UNIVERSITY OF LEEDS

Reliable and Accurate Brain-Computer Interface based on Steady-State Visual Evoked Potential

Yue Zhang

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

Faculty Engineering and Physical Sciences

School of Electronic and Electrical Engineering

February 2024

Intellectual Property

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

First and foremost, I am profoundly grateful to my supervisor, Dr. Zhi-Qiang Zhang, whose constant inspiration and guidance have motivated me beyond my expectations. It has been a delightful experience to do research under Dr. Zhang's supervision. I am immensely thankful for his constructive feedback on my research and daily life.

I would like to express my gratitude to Prof. Shane Xie and Dr. He Wang, the co-members of the supervision team, for their invaluable advice and consistent support throughout all stages of my research.

I would also like to express my gratitude to Dr.Zhenhong Li from the University of Manchester, Dr. Kun Qian from Heriot-Watt University, Dr. Yihui Zhao from the University of Bristol, and Tianzhe Bao from the University of Health and Rehabilitation Sciences for their assistance and support throughout my four years of research study.

My sincere gratitude and respect go to the research fellows in our group, Jie Zhang and Yue Shi, for their valuable advice and help. I also extend my thanks to my colleagues, Asma Alzaid, Chao Wang, Kaifeng Li, Haonan He, Shuhao Ma, Zhicheng Chen, Xin Wang, and Jiawei Jia for the fruitful collaboration and joyful moments we shared.

Lastly, I would like to convey my deepest gratitude to my family members, whose unwavering love and support have been the backbone of my study journey. I also want to thank all those who assisted me during this memorable four-year journey.

Yue Zhang

Leeds, 2024

Abstract

The brain-computer interface (BCI) facilitates a direct communication pathway between the human brain and external devices, bypassing the need for normal motor output pathways. Among various BCI methods, the steady-state visual evoked potential (SSVEP)-based BCI has gained significant attention due to its high signal-to-noise ratio (SNR), short training time, and rapid communication rate. It has been extensively explored in various applications, including assistive technologies, rehabilitation, communication, and entertainment. Despite some progress reported in recent literature, achieving reliable and accurate translations of user intentions in real-world scenarios remains highly challenging. This is mainly attributed to the instability of EEG signals and the disruptions encountered in practical situations. Besides, many existing systems only identify discrete commands, resulting in a gap between a user's cognitive intentions and the system's physical actions.

Within the framework of the SSVEP-based BCI system, the crucial significance of reliability and accuracy comes prominently to light. These attributes are pivotal for achieving precise command over external devices, simultaneously enriching the user experience and ensuring safety. Therefore, this thesis focuses on the development of a reliable and accurate BCI based on SSVEP technology, and the work is centered on four aspects: 1) To improve the recognition accuracy of the SSVEP signals; 2) To enhance the reliability of classification by rejecting low-confidence results; 3) To boost recognition performance for a new user through the incorporation of knowledge from existing users; 4) To apply the SSVEP-based BCI for controlling the velocity of the robotic arm according to the user's intention. To accomplish these objectives, the following efforts were undertaken: 1) A multi-objective

optimization-based high-pass spatial filtering method was proposed for improving SSVEP recognition accuracy. This approach has the potential to extract target-relevant features, reject target-irrelevant information, and mitigate the impact of volume conduction simultaneously. 2) A Bayesian-based classification confidence estimation method is proposed to improve recognition reliability. This method estimates the probability of correctness for each classification of the recognition system, allowing the identification and rejection of low-confidence results. The BCI system can make high-confidence decisions and mitigate potential errors by incorporating confidence information. 3) An inter-subject transfer learning method was proposed, leveraging SSVEP signals from source subjects to strengthen the recognition performance of a target subject. By transferring knowledge from existing users, this method enhances the adaptability of the BCI system, particularly for the new user with limited training data. 4) In pursuit of more natural and responsive control, a velocity modulation method based on stimulus brightness was integrated into the SSVEP-based BCI system. Unlike conventional approaches with fixed movement direction and speed, this method dynamically adjusts the movement direction and speed based on the subjects' intentions. Users can interactively modulate the robotic arm's movement by focusing on specific high- or low-brightness stimuli, leading to a more natural and intuitive control experience. In conclusion, this thesis contributes to the advancement of SSVEP-based BCI technology by developing a reliable and accurate system while exploring its practical applications.

Contents

List of Figures

List of Tables

Abbreviations

BCI	Brain-Computer Interface
EEG	Electroencephalography
SSVEP	Steady-State Visual Evoked Potential
SNR	Signal-to-Noise Ratio
ITR	Information Transfer Rate
CCA	Canonical Correlation Analysis
MsetCCA	Multiset Canonical Correlation Analysis
TRCA	Task-related Component Analysis
SSCOR	Sum of Squared Correlations
PSD	Power Spectral Density
JFPM	Joint Frequency and Phase Modulation
CCValue	Classification Confidence Value
GMM	Gaussian Mixture Model
SEM	Standard Error of Mean
ANOVA	Analysis of Variance
TW	Time Window
TAP	True Accept Proportion
TRP	True Reject Proportion
TL	Transfer Learning
DAV	Discrete Attenuated Velocity
DV	Discrete Velocity

Chapter 1

Introduction

1.1 Electroencephalography (EEG)-based Brain-Computer Interface (BCI)

The brain-computer interface (BCI) is a communication system that does not depend on the brain's normal output pathways of peripheral nerves and muscles [3]. In a BCI system, users modulate their brain activities to generate commands that interact with the environment instead of relying on body movements [2, 4]. Therefore, the BCI system offers an opportunity for paralyzed patients with severe motor disabilities to control external devices [5]. BCI systems were explored in a wide variety of applications, encompassing areas such as prosthetics, communication, entertainment, and smart home [6–8].

Brain signal recording can be either invasive or noninvasive. Invasive BCIs that require surgical intervention to implant electrodes and collect data from specific regions of brain tissue [2, 9]. Although implanted electrodes provide a higher signal quality, they restrict BCI's flexibility and convenience. On the contrary, noninvasive BCIs record brain signals with sensors over the scalp, which is healthier for the subject [10]. Noninvasive BCI systems normally rely on different modalities of functional neuro-imaging, such as electroencephalography (EEG) [11], functional near-infrared spectroscopy (fNIRS) [12], functional magnetic resonance imaging (fMRI) [13], and magnetoencephalography (MEG) [14]. Among the var-

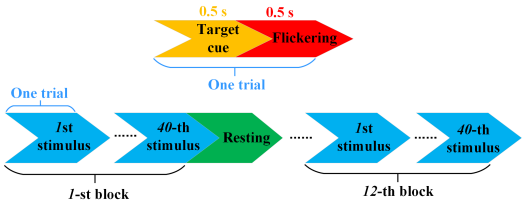
ious modalities, EEG is the most commonly used one due to its high temporal resolution, ease of use, and low cost [15–18]. EEG is the voltage variation on the scalp resulting from neurons' electrical activity in the cerebral cortex [2].

1.2 Steady-State Visual Evoked Potential (SSVEP)-based BCI

Four typical paradigms in the EEG signal, namely P300 event-related potential (ERP), slow cortical potential (SCP), sensorimotor rhythms (SMR), and steady-state visual evoked potential (SSVEP), are used for the analysis of brain activities [19]. SSVEP-based BCI has received extensive research interest in recent decades due to its fast communication rate, higher signal-to-noise ratio (SNR), and few user training requirements [20, 21].

The SSVEP-based BCI usually utilizes several visual oscillating stimuli, such as LEDs or boxes on a computer screen, which are generally modulated at different frequencies and phases [22, 23]. A typical experimental paradigm of the SSVEP-based BCI system contains M blocks, each containing N trials corresponding to N visual stimuli that flicker in random order. For example, Fig. 1.1(a) shows a typical stimulation experiment in [1], representing the setting for one of the highest numbers of stimuli. The user interface is a 5×8 matrix of visual stimuli, including 40 targets modulated by linearly increasing frequencies and phases, as shown in Fig. 1.1(b). In each experimental block, subjects were required to gaze at each visual stimulus for 0.5 s and completed 40 trials corresponding to all 40 targets. Each trial began with a 0.5 s visual cue that showed the target stimulus produced by the stimulus program. During the target cue period, users were required to shift their attention to the flickering target on the screen as quickly as possible. The subjects rested for a few minutes between two consecutive blocks to relieve visual and mental fatigue. Besides, subjects should avoid eye blinks during the experimental period to decrease artifacts generated by eye movements.

SSVEPs are periodic neural responses generated in occipital scalp areas of the brain, and the stimulus frequency will determine the response frequency content, which contains activities not only at the stimulus frequency but also at its higher harmonics [24]. The SSVEP can



(a) Scheme of the experimental paradigm.

Stimulus design of a 40-target high speed BCI

Freq. (Hz) / Phase (π)

8.0 / 0.00	9.0 / 1.75	10.0 / 1.50	11.0 / 1.25	12.0 / 1.00	13.0 / 0.00	14.0 / 0.50	15.0 / 0.25
8.2 / 0.35	9.2 / 0.10	10.2 / 0.00	11.2 / 1.60	12.2 / 1.35	13.2 / 1.10	14.2 / 0.85	15.2 / 0.60
8.4 / 0.70	9.4 / 0.45	10.4 / 0.2	11.4 / 1.95	12.4 / 1.70	13.4 / 1.45	14.4 / 1.20	15.4 / 0.95
8.6 / 1.05	9.6 / 0.80	10.6 / 0.55	11.6 / 0.30	12.6 / 0.05	13.6 / 1.80	14.6 / 1.55	15.6 / 1.30
8.8 / 1.40	9.8 / 1.15	10.8 / 0.90	11.8 / 0.65	12.8 / 0.40	13.8 / 0.15	14.8 / 1.90	15.8 / 1.65

(b) Stimulus design of the 40-target BCI system.

Figure 1.1: The redraw of the experimental paradigm and stimulus design in [1], which represented the setting for one of the highest numbers of stimuli.

be elicited up to at least 90 Hz [25]. Signal processing algorithms are applied to analyze the characteristics of SSVEP responses and identify the subject's intent to control the peripheral equipment. As a result, subjects can output desired commands by gazing at different target stimuli sequentially [22].

1.3 Research Motivations and Objectives

Within the context of SSVEP-based BCI systems, the critical importance of reliability and accuracy is prominently highlighted. These attributes play a pivotal role in attaining precise control over external devices, thereby enhancing the quality of the user experience and mitigating potential risks. Therefore, the core motivations of this thesis revolve around developing a reliable and accurate SSVEP-based BCI system. The following content will describe each motivation in detail.

The accuracy of recognition stands as a critical indicator to evaluate the effectiveness of SSVEP-based BCIs [26]. Research on recognition accuracy aids in reducing erroneous commands, enhancing the user experience, and improving communication efficacy. The target stimulus on which the user is focusing can be identified by analyzing the recorded SSVEP signals using various target identification methods. As the frequency components within SSVEP signals are determined by the stimulation frequency, employing frequency domain analysis facilitates the identification of the intended target stimulus [27, 28]. Additionally,

cortical source activities can be estimated by applying weighting coefficients to scalp EEG recordings collected from multiple electrodes [2]. Many methods have been introduced to extract spatial filters to reconstruct source activities, aiming to enhance the recognition accuracy of SSVEP-based BCI systems. However, previous spatial filter training commonly incorporated the signal from a single stimulus, and the knowledge from other stimuli was rarely included. In addition, EEG signals usually have low spatial resolution due to volume conduction. The majority of current research on volume conduction primarily concentrates on other EEG paradigms, neglecting its potential to enhance SSVEP recognition accuracy.

The effectiveness of many advanced recognition methods has been validated in SSVEP-based BCIs, but these methods generally rely on the largest correlation coefficient to determine the class. The classification performance can decline when the largest coefficient is insignificantly distinct from the remaining values. Wrong results can cause the external device to carry out the wrong actions, perhaps resulting in adverse incidents and serious physical harm to humans. In the practical usage scene, enhancing subjects' safety and security is essential, particularly in rehabilitation and assistive technology. Therefore, there arises a need to estimate the confidence level of classification results. Instances of low confidence should be rejected to enhance recognition reliability in SSVEP-based human-robot interaction.

To attain high performance in SSVEP-based BCIs, a large amount of data is typically required for model training. However, the process of data recording is time-consuming and physically exhausting, often leading to visual fatigue among the subjects [29, 30]. One promising method is to transfer data from one subject to another. Considering that each person has unique brain patterns [2], sharing EEG data directly from previous subjects to a new subject is not viable. To address this concern, transfer learning (TL) was developed, enabling the utilization of data from the source domain to enhance the recognition performance of the target domain [31]. By utilizing knowledge gained from previously recorded data, TL enhances the model's capacity to recognize SSVEP signals in the new domain, contributing to more efficient and reliable BCI applications.

A fundamental goal driving the evolution of SSVEP-based BCIs is to facilitate their integration into practical, real-world scenarios. Practical applications refer to situations where

the BCI system can provide timely responses based on the user's intentions or commands. SSVEP-based BCIs have been explored in various applications, such as wheelchair control, spelling assistance, and robotic arm manipulation. There remains a critical challenge to bridge the gap between a user's cognitive intentions and the system's physical actions. Specifically, the robot's movement direction and speed are generally fixed and prescribed, which neglects the user's actual requirements for the velocity adjustments. Users may prefer that the robot can change speed and direction according to their intentions in practical implementations. By enabling this level of adaptability, SSVEP-based BCIs can enhance the user experience and empower individuals to interact more naturally and effectively with the technology.

As a result, the four main objectives of this research are:

1. A target classification method aimed at enhancing SSVEP recognition accuracy will be introduced. This method utilizes multi-objective optimization high-pass spatial filtering techniques, with the goal of extracting target-related features, decreasing irrelevant information, and mitigating the negative effect of volume conduction.

2. A classification confidence estimation method based on Bayesian will be proposed for enhancing SSVEP detection reliability. The proposed method aims to enhance the system's reliability by accepting high-confidence results and rejecting low-confidence decisions.

3. A cross-subject transfer learning scheme will be proposed which incorporates SSVEP knowledge from the source subject to effectively strengthen the recognition performance for the target subject. By transferring this knowledge, the target subject can achieve optimized recognition performance without requiring extra calibration procedures.

4. A SSVEP-based BCI-controlled robotic platform with velocity modulation will be introduced. A stimulus brightness-based method will be designed to control robotic arm velocity, which has the potential to improve the robot's responsiveness to human intents and the user's interaction experience.

1.4 Thesis Outline

This thesis consists of 7 chapters. The initial two chapters encompass the introduction and literature review. Chapter 3 achieves the first objective by developing a multi-objective optimisation-based high-pass spatial filtering method for improving recognition accuracy. The second objective is achieved by Chapter 4, which proposes a Bayesian-based classification confidence estimation for enhancing SSVEP detection reliability. Chapter 5 establishes a cross-subject transfer learning method to accomplish the third objective. For the last objective of this research, a SSVEP-based BCI-controlled robotic platform with velocity modulation was proposed in Chapter 6. The last chapter presented the conclusion and future work. Specifically, this thesis is organized as follows.

Chapter 1 provides an introduction that elaborates on the SSVEP-based BCI system and offers an overview of its working principles. It highlights the research motivations underlying this thesis, provides the research objectives, and outlines the structure of the thesis, alongside presenting the list of publications.

Chapter 2 presents a systematic review of the SSVEP-based BCI system. Existing works on signal preprocessing and recognition are described in detail. Meanwhile, it encompasses an exploration of studies centered around classification confidence analysis, transfer learning methods, and SSVEP-based BCI-controlled robots.

Chapter 3 proposes a multi-objective optimisation-based high-pass spatial filtering method to improve the SSVEP identification performance. In this method, the filter was derived by maximising the correlation between the training signal and the individual template from the same target while minimising the correlation between signals from other targets and the template. Meanwhile, a constraint was imposed where the sum of spatial filter elements is zero. This configuration enables the passage of high spatial-frequency SSVEP signals while attenuating low spatial-frequency signals. Therefore, the proposed approach has the potential to extract the target-relevant features, reject the target-irrelevant information, and simultaneously diminish the impact of volume conduction.

Chapter 4 presents the research on classification confidence analysis in the SSVEP-based

BCI. This chapter proposed a Bayesian-based classification confidence estimation method for improving recognition reliability. Specifically, By leveraging the differences between the maximum correlation coefficient and the remaining values generated by a basic target identification method, a feature vector is constructed during the training phase. The Gaussian mixture model (GMM) is subsequently employed to estimate the probability density functions of these feature vectors from correct and incorrect classifications. During the testing procedure, Bayesian inference calculates the posterior probabilities for accurate and false classifications. Based on these probabilities, a classification confidence value (CCValue) is presented to estimate the confidence level in the classification. Finally, a decision-making rule is applied to determine whether the current classification result should be accepted or rejected.

Chapter 5 introduces an inter-subject transfer learning method for SSVEP recognition. This method leverages SSVEP signals from source subjects to effectively strengthen the recognition performance of a target subject. Specifically, the proposed method involves generating a multidimensional feature vector that combines the transferred spatial filter and the transferred SSVEP template from the source subject, along with the spatial filter of the target subject obtained through multiple-covariance maximization. A contribution score is assigned to each source subject by considering the distance between it and the target subject during the feature vector construction.

Chapter 6 proposes a velocity modulation method to control the robotic arm in the SSVEP-based BCI. The current control strategies generally offer velocities with fixed and prescribed directions and speeds during robot movements. This chapter introduced a velocity modulation method based on the subject's intention. The subject could change the velocity of the robotic arm by focusing on the flickers with different frequencies and brightnesses. For performance evaluation, online experiments involving single- and multi-target reaching tasks were conducted.

Chapter 7 concludes the efforts of this thesis. It also discusses some potential directions for future research exploration.

1.5 Publication List

The thesis contains materials that have been submitted or published in journals and conferences. The publication list is shown as follows:

Chapter 2

Zhang, Y., Xie, S. Q., Wang, H., & Zhang, Z. Q. (2020). Data analytics in steady-state visual evoked potential-based brain–computer interface: A review. IEEE Sensors Journal, 21(2), 1124-1138.

Chapter 3

Zhang, Y., Zhang, Z. Q, & Xie, S. (2021, July). Multi-objective optimisation for SSVEP detection. In 2021 IEEE 17th International Conference on Wearable and Implantable Body Sensor Networks (BSN) (pp. 1-4). IEEE.

Zhang, Y., Li, Z., Xie, S. Q., Wang, H., Yu, Z., & Zhang, Z. Q. (2022). Multi-objective optimization-based high-pass spatial filtering for SSVEP-based brain–computer interfaces. IEEE Transactions on Instrumentation and Measurement, 71, 1-9.

Chapter 4

Zhang, Y., Xie, S. Q., Wang, H., Shi, C., & Zhang, Z. Q. (2023). Bayesian-based Classification Confidence Estimation for Enhancing SSVEP Detection. IEEE Transactions on Instrumentation and Measurement, 72, 1-12.

Chapter 5

Zhang, Y., Xie, S. Q., Shi, C., Li, J., & Zhang, Z. Q. (2023). Cross-subject transfer learning for boosting recognition performance in SSVEP-Based BCIs. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 31, 1574-1583.

Chapter 6

Zhang, Y., Qian K., Xie, S. Q., & Zhang, Z. Q. (2023). SSVEP-based Brain-Computer Interface Controlled Robotic Platform with Velocity Modulation. IEEE Transactions on Neural Systems and Rehabilitation Engineering.

Case study

Zhang, Y., Xie, S. Q., Li, Z., Zhao, Y., Qian, K., & Zhang, Z. Q. (2022, September). CCA-based Spatio-temporal Filtering for Enhancing SSVEP Detection. In 2022 IEEE-EMBS International Conference on Wearable and Implantable Body Sensor Networks (BSN) (pp. 1-4). IEEE.

1.6 Chapter Summary

This chapter introduces the concept of SSVEP-based BCI. This thesis's main motivations are: 1) improving recognition accuracy; 2) enhancing classification reliability; 3) reducing time-consuming calibration; and 4) implementing velocity modulation within the practical scenario. The main objectives of this research include: 1) designing a multi-objective optimisation-based high-pass spatial filtering method for improving recognition accuracy; 2) developing a Bayesian-based classification confidence estimation method for enhancing SSVEP detection reliability; 3) introducing a cross-subject scheme to transfer SSVEP knowledge from the source subject and enhance the recognition performance of the target subject; and 4) presenting an SSVEP-actuated velocity modulation method for controlling a robotic arm. The thesis outline and publication list were also presented in this chapter.

Chapter 2

Literature Review of Steady-State Visual Evoked Potential-based Brain-Computer Interface

This chapter provides a systematic review of the SSVEP-based BCI system. It begins with signal preprocessing and SSVEP recognition methods, which encompass Fourier transform-based spectrum analysis, canonical correlation analysis (CCA)-based methods, nonlinear classification methods, spatial filtering methods, and so on. The chapter then explores the factors that influence recognition reliability and reviews recent advancements in classification confidence analysis. Subsequently, transfer learning methods are discussed, specifically focusing on template-based transfer and spatial filter-based transfer techniques. The integration of SSVEP-based BCIs with robotic systems is examined, covering communication aids, prosthetics control, and environmental control. At the end of this chapter, research gaps are summarized.

2.1 Signal Pre-processing Methods

The EEG potentials gathered by electrodes come from the brain, which can be easily contaminated by muscle activation, eye movement, and external artifacts [32]. Therefore, it is

necessary to pre-process the raw EEG signal to achieve a higher SNR before the SSVEP recognition step. Thus far, there are mainly two pre-processing methods: band-pass filtering and notch filter.

The band-pass filter is the most common pre-processing filtering to remove the noises whose frequencies are not overlapped with SSVEP responses. The band-pass is utilized to retain the pertinent parts of the EEG signal, which correspond to the stimulation frequencies and harmonics. Dividing the SSVEP signal into distinct frequency bands, the sub-band signal within the designated frequency range can be extracted for further analysis and interpretation. Many works about SSVEP-based BCI have adopted band-pass filter as the signal pre-processing algorithm, such as [1, 24, 33].

In most countries, power line interference is commonly concentrated near 50 Hz or 60 Hz [34,35]. The most prevalent method might be to filter the data in the respective frequency band with a notch filter (a band-stop filter with a narrow stop band) [36]. This filter effectively suppresses power within the specific stop band at 50 Hz or 60 Hz.

2.2 SSVEP Recognition Methods

The primary objective of pursuing high-performance BCI systems is to achieve accurate SSVEP recognition within a limited time window (TW). This subsection comprehensively explores various frequently employed SSVEP recognition algorithms, including Fourier transform -based spectrum analysis, signal decomposition-based analysis, basic spatial filtering methods, canonical correlation analysis-based methods, nonlinear classification methods, and recent spatial filtering methods. The technical details of these methods are further summarized in Table. 2.1. Moreover, many classifiers utilized in the SSVEP identification are also presented in this subsection.

2.2.1 Fourier Transform-based Spectrum Analysis Methods

One of the simplest detection approaches for SSVEP-based BCIs is power spectral density analysis (PSDA) which is based on the fast Fourier transform (FFT). By transforming

the time domain EEG signals to the frequency domain, the amplitude of each stimulation frequency can then be used for further classification [37]. The frequency corresponding to the peak is taken as the visual stimulus frequency. Many works [27, 38, 39] employ Fourier transform in SSVEP detection due to its small computation time and simplicity. The recognition methods based on PSDA rely on the fact that a periodic pattern with the same frequency as the stimulus frequency or one of its harmonics can be traced back in the brain signals. When the SSVEP is present in the brain signals, the amplitude of its periodic pattern is confined to a narrow frequency bandwidth, allowing for easy measurement in the frequency domain [28].

The PSDA method exhibits sensitivity to noise when using a single or bipolar channel, requiring a comparatively extended time window for accurate spectrum estimation with sufficient frequency resolution [40]. These limitations contribute to relatively reduced SSVEP recognition accuracy, especially when the time window is insufficient (e.g., less than 3 s) [41]. Despite aiming for high accuracy, the information transfer rate (ITR) is also affected. Some efforts attempted to improve the performance of PSDA methods through optimising parameters, such as electrode and time length selection. For example, Bin *et al* [20] used an exhaustive method to select the optimized bipolar lead to maximize the stimulus frequency SNR. Wang *et al* [42] proposed an adaptive data length regulation method. Real-time frequency detection is performed every short interval using the buffer data. If the same frequency is detected consecutively, a decision is made, followed by buffer clearing for new data input. This method improves the BCI system's ITR while reducing inter-user variation and bolstering system flexibility and reliability.

2.2.2 Signal Decomposition-based Analysis Methods

The FT summarizes the signal's spectrum over the entire data collection period [41]. The FT method demonstrates its efficacy in scenarios with stationary systems. However, in non-stationary SSVEP signals, FT cannot establish a direct correlation between the temporal signal changes and the spectrum's frequency features. The wavelet transform (WT) is a mathematical approach that gives the time-frequency representation of a signal with

the possibility to adjust the time-frequency resolution [43]. In other words, it not only reveals information about the frequency components within a signal but also the time of their occurrence. The WT is widely acknowledged as an effective time-frequency analysis tool for analyzing various physiological signals [44]. For example, Rejer *et al.* [45] employed wavelet analysis to detect SSVEP frequencies concerning the synchronization time and its strength. In the WT, a sample function known as the mother wavelet represents the SSVEP signal in the time and frequency domains. However, there is no universal mother wavelet function that fits all signals, rendering the selection of an appropriate mother wavelet function a potential challenge in WT-related studies. Sayilgan *et al* [46] investigated the effect of the mother wavelet function in classifying two distinct flickering frequencies for BCI applications. The findings indicated that the Haar wavelet function emerged as the most representative, while Symlet 4 performed the worst. Although the WT effectively processes non-stationary signals, it may be inappropriate for handling highly complex SSVEP signals that exhibit nonlinear dynamics and chaos [47]. This is because the basic idea of FT is to represent the EEG sequence as a linear superposition of sinusoidal waves (sines and cosines). In the WT, a sequence is represented as a linear combination of wavelets [48].

Huang [49] proposed the Hilbert-Huang transform (HHT), which encompasses Empirical Mode Decomposition (EMD) and the Hilbert Transform (HT). EMD, a non-linear technique, is adept at handling dynamic and complex signals. EMD adaptively decomposes signals into a group of intrinsic mode functions (IMFs), showcasing oscillation features in the non-stationary signals [34]. Subsequently, the HT is applied to each IMF to derive the amplitude distribution, known as the Hilbert spectrum. Many studies [50,51] have employed EMD successfully to achieve frequency recognition and enhance classification accuracy in SSVEP-based BCIs. For example, Zhao *et al* [52] used the Hilbert spectrum from the different frequency bands to construct energy vectors, which are then input to a classifier for SSVEP identification. Tello *et al* [53] calculated the PSD of the IMFs, and then a rule-based classifier identified the maximum values and determined the selected class through a majority vote. In addition, the EMD can also be integrated with CCA, in which the IMFs containing the highest energy levels are chosen and fed into CCA for SSVEP detection. [51].

