

Overview of recognition methods for SSVEP-based BCIs in World Robot Contest 2022: MATLAB undergraduate group

Chengzhi Yi^{1,§}, Yuxuan Wu^{1,§}, Fan Ye¹, Xinchen Zhang², Jingjing Chen^{3,4} (✉)

¹ Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, China

² College of Science, Beijing Forestry University, Beijing 100083, China

³ Department of Psychology, School of Social Sciences, Tsinghua University, Beijing 100084, China

⁴ Tsinghua Laboratory of Brain and Intelligence, Tsinghua University, Beijing 100084, China

[§] These authors contributed equally to this work.

ARTICLE INFO

Received: 14 April, 2023

Revised: 21 May, 2023

Accepted: 6 June, 2023

© The authors 2023. This article is published with open access at journals.sagepub.com/home/BSA



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<http://www.creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

KEYWORDS

brain–computer interfaces, electroencephalogram, steady-state visual evoked potential, BCI spellers, calibration-free, MATLAB

ABSTRACT

The steady-state visual evoked potential (SSVEP)-based speller has emerged as a widely adopted paradigm in current brain–computer interface (BCI) systems due to its rapid processing and consistent performance across different individuals. Calibration-free SSVEP algorithms, as opposed to their calibration-based counterparts, offer clear and intuitive mathematical principles, making them accessible to novice developers. During the World Robot Contest (WRC) 2022, participants in the undergraduate category utilized various approaches to accomplish target detection in the calibration-free setting, successfully implementing the algorithms using MATLAB. The winning approach achieved an average information transfer rate of 198.94 bits/min in the final test, which is notably high given the calibration-free scenario. This paper presents an introduction to the underlying principles of the selected methods, accompanied by a comparison of their effectiveness through analysis of results from both the final test and offline experiments. Additionally, we propose that the youth competition of WRC could serve as an ideal starting point for beginners interested in studying and developing their own BCI systems.

1 Introduction

Brain–computer interface (BCI) systems aim to establish a non-muscular communication and control pathway from the brain to external devices [1]. Among various methods for monitoring brain activities, electroencephalogram

(EEG) is the most widely used in BCIs due to its simplicity, non-invasiveness, low cost, and short response time to external stimuli [2]. Steady-state visual evoked potential (SSVEP), which is a brain response evoked by periodic flashing visual stimuli [3], is one of the most popular EEG control signals. In comparison to other paradigms,

Address correspondence to Jingjing Chen, chen-jj15@tsinghua.org.cn

SSVEP-based BCIs are usually easy to train and can achieve a relatively high information transfer rate (ITR) [4].

In an SSVEP-speller system, each key on the screen flickers at a unique frequency, which can evoke brain signals at the same frequency and higher harmonics in the subject's visual cortex [5]. Users type sentences by focusing on their intended letters one at a time. The main objective of target detection algorithms in SSVEP-based BCIs is to determine the predominant fundamental frequency of the evoked brain response. To achieve this goal, multiple approaches have been proposed [6–14]. Depending on whether training data is required, these approaches are further classified into calibration-based and calibration-free methods. Calibration-based approaches acquire subject-specific spatial filters and response templates from labeled training data. They usually yield higher ITR compared to calibration-free methods due to the utilization of prior knowledge [13, 14]. However, the calibration-free scenario represents an important milestone in this field. It requires no training phase, resulting in lower cost and enabling the development of “plug-and-play” systems, which is significant for practical applications. Additionally, the mathematical principles of calibration-free algorithms are often clearer and more intuitive, making them particularly suitable for novice developers to employ in their initial BCI systems.

In the BCI Controlled Robot Contest at the World Robot Contest 2022 (WRC2022), a special track called the “MATLAB undergraduate group” was established to encourage younger students to engage in BCI study and research [15]. To enhance accessibility for beginners, a prebuilt framework with standard interfaces was provided, and all algorithms were restricted to the calibration-free scenario. The competition comprised two rounds. Prior to the match, the

prebuilt framework and training data were made available online. In the first round, competitors implemented their core algorithms based on the framework and evaluated their performance on the training dataset. They could also submit their codes to the organizer and receive feedback on the test dataset. The top five competitors from the first round advanced to the second round (final) and were invited to the World Robot Conference. All participants were undergraduate students aged between 19 and 21 years. On the day of the conference, competitors attended the event in person to submit their codes, and the final ranking was based on the algorithm's performance on the test dataset.

Section 2 of this paper introduces three leading algorithms employed by undergraduate competitors in the final of WRC2022. Section 3 presents the performance of these algorithms in both the final and the offline experiments. Section 4 discusses potential avenues for algorithm improvement, while Section 5 offers concluding remarks.

2 Methods

In this section, we will begin by introducing the experimental paradigm employed in the contest. Subsequently, we will present three frequency recognition algorithms: spatiotemporal equalization, filter bank canonical correlation analysis (CCA), and online adaptation CCA. Additionally, we will introduce a dynamic window (DW) strategy designed to enhance the system's performance.

2.1 Experimental paradigm

The keyboard utilized in the experiment consisted of 40 targets or keys, with flashing frequencies ranging from 8 to 15.8 Hz. EEG data were collected in blocks, with each block comprising consecutive trials where stimuli were presented

in random order. Each individual trial lasted between 3.5 to 4.5 s. In the initial 0.5 s of each trial, a red square appeared on the screen, indicating the location of the target. This was followed by a stimulus phase lasting 3 to 4 s, during which all 40 targets flashed simultaneously at different frequencies. Participants were instructed to focus their attention on the cued target to generate a steady-state visually-evoked response in their EEG signals. A trigger was recorded at the beginning of each stimulus phase for every trial [15].

