtained by the standard CCA and the FBCCA with different number of subbands (R) for different subjects (Sub1 to Sub4). In both cases the accuracies with different data lengths (0.25 s to 2 s) are averaged.

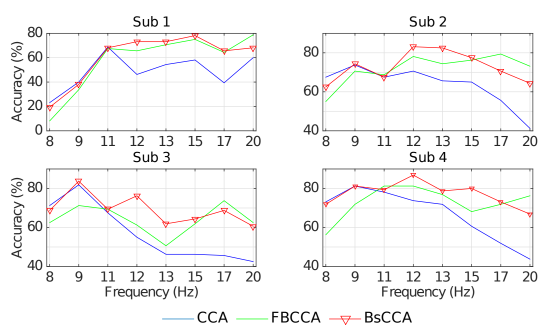

Fig. 5. Recognition accuracies of individual stimulus frequency obtained by the standard CCA, FBCCA, and BsCCA (with $\sigma = 4$) for different subjects. The accuracies with different data lengths (0.25 s to 2 s) are averaged.

is relatively steady with noticeable level of accuracy and hence improves the overall recognition rate.

The effect of the suppression parameter σ is also observed on individual subjects as illustrated in Fig. 4(L). It is noticed that the maximum accuracy is achieved with $\sigma = 4$ for each subject, hence the parameter can be treated as subject independent. In [12], they achieved the highest accuracy when using the standard CCA with

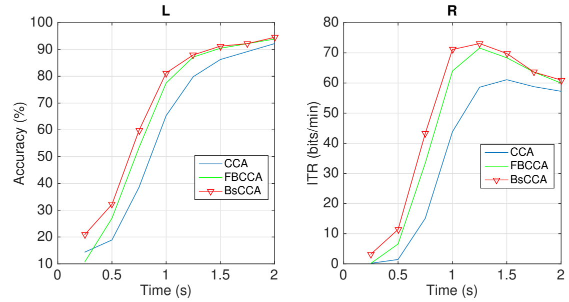

Fig. 6. Average frequency recognition accuracy (L) and ITR (R) across all subjects obtained by the standard CCA, FBCCA, and BsCCA (with $\sigma = 4$) as a function of signal length.

1.25 s-long data and 7 harmonics in reference signals. When they used FBCCA to 1.25 s-long data, 7 subbands showed the highest accuracy. The effects of the number of subbands of method (FBCCA) on each of the four subjects are also studied as shown in Fig. 4(R). The maximum recognition accuracy for individual subjects is achieved by using different number of subbands (7, 6, 5, and 7 subbands are required for subject 1, 2, 3, and 4 respectively) and hence the selection appropriate number of subbands is subject dependent, whereas the selection of proper value of σ is robust with respect to subjects.

The performance of the proposed method was compared with the standard CCA [14], and FBCCA [12] (using artificial reference signals with five harmonics). The average recognition accuracy is illustrated in Fig. 5 as a function of stimulus frequency. It is observed that our method consistently outperformed the other competing methods for all of four subjects. The accuracy of BsCCA is significantly increased at low and high frequencies with respect to other methods. It is noticed that our method performs better than that of the related recently proposed methods for a wide range the length of SSVEP signals as shown in Fig. 6(L). The ITR is an important parameter to measure the performance of frequency recognition of SSVEPs [15]. The average ITR (across the subjects) is examined and the comparative results as a function of SSVEP signal length are illustrated in Fig. 6(R). The value of ITR with BsCCA is always higher than the competitive methods from short data length to long time window. Importantly, BsCCA achieved comparable ITR to the highest value in FBCCA with shorter data length (BsCCA: Data length: 1 s, ITR 71.10 bits/min; FBCCA: Data length: 1.25 s, ITR 71.58 bits/min). This suggests that the BsCCA can save 0.25 s on visual stimulation and data analysis in online operation, leading to decrease of visual fatigues.