Compared with FFT and WT, HHT has better universality to handle nonlinear and non-stationary signals. It not only absorbs the advantages of multi-resolution in WT but also overcomes the difficulty of selecting an appropriate wavelet base.

2.2.3 Basic Spatial Filtering Methods

Spatial filtering plays a crucial role in enhancing performance by improving the SNR of SSVEP signals [2]. The cortical source activities can be estimated by applying weight coefficients to scalp EEG recordings from multiple electrodes. Built upon this concept, various approaches have been developed that extract optimal spatial filters (i.e., weight coefficients) to reconstruct source activities from scalp recordings [54].

Minimum energy combination (MEC) and maximum contrast combination (MCC) are common spatial filtering algorithms with different objective functions. In the MEC approach, noise components inherent to the original signals are initially extracted, followed by the attainment of a linear combination to suppress the noise signal [55]. Finally, this spatial filter is applied to the original signals, producing signals with reduced noise [56, 57]. The SNR values are measured for each stimulus frequency, and the frequency that yields the maximum SNR is considered the frequency of interest [58]. On the other hand, the MCC approach aims to construct a weight vector that maximizes the ratio between the power of SSVEP activity and that of non-SSVEP related activity [21, 55, 59].

Common spatial pattern (CSP) [60, 61] is another spatial filter to improve the distinction between EEG signals from the stimulus and non-stimulus situations. Two distributions are in a C-dimensional space, where C is the number of channels. CSP attempts to find projections minimizing the variance of one class while maximizing the variance of the other. Specifically, in the context of SSVEP-based BCIs, the presence of a stimulus leads to increased signal variance compared to the non-stimulus state. As a spatial filtering method to enhance SSVEP signals, CSP is commonly combined with separate steps for feature extraction and classification, allowing for the differentiation of various stimulation frequencies [21]. For example, in [60], amplitude estimations of filtered SSVEPs at different stimuli were extracted, followed by applying linear discriminant analysis (LDA) for classification.

Both MEC and MCC methods compute optimal spatial filters and noise power estimates for each new EEG data trial. While these strategies are advantageous since it accounts for the nonstationary nature of unwanted signals, they could lead to extended computational time during real-time operations, particularly when a substantial number of electrodes are involved [21].

2.2.4 Canonical Correlation Analysis (CCA)-based Methods

Canonical Correlation Analysis The CCA method is used to find the relationship between two sets of data, which can be used as a feature extraction algorithm in SSVEP-based BCIs. The CCA-based spatial filter, first presented by Lin *et al.* [38], has attracted much interest in recent years due to better SNR, higher recognition accuracy, and well usage of harmonic frequencies [21]. The CCA attempts to find a pair of linear combinations of the multi-channel signals and the artificial reference signals, generally sine and cosine waves, that have the correlation maximization at each stimulus frequency. Then, the frequency related to the maximal correlation coefficient is determined as the target [38, 62]. Nowadays, many improved CCA-based methods are proposed due to higher requirements of performance or the drawbacks of CCA, e.g., the artificial reference signals lack true information of EEG data, and multi-channel signals are easily influenced by background noise such as spontaneous EEG.

Multiway canonical correlation analysis (MwayCCA) Before introducing MwayCCA, the concept of tensor should be first referred to. A tensor is a multiway array of data, and its order is the number of dimensions, also called ways [63]. Tensor CCA is a standard CCA development that concentrates on calculating the correlation between two multiway data arrays rather than two sets of variables based on vector [64]. Based on this concept, MwayCCA optimizes the reference signals by maximizing the correlation between third-order EEG data tensor (channel $\times$ time $\times$ trial) and pre-constituted sine-cosine reference signal matrix (harmonic $\times$ time) [65]. Then, target frequency can be recognized by applying multiple linear regression (MLR) or CCA between test EEG data and optimized reference signals [65]. In MwayCCA, the EEG tensor is constructed by multiple trials where some

trials may contain more artifacts which generally have a negative contribution to the reference signal optimization. Therefore, L1-regularization is implemented on trial-way array optimization of MwayCCA to remove obstruction trials [66].

MwayCCA and its variation add a reference signal optimization procedure to enrich the reference signal with more real information about EEG signals, thereby improving the performance of standard CCA.

Multiset canonical correlation analysis (MsetCCA) The originally constructed reference signals with sine-cosine waves are generally short of real information of EEG data, which goes against SSVEP frequency recognition. Multiset canonical correlation analysis (MsetCCA), proposed by Zhang *et al.* [67], considers common features shared by EEG signals may be more real and natural than predefined signals. For a specific subject, some common characteristics are contained in a set of trials at a certain stimulus frequency, which can be used to construct optimal reference signals to achieve a higher detection accuracy. Specifically, MsetCCA learns multiple linear transforms that maximize the overall correlation among canonical variates from multiple sets of random variables [67]. Therefore, in the SSVEP-based BCIs, the optimal reference signals can be determined by MsetCCA through the joint spatial filtering of multiple sets of EEG training datasets for each stimulus frequency [68]. Jiao *et al.* [24] further presented a three-layer model based on MsetCCA, named multilayer correlation maximization (MCM), which adopts superiorities of both CCA and MsetCCA to avoid extracting the background noise as common features.

MsetCCA produces fully optimized reference signals based on the EEG signal training set. It turns out that the averaged classification accuracy and ITR of MsetCCA are better than those of MwayCCA and CCA [68]. However, one drawback is that it may treat background noises as common features, so it needs to be used with other denoising algorithms.

Filter bank canonical correlation analysis (FBCCA) Considering those harmonic SSVEP components are not employed for frequency recognition, Chen *et al.* [69] incorporated fundamental and harmonic frequency components to propose a new method called filter bank canonical correlation analysis (FBCCA). The FBCCA method contains three

steps. Firstly, a filter bank analysis implemented sub-band decomposition from EEG signals with multiple filters with different pass-bands. And then, CCA is employed to calculate the correlation between the sub-band components and the constructed reference signals with sine-cosine waves related to all stimulation frequencies. Finally, a weighted sum of squares of the correlation for all sub-band components is combined as the final feature for frequency identification.

FBCCA was often combined with current innovative methods in [70, 71], thereby further optimizing them and achieving higher detection performance. It can be seen that FBCCA is expected to become a new standard paradigm after CCA.

Individual template canonical correlation analysis (IT-CCA)-based methods The individual template-based CCA (IT-CCA) was first proposed in [72] to optimize the reference signals with sine-cosine waves by detecting temporal features of EEG data. The IT-CCA calculates the canonical correlation between test data and individual template signals acquired by averaging multiple training trials. Nakanishi *et al.* [73, 74] developed it and proposed a combination method of CCA and IT-CCA that applies three weight vectors as spatial filters for enhancing the target detection. They are spatial filter between test data and the individual template, spatial filter between test data and preconstructed reference signals, and spatial filter between the individual template and preconstructed reference signals, respectively. Then four correlation vectors as the above spatial filters obtain recognition features, and an ensemble classifier is employed to combine four vectors to form a weighted correlation coefficient as the final feature [68].

The individual template-based target identification has frequently been integrated with various spatial filtering techniques [71, 75, 76]. The diagram of the feature extraction is illustrated in Fig. 2.1. Spatial filters are obtained from the training data corresponding to each stimulus. These filters are subsequently applied to individual templates and input EEG data, thereby improving the SNR of SSVEP signals.

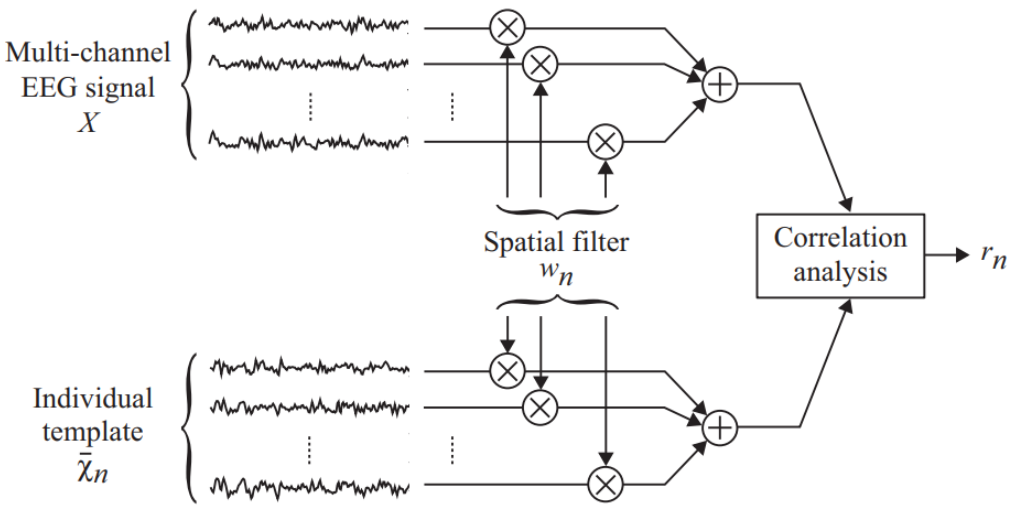


Figure 2.1: Framework of feature extraction including spatial filtering in the individual template-based method for detecting SSVEP signals [2].

2.2.5 Nonlinear Classification Methods

The transformation of CCA maximizes the mutual information between extracted multi-dimension features, but dealing with nonlinear relations in real signals is infeasible. Zhang *et al.* [77] presented a kernel CCA (KCCA)-based method for detecting idle states in SSVEP-based BCI systems. This method provided a practicable way to extract nonlinear characteristics of multi-dimension EEG signals. Through the utilization of a kernel function, both the EEG signal and reference signal are projected into a higher-dimensional feature space. Andrew *et al.* [78] developed deep CCA (DCCA) which processes input data through the deep network before calculating their correlations. DCCA is designed to learn complex nonlinear transformations of two data views, aiming to yield highly linearly correlated representations. Liu *et al.* [33] proposed an extension of DCCA, named deep multiset CCA (DMCCA) for SSVEP frequency recognition, that extracts the information within the real EEG signals to attain better detection accuracy.

The above CCA-based nonlinear frequency recognition algorithms align with the original EEG signals' nonlinear characteristics, leading to better results than the CCA. For KCCA, choosing the appropriate kernel is still a question worth considering. DMCCA achieved better recognition performance by combining the nonlinear method DCCA and linear method

MsetCCA, which provides a potential research direction.

The convolutional neural network (CNN) is another popular classifier for SSVEP-based BCIs. For instance, Kwak *et al.* [79] proposed a CNN-based classifier that uses frequency features as input for robust SSVEP detection. Guney *et al* [80] introduced a deep neural network (DNN) architecture designed to process multi-channel SSVEP data. This architecture employs convolutions across sub-bands of harmonics, channels, and time, ultimately conducting classification at the fully connected layer. With this background, neural network-based classifiers are more potential and efficient options to achieve higher accuracy with a mass of EEG data. Meanwhile, it is worth noting that broader knowledge and more time or more data are needed for adjusting related parameters and training feasible models [81].

2.2.6 Recent Spatial Filtering Methods

Multivariate synchronization Index (MSI)

Zhang *et al* [82] proposed a multivariate synchronization index (MSI) method for frequency recognition. MSI aims to estimate the synchronization between the observed mixed signals and the reference signals, similar to CCA. The synchronization index is quantified using the S-estimator [83], which relies on the entropy of the normalized eigenvalues of the correlation matrix of multivariate signals. Considering that the MSI method does not fully exploit SSVEP-related harmonic components, Qin *et al* [84] proposed a filter bank-driven MSI method (FBMSI) to improve the accuracy of SSVEP recognition further. Before the standard MSI method, the SSVEP signal was decomposed by the filter into multiple subbands. Extracting the harmonic components from each subband can achieve a better recognition effect than the standard MSI method.

Sum of Squared Correlations (SSCOR)

Rupnik *et al* [85] introduced the sum of squared correlations (SSCOR) to obtain a cross-lingual document similarity measure. G.R. *et al* [71] further employed it in SSVEP signal analysis. The SSCOR framework aims to find a mapping that projects the given EEG data onto a common SSVEP representation space. The method is based on the sum of squared

correlation formulation, where correlations between pairs (i.e., individual template and each training trial) are considered. SSCOR has been incorporated in many SSVEP recognition research papers [75, 86] to facilitate performance comparison.

Task-related component analysis (TRCA) Many techniques [87, 88] have been developed to extract task-related source signals from scalp recordings based on the idea that cortical source activities can be rebuilt through a weighted linear summation of EEG signals from multiple electrodes. Tanaka *et al.* [89, 90] proposed task-related component analysis (TRCA), which performs better than other task-related methods due to maximizing the reproducibility of time-locked activities across trials. In 2017, Nakanishi *et al.* [1] introduced TRCA-based analysis to EEG study especially SSVEP-based BCI systems, which successfully enhanced the SNR of EEG signals by eliminating the background noises and showed great capacity for different applications in communication and control. SSVEPs are time-locked photic-driving responses related to repetitive visual stimuli. Therefore, TRCA-based techniques can achieve higher SNR of EEG signals [2, 22].

Task-Discriminant Component Analysis (TDCA)

Liu *et al* [91] proposed a novel method called task-discriminant component analysis (TDCA) to improve the performance of individually calibrated SSVEP-BCI further. In contrast to the generative model of TRCA, where the projection direction is learned class by class, TDCA utilizes a discriminative model to learn the projection direction for all data classes. Furthermore, TDCA incorporates the temporal information of SSVEP into the training and test samples. TDCA optimizes the spatiotemporal filter by maximizing the between-class difference while minimizing the within-class difference.

2.2.7 Traditional Pattern Recognition Methods

In addition to the target identification methods as mentioned above, some traditional pattern recognition methods involving classic classifiers such as LDA, SVM, and k-nearest neighbour (kNN) are also usually used for SSVEP classification scheme [92, 93]. Features corresponding to different visual stimuli are regarded as the feature vector to train the

Table 2.1: SSVEP Recognition Method

Categories	Methods	Technical details	References
FT-based spectrum analysis	PSDA	PSDA is based on the FFT. By transforming the time domain EEG signals to frequency domain,amplitudes of each stimulation frequency are obtained, and then the frequency with the largest amplitude is determined as the target. FT shows poor performance when signals are non-linear and unstable.	[27, 38, 39]
	Channel Selection PSDA	An exhaustive method was used to select the optimized bipolar lead to maximize the stimulus frequency SNR	[20]
	Adaptive PSDA	Real-time frequency detection is performed every short interval using the buffer data. If the same frequency is detected consecutively, a decision is made.	[42]
Signal decomposition-based analysis	WT	WT gives the time-frequency representation of a signal with the possibility to adjust the time-frequency resolution	[45, 46]
	HHT	HHT includes EMD decomposition and Hilbert transformation. EMD adaptively decomposes signals into intrinsic mode functions (IMFs). The HT is applied to each IMF to derive the amplitude distribution.	[52]
	EMD+classifier	It calculated the PSD of the IMFs, and a rule-based classifier identified the maximum values and determined the selected class through a majority vote.	[53]
	EMD+CCA	The IMFs containing the highest energy levels are chosen and fed into CCA for SSVEP detection	[51]

Basic spatial filtering methods	MEC	MEC finds a spatial filter projecting the multi-channel signal to a low-dimensional combined one in order to weaken background noises.	[56, 57]
	MCC	MCC attempts to make the energy in the SSVEP frequencies is maximized through the computation of a weight matrix.	[21, 55, 59]
	CSP	It aims to maximize the SNR of SSVEP responses against the non-stimulus situation.	[60]
CCA-based linear methods	CCA	CCA tries to find a pair of linear combinations of multi-channel EEG signals and sine-cosine reference signals that have the maximum correlation with each other.	[38, 62]
	MwayCCA	Calculating the correlation between two multiway data arrays rather than vector variables.	[65]
	MsetCCA	MsetCCA extracts common features shared by the real EEG signals to optimize reference signals.	[67]
	MCM	In order to avoid extracting the background noise as common features, the MCM adopts superiorities of both CCA and Mset-CCA.	[24]
	FBCCA	It incorporates fundamental and harmonic frequency components together for target detection.	[69]
	IT-CCA	The reference signal is an individual template acquired by averaging multiple training trials.	[72]
	CCA +IT-CCA	Three weight vectors are applied as spatial filters which form four correlation vectors as recognition features.	[73, 74]

Nonlinear recognition methods	KCCA	The kernel is applied to project the data to high-dimension space to solve the problem that CCA is infeasible for nonlinear signals. How to choose the appropriate kernel is still a question worth thinking about.	[77, 94]
	DCCA	In DCCA, deep networks are used to process input data before the CCA.	[78]
	DMCCA	Combining the nonlinear method DCCA and linear method MsetCCA.	[33]
	CNN	A three-layer CNN-based classifier that uses frequency features as input for robust SSVEP detection.	[79]
	DNN	It employs convolutions across sub-bands of harmonics, channels, and time, ultimately conducting classification at the fully connected layer.	[80]
Recent spatial filtering methods	SSCOR	The spatial filter is based on the sum of squared correlation formulation.	[71]
	MSI	MSI aims to estimate the synchronization between the observed EEG signals and the reference signals. The synchronization index is quantified using the S-estimator, which relies on the entropy of the normalized eigenvalues of the correlation matrix of multivariate signals.	[82, 84]
	TRCA	TRCA extracts task-related components efficiently by maximizing the reproducibility of time-locked activities across trials.	[1]
	TDCA	TDCA utilizes a discriminative model to learn the projection direction for all data classes. It also uses the temporal information of SSVEP in the training and test samples.	[91]

classifier based on training data. Then, the experiment is conducted on the testing data with the trained classifier to determine targets. For example, in [95], the PSD in all possibly evoked frequency bands is extracted from the SSVEP responses to facilitate the discrimination task. In this work, three classifiers, namely LDA, SVM, and extreme learning machines (ELM), are performed at the target detection stage, and the ELM shows a more promising classification capacity in the context of SSVEP. Therefore, it proves the good generalization performance of neural network-based methods for SSVEP classification.

2.2.8 Discussion

As previously mentioned, several methods have been developed to classify the SSVEP signals, mainly including Fourier transform-based spectrum analysis, signal decomposition-based analysis, and the spatial filtering method. Among these methods, CCA-based techniques have received significant attention for enhancing the SNR of SSVEP signals, consequently leading to improved classification accuracy. Several subsequent studies have focused on exploring the correlation across training trials of each subject, resulting in the development of methods like SSCOR and TRCA. For example, in TRCA, weight coefficients are optimized by inter-trial covariance maximization. These methods have demonstrated advanced performance in SSVEP classification tasks. However, current approaches predominantly learn spatial filter parameters for a specific target using only the training data from the same stimulus. They often overlook incorporating information from other stimuli and the volume conduction problem during training. Consequently, further relevant efforts are necessary to improve the classification accuracy in SSVEP-based BCI systems.

2.3 Methods for Improving Recognition Reliability

The last section introduced many studies to improve recognition accuracy in SSVEP-based BCIs. Several of these methods determine the target class by finding the largest correlation coefficient. However, the classification performance usually degrades when the largest coefficient is not significantly different from the rest of the values. Additionally, wrong classifications of user intentions also exist due to various factors in real scenarios, such as

intra-subject variability and environmental distraction. In this context, improving recognition reliability is vital for SSVEP-based BCIs, as it ensures safety, enhances user experience, and prevents errors in practical applications. In this section, several representative disturbances that affect the system's robustness will be presented. Additionally, a literature review on the methods that enhance the reliability of SSVEP-based BCI systems also will be provided. The technical details were also shown in Table.2.2.

2.3.1 Disturbances in Practical Scenarios

Intra-subject Variability

As the experiment progresses, the subject may experience mental fatigue due to uncomfortable light twinkling and contrast changes in visual stimuli [96]. This can lead to variations in the subject's cognitive and electrophysiological state, potentially affecting the consistency of SSVEP responses. Moreover, the quality of SSVEP signals may be negatively affected, resulting in a degradation of the practical performance of the SSVEP-based system [97]. On the other hand, some user issues, like the body movement and the head's orientation, may also cause intra-subject variability during the experiment [47]. This may change the position of the electrodes relative to the scalp, leading to reduced signal quality.

Environmental Distraction

The current SSVEP-based BCI recognition methods in the controlled laboratory setting provide nearly satisfactory performance. In these circumstances, external noises-induced brain activities were largely avoided, and only the EEG signals evoked by the SSVEP experiments were recorded [98]. However, in practical scenarios, environmental distractions may divert the user's attention, compromising their ability to maintain focused and consistent SSVEP responses [99, 100]. On the other hand, the long-term quality of the contact of the electrode with the scalp skin is essential for EEG measurement. In many SSVEP applications, it is imperative to ensure a stable and low contact impedance sustained over extended periods [101]. However, the electrode-skin impedance may vary due to factors such as sweat or humidity, which can affect the reliability of SSVEP recognition.

2.3.2 Classification Confidence Analysis

To improve the reliability of SSVEP recognition in real-time scenarios, it is desirable to mitigate the negative impact of misclassification by analyzing model uncertainty. One approach to address uncertain decisions is through classification confidence analysis. By estimating the confidence level of the classification results, a post-decision operation can be implemented to handle uncertain decisions. For example, Lamti *et al* [100] presented a hybrid system for wheelchair control that integrates gaze data and SSVEP. In this system, the user distraction level was estimated using behavioral entropy. Based on the distraction levels and the issued commands, the validation module generates validated commands for the wheelchair control. Zhao *et al* [102] designed a decision-making module to integrate classification decisions of different frequency recognition methods based on CCA. To measure the uncertainty of each decision, feature vectors were extracted using the two methods' largest and second-largest correlation coefficients. The classification module classifies the feature vector into two classes, Method 1-false, and Methods 2-false, using a linear SVM classifier.

On the other hand, as pointed out by [103], it is essential to include a "no-control" state to handle situations where the user does not intend to generate any command. In such cases, it would likely be unreliable if the system still outputs a command. Similarly, mental fatigue can also lead to a "no-control" state. To solve this issue, Cecotti *et al* [103] incorporates a reject threshold into the system. The command corresponding to the frequency is produced if the highest probability is superior to a fixed threshold (in the experiments, it is 0.5). Similarly, Mora *et al* [104] introduced a confidence indicator, which enables the system to discriminate between accepted and rejected EEG epochs. The entire epoch is rejected when the indicator is less than a certain threshold, and no prediction is made. In the case of CCA-based methods, the difference between the maximum and second largest correlation coefficients could be assumed as the indicator. Kalunga *et al* [105] defined a probability threshold parameter to determine the acceptance or rejection of a classification decision. The predicted class that appeared most frequently in the last several classification outputs, with an occurrence probability surpassing the predefined threshold, is considered the class.

These techniques allow the BCI system to handle signals where the user's intention is unclear or when the user intends not to generate any specific command. By including a reject option, the system can effectively manage such situations and enhance the overall usability and reliability of the BCI system.

Another representative direction to enhance recognition reliability is adopting a dynamic window strategy. Specifically, by estimating the classification confidence, the data length used for each trial is dynamic to obtain a more reliable output. The main purpose of the dynamic window is to achieve high accuracy while minimizing the required data length. For example, Cecotti [106] compared the confidence score obtained by the normalized CCA correlation coefficient with a predefined threshold. If the confidence score is over the threshold, the decision is validated; otherwise, the decision is rejected, and the time segment should be increased accordingly. Chen *et al* [107] projected the correlation coefficients of FBCCA into probability space by softmax function and created a hypothesis testing model, in which cost function was used to evaluate the classification result's reliability. When the cost value of the rejection hypothesis is lower than the cost value of the recognition result, this algorithm believes that the current data are not enough to make a decision, and the applied system needs to collect more data. Jiang *et al* [108] presented two dynamic stopping strategies based on Bayes estimation and discriminant analysis. Specifically, the Bayes-based method tries to estimate the classification confidence by calculating the normalized maximum coefficient probability, leading to a correct prediction. Alternatively, the discriminant-based method employs the LDA algorithm to determine whether the predicted result is correct by the largest and second-largest coefficients. These methods use adaptively EEG data length that enhances the classification reliability.

2.3.3 Discussion

As mentioned earlier, the performances of the SSVEP-based BCI system can be negatively affected by different disturbances, such as intra-subject variation and environmental distractions. Consequently, even the advanced SSVEP recognition models can result in unintended classifications, potentially resulting in severe risks to the subject in practical scenarios. The

Table 2.2: Methods for improving recognition reliability in SSVEP-based BCIs

Categories	Method	Technical details	Reference
Confidence-based rejection method	Brain and gaze data fusion	A distraction level module estimates the intention of the user by behavioral entropy and validates/inhibits the command accordingly.	[100]
	Self-paced BCI	The proposed self-paced BCI supports the "no control" state. The command corresponding to the frequency is produced if and only if the highest probability is superior to a fixed threshold.	[103]
	CI-CCA	A confidence indicator (the difference between the maximum correlation coefficient and the second-largest coefficient in CCA), which enables the system to discriminate between accepted and rejected EEG epochs.	[104]
	MDRM	The predicted class that appeared most frequently in the last several classification outputs, with an occurrence probability surpassing the predefined threshold is considered to be the class. Otherwise, it is treated as a resting class.	[105]
Dynamic window strategy	CCA-DW	To shift the problem of the time segment to the choice of the threshold for determining if a response has been properly detected. If the confidence score is over the threshold, the decision is validated; otherwise, the decision is rejected and the time segment should be increased.	[106]

	FBCCA-DW	The purpose of the 'dynamic window' (DW) is to achieve high accuracy while minimizing the required data length. The correlation coefficients of FBCCA were transformed into probability space using the softmax function and a hypothesis-testing model was constructed, assessing the 'reliability' of the classification results through a risk function.	[107]
	Bayes-based DS Discriminant-based DS	Two dynamic stopping (DS) strategies were proposed, utilizing Bayes estimation and discriminant analysis. They aim to estimate classification confidence by considering the first two largest correlation coefficients.	[108]
Decision fusion	Decision-making selector (DMS)	The DMS method selects a decision more likely to be correct from two methods namely M1 and M2 by separating the M1-false and M2-false trials. To measure the uncertainty of each decision, feature vectors were extracted using the first two largest correlation coefficients.	[102]

classification confidence analysis is beneficial to enhance the reliability of decisions. Several methods have been proposed to estimate classification confidence and reduce incorrect classifications through the rejection module or adaptive time windows. Among these methods, the maximal coefficient or the largest two coefficients from the recognition method are commonly used as features [107, 108], but they may not be adequate for constructing informative confidence estimation features. Specifically, all correlation coefficients can reflect the confidence of decisions. Recognizing their efficacy in this aspect, a broader set of coefficients should be considered in classification confidence analysis.