2.2 Spatiotemporal equalization

The distinction between evoked signals and background noise poses a challenge for EEG target detection algorithms. To address this concern, equalization techniques are frequently employed to suppress frequency-specific interference. In accordance with communication theory [16], well-designed equalizers have the ability to adjust the amplitude-frequency characteristics of received signals and mitigate inter-symbol interference, resulting in nearly distortion-free transmission across all information channels. Drawing from this principle, Yang et al. [11] introduced a spatiotemporal equalization (STE) recognition method. This approach utilizes equalizers to transform colored background noise into approximately white noise, characterized by largely uncorrelated time-domain properties, thereby minimizing its detrimental effects.

Let $X \in \mathbb{R}^{L \times N}$ denote a single-trial EEG signal with L channels and N sampling points, S the SSVEP components and W the background noise. The data model of EEG signals is shown as

$$X = S + W \quad (1)$$

To simulate the generation and conduction process of background noise, a channel system is proposed. In this system, the noises are generated in the cortex and then travel through

the skull to the scalp. The STE method assumes that the aliasing process is non-stationary, while the conductive process is stationary. The stationary process has statistical properties that remain constant over time, while the non-stationary process has statistical properties that vary with time. Specifically, the original white noise is generated by L identically distributed Gaussian sources. It then undergoes a non-stationary zero-order channel, resulting in spatially correlated non-stationary noise that corresponds to the aliasing process of the source signal. Subsequently, the noise passes through a stationary high-order channel, transforming into noise that exhibits both spatial and temporal correlation, corresponding to the conductive process. Consequently, the background noise, denoted as W , encompasses both stationary and non-stationary features.

To reverse the aliasing and conductive process, both a stationary equalizer and a non-stationary equalizer are estimated based on multiple adjacent trials and the current trial, respectively. In each trial, the estimation of the EEG background noise, represented as \hat{W} , can be mathematically expressed as

$$\hat{W} = X - \hat{S} \quad (2)$$

where \hat{S} denotes the estimation of the SSVEP component, which is generally considered a linear mixture of multi-frequency sinusoidal signals that are independent of background noise. The estimation \hat{W} , obtained from multiple adjacent trials, represents the stationary component of the background noise and is used to estimate the stationary equalizer. In contrast, the non-stationary equalizer relies solely on EEG data from the current trial. Given that the signal-to-noise ratio (SNR) between the SSVEP component and the background noise is extremely low, the EEG data closely resemble the background noise data. Consequently, it is feasible to estimate the non-stationary equalizer using the EEG data

corresponding to the current trial instead of the background noise data.

The colored background noise W with spatial and temporal correlation is first subjected to the stationary equalizer, resulting in non-stationary noise with spatial correlation. Subsequently, the noise passes through the non-stationary equalizer, leading to equalized noise devoid of spatial or temporal correlation.

In most cases, there are fewer independent SSVEP components than channels. Therefore, optimizing the projection direction can effectively improve the SNR of the data. The minimum energy combination algorithm [17] is a method to determine the optimal projection direction, which is represented by a weight vector that minimizes the noise in the data. The optimization problem for the projection direction can be expressed as follows:

$$\lambda_k = \min_{\beta_k} \beta_k^T \theta \theta^T \beta_k \quad (3)$$

where θ is the equalized noise, λ_k is the smallest eigenvalue of $\theta \theta^T$, and β_k is the minimum energy projection direction. Finally, the frequency that enables the projected noise energy λ_k to reach its minimum is chosen as the final result:

$$\hat{k} = \arg \min_k \lambda_k, k = 1, \dots, 40 \quad (4)$$

Figure 1 presents the flowchart of STE. Note that to further enhance its performance, the algorithm is combined with the DW strategy, which will be introduced in Section 2.5.

2.3 Filter bank canonical correlation analysis

Canonical correlation analysis is a widely used target detection approach in SSVEP-based BCIs. Its effectiveness in enhancing the system's robustness against noise was first demonstrated by Lin et al. [6] who applied CCA as a multivariable statistical method. Subsequently, Bin et al. [7] employed CCA for online SSVEP-BCI, simplifying the process of channel selection and parameter optimization.

The main concept behind CCA is to maximize the Pearson's correlation between linear combinations of EEG signals and reference signals. EEG signals in the t -th trial are denoted as $X^{[t]}$, and the reference signals corresponding to the k -th frequency are denoted as Y_k . Specifically, CCA aims to solve the following problem:

$$\begin{aligned} \{u_k^{[t]}, v_k^{[t]}\} &= \text{CCA}(X^{[t]}, Y_k) \\ &= \max_{u,v} \frac{u^T (X^{[t]})^T Y_k v}{\sqrt{u^T (X^{[t]})^T X^{[t]} u \cdot v^T Y_k^T Y_k v}}, \\ 1 \leq k \leq N_f \end{aligned} \quad (5)$$

where $u_k^{[t]}$ and $v_k^{[t]}$ are spatial filters, $X^{[t]}u$ and

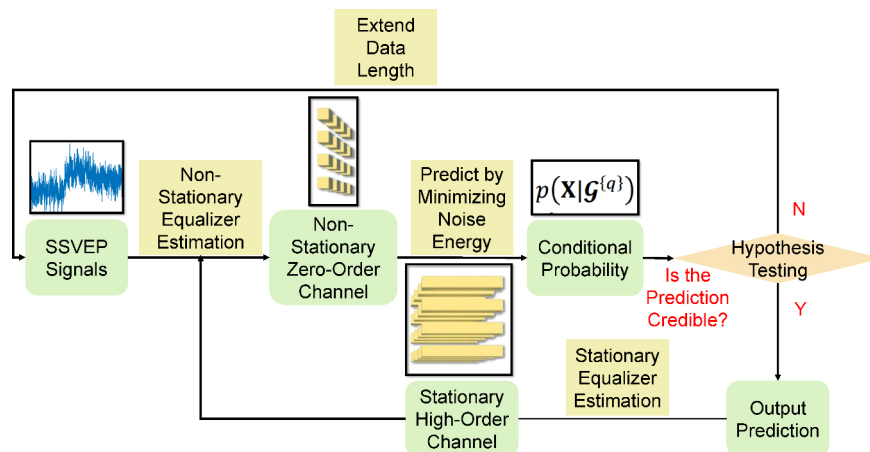


Fig. 1 Flowchart of spatiotemporal equalization with dynamic window.