4. CONCLUSIONS

This study developed a novel frequency recognition method for an SSVEP-based BCI by considering correlation between subband components of SSVEP signals and corresponding subset of reference signals. It was clarified by experimentation using the SSVEP data that the performance of the proposed BsCCA showed higher recognition accuracies than recently reported methods in all four subjects. Only two subbands are used in frequencies recognition that reduces the computational cost than recently reported FBCCA. The main concern in our future work will be to apply the BsCCA to the implementation of a online multi-command, real-time, portable, and visual fatigue-free BCI system.

5. REFERENCES

[1] M. Middendorf, G. McMillan, G. Calhoun, and K. Jones, "Brain-computer interfaces based on the steady-state visual-evoked response," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 211–214, Jun 2000.

[2] H. Bakardjian, T. Tanaka, and A. Cichocki, "Optimization of SSVEP brain responses with application to eight–command brain–computer interface," *Neuroscience Letters*, pp. 1138–1143, Nov 2010.

[3] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 6, pp. 1172–1176, June 2007.

[4] F.-C. Lin, J. Zao, K.-C. Tu, Y. Wang, Y.-P. Huang, C.-W. Chuang, H.-Y. Kuo, Y.-Y. Chien, C.-C. Chou, and T.-P. Jung, "SNR analysis of high-frequency steady-state visual evoked potentials from the foveal and extrafoveal regions of human retina," in *34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2012, pp. 1810–1814.

[5] Y. Wang, R. Wang, X. Gao, B. Hong, and S. Gao, "A practical VEP–based brain-computer interface," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,*, vol. 14, no. 2, pp. 234–240, 2006.

[6] T. Tanaka, C. Zhang, and H. Higashi, "SSVEP frequency detection methods considering background EEG," in *Joint 6th International Conference on Soft Computing and Intelligent Systems (SCIS) and 13th International Symposium on Advanced Intelligent Systems (ISIS)*, Nov 2012, pp. 1138–1143.

[7] Y. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Frequency recognition in SSVEP–based BCI using multiset canonical correlation analysis," *International Journal of Neural Systems*, vol. 24, no. 04, p. 1450013, 2014.

[8] Y. Zhang, G. Zhou, Q. Zhao, A. Onishi, J. Jin, X. Wang, and A. Cichocki, "Multiway canonical correlation analysis for frequency components recognition in SSVEP-based BCIs," in *Neural Information Processing*. Springer, 2011, pp. 287–295.

[9] M. Nakanishi, Y. Wang, Y.-T. Wang, Y. Mitsukura, and T.-P. Jung, "A high-speed brain speller using steady-state visual evoked potentials," *International Journal of Neural Systems*, vol. 24, no. 06, p. 1450019, 2014.

[10] M. R. Islam, T. Tanaka, N. Morikawa, and M. K. I. Molla, "Frequency recognition for SSVEP-based BCI with data adaptive reference signals," in *2015 IEEE International Conference on Digital Signal Processing (DSP)*, July 2015, pp. 799–803.

[11] M. Nakanishi, Y. Wang, Y.-T. Wang, Y. Mitsukura, and T.-P. Jung, "Enhancing unsupervised canonical correlation analysis-based frequency detection of SSVEPs by incorporating background EEG," in *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2014, pp. 3053–3056.

[12] X. Chen, Y. Wang, S. Gao, T.-P. Jung, and X. Gao, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based braincomputer interface," *Journal of Neural Engineering*, vol. 12, no. 4, p. 046008, 2015.

[13] M. Kleiner, D. Brainard, D. Pelli, A. Ingling, R. Murray, and C. Broussard, "What's new in Psychtoolbox-3," *Perception*, vol. 36, no. 14, p. 1, 2007.

[14] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method," *Journal of Neural Engineering*, vol. 6, no. 4, p. 046002, 2009.

[15] J. Wolpaw, N. Birbaumer, W. Heetderks, D. McFarland, P. Peckham, G. Schalk, E. Donchin, C. Robinson, and T. Vaughan, "Brain–computer interface technology: A review of the first international meeting," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 164–173, Jun 2000.