2.4 Transfer Learning in SSVEP-based BCIs

One significant drawback of SSVEP-based BCI is the long period of calibration. During the calibration process, the user is guided to engage in a predefined task, like focusing on a visual stimulus or imagining a specific movement. As a result, a labeled dataset is created, containing EEG segments paired with their respective correct labels. These EEG segments and labels are utilized to train a supervised machine-learning method. Once this training is complete, the BCI system can be effectively employed in the intended practical applications [2, 109].

Collecting a substantial amount of calibration data is typically necessary to train the model for the target user due to many reasons. First, the high-dimensional EEG signals are likely to be contaminated with noise and artifacts such as environmental interference, eye blinks, heart rate variations, and muscle movements [110,111]. If there are only a few training trials, extracting features from high-dimensional noisy EEG signals and predicting the right brain states is hard. Second, EEG signals are non-stationarity [112]. Non-stationarity can arise from various factors, including lack of concentration and mental fatigue, as described in Section 2.3. Measurement circumstances, such as changes in electrode impedance due to sweating [113], may also contribute to non-stationarity. Third, each individual has a unique brain wave pattern even for similar activities [114], so the direct transfer of EEG signals across subjects is not feasible. Each subject should participate in the experiments to ensure accurate calibration and personalized training for the BCI system.

To address this issue, transfer learning can be a promising solution. By leveraging information (such as raw data or features) from different domains, transfer learning can help overcome the lack of labeled data from the test subject [2]. As described in [30], each subject is required to undergo a burdensome calibration process to learn model parameters: the spatial filter and the SSVEP template. The following section will review relevant literature focusing on these two aspects. The technical details were also shown in Table.2.3.

2.4.1 Transfer Learning Methods

Template-based Transfer

As discussed in Section 2.1, it has been demonstrated in many studies [68,72] that individualized templates obtained from personalized calibration data can more effectively characterize user-specific SSVEP signals than artificial templates, i.e., sine-cosine waves. Therefore, some transfer learning studies focus on template-based SSVEP decoding to leverage data across multiple domains. For example, Yuan *et al* [115] proposed a training-free method known as transfer template CCA (tt-CCA), in which the transferred EEG template was generated by averaging SSVEP trials across source subjects. CCA was separately performed between the test data and the transferred EEG template, as well as between the test data and the reference signal. The Pearson correlation coefficients were calculated between these projected signals and summed for final recognition. The online transfer template method (ott-CCA) was further proposed to gradually update the transferred EEG templates. The greater the difference between the first and second largest coefficients, the higher the likelihood of accurately identifying the correct target. If the difference is above a pre-specified threshold, update the EEG template. Considering the individual difference, Nakanishi *et al* [116] employed a session-to-session transfer method (i.e., training data measured from the same subject on a different day) based on ott-CCA to reduce the training time for the next day. Waytowich *et al* [117] proposed the Adaptive Combined-CCA (Adaptive-C3A) method, which uses an adaptive transfer learner to extend the Combined-tCCA method. The template construction process exclusively considers high-confidence trials related to SSVEP signals. Confidence is quantified as the ratio between the highest-predicted and

second-best predicted classes' coefficients. Only trials from the source subjects that exceed a specific threshold are used to adapt the SSVEP template for the target subject. Similarly, Wang *et al* [118] proposed an inter- and intra-subject template-based multivariate synchronization index (IIST-MSI) method, in which high-confidence trials are specifically chosen for transfer learning. For instance, the inter-subject template is derived by averaging high-confidence trials across subjects, and the high-confidence trial selection relies on a threshold policy using MSI. These methods overlook the discrepancies in data distribution between subjects, which may hinder the capacity to transfer data from one subject to another.

Spatial filter-based Transfer

Spatial filter-based transfer learning has also emerged as a promising research direction for SSVEP-based BCIs, offering potential solutions to alleviate the calibration burden across various domains. Multiple BCI transfer learning studies cooperating on spatial filters to learn the common feature representations across different domains [2]. Notably, the derivation of the transferred spatial filter often involves individual templates. Nakanishi *et al* [119] introduced a novel method for transferring shared EEG responses across different devices. This approach first averages the training trials of SSVEPs acquired using a device to create individual templates for each visual stimulation tagged with a specific frequency/phase. Subsequently, spatial filters trained by CCA or TRCA are applied to the individual templates for transferring them to another device. The transferred templates and data from the target device were then integrated into the LST to derive the transferred spatial filter for SSVEP recognition. Liu *et al* [120] proposed a transfer learning framework named align and pool for EEG headset domain adaptation (ALPHA), which aligns the spatial pattern and the covariance for domain adaptation. To improve its performance and reduce the calibration effort for dry-electrode systems, cross-device transfer learning was used by exploiting auxiliary individual wet-electrode EEG. Liu *et al* [121] further explored within-subject (i.e., cross-day and cross-electrode type) transfer learning which improves the BCI performance and reduces the calibration burden via ALPHA.

Except for cross-device and within-subject transfer learning, most studies in SSVEP-based

BCIs focus on transfer across different subjects. Chiang *et al* [122] employed the least-squares transformation (LST) to achieve cross-subject transfer of SSVEP data, resulting in reduced calibration time and improved recognition performance for new users. Specifically, the LST method transforms the SSVEP data from existing subjects to fit the SSVEP templates of a new user using only a small number of the new user's signals. Based on their previous work, Chiang *et al* [123] propose a generalized framework of transfer learning that can leverage SSVEP data across multiple domains (sessions, subjects, and devices) toward a practical application. Yan *et al* [29] proposed a training-free cross-subject spatial filter transfer (CSSFT) method that leverages existing user data to compute spatial filters using CCA or FBCCA and transfers them to new user data. The core idea is inspired by the fact that the reference signal used by new and existing users is the same; both new and existing users aim to project their EEG data in the most relevant direction to the reference signal. Considering the amplitude and pattern of brain signals differed across trials, Yan *et al* [124] further introduced an enhanced version of CSSFT called Improved-CSSFT. This method replaced superposition averaging with a trial ensemble or expansion scheme to achieve better performance.

Compared with training-free methods, training-based methods tend to get higher recognition accuracy. Tanaka [125] extends the previous method TRCA by maximizing trial-by-trial reproducibility within single subjects and similarity across a group of subjects. This extension is referred to as group TRCA (gTRCA). The problem of maximizing the reproducibility of time series across trials and subjects is formulated as a generalized eigenvalue problem. Inspired by gTRCA, Wang *et al* [126] proposed an inter- and intra-subject maximal correlation (IISMC) based cross-subject assistance framework to enhance the SSVEP recognition. IISMC employs two sets of spatial filters: intra-subject filters that extract subject-specific information, and inter-subject filters that capture task-related similarities across individuals performing the same task. These filters are trained by maximizing correlation, leading to an effective method for extracting meaningful features from SSVEP data. Wei *et al* [127] introduced a spatial filtering method based on CCA, named intra- and inter-subject CCA (IISCCA). This method aims to train three types of spatial filters,

which the TDCCA estimated [128] with EEG data from the same domain or two different domains. The IISCCA not only extracts subject-specific information but also integrates similarities between subjects. To minimize the variability between subjects, an accuracy-based algorithm for subject selection (ASS) was incorporated with IISCCA. This algorithm is designed to identify source subjects whose data distributions are more similar to those of the target subject.

Some studies facilitate intra- and inter-subject transfer learning by investigating the relationships between different stimuli [129]. Wong *et al* [130] assumed that the training data corresponding to neighboring stimuli can be assigned with a common spatial filter. The key idea of the proposed multi-stimulus CCA (msCCA) is to train the spatial filter using data from the target stimulus and the neighboring stimuli. Based on this work, Wong *et al* [30] further proposed an intra-subject spatial filter and inter-subject SSVEP template transfer scheme. The subject transfer-based CCA (stCCA) approach introduced a new assumption that the target subject's spatially filtered SSVEP templates could be approximated by the weighted summation of spatially filtered SSVEP templates from source subjects. A multivariate linear regression (MLR) problem calculates the inter-subject spatial filter. Subsequently, Wong *et al* [131] explored the feasibility of transferring the model parameters, namely the spatial filters and the SSVEP templates, across two different groups of visual stimuli. The authors utilized the common impulse response to reconstruct the transferred SSVEP template, based on the theory that the steady-state response can be modeled as the convolution of the impulse response and the periodic stimulus [132].

2.4.2 Discussion

In the past decade, transfer learning has been widely investigated in SSVEP-based BCI, mainly including template-based transfer and spatial filter-based transfer across different domains (i.e., sessions, subjects, and devices). While some studies have been inclined towards direct knowledge transfers across domains [29, 115], others have investigated the correlations between different domains [126, 127]. Although continuous efforts have been made to alleviate the calibration effort while maintaining decent decoding accuracy [133],

Table 2.3: Transfer learning methods in SSVEP-based BCIs

Categories	Methods	Technical details	Reference
Template-based Transfer	tt-CCA	The transferred EEG template was generated by averaging SSVEP trials across the source subjects.	[115]
	ott-CCA	The template was updated depending on the difference between the first and second largest coefficients.	[115]
	SSott-CCA	A session-to-session transfer method, which uses templates obtained from datasets collected from the same subject on a different day.	[116]
	Adaptive-C3A	An extension of Combined-CCA with an adaptive template learner. High-confidence trials were averaged for transfer across subjects.	[117]
	IIST-MSI	The inter- and intra-subject templates were obtained by averaging high-confidence trials across subjects or within subject. The confidence score depends on a threshold policy with MSI.	[118]
	LST-based method	The LST method transforms the SSVEP data from the source domain to fit the SSVEP templates of the target domain.	[122, 123]
Spatial Filter-based Transfer	SLD	The transferring is done by projecting the scalp-channel EEG signals onto a shared latent domain (SLD) across devices.	[119]
	ALPHA	ALPHA aligns the spatial pattern and the covariance for domain adaptation.	[120, 121]

	CSSFT	It leverages existing user data to compute spatial filters using CCA or FBCCA and transfers them directly to new user data.	[29]
	Improved-CSSFT	It replaces superposition averaging with a trial ensemble or expansion scheme to achieve better performance.	[124]
	IISMC	The inter- and intra-subject spatial filters were trained by correlation maximization.	[126]
	IISCCA	The subject-specific information and similarities between subjects were extracted by TDCCA.	[127]
	ms-CCA	To train the spatial filter using data from not only the target stimulus but also the neighboring stimuli.	[130]
	stCCA	It introduces the intra-subject spatial filter (ms-CCA) and the inter-subject SSVEP template (MLR) for the CCA-based method.	[30]
	tlCCA	According to the theory that an SSVEP is a superposition of the impulse responses, tlCCA transfers the spatial filters and SSVEP templates across two groups of visual stimuli.	[131]

many open questions still need further investigation. For example, during the training of the transferred spatial filter and transferred template, the optimization problem rarely considers the pairwise relationships between the training trial and two types of templates (i.e., individual templates and artificially sine-cosine references) simultaneously. Additionally, in a cross-subject scenario, the distinct contributions of source subjects to the target subject have yet to be thoroughly studied.

2.5 SSVEP-based BCI-Controlled Robots

The SSVEP-based BCI has shown significant importance in the field of assistance and rehabilitation because it enables individuals with severe physical disabilities to regain control of their daily lives. We will describe the SSVEP-based BCI-controlled applications from the following three aspects, i.e., communication aids, prosthetics control, and environmental control. The classification of robots, along with their technical details and corresponding control commands, were presented in Table 2.4.

2.5.1 Communication Aids

The SSVEP-based BCI offers new communication methods for individuals with limited or no speech abilities. The speller is a typical application of SSVEP-based BCI, which allows users to express their thoughts and needs effectively [134]. Bremen Speller is one of the earliest and most well-known SSVEP-based BCI spellers that presents characters and stimuli on the screen, respectively [135]. The five flickers at different frequencies are used to control the cursor's movement and facilitate character selection.

In subsequent research, the stimuli usually directly represent characters, allowing users to output different characters by focusing their attention on specific visual cues [136]. Depending on whether a stimulus represents multiple characters or a single character, it can be further classified into two types, namely multi-layer spellers and one-layer spellers [134]. The multi-layer speller has two or more stages and built a screen sequence for selecting the letters [137–139]. For example, Cecotti *et al* [103] introduced a three-layer structure where each layer consisted of five flickering stimuli. A broad range of characters is initially pre-

sented to the user at the first layer. The system dynamically displays a narrowed selection on the second layer as the user directs their attention. Finally, at the third layer, the user can focus on selecting a specific character from the remaining options. Similarly, Nguyen *et al* [140] proposed a three-layer keyboard with five flickers in each layer and a total of 58 characters, facilitating precise character selection within the SSVEP-based system.

Unlike the multi-layer speller, the one-layer speller presents a one-to-one correspondence between commands and characters. This design enhances input efficiency and facilitates communication by streamlining the selection process. For example, Hwang *et al* [141] adopted a QWERTY style to lay out a keyboard with 30 LEDs flickering at different frequencies. The frequency band of 5-7.9 Hz with a span of 0.1 Hz was selected. Chen *et al* [142] introduced a 45-target BCI speller, where SSVEPs were induced by both low-frequency (7–15.8 Hz) and high-frequency (35.6- 44.4 Hz) stimuli. Chen *et al* [143] further proved that the discriminability of SSVEPs could be improved by incorporating both frequency and phase features. This work introduces two hybrid coding strategies: mixed frequency and phase modulation and joint frequency and phase modulation (JFPM). The mixed frequency and phase coding strategy involves eight frequencies (8–15 Hz with a 1Hz interval) and five phases (0, 0.4 π, 0.8 π, 1.2 π, and 1.6 π). The joint frequency and phase coding employs frequencies ranging from 8 to 15.8 Hz with a 0.2 Hz interval, and the phase interval between two adjacent frequencies is 0.5 π. Wang *et al* [23] presented a 40-target BCI speller using the same JFPM approach as described in [143]. The difference is that [23] involved more subjects, focused on collecting a benchmark dataset for the SSVEP-based BCI.

2.5.2 Prosthetics Control

The SSVEP-based BCI can be utilized to control robotic prosthetic limbs, allowing individuals with limb impairments to regain motor control and execute normal movements. Based on the applicable body part, these prosthetics can be categorized as upper limb prostheses and lower limb prostheses. The following will introduce the application of SSVEP-based BCIs in prosthetics control for both classes.

Upper Limb Prosthetics

SSVEP-based upper limb prosthetics are designed to assist or replace the user's arm or hand in performing essential tasks and facilitate the recovery and rehabilitation process through therapeutic exercises. In the field of assistance, SSVEP-based assistive robotic arms have been widely studied to help individuals perform practical series of actions, such as reaching and grasping movements [144]. For example, Chen *et al* [145] designed a 7-DOF robotic arm system to execute the move-grasp-lift task of a single object. This study utilized a set of 15 targets (8-15 Hz) that served as specific commands to control the movement and rotation of the robotic arm. In [146], a 25-target SSVEP-based BCI was designed for the pick and place tasks of the robotic arm. Each frequency within the range of 8-15.8 Hz corresponded to one of the 25 locations within the robotic arm's workspace. To deal with visual fatigue caused by flickering, a combination of a high-frequency SSVEP-based BCI with computer vision was proposed for controlling the robotic arm [147]. Computer vision was used to identify and locate the three objects within the workspace, and a four-stimulus interface was shown on the screen, where the three flickers were used to choose the corresponding object for the action, and the left one was used to undo the last operation. The augmented reality (AR) device was further incorporated to improve the portability of the SSVEP-based BCI system [148, 149]. For example, Chen *et al* [150] further developed the SSVEP system in [147] with augmented reality (AR) by a Microsoft HoloLens, where users can see the user interface of the BCI and the robotic arm simultaneously. The same issue was also studied in [151], where the flickering stimuli were attached to the robotic arm's gripper and moved with it. The experimental results demonstrated that the proposed paradigm outperformed the conventional approach that used fixed flickering stimuli. In the rehabilitative field, Muller-Putz *et al* [152] designed a hand prosthesis to restore the grasp function in spinal cord-injured people. Four mounted lights are attached to the fingers and the forearm, corresponding to four movements, i.e., turn left, turn right, hand open, and hand close. Ortner *et al* [153] designed a hand orthosis for persons with tetraplegia. The lights are also mounted on the orthosis for hand open and close tasks. Guo *et al* [154] designed a SSVEP-based BCI-controlled soft robotic glove for post-stroke hand function rehabilitation. The

two stimuli, displayed on both sides of the screen, were utilized to control the left or right glove, respectively. Zhao *et al* [155] introduced an integrated framework of SSVEP with functional electrical stimulation (FES) for motor rehabilitation. The movement intentions regarding the three operation speed modes are transformed into instructions to trigger FES, and muscles can be stimulated to induce arm movement. Sakurada *et al* [156] designed an occupational therapy assist suit for patients with upper cervical spinal cord injuries (SCIs). SSVEP signals were used to trigger the grasping-a-ball and carrying-the-ball movements.

Lower Limb Prosthetics

The SSVEP-based lower limb prosthetics can help individuals with lower limb disabilities walk and restore mobility. Wang *et al* [157] proposed a lower limb exoskeleton to help people to perform stand-up, sit down, and walk forward. In this study, assistance is required for the wearer to hold the display and view the stimulus. Kwak *et al* [158] developed a SSVEP-based lower limb exoskeleton control system in which users can achieve programmed motions (e.g., walking, turning, sitting, and standing). A visual stimulation device with five light-emitting diodes mounted to the exoskeleton stimulates the SSVEP. Qi *et al* [159] designed a lower limb exoskeleton with seven commands for effectively controlling individuals to perform leg rehabilitation. Zeng *et al* [160] introduced an ankle rehabilitation robot that allowed subjects' motion intentions (upper, bottom, left, and right) to trigger corresponding passive training. Wang *et al* [148] designed a portable AR-BCI system applied to rehabilitation exoskeleton. A four-class BCI was used to complete rehabilitation movements for the upper and lower limbs.

The SSVEP-based robotic prosthetic can provide assistance and support to individuals with disabilities. It can also help these people regain flexibility and perform daily movements.

2.5.3 Environmental Control

The SSVEP-based BCI can also be used to control wheelchair navigation, household appliances, and other devices, which enhances the quality of life for those with physical restrictions. Diez *et al* [161] presented a BCI system for operating a robotic wheelchair to

complete reaching tasks with or without obstacles. The LCD screen was installed on the wheelchair, and four stimuli were attached on the sides of the screen, corresponding to go forward, turn left, turn right, and stop control commands, respectively. Deng *et al* [7] designed a probabilistic model of human and robot control commands to provide the optimal control of a wheelchair robot. In this model, a multi-layer stimulation interface including four LEDS was employed, in which the first layer is used to determine the preliminary direction, and the second layer is used for more accurate angular control. Sakkalis *et al* [162] aimed to develop efficient electric wheelchair navigation combining high accuracy and comfort. A SSVEP-based control system featuring AR glasses was proposed to increase ease of use and patient acceptability. Rivera *et al* [163] incorporated compressive sensing (CS) in a robotic wheelchair control, which reduces the data size but preserves the signal's quality and increases the ITR. The benchmark data provided by Wang *et al* [23] were used for decoding the following four commands: reverse (from 8.0 to 9.8 Hz); right (from 10.0 to 11.8 Hz); left (from 12.0 to 13.8 Hz); and forward (from 14.0 to 15.8 Hz). Considering that only designated directional commands could be transmitted by the BCI system, Zhang *et al* [164] introduced a 2D navigation robot that utilizes the quantified value of brain commands to form an acceleration vector. It improves the flexibility, stability, and efficiency of BCIs.

SSVEP-based BCI systems also have made progress in the smart home scenarios, which provides disabled people with more direct interactions with the environment. It allows people to recognize various commands and control corresponding devices in their houses by watching different stimuli [165]. In [166], visual stimuli were displayed on AR smart glasses, which can be easily used to control household devices, such as lights, coffee machines, and elevators. QR codes were used to identify the equipment to be controlled with the BCI. Adams *et al* [167] introduced a more portable BCI system with a lightweight mobile EEG amplifier. Six devices in a real smart home environment can be controlled with this BCI system. The SSVEP-based BCI system also assists in reducing domestic pressure and improving home conditions by helping people accomplish heavy housework. For example, Shao *et al* [168] designed a novel EEG-based intelligent teleoperation system for a mobile

Table 2.4: SSVEP-based BCI controlled robots

Categories	Device	Technical details	Commands	Reference
Communication Aids	Speller	Bremen Speller: the five flickers at different frequencies are used to control the cursor to select characters.	Turn left, right, up, down, selection	[135]
	RC-speller	The row/column (RC) paradigm was used to increase the number of targets. The target is detected by subsequently determining the row and column coordinates.	Choose one row out of six, then one column out of six.	[136]
	Multi-layer speller	It has two or more stages and built a screen sequence for the selection of the letters	Choose a group of characters, then a specific one	[103, 137–140]
	Lone-layer speller	It presents a one-to-one correspondence between commands and characters.	Choose a specific character	[23, 141, 142]
Prosthetics Control	Robotic arm	Subjects control the 7-DOF robotic arm to reach a specified position in 3D space.	Move forward, backward, right, left, up, and down	[144]
	Robotic arm	A 15-target system allowed users to directly control the robotic arm to complete the move -grasp-lift task.	move up, down, left, right, forward, backward, rotation, and return	[145]
	Robotic arm	A 25-target system was developed to control a robotic arm for picking and placing a plastic block at the identified location.	Placement location	[146]
	Robotic arm + CV	Pick and place task: the computer vision (CV) can identify objects and locate positions, while the BCI allows the user to select one object to be acted upon by the robotic arm.	Choose the specific object	[147]
	Robotic arm + CV +AR	Augmented reality (AR) environment allowed users to see both the robotic arm and the user interface of the BCI.	Choose the specific object	[150]

Robotic Arm	The flickering stimuli were attached to the robotic arm's gripper and moved with it, so users do not need to switch their gaze between the flicker and the actual scene.	move forward, backward, left, and right	[151]
Hand prosthesis	Four mounted lights are attached to the fingers and the forearm, representing four movements.	turn left, turn right, hand open, and hand close	[152]
Hand orthosis	Two lights are mounted on the finger-flexion hand orthosis.	Hand open and close	[153]
Robotic glove	A soft robotic glove for post-stroke hand function rehabilitation	Choose left or right glove	[154]
FES device	An integrated framework of SSVEP with functional electrical stimulation (FES) for motor rehabilitation.	Speed mode (fast, medium, or slow)	[155]
Therapy assist suit	A two-class SSVEP-based occupational therapy assist suit for patients with upper cervical spinal cord injuries.	Grasp-a-ball and carry-the-ball movements	[156]
Lower limb exoskeleton	A lower-limb exoskeleton robot was developed for BCI-based control of exoskeletons. The wearer requires assistance to hold the display in place and view the stimulus.	Choose walk, sit, or stand	[157]
Lower limb exoskeleton	A visual stimulation unit consisting of five light-emitting diodes fixed to the exoskeleton.	Walk forward, turn left, stand, turn right, and sit	[158]
Lower limb exoskeleton	Each subject is required to wear protective equipment and exoskeleton to stand on the treadmill.	Acceleration, deceleration, weight loss, aggravation, left leg, right leg, stop	[159]
Ankle rehabilitation robot	Visual reality circumstance is a whack-a-mole game. Four hamsters are arranged in four directions as targets and a hammer is located in the center as the movable cursor.	Turn left, right, upper, and bottom	[160]

	Exoskeleton + AR	A sequential logic decoding method is proposed, which can realize 16 control commands by using a 4-class BCI.	1st layer: group number; 2nd layer: specific movement	[148]
Environmental Control	Wheelchair	To complete reaching tasks with/without obstacles. Stimuli were attached to the wheelchair.	Go forward, turn left, turn right, and stop	[161]
	Wheelchair	It designed a probabilistic model of human and robot control commands to provide the optimal control of a wheelchair.	1st layer: preliminary direction; 2nd layer: angular	[7]
	Wheelchair + AR	a SSVEP-based control system featuring AR glasses to achieve wheelchair navigation.	Forward, turn left, right, and back	[162]
	Wheelchair	It incorporated compressive sensing (CS), which reduces the data size but preserves the signal's quality.	Reverse, right, left, and forward	[163]
	Wheelchair	It introduced a 2D navigation robot that utilizes the quantified value of brain commands to form an acceleration vector.	Ten speed direction	[164]
	Household devices	The visual stimuli were displayed on AR smart glasses, which can be easily used to control household devices, such as lights, coffee machine, and elevators.	Light: on/off; elevator: floor number; coffee machine: coffee flavor	[166]
	Household devices	The visual stimuli for the BCI were placed on multiple screens in the smart home (such as the kitchen and living room).	Thirteen commands for controlling six devices	[167]
	Wall-crawling cleaning robot	The SSVEP system controls the robot to climb the wall and complete the tasks of cleaning.	Forward, backward, turn left, and right	[168]
	Meal-assist robot	SSVEP signals were used to decode the user's intention of food.	One of the five types of food	[169]

wall-crawling cleaning robot, which uses the crawler type instead of the traditional wheel type for window or floor cleaning. SSVEP-based BCI also has the potential to alleviate the requirement for caretakers [170]. For example, Ha *et al* [169] developed a hybrid BCI meal-assist robot system in which SSVEP signals were used to decode the user's intention of food. Five LEDs with different frequencies were connected to the tray with five kinds of food. The developments of SSVEP-based BCI in the smart environment field may offer the prospect of significantly improving the quality of life for disabled people in clinics and considerably increasing their independence, autonomy, and mobility, leading to reduced social costs.

2.5.4 Discussion

The SSVEP-based BCI has been widely utilized in various applications, effectively assisting disabled people in enhancing their communication and control capabilities with the external environment. In the SSVEP-controlled robot system, the subject can convey their motion intention by directing his or her focus toward different stimuli. In many research studies, the BCI system used to control robots often relies on predetermined directional commands, such as "turn left/right" or "go ahead" [145, 161], as well as instructions to reach specific locations at a fixed speed [146,154]. It neglects the subject's requirement for controlling the direction and speed of the robot. Several studies considered velocity control in the SSVEP paradigm, including examples like the upper extremity rehabilitation system in [155], and the robotic wheelchair in [171]. However, since different stimuli correspond to different velocity modes, the BCI still discretely accomplishes velocity modulation. Meanwhile, it has yet to explore the quantified value of brain commands fully.

2.6 SSVEP Dataset

This section outlines the frequency band applied in SSVEP data collection and the datasets utilized in this thesis for performance evaluation. They comprise one public benchmark dataset [23] and two datasets that were self-collected in the lab. For the datasets, detailed information is elaborated below.

2.6.1 SSVEP Frequency Band

Five distinct brain rhythms can be identified in EEG based on their frequency ranges: delta (1-4) Hz, theta [4-8) Hz, alpha [8-13) Hz, beta [13-32) Hz, and gamma [32-120) Hz [2,172]. To date, response amplitude is commonly used to guide the selection of stimulation frequencies [173]. Schielke et al. [172] highlighted that from 1954 to 2021 on PubMed, SSVEP research distribution across these bands was as follows: delta band accounted for 12 tests (5.26%), theta band 58 tests (25.44%), alpha band 103 tests (45.18%), beta band 51 tests (22.37%), and gamma band 4 tests (1.75%).