$Y_k v$ are the linear combinations, $r_k^{[t]}$ is the Pearson's correlation coefficient, and N_f is the number of frequencies. The reference signals Y_k are a set of harmonics:

$$Y_k = \begin{pmatrix} \sin(2\pi f_k t) \\ \cos(2\pi f_k t) \\ \vdots \\ \sin(2\pi N_h f_k t) \\ \cos(2\pi N_h f_k t) \end{pmatrix}, t = 1/S, 2/S, \dots, N_p/S \quad (6)$$

where f_k is the target frequency, N_h is the number of harmonics, S is the sampling rate, and N_p is the number of sample points. By repeating this process for each stimulus frequency, we obtain a series of coefficient $r_k^{[t]} (k=1, 2, \dots, N_f)$. Finally, the predicted target $\hat{k}^{[t]}$ is determined by the following equation:

$$\hat{k}^{[t]} = \arg \max_k r_k^{[t]} \quad (7)$$

Building upon CCA, Chen et al. introduced filter bank CCA (FBCCA) in their work [8]. Filter bank CCA was developed to incorporate both the fundamental frequency and higher harmonics, resulting in improved classification performance. The approach utilizes a set of band-pass filters to decompose the SSVEP signals into N_{band} sub-band components. There are various methods available for designing the filter bank. A commonly

employed choice is to utilize Chebyshev Type I infinite impulse response filters, where each filter (indexed by $n = 1, 2, \dots, 7$) has a passband in the range $[(8n - 2) \text{ Hz}, 88 \text{ Hz}]$.

Subsequently, CCA is performed independently between each sub-band component and the reference signals, yielding the correlation coefficient $r_{k,n}^{[t]}$. The weighted sum of squares is subsequently calculated as the feature for identification:

$$\tilde{r}_k^{[t]} = \sum_{n=1}^{N_{\text{band}}} w(n) r_{k,n}^{[t]2} \quad (8)$$

According to Chen et al. [8], $w(n)$ is determined by

$$w(n) = n^{-a} + b, n = 1, 2, \dots, N_{\text{band}} \quad (9)$$

where a and b are constants, selected as 1.25 and 0.25, respectively, in Ref [8] by grid search. Consistent with CCA, the final target is determined by

$$\hat{k}^{[t]} = \arg \max_k \tilde{r}_k^{[t]} \quad (10)$$

Figure 2 presents the flowchart of FBCCA. The algorithm is also combined with the DW strategy.

2.4 Online adaptation canonical correlation analysis

The performance disparity between calibration-

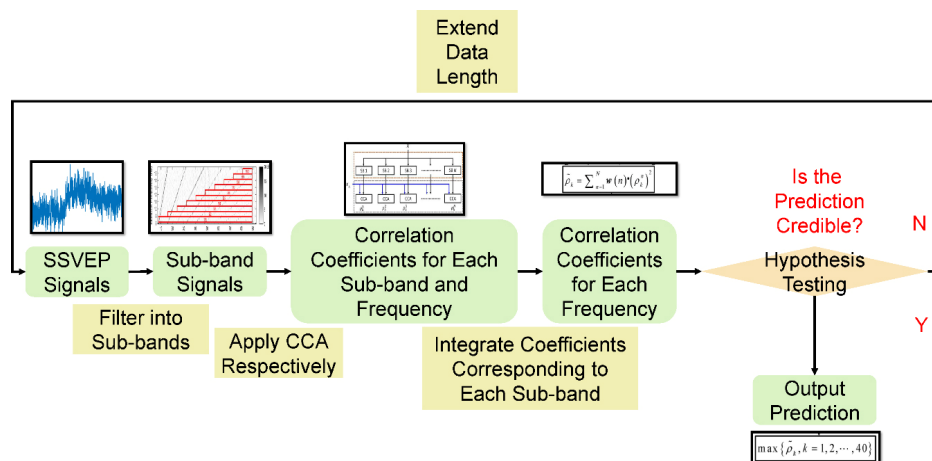


Fig. 2 Flowchart of FBCCA with dynamic window.

based and calibration-free approaches primarily stems from the inclusion of personalized information in the parameters used by calibration-based algorithms (e.g., subject-specific response EEG templates in task-related component analysis [14]), which aids in recognition. In contrast, calibration-free algorithms employ parameters (e.g., sine-cosine templates and CCA spatial filters in FBCCA) that lack such personalized knowledge. To bridge this gap, researchers have introduced subject-specific information into calibration-free algorithms. For example, spatial filters [9] and SSVEP templates [18] in CCA have been continuously updated using unlabeled SSVEP data collected from the same subject. Additionally, other parameters such as data length (or time-window length, TW) and stimulus size can be adjusted online to enhance algorithm performance. Chen et al. [19] dynamically adjusted the target identification TW to strike a balance between accuracy and time consumption in their study. Volosyak [20] varied the stimulus size based on SSVEP amplitude to provide better visual feedback. Building upon this concept, Wong et al. [12] proposed online adaptation CCA (OACCA), which incorporated online adapted spatial filters into the FBCCA algorithm. According to Kuś et al. [21], SSVEPs of the same subject exhibit similar spatial patterns across different frequencies. Exploiting this assumption, OACCA combines both CCA with prototype spatial filter (PSF) proposed by Lao et al. [10] and online multi-stimulus CCA (OMSCCA) proposed by Wong et al. [9] to incorporate common patterns among different stimulus frequencies.