Over the past few decades, continuous efforts have been made to contribute to the public SSVEP database. These datasets typically focus on frequency bands that evoke the strongest SSVEP responses. Nakanishi et al. [68] introduced a 12-class dataset collected from 10 subjects, featuring frequencies ranging from 9.25 Hz to 14.75 Hz in 0.5 Hz increments. Wang et al. [23] introduced a 40-target database, comprising 64-channel SSVEP trials from 35 subjects who performed an offline cue-spelling task. The stimulation frequencies ranged from 8 Hz to 15.8 Hz. This dataset has been extensively utilized in numerous SSVEP studies to compare the performance of various algorithms [80,91,174]. Liu et al. [175] further presented the Benchmark database Towards BCI Application (BETA), which comprises data from 70 subjects performing a 40-target cued-spelling task. The frequency range for this database extends from 8 to 15.8 Hz with an interval of 0.2 Hz.

2.6.2 Public Dataset

The benchmark dataset in [23] was introduced in this subsection. This dataset was collected from 35 healthy subjects with 40 visual stimulation modulated at different frequencies (8–15.8 Hz with an interval of 0.2 Hz), while the phase difference between two neighboring targets is 0.5 π. The stimulation interface of the 40-class BCI is shown in Fig. 2.2. In the experiment, each participant underwent six blocks, with each block comprising 40 trials that corresponded to 40 stimuli. Each trial in the experiment lasted for 6 s, starting with a 0.5 s cue time during which the subject was instructed to shift his or her gaze to the target stimulus as quickly as possible. After the cue, all stimuli started to flicker concurrently for

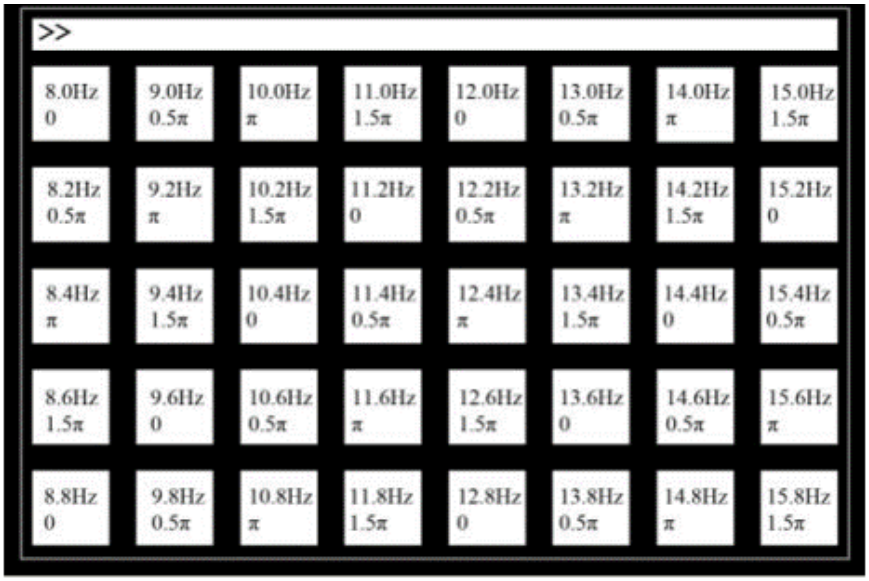


Figure 2.2: The stimulation interface of the 40-target BCI.

5 s. The trial concluded with a 0.5 s rest period before commencing the next trial. More information about this dataset can be found in [23].

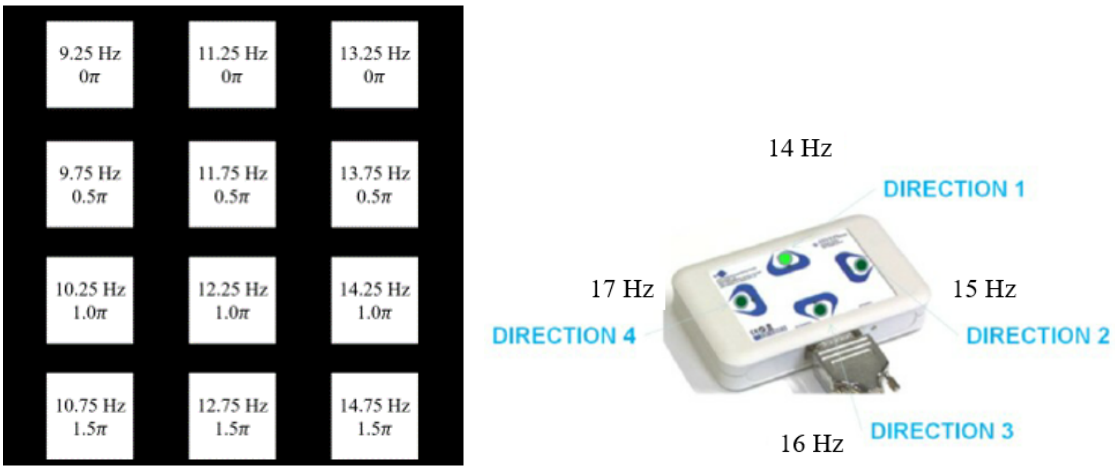
2.6.3 Self-collected Dataset

Participants

In Dataset I, SSVEP signals were recorded from 11 subjects (four females and six males, mean age: twenty-five years). In Dataset II, eight subjects (four females and four males, mean age: twenty-seven years) participated in the SSVEP-based BCI experiment. All subjects were in good health, with normal or corrected-to-normal vision. Each participant was seated in a comfortable chair, facing the visual simulation. The Research Ethics Committee of the University of Leeds has approved the experiment. All subjects read and signed the participant consent form before the SSVEP experiment.

Stimulus Design

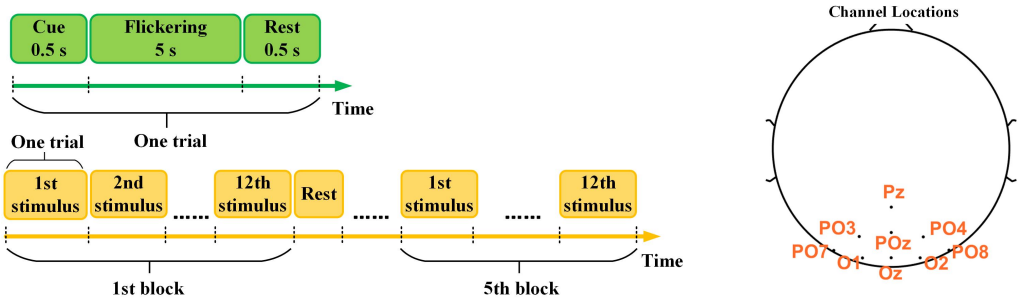
In Dataset I, the visual stimulation is designed as a 4×3 matrix coded by the JFPM method on a 23.6-inch LCD monitor. Its resolution and refresh rate are 1920×1080 pixels and 60 Hz. The 12 visual stimuli flashed at frequencies ranging from 9.25 Hz to 14.75



(a) Dataset I (b) Dataset II

Figure 2.3: The stimulation interface of (a) 12-class dataset and (b) 4-class dataset.

Hz with an interval of 0.5 Hz. The phase started from 0π to 1.5π in steps of 0.5π. The stimulation interface was shown in Fig. 2.3(a). The experiment consisted of 5 blocks for each subject. Each block includes 12 trials corresponding to 12 stimuli. Each trial began with a 0.5 s red dot cue which indicates the target flicker. Later, 12 stimuli started to flash for 5 s simultaneously, during which the subject should stare at the target stimulus and avoid eye movements. Following this, there is a 0.5 s rest period before the next trial commences. Additionally, there was a 0.5 s break between two consecutive blocks. The experimental paradigm is shown in Fig.2.4(a). In Dataset II, the g.SSVEPbox was adopted for stimulation design, as shown in Fig.2.3(b). It contains 4 LEDs flickering at various frequencies, namely 14 Hz, 15 Hz, 16 Hz, and 17 Hz. The experiment of each participant also consists of 5 blocks, and there are 4 trials corresponding to four visual stimulation in each block. A small green light appeared about 2 mm from one of the LEDs for cue-guided purposes. In each trial, four stimuli flash in white for 5 s concurrently. The frequency ranges of [9.25-14.75] Hz and [14-17] Hz were selected as they fall within the theta and alpha ranges (8-32 Hz), which are known to elicit more responsive and stronger SSVEP signals [172,173]. Even though the 12-class SSVEP dataset has already been gathered in prior research [68], replicating this experiment is valuable not only for validating existing results but also for gaining a more comprehensive understanding of the SSVEP data collection process.



(a) Experimental paradigm

(b) Channel location representation

Figure 2.4: (a) experimental paradigm and (b) channel location of SSVEP recording.

EEG Recording

For the benchmark dataset, all epochs were simply down-sampled to 250 Hz. Nine channel EEG data (i.e., Pz, PO5, PO3, POz, PO4, PO6, O1, Oz, and O2) was selected for analysis. In the self-collected datasets, SSVEP signals were recorded by the equipment from g.tec medical engineering GmbH. The g.USBamp amplifier was applied to sample data at 256 Hz. For Datasets I and II, nine channels, including Pz, PO3, POz, PO4, PO7, O1, Oz, O2, and PO8, were selected for EEG data collection. The ground electrode was located over electrode FPz, and the right earlobe was the reference channel. The channel location was shown in Fig. 2.4(b).

2.7 Research Gap

To conclude, the existing research gaps are summarized as:

- One of the main directions for improving the performance of SSVEP-based BCIs is to enhance target identification algorithms. Initially, FFT-based methods were widely used to detect the target frequency of SSVEP signals [20, 27]. Additionally, more efficient methods have been developed by incorporating temporal features of SSVEP signals, such as CCA and its variations [38, 67]. Later, the inter-trial correlation was considered in the training process [1, 71]. Among these methods, spatial filters are typically trained using the signal from a single stimulus, and information from other stimuli is seldom considered. Additionally, the volume conduction artifact, a common

problem arising from the coherence among various electrodes [176], does not receive much attention in SSVEP classification methods. Addressing these limitations is a crucial step toward further enhancing the performance of SSVEP-based BCIs.

- Many existing SSVEP recognition methods rely on determining the target class by finding the largest correlation coefficient [69, 72]. However, the classification performance degrades when the largest coefficient is not significantly different from the rest of the values. Therefore, classification confidence analysis becomes essential for SSVEP recognition, as it enables the rejection of unintended results and reduces risks in practical scenarios. This, in turn, significantly enhances the reliability of SSVEP-based BCI systems. Currently, research on classification confidence estimation typically focuses on considering the largest two correlation coefficients [104, 108]. However, more is needed to construct informative features for enhancing SSVEP detection. Therefore, thoroughly investigating all correlation coefficients becomes crucial to improving the system's reliability, ultimately enhancing users' safety and security.

- Collecting a substantial amount of calibration data is typically necessary to train the model for each subject [122]. This process is often time-consuming and troublesome [177], potentially affecting the quality of the signals and, consequently, impacting the recognition reliability and accuracy. Transfer learning has gained much attention as a potential solution to address this challenge [2]. However, existing transfer learning papers often design a cross-subject scheme, neglecting the full consideration of the relation between training data and different types of templates (i.e., the artificial template, and the reference signal) in the construction of transferred spatial templates and transferred spatial filters. Additionally, the distinct contributions of source subjects to the target subject are rarely taken into account in the recognition process.

- SSVEP-based BCIs have been extensively implemented in various applications by classifying EEG signals and mapping them to different commands [135,147,161]. Conventional control methods usually provide commands with fixed and predefined movement directions and speeds [144,145], neglecting the user's requirement for velocity changes during real-world implementations. This limitation could restrict the user's control

precision and impede the overall user experience. To address this challenge, a potential solution lies in incorporating velocity modulation based on the user's intent. The BCI system can respond to their intended movements by allowing the user to modulate the velocity, providing a more intuitive and interactive control experience.

2.8 Research Contributions

Considering the research objectives and the gaps in existing research, the contributions of this thesis can be summarized as follows:

- Introducing a novel target classification method using multi-objective optimization high-pass spatial filtering techniques to enhance SSVEP recognition accuracy by extracting target-related features and reducing volume conduction effects.
- Proposing a Bayesian-based classification confidence estimation method to improve SSVEP detection reliability by accepting high-confidence results and rejecting low-confidence decisions, enhancing the overall system reliability.
- Introducing a novel cross-subject transfer learning scheme for SSVEP-based BCI, incorporating knowledge from source subjects to effectively strengthen recognition performance in a target subject, reducing the need for extra extensive individual calibration procedures.
- Implementing an SSVEP-based BCI-controlled robotic platform with a stimulus brightness-based method for velocity modulation, providing adaptability in robotic movement to better align with users' intentions, thereby enhancing the user experience and practical applicability of SSVEP-based BCIs.

2.9 Chapter Summary

This chapter presents a comprehensive survey of previous works on SSVEP-based BCI, with a particular focus on 1) SSVEP signal preprocessing and recognition techniques, 2) recognition confidence analysis, 3) transfer learning methods, and 4) SSVEP-based BCI-

actuated robots. The chapter also highlights some research gaps and challenges in the field. These insights guide the subsequent sections, which aim to address these issues and develop a reliable and accurate SSVEP-based BCI system.

Chapter 3

Multi-objective Optimisation-based High-pass Spatial Filtering for SSVEP Detection

Many spatial filtering methods have been proposed to enhance the target identification performance of the SSVEP-based BCI. The existing approaches tend to learn the spatial filter parameters of a certain target using only the training data from the same stimulus, and they rarely consider the information from other stimuli or the volume conduction problem during the training process. This paper proposes a novel multi-objective optimisation-based high-pass spatial filtering method to improve SSVEP detection accuracy and robustness. The filters are derived by maximising the correlation between the training signal and the individual template from the same target while minimising the correlation between the signal from other targets and the template. The optimisation will also be subject to the constraint that the sum of filter elements is zero. The evaluation study on two self-collected SSVEP datasets (12 and 4 frequencies, respectively) shows that the proposed method outperformed the compared methods, such as CCA, MsetCCA, SSCOR, and TRCA. The proposed method was also verified on a public 40-class SSVEP benchmark dataset from 35 subjects. The experimental results have demonstrated the effectiveness of the proposed

approach for enhancing the SSVEP detection performance.

3.1 Introduction

The BCI establishes a direct communication path between the human brain and external devices without relying on normal motor output pathways [178–180]. The SSVEP-based BCI has been widely explored in various applications, such as robot arm [150], communication [1] and augmented reality (AR) glasses [181] [182] due to its high SNR, minimal training time, and fast communication rate [183–185].

The main task of the SSVEP-based BCI system is to identify target stimuli and then translate them into corresponding commands for subsequent control actions. Thus far, various spatial filters applied transformations in the channel domain, which enhanced the SSVEP identification effectively via removing background artifacts [185–187]. As one of the most popular spatial filtering methods, CCA seeks a pair of weights to maximise the correlation between SSVEP signals and sine-cosine reference signals [38]. Some extended versions of the standard CCA were proposed to improve the performance of SSVEP detection further. In these methods, one branch starts with brain activity processing, such as FBCCA [69]. As it decomposes SSVEP signals into various sub-band components, the specific information contained in harmonic components can be employed for frequency recognition. Another branch is to optimise the predefined artificial reference signal by incorporating individual calibration data. For example, Zhang *et al.* presented MwayCCA [65] and L1-regularized MwayCCA [66] to construct the new reference template via the third-order EEG data tensor. They later proposed the MsetCCA, which extracts common features shared by multiple calibration signals to create optimised references, and it outperforms MwayCCA in classification [67] [68]. The latest SSVEP detection methods tend to learn the spatial filter by incorporating a training stage. Besides, these methods are commonly based on template matching, in which the individual template is acquired by averaging multiple training trials [2]. For instance, Wei *et al.* [128] proposed a training data-driven CCA algorithm that yields more robust spatial filters to enhance the SNR of SSVEP signals. However, the inter-trial relationship from the same target was not fully considered. In another study, Nakanishi

et al. [1] presented TRCA based on the idea that the source signal can be reconstructed through a linear summation of multi-channel EEG signals. The TRCA-based spatial filter is learned by inter-trial covariance maximisation. However, during the spatial filter training process, the methods mentioned above only utilised the signal from a single stimulus, and they rarely considered the information from other stimuli.

Although multi-channel scalp EEG activities are beneficial to achieve high-quality spatial filters, the volume conduction artifact may also be introduced due to the coherence among various EEG channels [188]. It is well known that the measurement of each EEG electrode is a linear mixture of concurrently multiple brain source activities, not just the activity of brain source in its vicinity [189]. In other words, since the signal from each brain source spreads among various EEG channels when passing through the intracerebral area and the scalp, it is infeasible to connect each electrode with a single brain region [190]. Therefore, the volume conduction decreases the spatial precision of EEG signals when spatially broad features are shared among multiple electrodes [176]. The high-pass spatial filter is commonly used to reduce the effect of spatial blurring for EEG signals [191]. For example, the surface Laplacian has been implemented to diminish volume conduction effectively by attenuating low-spatial-frequency activities whereas highlighting high-spatial-frequency signals [176] [192]. The surface Laplacian filter has some advantages in EEG signal analysis, including ease of use and conceptually simple. However, it is still a kind of "stationary" spatial filter since it assumes that the EEG activity is not variable across time [193] [194]. To solve this problem, Lu *et.al* [195] designed an adaptive Laplace filter for the sensorimotor rhythms (SMR) analysis, which employs a Gaussian kernel to construct parameters of the spatial filter. Currently, most existing research on volume conduction focused on other paradigms of EEG, like event-related potential (ERP) [193] [192] and SMR [195] [196], rather than SSVEP signals. Therefore, there is still a lack of sufficient studies about the high-pass spatial filter on the volume conduction phenomenon in SSVEP signals. The feasibility of this filter to enhance target detection in the SSVEP paradigm remains unclear.

In this study, a multi-objective optimisation-based high-pass spatial filtering method was proposed to improve the SSVEP identification performance. The filter was derived by max-

imising the correlation between the training signal and the individual template from the same target while minimising the correlation between signals from other targets and the template. Meanwhile, the constraint condition is that the sum of spatial filter elements is zero. This setting allows the high spatial-frequency SSVEP signals to pass and attenuates low spatial-frequency signals. Therefore, the proposed approach has the potential to extract the target-relevant features, reject the target-irrelevant information, and reduce the effect of volume conduction simultaneously. The method's performance was evaluated on two self-collected SSVEP datasets containing 12 and 4 visual stimulation, respectively. The effectiveness of the proposed approach was verified with an averaged SSVEP recognition accuracy of 87.58% in Dataset I and 87.5% in Dataset II using a 1s-long data epoch. Many well-known and state-of-the-art methods such as CCA, MsetCCA, the SSCOR [71], and TRCA was implemented for extensive comparisons. In addition, a 40-class SSVEP benchmark dataset [23] recorded from 35 subjects was also employed to evaluate the feasibility of the proposed model. The experimental results demonstrate the outperformance of the proposed method in terms of classification accuracy and ITRs, which is particularly true when the number of stimuli is small. As the number of stimuli increases, the performance of the proposed method will downgrade slightly, but it could still achieve better performance than compared methods.

The remaining paper is arranged as follows: The experiment description and multi-objective optimisation-based high-pass spatial filtering methodology are described in Section 3.2. The results and discussion are provided in Section 3.3. Section 3.4 presents the conclusion.

3.2 Materials

3.2.1 Dataset

A public benchmark SSVEP dataset [23] and two self-collected SSVEP datasets, i.e., Dataset I and Dataset II, are employed in this chapter.

3.2.2 Data Pre-processing

Considering the latency delay in the visual system of each participant, the data epoch after 0.14 s is extracted for analysis. This study incorporated the Chebyshev Type I Infinite Impulse Response (IIR) filter to provide a band-pass filter with a band [8 40] Hz for Dataset I and [13 40] Hz for Dataset II. The frequency band encompasses both the fundamental wave and the second harmonic, which exhibit a more pronounced amplitude compared to higher harmonics. The cheb1ord function in MATLAB determines the lowest order for a Chebyshev Type I filter, ensuring no more than 3 dB of loss in the passband and achieving at least 20 dB attenuation in the stopband. The zero-phase digital filtering was performed in both the forward and reverse directions.

3.3 Multi-objective Optimization-based High-pass Spatial Filtering Method

This subsection introduces a novel multi-objective optimisation-based high-pass spatial filter method. The proposed algorithm aims to extract target-relevant features and minimise target-irrelevant information while reducing the volume conduction artifact.

Denote the single-trial individual calibration signal as $\boldsymbol{\chi}_i^h \in \mathbb{R}^{N_c \times N_s}$ ($h = 1, 2, ..., N_t$, $i = 1, 2, ..., N_f$), where h and i refer to the index of training trial and the stimulus, respectively. Hereafter, N_c represents the number of channels, N_s is the number of samples, N_t is the number of training trials, and N_f is the number of visual stimulation. Therefore, the continuous training signal of i-th stimulus yielded by concatenating N_t training trials is represented as $\boldsymbol{\chi}_i = [\boldsymbol{\chi}_i^1, \boldsymbol{\chi}_i^2, ..., \boldsymbol{\chi}_i^{N_t}] \in \mathbb{R}^{N_c \times (N_s \cdot N_t)}$. The single-trial individual template signal is denoted as $\overline{\boldsymbol{\chi}}_i = \frac{1}{N_t} \sum_{h=1}^{N_t} \boldsymbol{\chi}_i^h \in \mathbb{R}^{N_c \times N_s}$ which obtained by averaging multiple training trials. Therefore, the continuous individual template signal is defined as $\overline{\boldsymbol{X}}_i = [\overline{\boldsymbol{\chi}}_i, \overline{\boldsymbol{\chi}}_i, ..., \overline{\boldsymbol{\chi}}_i] \in \mathbb{R}^{N_c \times (N_s \cdot N_t)}$. Compared with the artificial reference with sine and cosine waves, the averaged SSVEP signal can effectively enhance the SNR of EEG data [68].

The multi-objective optimisation-based high-pass spatial filtering method designs objective functions for each visual stimulation using training signals from not only the same stimulus

but also others. To be specific, as the model trains the spatial filter $\hat{\boldsymbol{w}}_i \in \mathbb{R}^{N_c}$, i-th stimulus is regarded as "aim" and the remaining $(N_f - 1)$ stimuli represent "non-aim". For ease of reference, signals of $(N_f - 1)$ "non-aim" form a new dataset $(\gamma_1, \gamma_2, ..., \gamma_{N_f-1})$. The proposed method aims to maximise the correlation coefficient between the continuous training signal $\boldsymbol{\chi}_i$ from "aim" and its continuous individual template $\overline{\boldsymbol{X}}_i$, whereas minimise the correlation between the continuous training signal from each "non-aim" γ_j $(j = 1, 2, ..., N_f - 1)$ and $\overline{\boldsymbol{X}}_i$. Therefore, N_f objective functions are designed as:

$$\begin{aligned} f_1(\boldsymbol{w}_i) &= -\rho(\boldsymbol{\chi}_i^{\mathsf{T}}\boldsymbol{w}_i, \overline{\boldsymbol{X}}_i^{\mathsf{T}}\boldsymbol{w}_i), \\ f_2(\boldsymbol{w}_i) &= \rho(\boldsymbol{\gamma}_1^{\mathsf{T}}\boldsymbol{w}_i, \overline{\boldsymbol{X}}_i^{\mathsf{T}}\boldsymbol{w}_i), \\ f_3(\boldsymbol{w}_i) &= \rho(\boldsymbol{\gamma}_2^{\mathsf{T}}\boldsymbol{w}_i, \overline{\boldsymbol{X}}_i^{\mathsf{T}}\boldsymbol{w}_i), \\ &\vdots \\ f_{N_f}(\boldsymbol{w}_i) &= \rho(\boldsymbol{\gamma}_{N_f-1}^{\mathsf{T}}\boldsymbol{w}_i, \overline{\boldsymbol{X}}_i^{\mathsf{T}}\boldsymbol{w}_i) \end{aligned} \tag{3.1}$$

where $\rho(\boldsymbol{a}, \boldsymbol{b})$ refers the Pearson correlation coefficient between vector $\boldsymbol{a}$ and vector $\boldsymbol{b}$. In this case, all objective functions are to be minimised. Therefore, the feature related to the i-th stimulus is maximally extracted. On the contrary, the information relevant to "non-aim" is minimally included in the spatial filter $\hat{\boldsymbol{w}}_i$. Simultaneously, considering the effect of volume conduction in SSVEP signals, the model is subject to a linear equality constraint. That is, the sum of spatial filter elements is zero. It aims to reduce the low spatial-frequency signal while retaining the high spatial-frequency. Therefore, the multi-objective optimisation problem is formulated by the following statement:

$$\begin{aligned} &\underset{\boldsymbol{w_i}}{\text{minimise}}\ F(\boldsymbol{w}_i) = [f_1(\boldsymbol{w}_i), f_2(\boldsymbol{w}_i), ..., f_{N_f}(\boldsymbol{w}_i)], \\ &\text{subject to}\ \sum_{c=1}^{N_c} w_i^c = 0, \boldsymbol{w}_i \in \boldsymbol{W} \end{aligned} \tag{3.2}$$

where c is the channel index and w_i^c refers to the c-th element of spatial filter vector $\boldsymbol{w}_i$. The $\boldsymbol{W} \subseteq \mathbb{R}^{N_c}$ is the feasible set of solution vectors. Therefore, the spatial filter $\hat{\boldsymbol{w}}_i$ can be

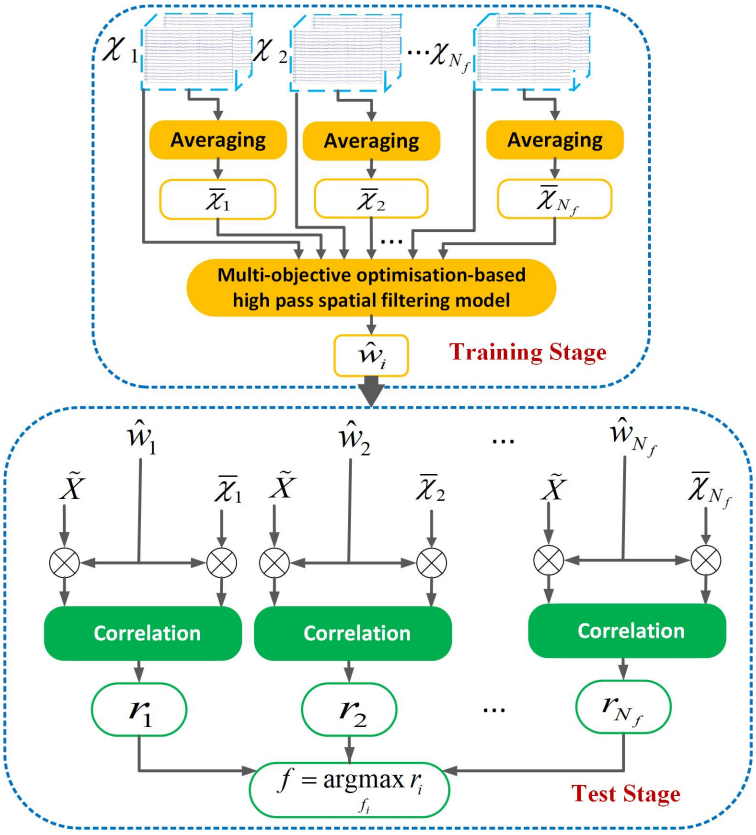


Figure 3.1: Flowchart of the proposed SSVEP recognition model. In the training stage, the spatial filter for each stimulus $\hat{\boldsymbol{w}}_i, i = 1, 2, ..., N_f$, is generated with formulas (3.1)-(3.3), and the new reference signal, i.e. $\overline{\boldsymbol{X}}_1, \overline{\boldsymbol{X}}_2, ..., \overline{\boldsymbol{X}}_{N_f}$, is obtained by averaging across multiple training trials. In the test stage, with spatial filters $\hat{\boldsymbol{w}}_1, \hat{\boldsymbol{w}}_2, ..., \hat{\boldsymbol{w}}_{N_f}$, the correlation between a test trial $\widetilde{\boldsymbol{X}}$ and each individual template $\overline{\boldsymbol{X}}_i, i = 1, 2, ..., N_f$, is computed by formula (3.4). The frequency of the template signal with the maximum correlation coefficient is determined as that of the test sample by formula (3.5).

acquired as follows:

$$\hat{\boldsymbol{w}}_i = \arg\min_{\boldsymbol{w}_i} F(\boldsymbol{w}_i). \tag{3.3}$$

This constrained multi-objective optimisation problem can be solved by the fgoalattain() function in Matlab. In the test phase, the signal-trial test SSVEP data $\widetilde{\boldsymbol{X}} \in \mathbb{R}^{N_c \times N_s}$ and the single-trial template signal $\overline{\boldsymbol{X}}_i$ are both spatially filtered with the optimal solution vector

$\hat{\boldsymbol{w}}_i$. The correlation coefficient calculated between two optimised signals is shown as follows:

$$r_i = \rho(\widetilde{\boldsymbol{X}}^{\mathsf{T}}\hat{\boldsymbol{w}}_i, \overline{\boldsymbol{\chi}}_i^{\mathsf{T}}\hat{\boldsymbol{w}}_i),\ i = 1, 2, ..., N_f \tag{3.4}$$

The N_f coefficients can be calculated by this formula. The frequency of the template signal corresponding to the maximal correlation coefficient value is determined as the frequency f of the SSVEP test trial $\widetilde{\boldsymbol{X}}$, and it is can be represented as:

$$f = \arg\max_{f_i} r_i, i = 1, 2, ..., N_f \tag{3.5}$$

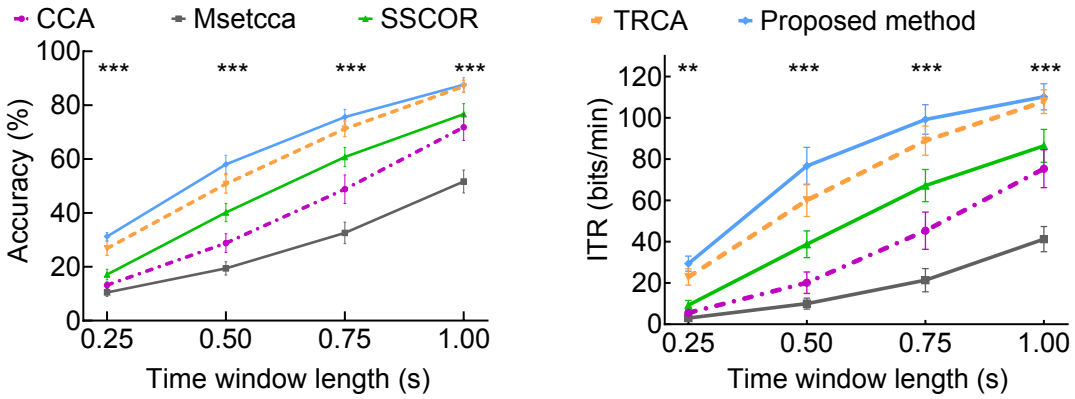
The diagram of the whole SSVEP recognition model is shown in Fig.3.1. It includes two stages, namely training and testing. For each stimulus, the training part provides its spatial filter and the corresponding template. The test part aims to classify the single-trial test signal to a specific visual stimulus based on the outputs of the training phase.

3.4 Results and Discussion

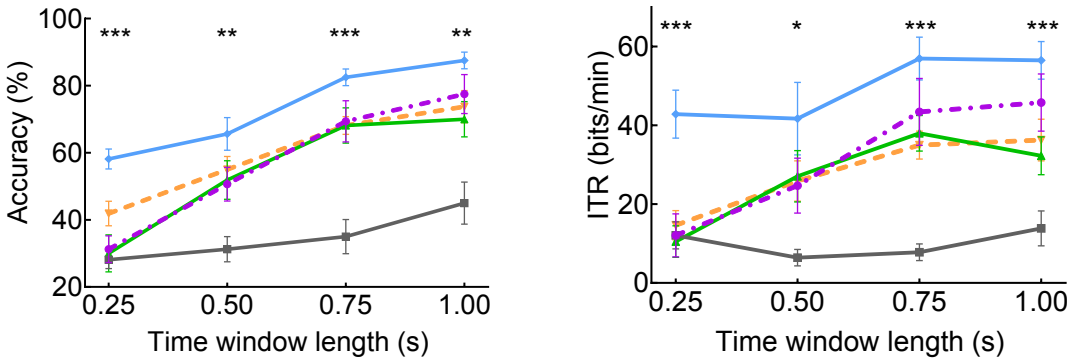
In this section, the performance of the proposed multi-objective optimisation-based high-pass spatial filtering method was first evaluated on two self-collected SSVEP datasets and a 40-target benchmark dataset [23]. In Dataset I, the parameters are set as $N_t = 4, N_f = 12$; in Dataset II, they are $N_t = 4, N_f = 4$; and in the benchmark dataset, the parameters are $N_t = 5, N_f = 40$. Extensive comparisons were implemented between the proposed method and many state-of-the-art SSVEP recognition methods. The ensemble-based methods were also implemented in this section for comparison purposes. The influences of different parameters such as the number of electrodes, the number of training blocks, and the number of frequencies on the performance were also discussed.

3.4.1 Performance Evaluation

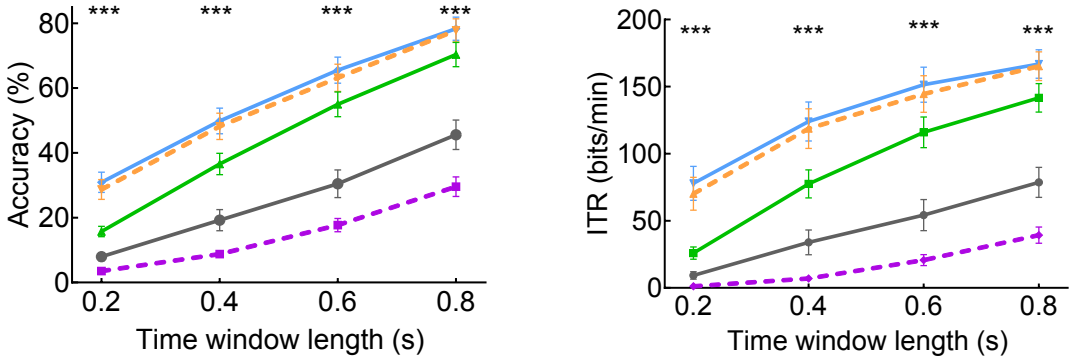
The performance comparison was conducted between the proposed model and many recognition methods such as CCA, MsetCCA, SSCOR, and TRCA. The classification accuracy



(a) Dataset I

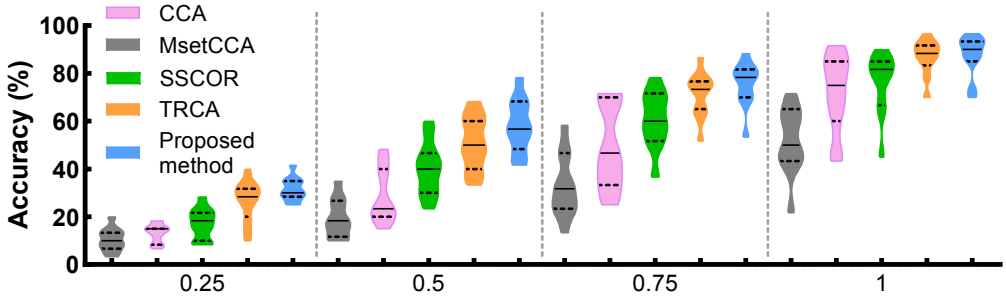


(b) Dataset II

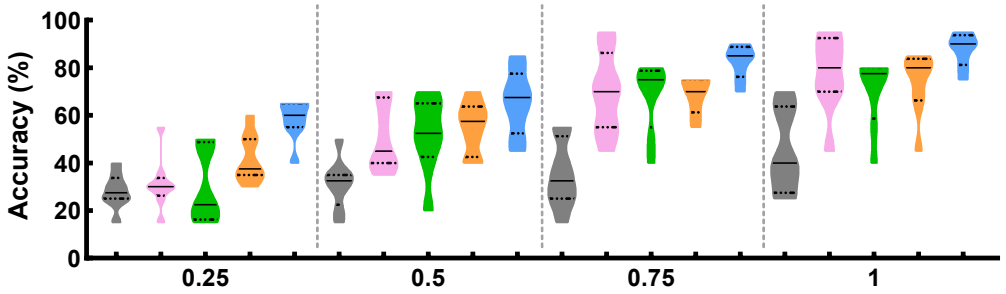


(c) Benchmark dataset

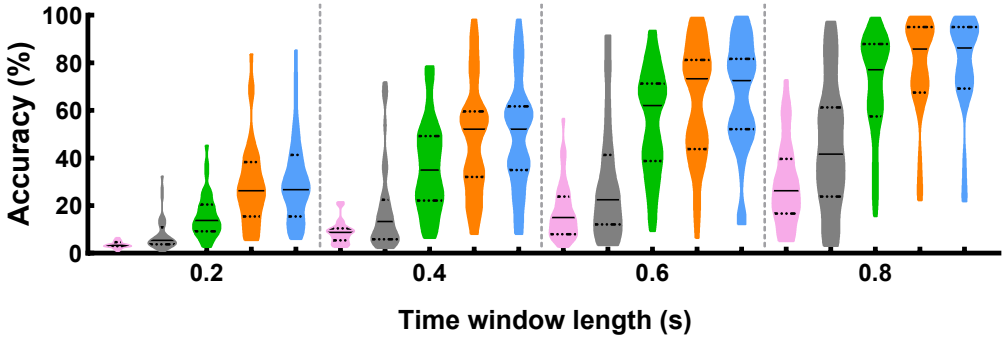
Figure 3.2: Averaged recognition accuracy and ITRs across subjects of various methods using different time windows on (a) Dataset I, (b) Dataset II, and (c) benchmark dataset. The error bars represent standard error of mean (SEM). The asterisks indicate significant differences between the five methods obtained by one-way repeated-measures ANOVA ($*$ $p<0.05$, $**p<0.01$, $***p<0.001$).



(a) Dataset I



(b) Dataset II

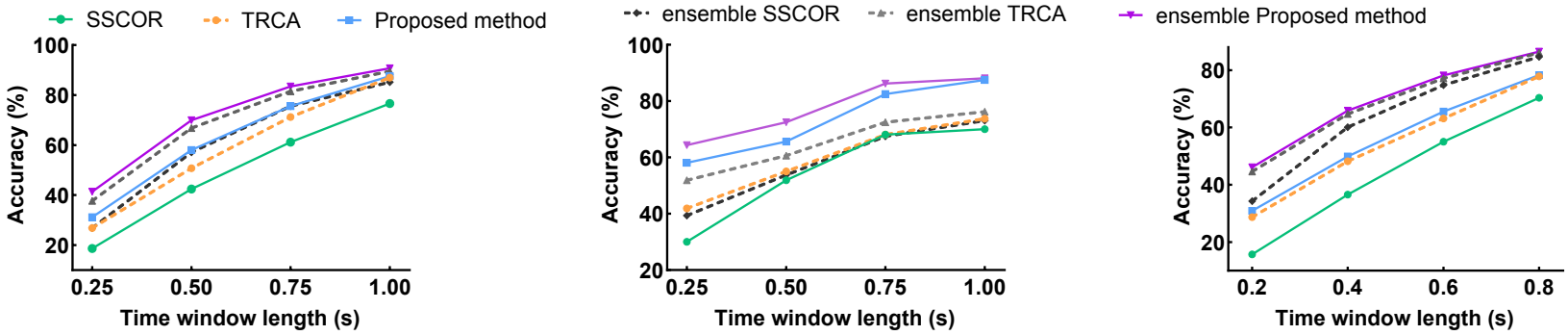


(c) benchmark dataset

Figure 3.3: Violin plots represent the distributions of classification accuracy of all subjects achieved by the five methods with various TWs on (a) Dataset I, (b) Dataset II, and (c) benchmark dataset. The black solid line in each violin indicates the median and two black dotted lines represent interquartile ranges (25% and 75% percentiles).

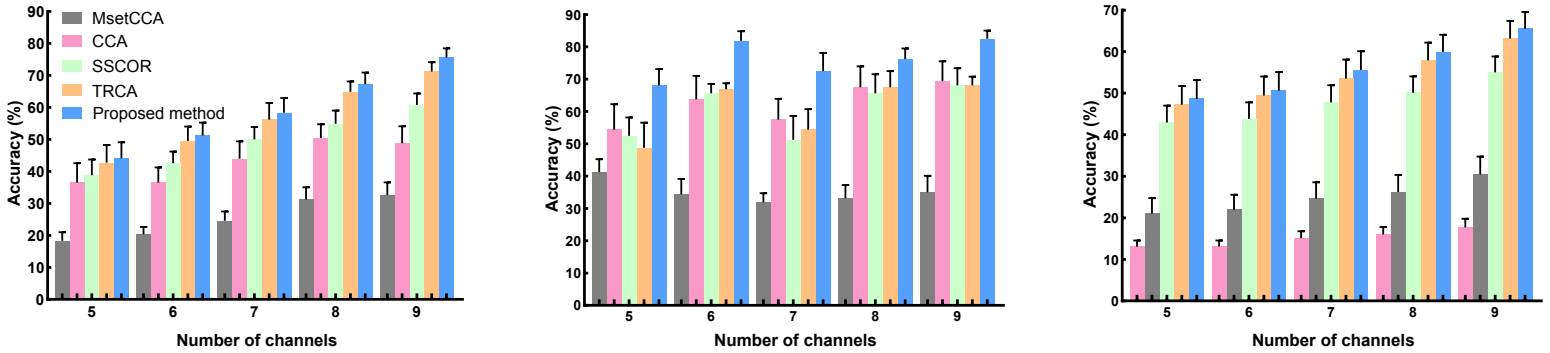
and ITRs were calculated by leave-one-block-out cross-validation to evaluate the recognition performance of these algorithms. Specifically, the SSVEP signals from four (in Dataset I and II) or five (in benchmark dataset) blocks were employed as the training data, and the left-out block was used as test data. Fig. 3.2 illustrates the averaged classification accuracy and ITRs across all subjects in three datasets with different time windows (TWs). Due to the different sampling rates in the self-collected datasets and the benchmark dataset, the data lengths were varied accordingly to ensure the number of samples remained as whole numbers, without decimals. It is evident that the proposed method can in general achieve the highest accuracy and ITRs with different data lengths. One-way repeated-measures analysis of variance (ANOVA) was conducted to investigate the similarity of accuracy or ITRs of different SSVEP detection approaches. The results indicate that their accuracy and ITRs have a significant difference in most TWs on the three datasets. For example, in Fig. 3.2(a), the proposed method improved by 15.78% for CCA, 35.98% for MsetCCA, 10.98% for SSCOR, and 0.68% for TRCA with 1s data length.

In addition to the averaged values, the distribution of numeric data across multiple methods was further explored. Fig. 3.3 adopts the violin plot to show the probability density of recognition accuracy of all subjects on (a) Dataset I, (b) Dataset II, and (c) benchmark dataset, respectively. The plots are based on five methods with different data lengths. Each density curve can be compared to see the similarities or differences between the methods. Fig. 3.3(a) and Fig. 3.3(b) show that this method achieves the highest median and most concentrated distribution with most data lengths. Therefore, to a large extent, the proposed method provides higher and more consistent accuracy across subjects. In Fig. 3.3(c), all methods show a more scattered distribution compared with Fig. 3.3(a) and Fig. 3.3(b), because there are significantly more subjects in the benchmark dataset. The proposed method provides medians similar to TRCA, but obviously higher than those of other methods.



(a) Dataset I (b) Dataset II (c) benchmark dataset

Figure 3.4: Averaged recognition accuracy comparison between ensemble methods and standard methods on (a) Dataset I, (b) Dataset II, and (c) benchmark dataset.



(a) Dataset I (b) Dataset II (c) benchmark dataset

Figure 3.5: Barcharts of the five methods' classification accuracy with different number of channels on (a) Dataset I, (b) Dataset II, and (c) benchmark dataset. The error bars represent SEM.

3.4.2 Ensemble-based Method Evaluation

This work extends the proposed method with an ensemble version. Ideally, the transformation coefficients from SSVEP source to the scalp could be treated similarly within the used frequency range [1] [197] [71]. Hence, N_f different spatial filters should be similar. The ensemble-based method concatenates all N_f spatial filters to construct an ensemble spatial filter $\boldsymbol{W} \in \mathbb{R}^{N_c \times N_f}$,

$$\boldsymbol{W} = [\hat{\boldsymbol{w}}_1, \hat{\boldsymbol{w}}_1, ..., \hat{\boldsymbol{w}}_{N_f}] \tag{3.6}$$

Then the (3.4) can be re-defined as follows:

$$r_i = \rho(\widetilde{\boldsymbol{X}}^{\mathsf{T}}\boldsymbol{W}, \overline{\boldsymbol{\chi}}_i^{\mathsf{T}}\boldsymbol{W}),\ i = 1, 2, ..., N_f \tag{3.7}$$

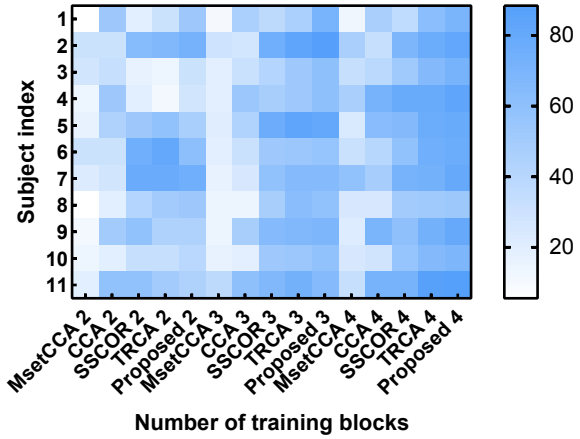
The intended target is determined via (3.5). Fig. 3.4 shows the performance comparison between different standard methods and ensemble methods on (a) Dataset I, (b) Dataset II, and (c) benchmark dataset. As shown by the blue and purple lines in Fig. 3.4, the performance of the proposed method is improved via the ensemble version in three datasets. Meanwhile, in Fig. 3.4(a) and Fig. 3.4(c), the classification accuracy of the ensemble proposed method is slightly better than that of ensemble TRCA, but the gap with ensemble SSCOR is even greater. In Fig. 3.4(b), the standard proposed method shows superiority compared with ensemble TRCA and ensemble SSCOR with all data lengths.

3.4.3 The Influence of Parameters on Performance

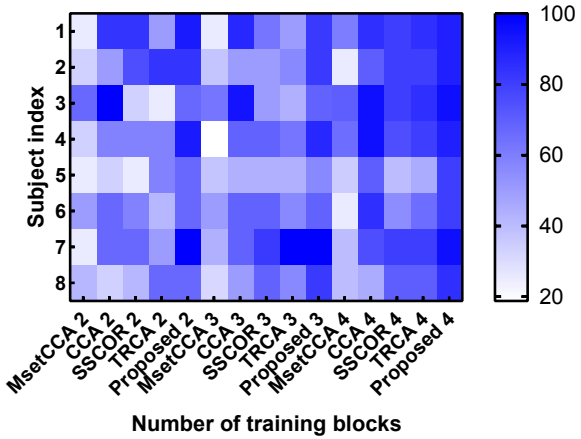
To further assess the performance of the five approaches, an evaluation was conducted on how the number of electrodes, training blocks, and stimulation levels impact SSVEP recognition accuracy.

The number of electrodes

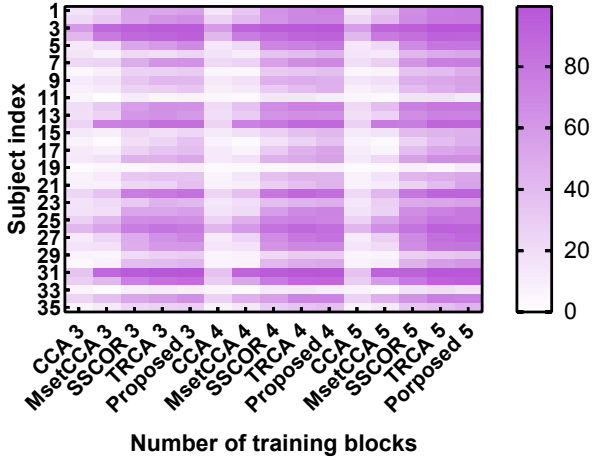
Fig. 3.5 depicts the averaged SSVEP classification accuracy with various numbers of channels using 0.75s-long data on (a) Dataset I, (b) Dataset II, and (c) benchmark dataset. In Fig. 3.5, the accuracy of each method generally declines as the number of channels decreases.



(a) Dataset I



(b) Dataset II



(c) benchmark dataset

Figure 3.6: Heat maps of the classification accuracy of five methods under different number of training blocks on (a) Dataset I, (b) Dataset II, and (c) benchmark dataset.

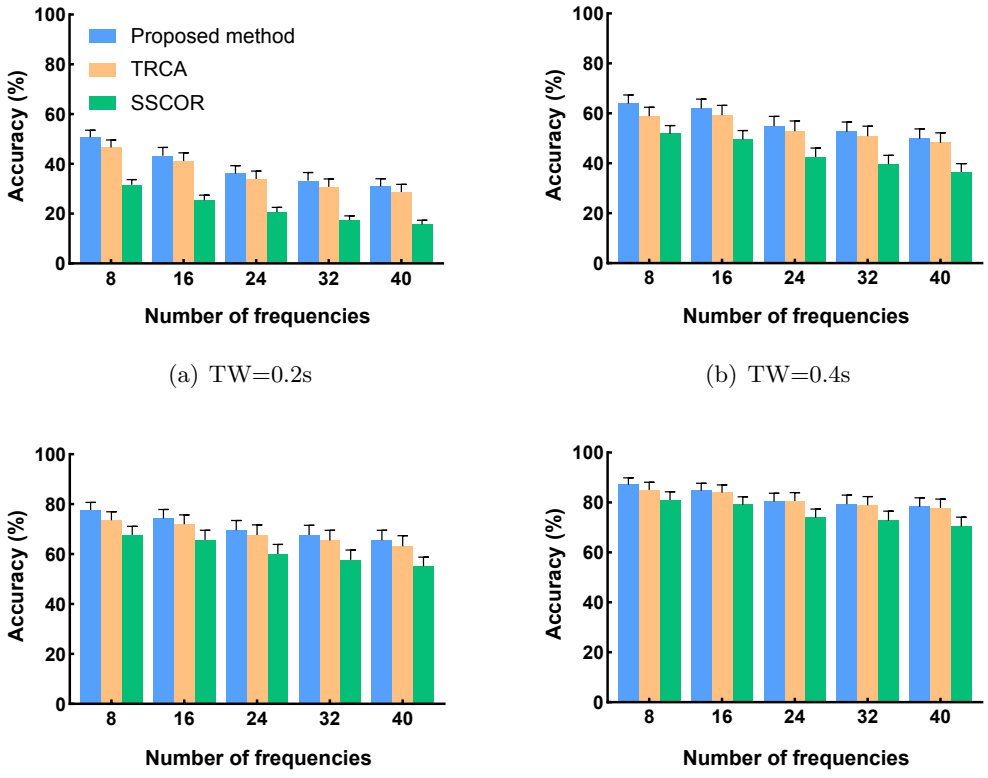
It indicates that the number of channels affects the performance of various SSVEP recognition methods. Specifically, the method with more channels can commonly achieve better performance. It is worth mentioning that for channels = 5, 6, 7, 8, and 9, the proposed model (bars in blue) always outperforms the other four methods on all datasets.

The number of training blocks

Additionally, an exploration was conducted to assess how the number of training blocks affects the target identification performance of five algorithms on three datasets. The heat maps in Fig. 3.6 illustrate comparison results with 0.75 s TWs. The x-axis refers to the method with a different number of training blocks. The y-axis is the subject index, and the color indicates the corresponding performance. The maximal SSVEP detection accuracy shows the deepest color. For all methods in the three datasets, the color obtained with more training blocks is usually darker than that obtained with fewer training blocks. In addition, the proposed method generally shows deeper color compared with the other four methods for most subjects under the different number of training blocks. The results indicate that, to some extent, this method is superior to other methods regardless of the number of training blocks.

The number of frequencies

In this study, an investigation was also conducted into the effect of varying target numbers on the proposed method. In order to explore more types of target numbers, 40-class benchmark dataset was employed in this specific subsection. SSCOR and TRCA significantly outperform CCA and MsetCCA on the benchmark dataset as shown in Fig. 3.2, so these two methods are utilized for performance comparison. Fig. 3.7 shows the averaged classification accuracy of three methods to classify 8, 16, 24, 32, and 40 stimulation, respectively. The choice of targets is random. The comparison result indicates that the performance of this method tends to decrease slightly with the increasing number of stimulation. Under the same setting, TRCA and SSCOR also show a similar declining trend. However, even with a large number of targets, the recognition accuracy of the proposed method is always better than other models. Similar comparison results can also be found in Fig. 3.2(a) with



(a) TW=0.2s

(b) TW=0.4s

(c) TW=0.6s

(d) TW=0.8s

Figure 3.7: Performance comparison with various numbers of targets, i.e. 8, 16, 24, 32, and 40 using different data lengths (a) TW=0.2s, (b) TW=0.4s, (c) TW=0.6s, (d) TW=0.8s. The error bars represent SEM.