In CCA with PSF [9], the PSF in each trial is defined as a spatial filter that exhibits maximal cosine similarity with the normalized spatial filters from all previous trials. This enables the extraction of subject-specific and frequency-invariant spatial patterns. Assuming that the normalized spatial filter in the t -th trial is defined

as $\tilde{\mathbf{u}}^{[t]} = \frac{\mathbf{u}_{\hat{k}^{[t]}}^{[t]}}{|\mathbf{u}_{\hat{k}^{[t]}}^{[t]}|}$, PSF in the $(t + 1)$ -th trial is described by

$$\begin{aligned}\mathbf{u}_{\text{PSF}}^{[t+1]} &= \arg \max_{\mathbf{u}} \sum_{m=1}^t \left(\frac{\mathbf{u}^T \tilde{\mathbf{u}}^{[m]}}{|\mathbf{u}| \cdot |\mathbf{u}^{[m]}|} \right)^2 \\ &= \arg \max_{\mathbf{u}} \frac{\mathbf{u}^T \sum_{m=1}^t \tilde{\mathbf{u}}^{[m]} (\tilde{\mathbf{u}}^{[m]})^T \mathbf{u}}{\mathbf{u}^T \mathbf{u}} \\ &= \arg \max_{\mathbf{u}} \frac{\mathbf{u}^T \mathbf{S}^{[t]} \mathbf{u}}{\mathbf{u}^T \mathbf{u}}\end{aligned}\quad (11)$$

where $\mathbf{S}^{[t]}$ is updated iteratively by

$$\mathbf{S}^{[t]} = \mathbf{S}^{[t-1]} + \tilde{\mathbf{u}}^{[t]} (\tilde{\mathbf{u}}^{[t]})^T \quad (12)$$

The correlation coefficient output of CCA with PSF is defined as follows:

$$\hat{r}_k^{[t]} = \text{CCA}(\mathbf{X}^{[t]} \mathbf{u}_{\text{PSF}}^{[t]}, \mathbf{Y}_k) \quad (13)$$

In OMSCCA [9], spatial filters aim to maximize the correlation between projections of EEG data and the corresponding sine-cosine templates from all previous trials:

$$\begin{aligned}\{\mathbf{w}_x^{[t+1]}, \mathbf{w}_y^{[t+1]}\} &= \arg \max_{\mathbf{u}, \mathbf{v}} \frac{\mathbf{u}^T \sum_{m=1}^t (\mathbf{X}^{[m]})^T \mathbf{Y}_{\hat{k}^{[m]}} \mathbf{v}}{\sqrt{\mathbf{u}^T \sum_{m=1}^t (\mathbf{X}^{[m]})^T \mathbf{X}^{[m]} \mathbf{u} \cdot \mathbf{v}^T \mathbf{v}}} \\ &= \arg \max_{\mathbf{u}, \mathbf{v}} \frac{\mathbf{u}^T \mathbf{C}_{XY}^{[t]} \mathbf{v}}{\sqrt{\mathbf{u}^T \mathbf{C}_{XX}^{[t]} \mathbf{u} \cdot \mathbf{v}^T \mathbf{v}}}\end{aligned}\quad (14)$$

where $\mathbf{C}_{XY}^{[t]}$ and $\mathbf{C}_{XX}^{[t]}$ are the summations of the covariance matrices updated iteratively by

$$\mathbf{C}_{XY}^{[t]} = \mathbf{C}_{XY}^{[t-1]} + (\mathbf{X}^{[t]})^T \mathbf{Y}_{\hat{k}^{[t]}} \quad (15)$$

$$\mathbf{C}_{XX}^{[t]} = \mathbf{C}_{XX}^{[t-1]} + (\mathbf{X}^{[t]})^T \mathbf{X}^{[t]} \quad (16)$$

The correlation coefficient output of OMSCCA is defined as

$$\hat{r}_k^{[t]} = \text{corr}(\mathbf{X}^{[t]} \mathbf{w}_x^{[t]}, \mathbf{Y}_k \mathbf{w}_y^{[t]}) \quad (17)$$

Finally, the ultimate correlation coefficient

is the summation of outputs yielded by CCA without online adaptation, CCA with PSF, and OMSCCA combined:

$$\rho_k^{[t]} = r_k^{[t]} + \dot{r}_k^{[t]} + \ddot{r}_k^{[t]} \quad (18)$$

Summing up the three correlation coefficients allows the system to leverage the strengths of each algorithm. The OACCA algorithm, developed by Wong et al. [12], utilizes the filter bank technique in combination with CCA. In OACCA, the online adaptation process described earlier is applied individually to each sub-band component, and then the weighted sum of squares is calculated for target detection. Therefore, OACCA integrates naive CCA, CCA with PSF, and OMSCCA with the filter bank technique. During the initial stage, when stable performance through online adaptation has not yet been achieved, the final output of the system is largely determined by the performance of naive FBCCA, which does not rely on any previous data.

2.5 Dynamic window

Due to the predominant individual and temporal variations in EEG, the recognition performance in the SSVEP paradigm varies among subjects and trials. Therefore, optimizing parameters such as the TW based on the difficulty of each trial can improve the overall performance of the algorithm [19]. Typically, increasing the TW is associated with higher accuracy in target identification. To maximize the ITR, developers can initially set a small TW value for each trial. If the algorithm's output prediction has a low confidence level, extending the TW can significantly improve the likelihood of correct identification. Conversely, if the prediction is deemed reliable, the algorithm directly accepts it to save time.

Chen et al. [19] introduced a hypothesis testing approach to assess the credibility of predictions. Specifically, null hypothesis H_0 is defined as

“current data are insufficient for the algorithm to output a credible label”, whereas alternative hypothesis H_i is defined as “label corresponding to the i -th frequency is credible to be the ground truth”. The posterior probability is quantified by a SoftMax function:

$$p(H_i | \rho^{[t]}) = \frac{e^{\rho_i^{[t]}}}{\sum_{j=1}^{N_f} e^{\rho_j^{[t]}}}, \quad 1 \leq i \leq N_f \quad (19)$$

The cost for each hypothesis is measured by the cross-entropy loss function:

$$\text{Cost}(H_i) = -\sum_{j=1}^{N_f} C(H_j, H_i) \log p(H_j | \rho^{[t]}) \quad (20)$$

For null hypothesis H_0 , function $C(\cdot)$ is defined as a constant parameter ε :

$$C(H_i, H_0) = \varepsilon, \quad 1 \leq i \leq N_f \quad (21)$$

Suppose $\hat{k}^{[t]}$ is the label with the largest probability, while $\hat{k}_{2\text{nd}}^{[t]}$ is the second largest. For the largest probability hypothesis $H_{\hat{k}^{[t]}}$, function $C(\cdot)$ is defined as follows:

$$C(H_i, H_{\hat{k}^{[t]}}) = \begin{cases} 1, & i = \hat{k}^{[t]} \\ -1, & i = \hat{k}_{2\text{nd}}^{[t]} \\ 0, & \text{otherwise} \end{cases}, \quad 1 \leq i \leq N_f \quad (22)$$

Combining and unfolding the equations above, the cost for the null hypothesis and the largest probability hypothesis are as follows:

$$\text{Cost}(H_0) = -\varepsilon \left[\sum_{k=1}^{N_f} \rho_k^{[t]} - N_f \log \left(\sum_{k=1}^{N_f} e^{\rho_k^{[t]}} \right) \right] \quad (23)$$

$$\text{Cost}(H_{\hat{k}^{[t]}}) = -(\rho_{\hat{k}^{[t]}} - \rho_{\hat{k}_{2\text{nd}}^{[t]}}) \quad (24)$$

If the cost associated with the null hypothesis is lower than the cost associated with the identification result, the algorithm will reject the current prediction and continue gathering data until the prediction is deemed credible or the TW

reaches a predefined upper limit. The conditions for accepting the prediction can be expressed as follows:

$$\varepsilon \leq \frac{-\left(\rho_{k^{[t]}} - \rho_{k_{2nd}^{[t]}}\right)}{-\left(\sum_{k=1}^{N_f} \rho_k^{[t]} - N_f \log\left(\sum_{k=1}^{N_f} e^{\rho_k^{[t]}}\right)\right)} \quad (25)$$

Figure 3 presents the flowchart of OACCA combined with the DW strategy.

3 Results

3.1 EEG data

The EEG signals were recorded using a 64-lead EEG amplifier developed by Neuracle, CN, with CPz set as the reference electrode and AFz set as the ground electrode. The original sampling rate was 1000 Hz. Subsequently, the signal was down-sampled to 250 Hz and passed through a 50 Hz notch filter to eliminate power line noise. For analysis purposes, only 9 channels located in the occipital region (Pz, PO5, PO3, POz, PO4, PO6, O1, Oz, and O2) were considered in both the offline experiments and the final competition.

A total of 15 subjects took part in the offline experiment. Each subject completed 4 blocks, with each block comprising 40 trials corresponding to 40 targets. Rest intervals were provided between

two blocks to allow subjects to take a break. The data used in the final competition were largely similar, except that they were collected from a different set of 15 subjects. Furthermore, the data from different blocks in the final competition were merged into a single block.

3.2 Performance evaluation

In the final of WRC 2022, we analyzed three algorithms employed by three competitors, namely STE, FBCCA, and OACCA, as previously introduced. All three algorithms were integrated with the DW strategy. For brevity, we will refer to them as STE + DW, FBCCA + DW, and OACCA + DW, respectively. We subsequently assessed the performance of each algorithm by considering identification accuracy, TW, and ITR for each block.

3.3 Offline experiments

Figure 4 displays the performance of each algorithm in the offline experiments. Among the three algorithms, OACCA + DW demonstrated the highest average accuracy and the shortest TW, resulting in the highest average ITR of 237.65 bits/min. This ITR was 38.5% higher than that of FBCCA + DW (171.56 bits/min) and 100.0% higher than that of STE + DW (118.80 bits/min). Throughout all blocks, OACCA consistently

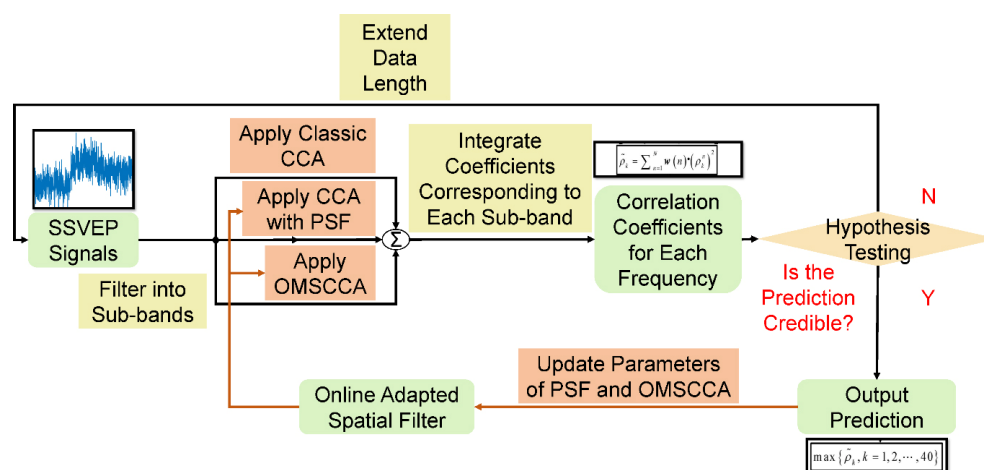


Fig. 3 Flowchart of OACCA with dynamic window.