12-class Dataset I and Fig. 3.2(b) with 4-class Dataset II.

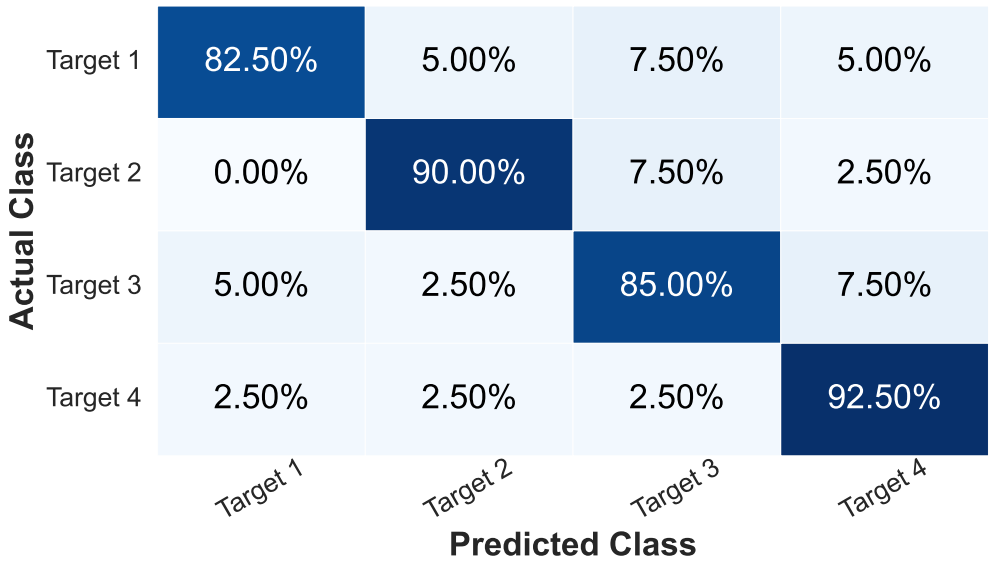
3.4.4 Discussion

In order to learn the class-specific spatial filter, the traditional SSVEP recognition scheme generally considers the correlation among training trials from the same class [1] [67] [71]. For example, TRCA trains the spatial filter by maximizing the sum of covariance of all possible combinations of training trials from the same stimulation [1]. For the proposed model, it constructs multi-objective functions creatively not only to decrease the distance between the training signal and the template from the same class but also to increase the distance between SSVEP data from other classes and the template, as shown in formulas (3.1)-(3.3). Therefore, the trained spatial filters have the potential to extract target-related features and

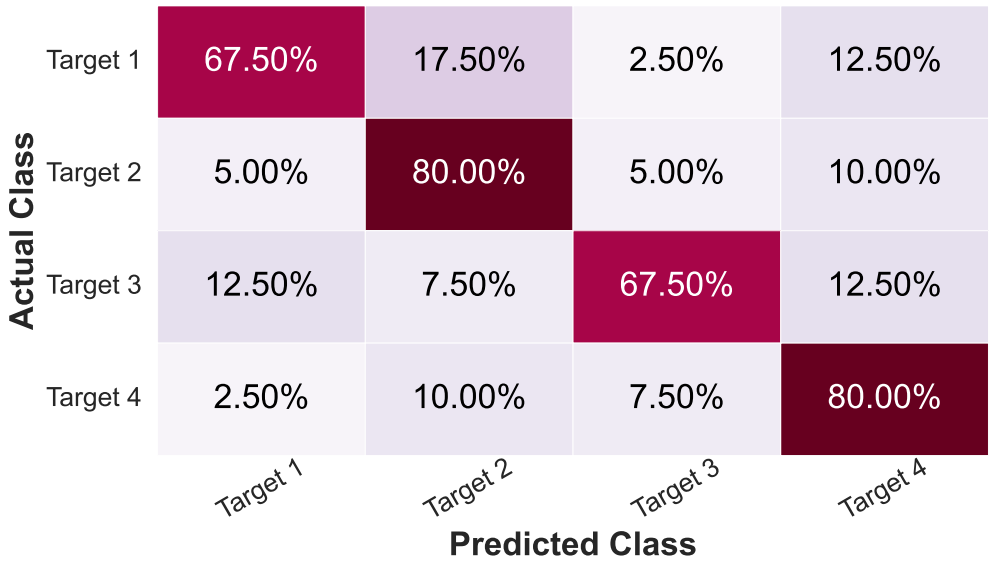
decrease target-unrelated information. Besides, the optimisation problem is subject to the constraint that the sum of filter elements is zero so that low spatial-frequency signals can be alleviated. With the spatial filters, the proposed model could better distinguish between target and non-target stimulation. The false positive rate (FPR) for a multi-target system is the percentage of non-targets that are wrongly categorised as the true target [128]. As a representative example, the confusion matrix in Fig. 3.8 provides the averaged FPR across all subjects and targets. Clearly, the averaged FPRs of the proposed method and TRCA are 12.5% and 26.25%, respectively. Furthermore, the probability density of detection accuracy for the proposed method was investigated, as depicted in Fig. 3.3. The frequency of data points in each region correlates to the width of each curve. The thicker part means that the corresponding value has a higher frequency. As shown in Fig. 3.3, when the median value among methods is approximated (e.g. 1 s data length in Dataset II), the proposed method has a more concentrated distribution. It implies that the proposed method could achieve a more stable and consistent performance across subjects.

In order to further clarify the effect of stimulation mechanisms on the performance, visual stimuli are displayed in different ways, i.e. LCD monitor in Dataset I and LEDs in Dataset II, respectively. As shown by the blue lines in Fig. 3.2, the proposed method always exhibits the best classification accuracy with various TWs. Blue violins in Fig. 3.3 reflect the similar result via another form. It illustrates that different choices of stimulation mechanisms do not affect the performance of the proposed method too much. The choice of varying frequencies in Dataset I and Dataset II was intentional to broaden the explored frequency band and investigate performance across a wider spectrum.

As a classical method, CCA has been widely validated by many studies in detecting SSVEPs [38] [69]. Traditionally, its reference signals, constructed by sine-cosine waves, are employed to model the visual stimuli [130]. Fig. 3.2 shows that CCA could reach an average recognition accuracy of 71.81% in Dataset I, 77.5% in Dataset II, and 29.57% in the Benchmark dataset. It does not achieve as high accuracy as other methods at some TWs. A potential problem with this method is that the artificially predefined reference signal cannot reflect subject-specific features [67] [74]. This method adopts the individual training data to con-



(a) Confusion matrix generated by the proposed model



(b) Confusion matrix generated by TRCA

Figure 3.8: Confusion matrices of target recognition on Dataset II achieved by (a) multi-objective optimisation-based high-pass spatial filtering model and (b) TRCA with 1s data length. Each cell shows the ratio of the number of observations to the total number of test trials per target.

struct subject-specific templates. As illustrated in [68], one of the factors in the target recognition method leads to class and non-class SSVEP signals being more distinguishable as individual templates.

Although this proposed method provides good recognition performance in SSVEP-based BCIs, it still has some potential improvement directions. First, as shown in Fig. 3.7, as the number of stimuli increases, the performance of this approach tends to decline slightly. In fact, as the number of objectives in a multi-objective optimisation problem increases, the dimensionality of the model also increases, which makes it difficult to converge [198]. In addition, the interaction among various objective functions makes the optimisation problem more intricate. It has been observed that the multi-objective optimisation technique experiences performance degradation when applied to a problem with a high number of objective functions [199]. Therefore, the performance comparisons do not show a huge difference when the proposed approach is applied to an optimisation problem based on a large number of objective functions. Future work will thus be required to 1) optimise the selection of non-aim stimuli to reduce the number of minimisation objectives and 2) explore other algorithms to solve complex multi-objective optimisation problems with a large number of objectives for the SSVEP-based BCI system. Second, as Fig. 3.6 demonstrated, the training blocks may not be sufficient for some subjects, resulting in relatively low classification accuracy. Therefore, future work would transfer information from other subjects to solve the data insufficiency problem.

3.5 Conclusion

In this study, a novel multiple-objective optimisation-based high-pass spatial filtering method was proposed to improve the target recognition performance for the SSVEP-based BCI system. this method defined multi-objective functions for each target by the correlation between the training signal and the individual template from the same stimulation maximisation, whereas the correlation between SSVEP data from others and the template minimisation. It aims to effectively extract target-relevant information while decreasing target-irrelevant features. Meanwhile, the proposed model has the constraint that the sum of filter

elements is zero. This setting could reduce the negative effect of volume conduction by alleviating low spatial-frequency signals. Experimental results based on two self-collected SSVEP datasets and a public benchmark database showed that the proposed model achieved better classification performance than some state-of-the-art methods.

Chapter 4

Bayesian-based Classification Confidence Estimation for Enhancing SSVEP Detection

The BCI enables paralyzed people to communicate directly with and operate peripheral equipment. The SSVEP-based BCI system has been extensively investigated in recent years due to its fast communication rate and high signal-to-noise ratio. Many present SSVEP recognition methods determine the target class by finding the largest correlation coefficient. However, the classification performance usually degrades when the largest coefficient is not significantly different from the rest of the values. This study proposed a Bayesian-based classification confidence estimation method to enhance the target recognition performance of SSVEP-based BCI systems. This method uses the differences between the largest and the other values generated by a basic target identification method to define a feature vector during the training process. The Gaussian mixture model (GMM) is then employed to estimate the probability density functions of feature vectors for both correct and wrong classifications. Subsequently, the posterior probabilities of being an accurate and false classification are calculated via Bayesian inference in the test procedure. A classification confidence value (CCValue) is presented based on two posterior probabilities to estimate

the classification confidence. Finally, the decision-making rule can determine whether the present classification result should be accepted or rejected. Extensive evaluation studies were performed on an open-access benchmark dataset and a self-collected dataset. The experimental results demonstrated the effectiveness and feasibility of the proposed method for improving the reliability of SSVEP-based BCI systems.

4.1 Introduction

BCI systems can detect brain activity and then translate neural signals directly into commands to operate external devices without relying on peripheral nerves and muscles [75, 178, 200]. The EEG-based BCI is a popular non-invasive technique due to portability, low cost, and high temporal resolution [185, 201, 202]. Three paradigms in the EEG signal are most widely explored, namely, the SSVEP, the P300 event-related potential (ERP), and the event-related desynchronization (ERD) [203]. These paradigms have come to light in several practical applications, including assistance robots [7], AR glasses [181, 182], and entertainment [204, 205]. Among these paradigms, SSVEP-based BCI systems have attracted extensive research attention because of their advantages of high ITR and SNR [86, 183, 184, 206].

In recent decades, many target recognition methods have been proposed to analyze the SSVEP features and detect the subject's intent to operate the peripheral device [186]. In particular, CCA is the most popular target detection method because of its simplicity of use and robustness [2, 68]. However, the performance of CCA is still influenced by the interference from spontaneous EEG signals [1]. In recent years, many improved approaches have been proposed for SSVEP detection. Generally, the literature presents three major optimization directions, i.e., individual templates [24, 67], time filters [207], and spatial filters [1, 71]. Among many methods, the SSCOR [71] and TRCA [1] have attained nice performance in SSVEP detection. In the recognition stage of the methods above, the largest correlation coefficient identifies the target class. It may lead to misclassification when the maximum coefficient has a low confidence level. The detection performance may deteriorate if the maximum value is not remarkably different from the other values. Therefore, evaluating the reliability of classification results is another crucial direction for enhancing

the performance and applicability of SSVEP-based BCI systems.

The classification confidence analysis process could facilitate detection methods to reject results with a low level of confidence [208]. In recent research, many confidence evaluation methods for the SSVEP-based BCI system have been introduced. For instance, Zhao *et. al* [102] designed a decision-making selector to select a more reliable result from a pair of CCA-based methods. The overall recognition performance was enhanced by the fusion strategy, but the average detection time increased accordingly. Currently, many researchers have focused on confidence estimation based on a single decision. Chen *et.al* [107] created a hypothesis testing model for evaluating the credibility of results using the coefficients of filter-bank CCA. Cecotti [106] investigated the impact of different dynamic time segment selections on the confidence of CCA's outputs. Similarly, Jiang *et.al* [108] estimated the classification confidence based on the largest two coefficients and then determined the optimal data length. According to several previous studies, the difference between the first and the second-largest feature values provides useful information for the classification estimation [115]. In general, the probability of correct recognition is higher as this difference is larger [107]. However, these methods simply exploit the first two coefficients or their difference, which is insufficient to construct informative features for enhancing SSVEP detection.

In this paper, the Bayesian-based classification confidence estimation method was proposed for improving the recognition reliability of SSVEP-based BCI systems, which is crucial for SSVEP-based human-robot interaction [209, 210]. Wrong classifications can cause the external device to carry out the wrong actions, perhaps resulting in adverse incidents and serious physical harm to humans. In the practical usage scene, enhancing subjects' safety and security is essential, particularly in rehabilitation and assistive technology. The main contributions of this work include: 1) In the training step, the feature vector involving the differences between the largest correlation coefficient and the other values was constructed. Gaussian mixture model (GMM) was used to estimate feature vectors' conditional probability density functions, given correct and wrong results. 2) In the test step, Bayesian inference was used to calculate the posterior probabilities of being a correct and wrong classification using the newly obtained feature vector. A classification confidence value (CCValue) was

then presented to estimate the classification confidence. 3) The decision-making rule decides whether the present classification result should be accepted or rejected.

For this study, SSCOR and TRCA were selected as the basic target recognition methods. The proposed methods incorporating CCValue estimation based on SSCOR/TRCA are named SSCOR+CCValue and TRCA+CCValue, respectively. The performance was assessed on a 40-class publicly available benchmark dataset [23] and a 12-class self-collected dataset. Extensive comparisons were performed among the four methods. The effectiveness and reliability of SSCOR+CCValue and TRCA+CCValue were demonstrated via experimental evaluation studies on two datasets.

4.2 Materials

4.2.1 Datasets

This chapter utilized an open-access dataset [23] and a 12-class self-collected SSVEP dataset, i.e., benchmark dataset to evaluate the proposed method's performance.

4.2.2 Data Pre-processing

To account for the latency delay in the human visual system, the EEG signal in [0.14 s $0.14 + d$ s] was extracted for method performance evaluation [211]. The variable d in this context refers to the length of the data that is being used for analysis. The Chebyshev Type I IIR filter was applied in this work to create band-pass filters. The data was filtered between eight Hz and eighty-eight Hz for the benchmark dataset. The data was filtered between eight Hz and forty Hz for Dataset I. In addition, a data standardization step was performed on both datasets [1].

4.3 Bayesian-based SSVEP Classification Confidence Estimation Method

A Bayesian-based classification confidence estimation method was proposed for improving SSVEP recognition reliability. As shown in Fig. 4.1, it includes four modules, namely

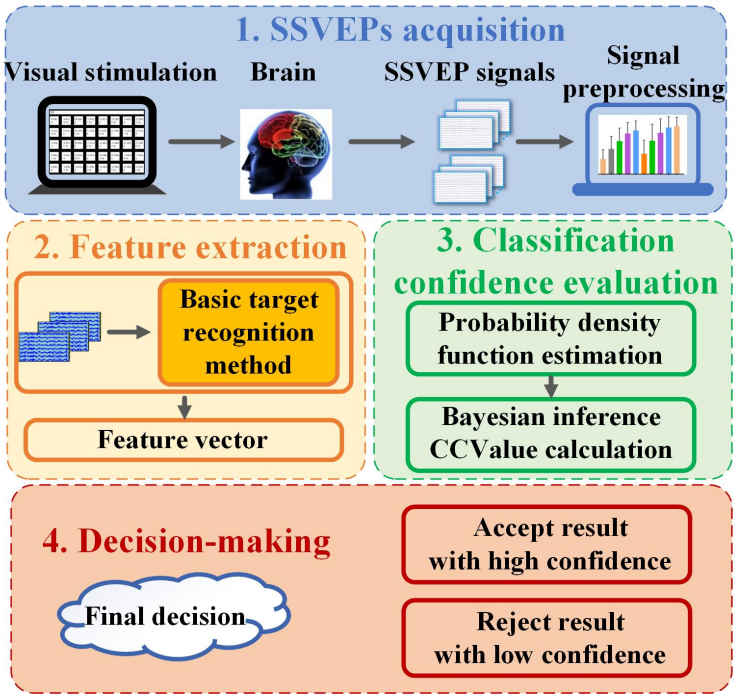


Figure 4.1: Diagram of the Bayesian-based SSVEP classification confidence estimation method.

EEG signal acquisition, feature extraction, classification confidence evaluation, as well as decision making. The dataset descriptions have been given in the previous section 4.2.1. In the following subsections, the work procedures of the remaining three modules will be explained in detail.

Feature Extraction

Denote a four-way tensor $\boldsymbol{\chi} \in \mathbb{R}^{N_f \times N_c \times N_s \times N_t}$, where N_f indicates the number of stimuli, N_c represents the number of channels, N_s is the number of samples, and N_t is the number of training trials. Hereafter, i refers to the stimulus index, j refers to the channel index, m refers to the sample index, and h refers to the index of training trials. Therefore, the recorded individual calibration signal for i-th stimulus is $\boldsymbol{\chi}_i \in \mathbb{R}^{N_c \times N_s \times N_t}$. The spatial filter $\boldsymbol{w}_i \in \mathbb{R}^{N_c}$ for i-th stimulus can be constructed as $\boldsymbol{w}_i = f(\boldsymbol{\chi}_i)$ by a basic target recognition method in SSVEP-based BCIs. $f(\cdot)$ represents the spatial filtering method. In this study, TRCA and SSCOR were selected. In TRCA [1], weight coefficients are optimized to maximize inter-trial covariance of brain activities. SSCOR spatial filter learns

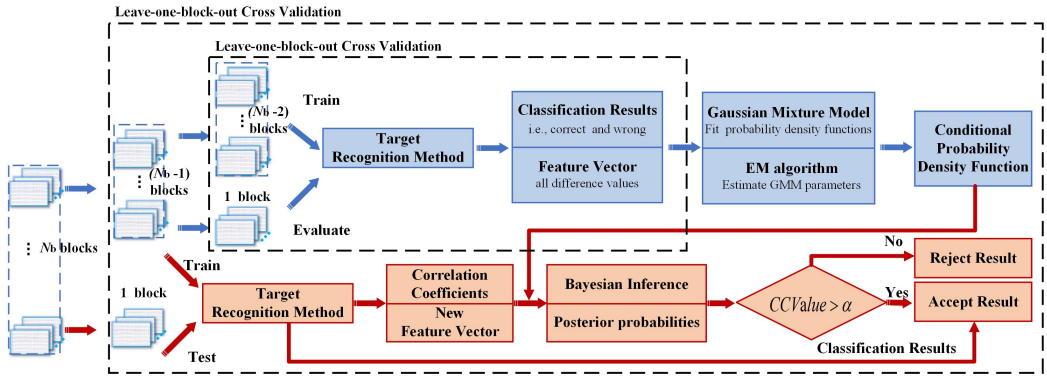


Figure 4.2: The detailed framework of the Bayesian-based classification confidence estimation method for SSVEP detection. The leave-one-block-out cross-validation was performed in the experiment evaluation.

a common SSVEP representation space through the optimization of the individual SSVEP templates. It could improve the SNR of the SSVEP components embedded in the recorded EEG data [71].

The single-trial individual template signal is denoted as $\overline{\boldsymbol{\chi}}_i = \frac{1}{N_t}\sum_{h=1}^{N_t}\boldsymbol{\chi}_{ih} \in \mathbb{R}^{N_c \times N_s}$. Once the spatial filter $\boldsymbol{w}_i$ is produced, the evaluation SSVEP data $\widetilde{\boldsymbol{X}} \in \mathbb{R}^{N_c \times N_s}$ and individual template signal $\overline{\boldsymbol{\chi}}_i$ can both be optimised. Therefore, the SSVEP feature was further extracted from recorded EEG signals. The correlation coefficient between the two spatially filtered signals corresponding to each stimulus is shown as follows:

$$r_i = \rho(\widetilde{\boldsymbol{X}}^{\mathsf{T}}\boldsymbol{w}_i, \overline{\boldsymbol{\chi}}_i^{\mathsf{T}}\boldsymbol{w}_i),\ i = 1, 2, ..., N_f \tag{4.1}$$

where $\rho(\boldsymbol{a}, \boldsymbol{b})$ refers the Pearson correlation coefficient between vector $\boldsymbol{a}$ and vector $\boldsymbol{b}$. The frequency of the individual template related to the largest correlation coefficient is decided as the frequency f of the test signal:

$$f = \arg\max_{f_i} r_i, i = 1, 2, ..., N_f \tag{4.2}$$

Considering this type of decision-making rule may result in poor classification performance when the maximal coefficient is not much different from others, a Bayesian-based classification confidence estimation method was proposed in this study. The coefficient vector

calculated by (3.4) is denoted as $\mathbf{\Phi} = [r_1, r_2, ..., r_{N_f}]$. The coefficient vector was rearranged in descending order, resulting in a new vector $\widetilde{\mathbf{\Phi}} = [\widetilde{r}_1, \widetilde{r}_2, ..., \widetilde{r}_{N_f}]$. It means that the largest coefficient is $\widetilde{r}_1$, and the smallest one is $\widetilde{r}_{N_f}$. Subsequently, it is possible to calculate the differences between the largest coefficient $\widetilde{r}_1$ and other values $\widetilde{r}_j, (j = 2, 3, \ldots, N_f)$, and thus yield $(N_f - 1)$ differences. Therefore, the difference values $\Delta r_i, (i = 1, 2, \ldots, N_{f-1})$ can be expressed as:

$$\begin{aligned} \Delta r_1 &= \widetilde{r}_1 - \widetilde{r}_2 \\ \Delta r_2 &= \widetilde{r}_1 - \widetilde{r}_3 \\ &\vdots \\ \Delta r_{N_{f-1}} &= \widetilde{r}_1 - \widetilde{r}_{N_f}. \end{aligned} \tag{4.3}$$

The final feature vector can be expressed as $\boldsymbol{F} = [\Delta r_1, \Delta r_2, \ldots, \Delta r_{N_{f-1}}]$ by $(N_f - 1)$ differences.

Bayesian-based Classification Confidence Evaluation

As illustrated in Fig. 4.2, performance was assessed by the leave-one-block-out cross-validation. Specifically, for N_b blocks of EEG signals, $(N_b - 1)$ blocks were selected for training conditional probability density functions, and one block was used for testing. Moreover, in the training process, leave-one-block-out cross-validation was again employed to collect classification results and construct feature vectors (blue part in Fig. 4.2). Specifically, $(N_b - 2)$ blocks were selected to train the target recognition method, and the left-out block was used as evaluation data. The signal in each block is represented as $\boldsymbol{\chi}_h \in \mathbb{R}^{N_f \times N_c \times N_s}$. Therefore, there are total $(N_b - 1) \times N_f$ trials that can be evaluated. Thus classification results and feature vectors can be collected to train GMM accordingly.

The classification results were subsequently separated into two groups. Suppose the correct classification is represented as C_1, and the corresponding feature vectors are $\boldsymbol{F}_c$. The wrong classification is denoted as C_0, and the corresponding feature vectors are $\boldsymbol{F}_w$. The probability density functions of the feature vector for correct and wrong classifications are represented as $p(\boldsymbol{F}|C_1)$ and $p(\boldsymbol{F}|C_0)$, respectively. For ease of reference, they can also be

Table 4.1: A confusion matrix of explanation about four parameters, i.e., TR, FA, FR, and TA.

		The proposed method's decision	
		Rejection	**Acceptance**
Target identification method's decision (e.g. TRCA, SSCOR)	**Wrong**	True Rejection (TR)	False Acceptance (FA)
	Correct	False Rejection (FR)	True Acceptance (TA)

written as $p(\boldsymbol{F}_c)$ and $p(\boldsymbol{F}_w)$. In this study, GMM was applied to fit feature vectors from correct and wrong classifications. The GMM is a versatile and efficient probabilistic model that can build any complicated probability distribution function [212]. Therefore, the two probability distribution functions can be expressed as follows:

$$\begin{aligned} p(\boldsymbol{F}|C_1) = p(\boldsymbol{F}_c) &= \sum_{k=1}^{K} \lambda_k \mathcal{N}(\boldsymbol{F}_c|\boldsymbol{\theta}_k) \\ p(\boldsymbol{F}|C_0) = p(\boldsymbol{F}_w) &= \sum_{k=1}^{K} \eta_k \mathcal{N}(\boldsymbol{F}_w|\boldsymbol{\psi}_k) \end{aligned} \tag{4.4}$$

where K is the number of mixed components. The $\lambda_k \in [0,1]$ and $\eta_k \in [0,1]$ are the mixture component weights for the k-th component, with the constraint that $\sum_{k=1}^{K} \lambda_k = 1$ and $\sum_{k=1}^{K} \eta_k = 1$. The Gaussian density functions $\mathcal{N}(\boldsymbol{F}_c)$ and $\mathcal{N}(\boldsymbol{F}_w)$ are determined by the parameter $\boldsymbol{\theta}_k = (\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ and $\boldsymbol{\psi}_k = (\boldsymbol{\nu}_k, \boldsymbol{\Gamma}_k)$, where $\boldsymbol{\mu}_k$ and $\boldsymbol{\nu}_k$ refer to the mean, while $\boldsymbol{\Sigma}_k$ and $\boldsymbol{\Gamma}_k$ are the covariance matrix, respectively. The GMM parameters, namely, $\lambda_k, \eta_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \boldsymbol{\nu}_k$ and $\boldsymbol{\Gamma}_k (k = 1, 2, \ldots, K)$, were estimated by the Expectation-Maximization (EM) algorithm in this study. The EM algorithm is an iterative method for estimating parameters in statistical models [213]. Each iteration of this algorithm involves two steps: the expectation (E) step and the maximization (M) step.

Consider the case of $p(\boldsymbol{F}_c)$, assuming that there are N_{co} accurate classifications and $\boldsymbol{F}_c^t, (t = 1, 2, ..., N_{co})$ is the feature vector corresponding to t-th accurate result. A latent variable

$\gamma_k^t, (t = 1, 2, ..., N_{co}; k = 1, 2, ..., K)$ was defined, and its expression is:

$$\gamma_k^t = \begin{cases} 1, & \boldsymbol{F}_c^t \text{ is from } k\text{-th mixed component} \\ 0, & \text{otherwise} \end{cases} \tag{4.5}$$

Therefore, the complete data is $(\boldsymbol{F}_c^t, \gamma_1^t, \gamma_2^t, ..., \gamma_K^t)$.

E step is to determine the Q function, which is the expectation of the log-likelihood function for complete data:

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(s)}) = E[\log p(\boldsymbol{F}_c, \boldsymbol{\gamma}|\boldsymbol{\theta})|\boldsymbol{F}_c, \boldsymbol{\theta}^{(s)}] \tag{4.6}$$

$\boldsymbol{\theta}^{(s)}$ represents the parameters obtained by the s-th iteration.

M step is to find the model parameter corresponding to the maximum value of the Q function:

$$\boldsymbol{\theta}^{(s+1)} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(s)}) \tag{4.7}$$

The updated model parameters $\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \lambda_k, (k = 1, 2, ..., K)$ are [214]:

$$\boldsymbol{\mu}_k^{(s+1)} = \frac{\sum_{t=1}^{N_{co}} \hat{\gamma}_k^t \boldsymbol{F}_c^t}{\sum_{t=1}^{N_{co}} \hat{\gamma}_k^t} \tag{4.8}$$

$$\boldsymbol{\Sigma}_k^{(s+1)} = \frac{\sum_{t=1}^{N_{co}} \hat{\gamma}_k^t (\boldsymbol{F}_c^t - \boldsymbol{\mu}_k^{(s+1)})(\boldsymbol{F}_c^t - \boldsymbol{\mu}_k^{(s+1)})^{\mathsf{T}}}{\sum_{t=1}^{N_{co}} \hat{\gamma}_k^t} \tag{4.9}$$

$$\lambda_k^{(s+1)} = \frac{1}{N_{co}} \sum_{t=1}^{N_{co}} \hat{\gamma}_k^t \tag{4.10}$$

where $\hat{\gamma}_k^t$ is the probability that t-th feature vector $\boldsymbol{F}_c^t$ belongs to k-th mixed component. $\hat{\gamma}_k^t, (t = 1, 2, ..., N_{co}; k = 1, 2, ..., K)$ can be calculated via the following equation:

$$\hat{\gamma}_k^t = E(\gamma_k^t|\boldsymbol{F}_c, \boldsymbol{\theta}) = \frac{\lambda_k \mathcal{N}(\boldsymbol{F}_c^t|\boldsymbol{\theta}_k)}{\sum_{k=1}^{K} \lambda_k \mathcal{N}(\boldsymbol{F}_c^t|\boldsymbol{\theta}_k)} \tag{4.11}$$

The iteration between the E-step and M-step continues until convergence. Finally, $p(\boldsymbol{F}_c)$, also known as $p(\boldsymbol{F}|C_1)$ can be obtained. Accordingly, the parameters of the probability density function $p(\boldsymbol{F}|C_0)$ can also be calculated by the EM iterations. The distinction is that the $\boldsymbol{F}$ here refers to the feature vector $\boldsymbol{F}_w$ associated with the wrong classifications.

The prior probabilities of the correct and wrong classifications can be formulated as follows:

$$\begin{aligned} P(C_1) &= \frac{N_{co}}{N_{co}+N_{wr}} \\ P(C_0) &= \frac{N_{wr}}{N_{co}+N_{wr}} \end{aligned} \tag{4.12}$$

where N_{wr} indicates the number of wrong classification results. The target recognition method is then trained using (N_b-1) blocks of SSVEP signals, and the trained model is tested using the left-out block. According to the newly obtained feature vector $\hat{\boldsymbol{F}} \in \mathbb{R}^{(N_f-1)}$, Bayesian inference is used to calculate the posterior probabilities of being a correct classification $P(C_1|\hat{\boldsymbol{F}})$ and a wrong classification $P(C_0|\hat{\boldsymbol{F}})$:

$$\begin{aligned} P(C_1|\hat{\boldsymbol{F}}) &= \frac{p(\hat{\boldsymbol{F}}|C_1)P(C_1)}{p(\hat{\boldsymbol{F}}|C_1)P(C_1)+p(\hat{\boldsymbol{F}}|C_0)P(C_0)} \\ P(C_0|\hat{\boldsymbol{F}}) &= \frac{p(\hat{\boldsymbol{F}}|C_0)P(C_0)}{p(\hat{\boldsymbol{F}}|C_1)P(C_1)+p(\hat{\boldsymbol{F}}|C_0)P(C_0)} \end{aligned} \tag{4.13}$$

Based on (4.13), the classification confidence value (CCValue) can be defined as:

$$\text{CCValue}(\hat{\boldsymbol{F}}) = P(C_1|\hat{\boldsymbol{F}}) - P(C_0|\hat{\boldsymbol{F}}) \tag{4.14}$$

Decision-Making Rule

In the decision-making module, the CCValue needs to be compared with a threshold α. The classification result should be accepted if the CCValue is greater than α. Otherwise, this module should reject the classification result. Therefore, the decision-making rule can

be written as:

$$D_{final}(\hat{\boldsymbol{F}}) = \begin{cases} \text{Accept}, & \text{if CCValue}(\hat{\boldsymbol{F}}) > \alpha \\ \text{Reject}, & \text{if CCValue}(\hat{\boldsymbol{F}}) \leq \alpha \end{cases} \tag{4.15}$$

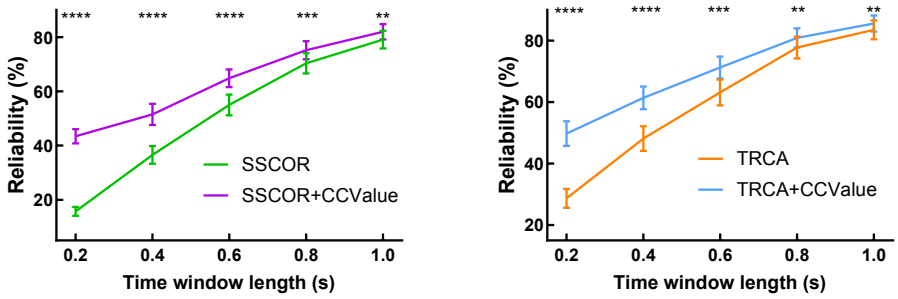
As shown in (4.15), $D_{final}(\hat{\boldsymbol{F}})$ works as a binary classifier. The grid-search method was used to determine α via $(N_b - 1)$ blocks of training data, with the remaining block reserved for testing. The range of α is specified as [-1, 1] according to (4.14). An exhaustive search is performed on the threshold values of the method with an interval of 0.1. In the search process, leave-one-block-out cross-validation was employed. Finally, the value that provides the highest average classification reliability across subjects was determined as α. This trained threshold value was subsequently applied to the untouched test data.

TRCA/SSCOR's classification result will be compared with the label of this classification. The classification results will be given a new label, i.e., "correct" or "wrong", which represents the ground truth. It is a gold standard that can be used to compare and evaluate the proposed method's results. If the proposed detection method can accept the "correct" classification or reject the "wrong" classification successfully, it means that the proposed method is effective.

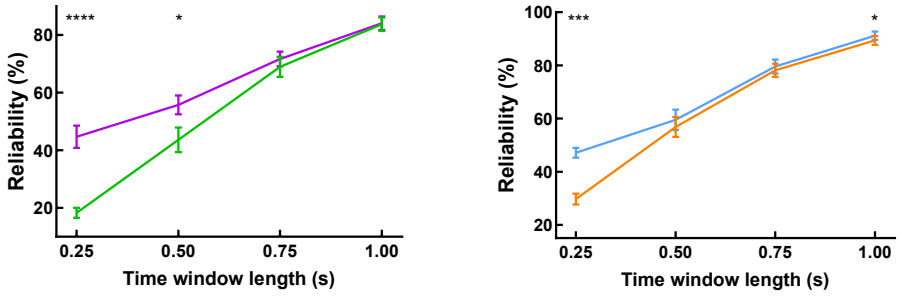
The details of the proposed Bayesian-based classification confidence estimation method are shown in Fig. 4.2. This framework aims to reduce the number of low-confidence results and thus improve recognition reliability.

4.4 Results

In this section, the proposed Bayesian-based classification confidence estimation method was applied to a 40-target benchmark dataset [23] as well as a 12-target self-collected dataset. TRCA+CCValue, SSCOR+CCValue, TRCA, and SSCOR are compared extensively. The number of channels and training blocks were set to nine, and five for benchmark dataset and nine, and four for Dataset I, respectively. The two datasets have different numbers of training blocks because of their different sizes. The selections of these hyperparameters



(a) benchmark dataset



(b) Dataset I

Figure 4.3: Average recognition reliability across subjects of various methods (i.e., SSCOR, TRCA, SSCOR+CCValue, and TRCA+CCValue) using different time windows (TWs) on (a) benchmark dataset and (b) Dataset I. The error bars represent SEM, $\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$ where $\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}}$. x_i is the classification reliability of i-th subject, $\overline{x}$ is the mean of samples, and n is the number of subjects. The asterisks indicate a significant difference between the two methods obtained by paired t-test analysis ($*$: p<0.05, $**$: p<0.01, $***$: p<0.001, $****$: p<0.0001).

Table 4.2: Reliability comparison between four methods in benchmark dataset

Methods	Averaged Recognition Reliability ± SEM (%)				
	0.2 s	0.4s	0.6 s	0.8 s	1 s
CCA	3.76 ± 0.21	9.1 ± 0.8	17.8 ± 2.1	30 ± 3	41 ± 4
Msetcca	8.0 ± 1.2	19 ± 3	31 ± 4	46 ± 5	55 ± 5
SSCOR	15.7 ± 1.6	37 ± 3	55 ± 4	70 ± 4	79 ± 3
TRCA	29 ± 3	48 ± 4	63 ± 4	79 ± 4	84 ± 3
SSCOR+CCValue	43.5 ± 2.6	52 ± 4	65 ± 3	75 ± 3	82.0 ± 2.8
TRCA+CCValue	45 ± 4	61 ± 4	71 ± 4	81 ± 3	85.6 ± 2.7
P-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

were made to ensure that the model had access to all the available information and to facilitate the training process for each dataset. The number of Gaussian mixture components was set to two. The optimal number of components in the GMM was selected using the Akaike information criterion (AIC), which provides a trade-off between the goodness of fit of the model and its complexity. The effects of parameters, such as the number of channels, training blocks and correlation coefficients on recognition performance were further investigated.

4.4.1 Performance Evaluation

Table. 4.1 introduced four measures: true rejection (TR), false acceptance (FA), false rejection (FR), and true acceptance (TA). As indicated in the Table, the wrong classification results of TRCA/SSCOR can be divided into TR and FA. The correct results of TRCA/SSCOR can be divided into FR and TA. The significance of the proposed method is that low-confidence decisions can be detected and rejected so the system can be more robust and reliable. Therefore, the accuracy of the confusion matrix (i.e., Table. 4.1) [215], also expressed as the recognition reliability (%) of the proposed method, can be defined as follows:

$$\text{Reliability} = \frac{TA + TR}{TA + FA + TR + FR} \times 100\% \tag{4.16}$$

Fig. 4.3 shows the average classification reliability of SSCOR, TRCA, SSCOR+CCValue, and TRCA+CCValue on (a) benchmark dataset and (b) Dataset I. The sampling rates are different in the two datasets, so different data lengths were used to keep the number of samples without decimals. To depict the improvement more intuitively, a pairwise comparison was performed between SSCOR and SSCOR+CCValue, as well as TRCA and TRCA+CCValue. The proposed method can attain higher reliability across a wide range of data lengths. Specifically, SSCOR+CCValue improved SSCOR by 2.90% $\sim$ 27.74% and TRCA+CCValue increased TRCA by 2.04% $\sim$ 21.07% in Dataset I. The classification reliability of SSCOR+CCValue is greater than that of SSCOR by 0.30% $\sim$ 26.37% in Dataset I. Similarly, TRCA+CCValue improved TRCA by 1.37% $\sim$ 17.42%. The paired t-test was conducted to explore the similarity of reliability between the basic recognition method and

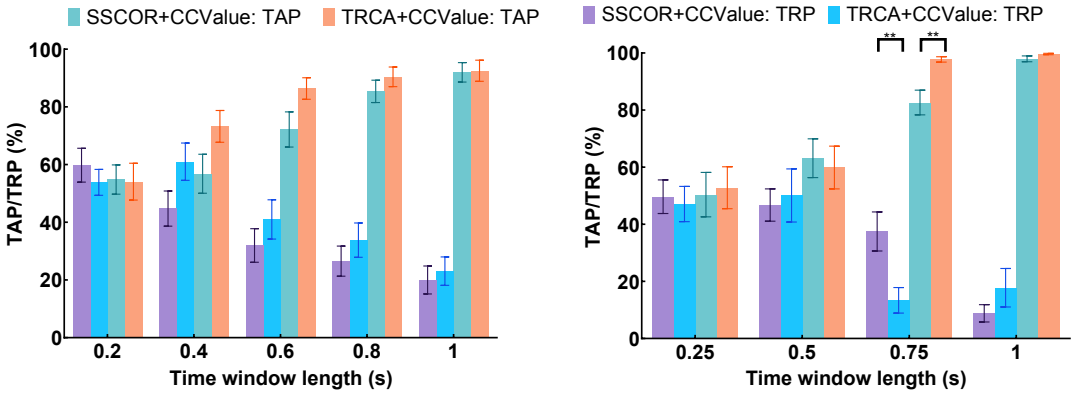
Table 4.3: Reliability comparison between four methods in Dataset I

Methods	Averaged Recognition Reliability ± SEM (%)			
	0.25 s	0.5 s	0.75 s	1 s
CCA	13.2 ± 1.2	29 ± 4	49 ± 5	72 ± 5
Msetcca	10.5 ± 1.8	20 ± 4	33 ± 4	51.7 ± 2.4
SSCOR	18.3 ± 1.8	44 ± 4	69 ± 4	83.8 ± 2.4
TRCA	29.7 ± 2.1	57 ± 4	78.2 ± 2.5	89.4 ± 1.7
SSCOR+CCValue	45 ± 4	56 ± 3	71.7 ± 2.6	84.1 ± 2.3
TRCA+CCValue	47.1 ± 1.8	60 ± 4	79.6 ± 2.6	91.2 ± 1.6
P-value	<0.0001	<0.0001	<0.0001	<0.0001

the corresponding proposed method. Statistical analysis shows that the reliability of SSCOR is significantly different from that of SSCOR+CCValue for almost all data lengths. This conclusion also applies to TRCA and TRCA+CCValue.

Table. 4.2 and Table. 4.3 provides the intuitional numerical results for comparing methods more clearly [216]. As shown in Tables, TRCA+CCValue always achieves the best performance with various data lengths in each dataset. SSCOR consistently performs worse than TRCA, whereas the performance of SSCOR+CCValue was improved after accounting for classification confidence estimation, and finally, SSCOR+CCValue outperforms TRCA at some TWs. Two popular recognition methods, i.e., CCA and Msetcca, were included for comparison. It is obvious that the proposed method achieves much higher recognition reliability than the two methods. A One-way repeated-measures ANOVA was conducted to investigate the similarity of classification reliability among these methods. The P-value is always < 0.0001, indicating statistically significant differences between the reliability of these methods at each TW.

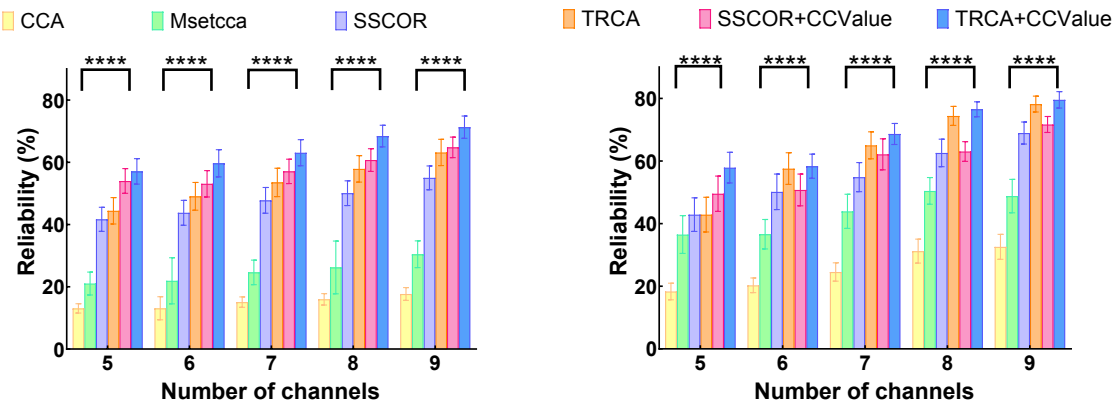
The proposed method enhances recognition performance by accepting highly-trustworthy results and rejecting unconfident ones. Therefore, the method was further assessed regarding two other indicators, i.e., the true accept proportion (TAP) and the true reject proportion (TRP). TAP is defined as the proportion of correct target identification method decisions to be accepted by $D_{final}(\hat{\boldsymbol{F}})$. TRP is defined as the proportion of wrong decisions rejected by the proposed method. Therefore, TRP indicates the rejection efficiency, whereas TAP



(a) benchmark dataset (b) Dataset I

Figure 4.4: TAP and TRP of SSCOR+CCValue and TRCA+CCValue on (a) benchmark dataset and (b) Dataset I with different data lengths. The error bars represent SEM. The asterisks indicate a significant difference between methods obtained by t-test analysis.

relates to the cost [208]. Fig. 4.4 displays the TAP and TRP of SSCOR+CCValue and TRCA+CCValue on (a) benchmark dataset and (b) Dataset I using different data lengths. In Dataset I, with increasing data lengths, SSCOR+CCValue's TAP increases from 54.81% to 91.98%, while its TRP decreases from 53.84% to 23.06%. For TRCA+CCValue, TAP rises from 54.08% to 92.53%, and its TRP changes from 59.79% to 20.01%. Dataset I exhibits similar results as well. Similarly, in Dataset I, SSCOR+CCValue's TAP increases from 50.39% to 97.96%, while its TRP changes from 49.66% to 8.82% with longer TWs. For TRCA+CCValue, TAP rises from 52.80% to 99.64%, and its TRP changes from 47.10% to 17.78%. The underlying reason is that SSCOR/TRCA provides more correct classification results as the TW increases. So SSCOR+CCValue/TRCA+CCValue is more inclined to accept the results of SSCOR/TRCA. It is worth noting that although TRP has dropped, the False Acceptance in TABLE. 4.1 generally did not increase due to a decrease in the number of wrong classifications from SSCOR/TRCA. The t-test was used to perform statistical analysis between TAP or TRP of different methods. The result shows no significant difference in almost all data lengths. It indicates that the proposed method has similar effectiveness for TRCA and SSCOR on datasets of different scales.



(a) benchmark dataset

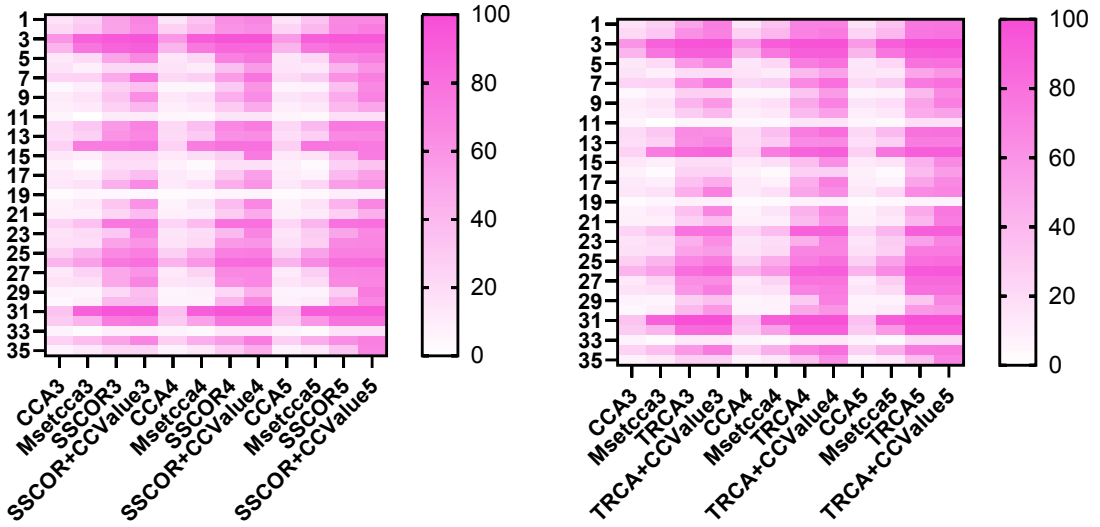
(b) Dataset I

Figure 4.5: Barchart of the classification reliability of six methods with different numbers of electrodes on (a) benchmark dataset and (b) Dataset I. The error bars represent SEM. The asterisks indicate significant differences between the four methods obtained by one-way repeated-measures ANOVA.

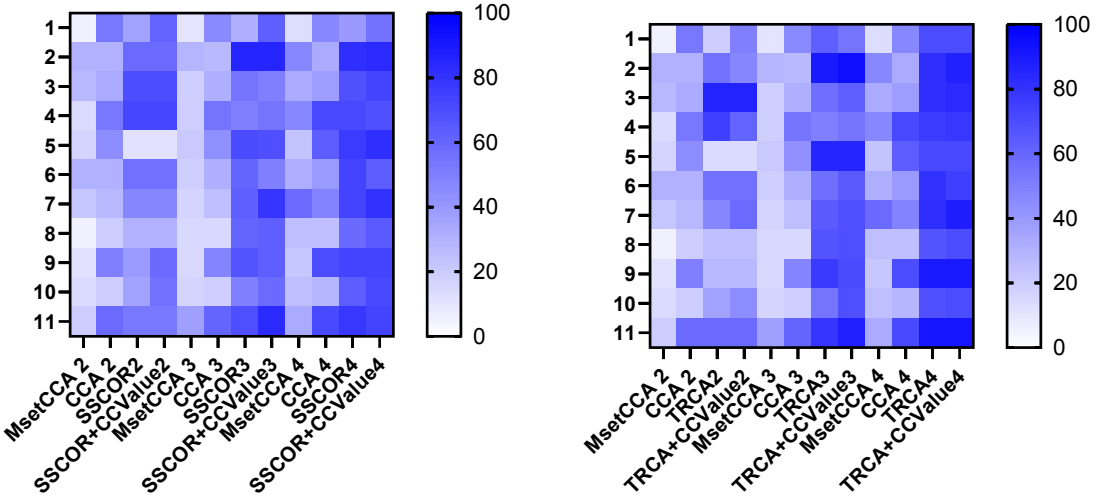
4.4.2 The Influence of Parameters

The Number of Channels

Fig. 4.5 shows the average classification reliability rate of four methods with different numbers of electrodes using 0.6 s-long data on (a) benchmark dataset and 0.75 s-long data on (b) Dataset I. The number of training blocks is set to five for benchmark dataset and four for Dataset I, respectively. Generally, the performance of each method improved as the number of electrodes increased. For N_c = 5, 6, 7, 8, and 9, it is obvious that the proposed SSCOR+CCValue always outperforms SSCOR. TRCA+CCValue shows higher recognition reliability compared with TRCA. Meanwhile, these four methods all achieve better performance than CCA and Msetcca. A one-way repeated-measures ANOVA showed significant differences between the six methods at each TW on two datasets. The results in Fig. 4.5 demonstrate that, to some extent, this method is superior to some existing advanced methods, irrespective of the number of electrodes. Specifically, SSCOR+CCValue improved SSCOR by 9.31% $\sim$ 12.32% and TRCA+CCValue increased TRCA by 8.12% $\sim$ 12.65% in Dataset I. The classification reliability of SSCOR+CCValue is greater than that of SSCOR by 0.45% $\sim$ 7.27% in Dataset I. Similarly, TRCA+CCValue improved TRCA by 0.76% $\sim$ 15.00%.



(a) benchmark dataset



(b) Dataset I

Figure 4.6: Heat maps of the classification reliability of four methods under different numbers of training blocks on (a) benchmark dataset and (b) Dataset I.

Table 4.4: Reliability comparison between SSCOR and SSCOR+CCValue with different numbers of training blocks

Methods	Reliability with different number of training blocks					
	benchmark dataset			Dataset I		
	3	4	5	2	3	4
SSCOR	45.23	51.01	55.00	46.21	60.04	68.94
SSCOR+CCValue	55.36	62.07	64.80	51.94	65.15	71.67
P-value	<0.0001	<0.0001	<0.0001	0.0843	0.1848	0.2255

The Number of Training Blocks

It also investigated how the number of training blocks affects the classification reliability of six different methods. The heat map is a valuable data visualization tool for displaying an indicator in color in two dimensions. It offers a method for understanding numerical numbers visually. The heat maps in Fig. 4.6 show the reliability comparison between SSCOR and SSCOR+CCValue, as well as between TRCA and TRCA+CCValue, on two datasets using 0.6 s or 0.75 s data length. For the sake of comparison, the performance of two other algorithms, namely CCA and Msetcca, were also included. In a heat map, the x-axis indicates recognition methods with varying numbers of training blocks, and the y-axis represents the index of the subject. The range of the number of training blocks was $[3, 5]$ for benchmark dataset and $[2, 4]$ for Dataset I. The shade of color indicates the level of classification reliability. The darkest color is always displayed at its maximum value. As demonstrated in Fig. 4.6, with varying numbers of training blocks, the color squares generated by SSCOR+CCValue and TRCA+CCValue are generally more profound than those created by SSCOR and TRCA, and notably darker than those generated by CCA and Msetcca. This indicates that the proposed method produces more reliable and consistent results compared to the other methods. Furthermore, the color squares generally get darker as the number of training blocks increases.

Table. 4.4 presents the numerical classification reliability of SSCOR and SSCOR+CCValue, and the corresponding paired t-test analysis results. Similarly, Table. 4.5 shows the outcome of TRCA and TRCA+CCValue. The average classification reliability of SSCOR+CCValue is higher than that of SSCOR by 10.33% across different numbers of training blocks in

Table 4.5: Reliability comparison between TRCA and TRCA+CCValue with different numbers of training blocks

Methods	Reliability with different number of training blocks					
	benchmark dataset			Dataset I		
	3	4	5	2	3	4
TRCA	49.98	57.67	63.17	45.45	67.42	78.18
TRCA+CCValue	57.54	67.64	71.29	47.98	70.83	79.55
P-value	<0.0001	<0.0001	0.0002	0.4825	0.1145	0.1698

benchmark dataset and by 8.55% in Dataset I. Meanwhile, TRCA+CCValue improved TRCA by 4.52% in benchmark dataset and by 2.44% in Dataset I. The paired t-test analysis results revealed a statistically significant difference (i.e., $P < 0.0001$) between the compared method and the proposed method with all numbers of training blocks for Dataset I. For Dataset I, although the significant difference is not as large as for Dataset I, the proposed methods still provide higher recognition reliability than SSCOR and TRCA, respectively. In conclusion, the effectiveness of the proposed method is further demonstrated by the two tables, which offer additional quantitative evidence across varying amounts of training data.