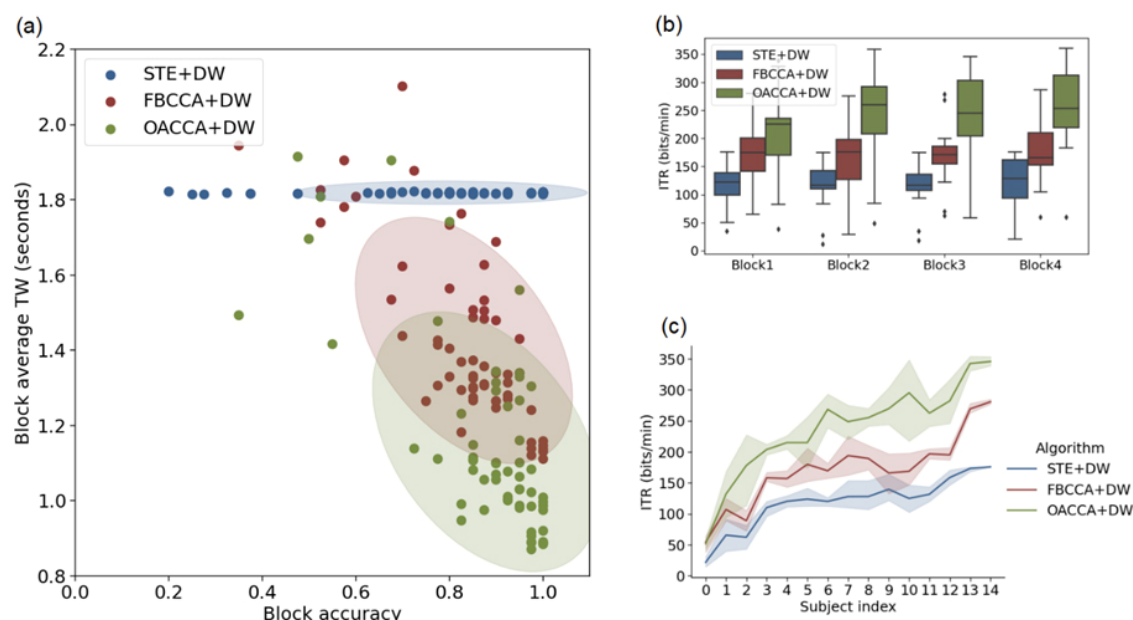


Fig. 4 Results of the offline experiments: (a) Scatter diagram of the block average identification accuracy and TW of each algorithm. Each dot represents the algorithm performance on a single block of a single subject. The shaded regions indicate half the standard deviation (SD) of the encoding models across all blocks and all subjects. (b) Average ITR in each block. (c) Average ITR for each subject. The plot shows the mean and 95% confidence interval for each algorithm. The subject indices are arranged based on the average ITR for all three algorithms.

achieved a significantly higher ITR compared to the other algorithms, particularly in Block 2, Block 3, and Block 4. Additionally, OACCA displayed better performance across most of the subjects.

3.4 Results in the final

Figure 5 illustrates the performance of each algorithm in the final competition. Similar to the offline experiments, OACCA + DW exhibited higher identification accuracy and shorter TW compared to the other algorithms. However, the ITRs of all three algorithms were slightly lower in the final competition compared to the offline experiments. OACCA + DW maintained the highest average ITR of 198.94 bits/min in the final competition. This ITR was 36.6% higher than that of FBCCA + DW (145.68 bits/min) and 96.2% higher than that of STE+DW (101.39 bits/min).

4 Discussion

In this paper, our focus was on three representative

algorithms used in the final of WRC2022 in a calibration-free scenario. We presented the underlying principles of these algorithms and evaluated their performance on two separate datasets. Our findings revealed that CCA-based methods outperformed non-CCA approaches, indicating that CCA has emerged as a fundamental technique for SSVEP target recognition. Furthermore, by incorporating subject-specific information, OACCA further enhanced the algorithm's performance.

The winning approach in this year's competition achieved an average ITR of 237.65 bits/min in the offline experiments, which is comparable to the highest ITR reported in a complete calibration-free setting (243.49 bits/min) [22] and also similar to that of an effective calibration-based algorithm (244.34 bits/min) [13]. In the final competition, the ITR reached 198.94 bits/min, surpassing the performance of the winning approach from the previous year (170.11 bits/min) [23] by more than 15%.

We observed a slightly lower ITR in the first

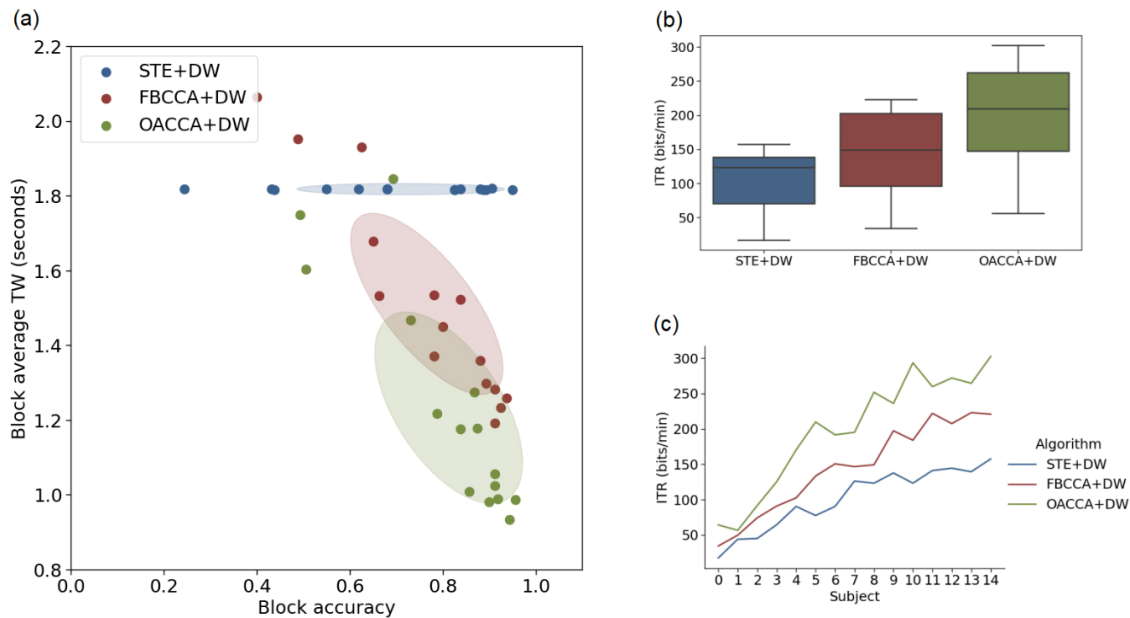


Fig. 5 Results in the final. (a) Scatter diagram of the block average identification accuracy and the TW of each algorithm. Each dot represents the algorithm performance on a single subject. The shaded regions indicate half the SD of the encoding models across all subjects. (b) Average ITR of each algorithm. (c) Average ITR for each subject. The subject indices are arranged based on the average ITR for all three algorithms.

block for OACCA + DW. This can be attributed to the trials required for online adaptation of parameters, such as spatial filters in PSF and OMSCCA, to converge to optimal values. To address this, we propose assigning a higher weight to FBCCA and using a higher accuracy threshold initially, and gradually reducing them as the performance of PSF and OMSCCA becomes more stable. Additionally, in real-world applications, subjects may initially have limited skill, resulting in changes in their optimal spatial filters over time. Therefore, when updating adaptive parameters, it may be beneficial to incorporate a decay rate γ as described below:

$$\mathbf{X}^{[t+1]} = \gamma \mathbf{X}^{[t]} + \delta, 0 < \gamma < 1 \quad (26)$$

The success of OACCA highlights the importance of incorporating prior knowledge into SSVEP-based BCIs. However, it is worth noting that other forms of prior information can also be valuable. For instance, algorithms based on transfer learning, such as domain generalization and domain adaptation, leverage labeled data

from other subjects to enhance performance on unseen subjects [24]. Additionally, analyzing data from previous trials of the same subject can help identify patterned noise (e.g., EOG, ECG, and EMG) and improve data quality.

All three competitors in the WRC2022 final utilized the DW strategy to enhance identification performance. Typically, a smaller time step leads to a higher ITR. However, it is important to consider the computational speed of the system in real-world applications. If the time step is shorter than the time required for calculations, the system may fail to provide real-time feedback, resulting in performance degradation. Therefore, when developing real-time BCI systems, it is crucial to consider the arithmetic speed of the device and simplify the algorithm if necessary.

The undergraduate group in WRC2022 utilized MATLAB as the basis for their framework. MATLAB, developed by MathWorks, is well-known for its strength in data analysis and signal processing. It offers a wide range of statistical, machine learning, and deep learning

algorithms for tasks such as categorization, prediction, and clustering. The processing power can be further enhanced by utilizing graphics cards on personal computers, workstations, or remote clusters. Moreover, MATLAB allows seamless data streaming between various hardware devices with sub-millisecond accuracy, which is crucial for real-time BCI systems [25]. Its user-friendly features, such as the “run section” and workspace functionality, make it accessible for beginner programmers.

The field of SSVEP-based BCIs has a relatively mature technical environment compared to other paradigms. With well-developed theories, frameworks, and datasets [26], it provides a conducive learning environment for beginners to study EEG signal features and their application in BCI systems through practical experience. The success of the first youth competition in WRC2022 demonstrates that with proper support and positive feedback, novice developers can quickly make progress and implement high-performance algorithms. Additionally, the youth competition has the potential to benefit BCI research by expanding the influence of BCIs and providing a talent pool for the research community. Therefore, we encourage more students interested in BCI systems to participate in such competitions and contribute to the advancement of related fields in the future.

5 Conclusion

This paper presents three algorithms (STE + DW, FBCCA + DW, and OACCA + DW) utilized by competitors in the MATLAB undergraduate group of WRC2022. Through a comparative analysis of the experimental results, we demonstrate the superiority of CCA-based methods among the three frequency recognition approaches. Additionally, we highlight the

effectiveness of incorporating higher harmonics and leveraging previous data to improve performance in the calibration-free scenario. Furthermore, we propose that the youth competition of WRC serves as an excellent platform for beginners to initiate their exploration and development of BCI systems.

Conflict of interests

All contributing authors report no conflict of interest in this work.

Funding

This work is granted by Open Project of Key Laboratory of Intelligent Computing & Signal Processing, Ministry of Education (Grant No. 2020A005).

Authors' contribution

X. Z., F. Y., and Y. W. implemented the algorithms under the guidance of J. C.; F. Y. performed the experiments and collected the data; C.Y. analyzed the data and interpreted the results; J. C. and Y.W. edited the paper with constructive discussions. All authors contributed significantly to manuscript preparation and approved the final version.

References

- [1] Gao XR, Wang YJ, Chen XG, et al. Interface, interaction, and intelligence in generalized brain–computer interfaces. *Trends Cogn Sci* 2021, **25**(8): 671–684.
- [2] Wolpaw JR, Birbaumer N, McFarland DJ, et al. Brain–computer interfaces for communication and control. *Clin Neurophysiol* 2002, **113**(6): 767–791.
- [3] Norcia AM, Appelbaum LG, Ales JM, et al. The steady-state visual evoked potential in vision research: A review. *J Vis* 2015, **15**(6): 4.

- [4] Cecotti H. Spelling with non-invasive brain–computer interfaces—Current and future trends. *J Physiol Paris* 2011, **105**(1/2/3): 106–114.
- [5] Müller-Putz GR, Scherer R, Brauneis C, et al. Steady-state visual evoked potential (SSVEP)-based communication: Impact of harmonic frequency components. *J Neural Eng* 2005, **2**(4): 123–130.
- [6] Lin ZL, Zhang CS, Wu W, et al. Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs. *IEEE Trans Biomed Eng* 2006, **53**(12 Pt 2): 2610–2614.
- [7] Bin GY, Gao XR, Yan Z, et al. An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method. *J Neural Eng* 2009, **6**(4): 046002.
- [8] Chen XG, Wang YJ, Gao SK, et al. Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface. *J Neural Eng* 2015, **12**(4): 046008.
- [9] Wong CM, Wan F, Wang BY, et al. Learning across multi-stimulus enhances target recognition methods in SSVEP-based BCIs. *J Neural Eng* 2020, **17**(1): 016026.
- [10] Lao KF, Wong CM, Wang Z, et al. Learning prototype spatial filters for subject-independent SSVEP-based brain-computer interface. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. Miyazaki, Japan, 2019, pp 485–490.
- [11] Yang C, Han X, Wang YJ, et al. A dynamic window recognition algorithm for SSVEP-based brain-computer interfaces using a spatio-temporal equalizer. *Int J Neural Syst* 2018, **28**(10): 1850028.
- [12] Wong CM, Wang Z, Nakanishi M, et al. Online adaptation boosts SSVEP-based BCI performance. *IEEE Trans Biomed Eng* 2022, **69**(6): 2018–2028.
- [13] Liu BC, Chen XG, Shi NL, et al. Improving the performance of individually calibrated SSVEP-BCI by task-discriminant component analysis. *IEEE Trans Neural Syst Rehabil Eng* 2021, **29**: 1998–2007.
- [14] Nakanishi M, Wang YJ, Chen XG, et al. Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis. *IEEE Trans Biomed Eng* 2018, **65**(1): 104–112.
- [15] 2022 World Robot Contest. <http://www.worldrobotconference.com/cn/about/138.html> (Accessed February 11, 2023).
- [16] Rappaport T. *Wireless Communications: Principles and Practice (2nd edition)*. New Jersey, USA: Pearson, 2001.
- [17] Friman O, Volosyak I, Gräser A. Multiple channel detection of steady-state visual evoked potentials for brain–computer interfaces. *IEEE Trans Biomed Eng* 2007, **54**(4): 742–750.
- [18] Yuan P, Chen XG, Wang YJ, et al. Enhancing performances of SSVEP-based brain–computer interfaces via exploiting inter-subject information. *J Neural Eng* 2015, **12**(4): 046006.
- [19] Chen YH, Yang C, Chen XG, et al. A novel training-free recognition method for SSVEP-based BCIs using dynamic window strategy. *J Neural Eng* 2021, **18**(3): 10.1088/1741–10.1088/2552/ab914e.
- [20] Volosyak I. SSVEP-based Bremen-BCI interface—boosting information transfer rates. *J Neural Eng* 2011, **8**(3): 036020.
- [21] Kuś R, Duszyk A, Milanowski P, et al. On the quantification of SSVEP frequency responses in human EEG in realistic BCI conditions. *PLoS One* 2013, **8**(10): e77536.
- [22] Shi NL, Li X, Liu BC, et al. Representative-based cold start for adaptive SSVEP-BCI. *IEEE Trans Neural Syst Rehabil Eng* 2023, **31**: 1521–1531.
- [23] Bian R, Wu DR. Overview of the winning approaches in BCI Controlled Robot Contest in World Robot Contest 2021: Calibration-free SSVEP. *Brain Science Advances* 2022, **8**(2): 99–110.
- [24] Bassi PRAS, Rampazzo W, Attux R. Transfer learning and SpecAugment applied to SSVEP based BCI classification. *Biomed Signal Process Contr* 2021, **67**: 102542.
- [25] Neuroscience - MATLAB and Simulink solutions. <https://ww2.mathworks.cn/solutions/neuroscience.html> (Accessed March 19, 2023).
- [26] Liu BC, Huang XS, Wang YJ, et al. BETA: a large benchmark database toward SSVEP-BCI application. *Front Neurosci* 2020, **14**: 627.



Chengzhi Yi is pursuing a B.Eng. degree in the School of Medicine, Tsinghua University, China. His current research interest focuses on language processing in the human brain. E-mail: ycz19@mails.tsinghua.edu.cn



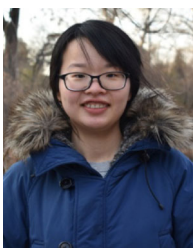
Yuxuan Wu received her Bachelor's degree from the Department of Biomedical Engineering, School of Medicine, Tsinghua University, China (2023). She is going to be a Ph.D. student starting in September 2023. She has a broad interest in neuro-engineering and computational neuroscience. E-mail: wyx@stanford.edu



Fan Ye received his Bachelor's degree from the Department of Biomedical Engineering, Tsinghua University, China (2023). He will be a Ph.D. candidate in Tsinghua University. His research focuses on computational neuroscience. E-mail: yef02thu@gmail.com



Xinchen Zhang is an undergraduate student majoring in mathematics and applied mathematics at the School of Science, Beijing Forestry University. Her research interests include brain-computer interfaces, steady-state visual evoked potential, and computer vision. E-mail: zhangxc@bjfu.edu.cn



Jingjing Chen received her Ph.D. degree from Department of Biomedical Engineering, Tsinghua University, China (2015). She is now a postdoctoral in Department of Psychology, Tsinghua University, China. Her current research interests focus on EEG-based affective computing. E-mail: chen-jj15@tsinghua.org.cn