The Number of Correlation Coefficients Incorporated in the Feature Vector

In this study, it is also explored how the number of correlation coefficients used for constructing the feature vector affects the classification performance. The aforementioned performance evaluation figures were all generated by $(N_f - 1)$-dimensional feature vectors. It implies that the feature vector was constructed using N_f correlation coefficients via (4.3). In this subsection, a 40-class benchmark dataset was used to evaluate more types of coefficient numbers. The correlation coefficients were sorted in descending order. The top two, four, eight, sixteen, thirty-two, or forty values were chosen to construct the feature vector via (4.3). Here, the number of electrodes and training trials are set to be nine and five, respectively. Fig. 4.7(a) shows TRP, TAP, and classification reliability of SSCOR+CCValue for various numbers of correlation coefficients. The reliability of SSCOR (blue bars) was also incorporated into the performance comparison. The number of correlation coefficients does not affect the performance of SSCOR. Hence, the corresponding reliability remains constant

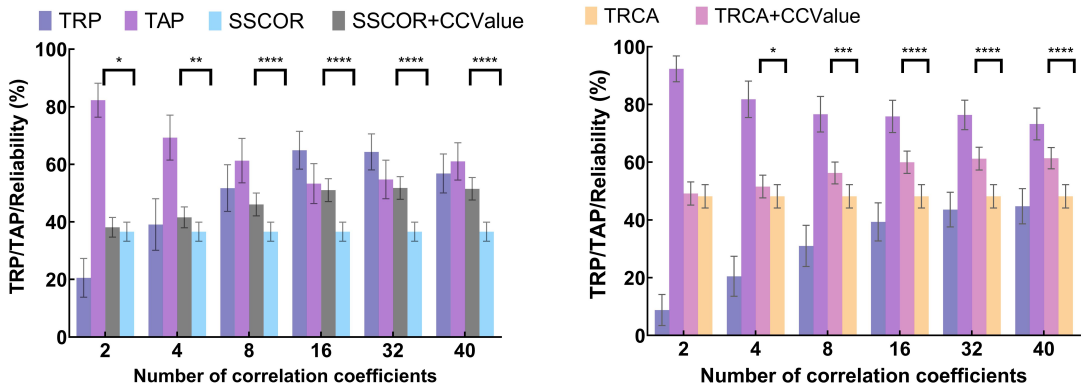
(i.e., SSCOR: 36.58%). Similarly, Fig. 4.7(b) shows the evaluation results of TRCA and TRCA+CCValue. The reliability of TRCA is represented by the orange bars at 48.18%.

With an increasing number of correlation coefficients, TRP generally climbs fast and then lowers slightly. TAP gradually decreased and then increased. TRP and TAP are both critical indicators for a classification confidence evaluation model. TRP indicates the model's rejection effectiveness, whereas TAP relates to the cost. As a result, it is preferable to keep them both at a relatively high level. TAPs have achieved the largest values for two coefficients in Fig. 4.7(a) and Fig. 4.7(b), but TRPs reached relatively low values. Therefore, two is not an ideal number of coefficients for this dataset. It is worth mentioning that TRP and TAP for forty coefficients were both within a satisfactory range, and their difference was not as large as that of other coefficients. Moreover, the reliability also reached superior values for forty coefficients (i.e., TRCA+CCValue: 61.39% and SSCOR+CCValue: 51.48% with 0.4 s-long data). The TRCA+CCValue and SSCOR+CCValue provide consistently higher reliability than TRCA and SSCOR, regardless of the number of correlation coefficients. TRCA+CCValue improves TRCA by 3.84% $\sim$ 15.84%, and SSCOR+CCValue increases SSCOR by 1.42% $\sim$ 12.71%. Moreover, paired t-test analysis showed that statistical differences between the compared algorithms become more significant as the number of coefficients increases.

4.5 Discussion

4.5.1 Performance of SSCOR+CCValue and TRCA+CCValue

Almost all existing advanced SSVEP recognition methods determine the signal triggered by which stimulus via the largest correlation coefficient, such as CCA [68], MsetCCA [67], SSCOR [71], and TRCA [1]. It can easily lead to erroneous results when the maximum coefficient is slightly larger than the other values. In this study, a classification confidence estimation method based on Bayesian theory was proposed to improve the SSVEP recognition performance. The feature vector was constructed by differences between the largest coefficient and the remaining values. This kind of design can make full use of all the co-



(a) SSCOR and SSCOR+CCValue

(b) TRCA and TRCA+CCValue

Figure 4.7: Barchart of the TRP, TAP and classification reliability of SSCOR+CCValue and TRCA+CCValue with different numbers of correlation coefficients. The error bars represent SEM. The asterisks indicate significant differences between methods obtained by paired t-test analysis. The reliability of SSCOR and TRCA were used as a comparison.

efficient information. As a consequence, the proposed method can accept high-confidence classification results while rejecting results with low confidence. As shown in Fig. 4.3(a), TRCA+CCValue and SSCOR+CCValue obtained the highest reliability of 85.57% and 81.98% for the data length of 1 s. TRCA+CCValue and SSCOR+CCValue both improved the performance of the basic target recognition methods.

In Fig. 4.3, the performance of TRCA+CCValue is slightly better than that of TRCA at 1 s TW. Besides, a similar situation is reflected in SSCOR+CCValue and SSCOR. The underlying reason is that long-length signals generally contain more EEG information and are thus more likely to lead to correct classification results. The proposed methods can accept results with high confidence and reject results with low confidence. Therefore, the proposed method accepted more reliable results and achieved reliability comparable to basic methods at 1 s TW.

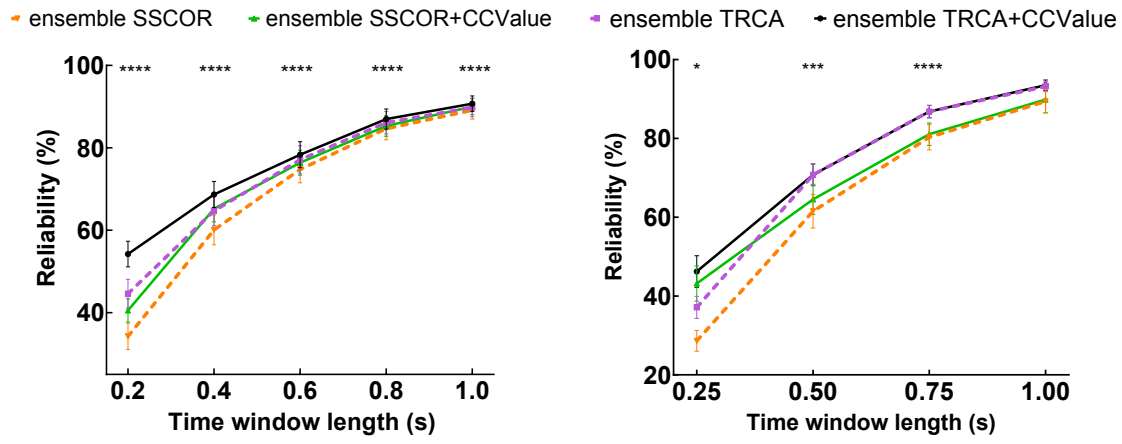
Although the experiment was conducted in a relatively quiet environment and the subjects were typically requested to avoid movements during signal recording, complete elimination of environmental and body noises is difficult to achieve. Noise is usually present due to a variety of factors, including muscle movements, eye blinks, and external sources such as traffic or other environmental factors. In this study, the fundamental target recogni-

tion methods employed for feature extraction are TRCA and SSCOR. These two methods can reduce background EEG activities in different ways [1, 71]. For example, TRCA is a spatial filtering method, in which weight coefficients are optimized to maximize inter-trial covariance of brain activities. It can be used for removing background EEG activities from scalp recordings [1]. TRCA+CCValue and SSCOR+CCValue can improve classification performance via confidence estimation and take advantage of TRCA and SSCOR to decrease background noises. To evaluate the performance of the proposed method, CCA and Msetcca were used in this study for extensive comparison, and the numerical results of the six methods are shown in Table. 4.2 and Table. 4.3. Additional experiments were also conducted to compare the performance of the six methods under various parameters, such as the number of electrodes and the number of training blocks, as shown in Fig. 4.5 and Fig. 4.6. The evaluation results indicate that the proposed method provides better recognition performance than the other four methods across a range of different parameter settings. For example, SSCOR+CCValue improved CCA and Msetcca by 40.00% $\sim$ 47.10% and 31.17% $\sim$ 34.48% with the different number of channels. For different number of training blocks, TRCA+CCValue increased them by 39.88% $\sim$ 53.59% and 31.34% $\sim$ 40.84%, respectively.

In the presented method, leave-one-block-out cross-validation was performed in the experiments. The detailed process was shown in Fig. 4.2. Cross-validation is a widely used technique in machine learning and statistical modeling to estimate the performance of a model and prevent over-fitting. Cross-validation provides an accurate evaluation of the performance of the proposed method because it uses all the available data for both training and testing. Therefore, it helps improve the reliability and generalization of the experimental results.

4.5.2 Ensemble-based Methods Comparison

In the previous sections, the effectiveness and superiority of the proposed method were demonstrated by comparing TRCA+CCValue and SSCOR+CCValue with the basic target recognition methods. In this part, the performance comparison of ensemble-based methods was carried out. Specifically, the target recognition method was enhanced by utilizing an



(a) benchmark dataset (b) Dataset I

Figure 4.8: Comparison of average recognition reliability among ensemble methods on (a) benchmark dataset and (b) Dataset I. The asterisks indicate significant differences between the four methods obtained by one-way repeated-measures ANOVA.

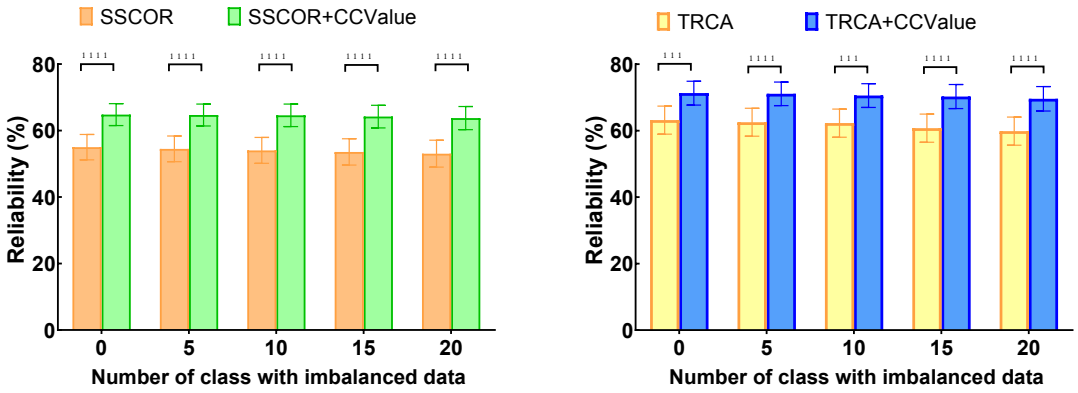
ensemble approach, in which N_f spatial filters were concatenated to create an ensemble spatial filter $\boldsymbol{W} \in \mathbb{R}^{N_c \times N_f}$:

$$\boldsymbol{W} = [\boldsymbol{w}_1, \boldsymbol{w}_1, ..., \boldsymbol{w}_{N_f}] \tag{4.17}$$

The correlation coefficient in (3.4) can be re-defined as follows:

$$r_i = \rho(\widetilde{\boldsymbol{X}}^{\mathsf{T}}\boldsymbol{W}, \overline{\boldsymbol{\chi}}_i^{\mathsf{T}}\boldsymbol{W}),\ i = 1, 2, ..., N_f \tag{4.18}$$

The feature extraction, classification confidence evaluation, and decision-making steps are the same as described in the previous section. Fig. 4.8 shows the classification reliability comparison between several ensemble-based methods on (a) benchmark dataset and (b) Dataset I. As shown by the black line and the purple dotted line, the ensemble TRCA+CCValue achieved higher reliability than the ensemble TRCA, with almost TWs on two datasets. The ensemble SSCOR+CCValue also exhibits a superior performance than SSCOR at all TWs. For example, ensemble TRCA+CCValue improved ensemble TRCA by 3.96%, and ensemble SSCOR+CCValue increased ensemble SSCOR by 5.13% with 0.4s data length in Dataset I. Similarly, the classification reliability of SSCOR+CCValue is greater than that of SSCOR by 14.55%, and TRCA+CCValue improved TRCA by 9.09% with 0.25s data



(a) SSCOR and SSCOR+CCValue

(b) TRCA and TRCA+CCValue

Figure 4.9: Barchart of the classification reliability of SSCOR, TRCA, SSCOR+CCValue and TRCA+CCValue with different numbers of classes considering imbalanced data. The error bars represent SEM. The asterisks indicate significant differences between methods obtained by paired t-test analysis.

length in Dataset I. A one-way repeated measures ANOVA revealed a statistically significant difference between the compared methods with various data lengths. As a result, the proposed method can improve the performance of both basic and ensemble-based SSVEP detection methods.

4.5.3 Feature Vector Construction

Recently, some studies have also focused on estimating classification confidence based on correlation coefficients for SSVEP-based BCI. These works usually use the largest and the second-largest values or their difference, such as [106–108]. In this study, N_f correlation coefficients were incorporated, and then a (N_f-1)-dimensional feature vector was formed by calculating the differences between the maximum value and the other values. The higher-dimensional features are beneficial to improving SSVEP detection, which was confirmed in Fig. 4.7. The N_f of the benchmark dataset is forty. Compared with other numbers of correlation coefficients, TAP, TRP, and classification reliability generated by the feature vector with forty correlation coefficients achieve high values. For example, TRCA+CCValue reached the highest reliability of 61.39%, and SSCOR+CCValue reached the reliability of 51.49% (highest value: 51.79% with thirty-two correlation coefficients) with 0.4 s TW. For those cases with similar reliability, the gap between TAP and TRP provided by the proposed

method is relatively smaller. For example, the gap is 4.19% for forty coefficients but 9.56% for thirty-two coefficients for SSCOR+CCValue. Therefore, it indicates that the proposed method can achieve high classification reliability while maintaining a better balance between the model's rejection efficiency (TRP) and the cost (TAP).

4.5.4 Data Imbalance

Data imbalance is a common issue in real-world datasets, and it occurs when the distribution of classes in a dataset is uneven. Therefore, it is important to evaluate the effectiveness of the proposed method on unbalanced datasets to further validate its reliability in real SSVEP-based BCI systems. In Fig. 4.9, the classification reliability of four methods is shown under different numbers of classes with imbalanced data. For instance, when the x-axis is five, it means that five classes are randomly selected with insufficient training data (i.e., four training blocks), while the other thirty classes have sufficient training data (i.e., five training blocks). The evaluation results indicate that the proposed method achieves consistently better performance. The paired t-test was used to perform statistical analysis of the recognition performance of different methods. Statistical analysis shows that the reliability of SSCOR is significantly different from that of SSCOR+CCValue regardless of the number of imbalanced classes. The same conclusion applies to TRCA and TRCA+CCValue. In addition, the performance of the proposed method does not show much difference between datasets with many imbalanced classes and those without any imbalanced classes. For example, the recognition reliability of TRCA+CCValue and SSCOR+CCValue are 69.59% and 63.76% when tested on a dataset with twenty imbalanced classes, while on a dataset with zero imbalanced class, the recognition reliability of TRCA+CCValue is 71.29% and SSCOR+CCValue is 64.80%. This suggests that TRCA+CCValue and SSCOR+CCValue are robust to the number of imbalanced classes in the dataset, indicating their potential for handling imbalanced datasets in practical situations. Additionally, Fig. 4.6, TABLE. 4.4, and TABLE. 4.5 in Section 4.3.2 show the experimental evaluation results after balancing the dataset. It involves adjusting the class distribution so that each class has an equal number of examples.

4.6 Conclusion

In this study, a Bayesian-based classification confidence estimation method was proposed for enhancing the SSVEP recognition performance. The differences between the largest correlation coefficient and the other values were used to define the feature vector. The probability density functions of feature vectors given correct and wrong classifications were then estimated using the GMM model. In the test process, the posterior probabilities of an accurate and wrong recognition can be calculated using Bayesian inference with the newly obtained feature vector. The CCValue, the difference between two posterior probabilities, was applied to evaluate the confidence of the classification result. Eventually, the decision-making process can determine whether to accept trustworthy results or reject unconfident results. This method was evaluated on a publicly available benchmark dataset and a self-collected dataset. The experimental results demonstrated the effectiveness and feasibility of the proposed method in the SSVEP-based BCIs.

Chapter 5

Cross-Subject Transfer Learning for Boosting SSVEP Recognition Performance

SSVEP-based BCIs have been substantially studied in recent years due to their fast communication rate and high signal-to-noise ratio. The transfer learning is typically utilized to improve the performance of SSVEP-based BCIs with auxiliary data from the source domain. This study proposed an inter-subject transfer learning method for enhancing SSVEP recognition performance through transferred templates and transferred spatial filters. In this method, the spatial filter was trained via multiple covariance maximization to extract SSVEP-related information. The relationships between the training trial, the individual template, and the artificially constructed reference are involved in the training process. The spatial filters are applied to the above templates to form two new transferred templates, and the transferred spatial filters are obtained accordingly via the least-square regression. The contribution scores of different source subjects can be calculated based on the distance between the source subject and the target subject. Finally, a four-dimensional feature vector is constructed for SSVEP detection. To demonstrate the effectiveness of the proposed method, a publicly available dataset, and a self-collected dataset were employed for per-

formance evaluation. The extensive experimental results validated the feasibility of the proposed method for improving SSVEP detection.

5.1 Introduction

EEG-based BCIs provide humans a direct communication path between brain activities and external equipment without the need to move peripheral nerves or muscles [200, 201, 217]. SSVEP is one of the most popular paradigms in the research area of BCI due to its high SNR, reliability, and minimal set-up requirement [75, 86, 211, 218]. SSVEP-based BCI has been broadly employed in various applications, such as communication [218], robot [7, 150], and smart home [219].

By analyzing the information from the measured SSVEP signals, the visual stimulus that the user is gazing at can be detected, and the corresponding control command can be output accordingly [185]. In recent years, many target recognition methods have been proposed for SSVEP-based BCI systems. CCA is the most popular method to classify stimuli due to its ease of use and robustness [2, 68]. However, as a training-free method, its performance is easily influenced by interference from spontaneous brain activities. To alleviate this issue, many improved approaches have been proposed for SSVEP detection. In the direction of template optimization, to name a few, the L1-regularized multiway CCA (L1-MwayCCA) [66], MsetCCA [67], ITCCA [72] and MCM [24]. Alternatively, several spatial filtering methods have also been reported to lower the misclassification rate in SSVEP detection, such as a combination method of CCA and ITCCA [74], the SSCOR [71], and TRCA [1].

Although the performance of SSVEP-based BCI systems was significantly boosted by these templates- or spatial filter-based methods, EEG usually suffers from inter-subject variability problems [31]. Therefore, trained templates or spatial filters can only be used for a single subject and it is difficult to transmit knowledge directly across subjects. It would prevent the broad and practical usage of BCIs in our real lives. Recently, the transfer learning (TL) technique was explored in BCIs to transfer knowledge from old sessions or subjects (the source domain) to new sessions or subjects (the target domain) so that the performance

of the target domain can be boosted [120, 220]. As one of the research directions, training data is usually transferred across different domains to augment the size of calibration data for new users [122, 123]. Template-based transfer learning is also a popular study area, and several approaches are listed, such as transfer template-based canonical correlation analysis (tt-CCA) [115], adaptive combined-CCA (Adaptive-C3A) [117] and inter- and intra-subject template-based multivariate synchronization index (IIST-MSI) [118]. In these methods, the transferred template is simply generated by averaging multiple trials from source subjects, which may not contain sufficient SSVEP features. Alternatively, there are multiple BCI transfer learning studies cooperating on spatial filters to learn the common feature representations across different domains [2, 119]. Liu *et.al* [221] proposed an all-to-one method to use data from all source subjects to train TRCA-based spatial filters. Wang *et.al* [126] presented an inter-subject maximal correlation method to improve the robustness of SSVEP classification. Wong *et.al* [30] proposed a subject transfer based CCA method which utilizes the knowledge within-subject and between subjects. However, these methods rarely consider the correlation among the training data, the individual template, and the predefined sine-cosine signal simultaneously to enhance the effectiveness of the spatial filter [222].

In this study, the objective was to investigate and leverage a transfer learning architecture for enhancing recognition performance within the SSVEP-based BCI system. The main contributions of this paper are as follows: 1) a cross-subject scheme is proposed which incorporates SSVEP knowledge from the source subject to effectively strengthen the recognition performance for the target subject. 2) a powerful and informative feature vector is constructed under this scheme. The multidimensional feature vector is driven partly by the transferred spatial filter and the transferred SSVEP template from the source subject, and partly by the spatial filter of the target subject obtained by multiple covariance maximization. 3) a contribution score is introduced to each source subject by further exploring the distance between the source subject and the target subject. Validation of the performance of the proposed method was performed on a publicly available 40-class dataset [23] and a self-collected 12-class dataset. Extensive evaluations were conducted to demonstrate its effectiveness in comparison to some well-known methods. The efficiency and reliability

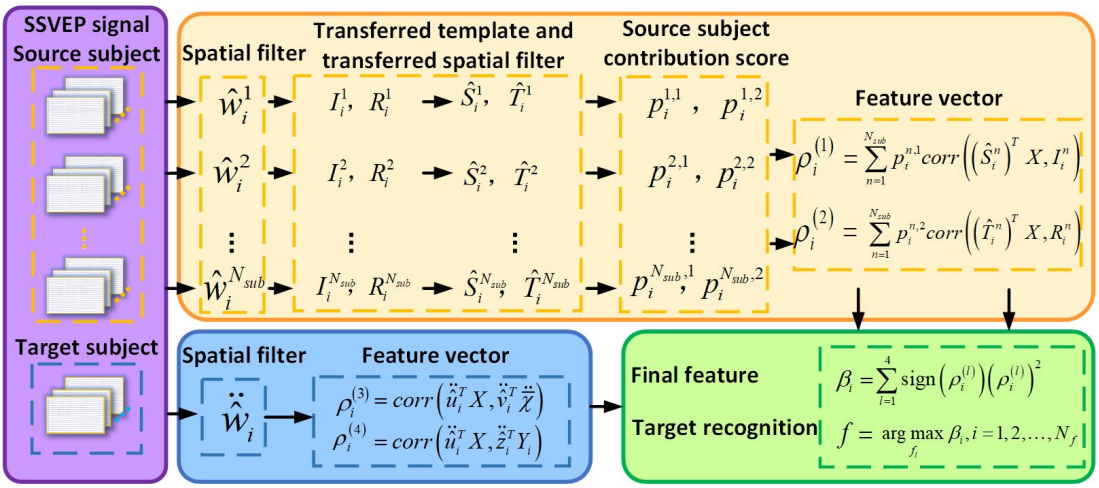


Figure 5.1: The diagram of the cross-subject transfer learning method for enhancing SSVEP detection. For i-th stimulus, the spatial filter for n-th source subject $\hat{\boldsymbol{w}}_i^n$ and for the target subject $\ddot{\hat{\boldsymbol{w}}}_i$ are firstly calculated based on the correlation maximization between any two of the three kinds of signals (training trials, the individual template, and the reference signal) as well as themselves via (5.1) - (5.13). The transferred template $\boldsymbol{I}_i^n$, $\boldsymbol{R}_i^n$ and transferred spatial filter $\hat{\boldsymbol{S}}_i^n$, $\hat{\boldsymbol{T}}_i^n$ are then be obtained via (5.14) - (5.19). The contribution score $p_i^{n,1}$, $p_i^{n,2}$ are assigned to correlation coefficients of n-th source subject via (5.22) - (5.25). Finally, four-dimensional feature vector ρ_i can be formed by (5.26) and recognition results are determined via (5.27) - (5.28).

were demonstrated with an average classification accuracy of 89.98% and 94.61% on the two datasets, respectively.

5.2 Materials

5.2.1 Datasets

In this study, the proposed method and comparing methods were evaluated on a publicly available benchmark dataset [23] and a 12-class self-collected SSVEP dataset.

5.2.2 Data Pre-processing

Due to the effect of visual latency in the human visual system, the data was extracted in $[0.14\ (0.14+d)]s$, where d refers to the data length selected for performance analysis. The data were filtered by the Chebyshev Type I IIR filter to pass signals between eight Hz and forty Hz for Dataset I.

5.3 Inter-subject Transfer Learning Method

Assume that the four-dimensional EEG signal is denoted as $\boldsymbol{\chi} \in \mathbb{R}^{N_t \times N_f \times N_c \times N_s}$, where N_t represents the number of training trials, N_f indicates the number of visual stimuli, N_c is the number of channels, and N_s is the number of samples. Hereafter, i and j refer to the index of stimulus and training trial, respectively. Therefore, the two-way tensor $\boldsymbol{\chi}_{i,j} \in \mathbb{R}^{N_c \times N_s}$ represents the individual EEG signal for the i-th stimulus and the j-th training trial. The continuous training data is denoted as $\hat{\boldsymbol{\chi}}_i = [\boldsymbol{\chi}_{i,1}, \boldsymbol{\chi}_{i,2}, ..., \boldsymbol{\chi}_{i,N_t}] \in \mathbb{R}^{N_c \times (N_t \cdot N_s)}$ which is constructed by concatenating N_t training trials. The single-trial individual template is obtained by averaging multiple training trails, i.e., $\overline{\boldsymbol{\chi}}_i = \frac{1}{N_t} \sum_{j=1}^{N_t} \boldsymbol{\chi}_{i,j} \in \mathbb{R}^{N_c \times N_s}$. SSVEP signals can also be characterized by a series of artificial sine-cosine waves, so the reference signal $\boldsymbol{Y}_i \in \mathbb{R}^{2N_h \times N_s}$ is defined as:

$$\boldsymbol{Y}_i = \begin{bmatrix} \sin(2\pi ft) \\ \cos(2\pi ft) \\ \vdots \\ \sin(2\pi N_h ft) \\ \cos(2\pi N_h ft) \end{bmatrix}, \; t = [1/F_s, 2/F_s, ..., N_s/F_s] \tag{5.1}$$

where N_h is the number of harmonics, F_s represents the sampling rate, and f is the visual stimulation frequency.

The spatial filter $\boldsymbol{w}_i = [\boldsymbol{u}_i^T, \boldsymbol{v}_i^T, \boldsymbol{z}_i^T]^T \in \mathbb{R}^{2(N_c+N_h)\times 1}$ can be computed by maximizing the inter-trial covariance, the covariance between training trials and individual template, the covariance between training trials and artificial reference, as well as the covariance between the individual template and artificial reference. Therefore, the covariance matrix $\boldsymbol{C}$ could be represented as:

$$\boldsymbol{C} = \begin{pmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} & \boldsymbol{C}_{13} \\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} & \boldsymbol{C}_{23} \\ \boldsymbol{C}_{31} & \boldsymbol{C}_{32} & \boldsymbol{C}_{33} \end{pmatrix} \tag{5.2}$$

where $\boldsymbol{C}_{11}$ is denoted as the inter-trial covariance:

$$\boldsymbol{C}_{11} = \sum_{j,h=1,i\neq h}^{N_t} cov(\boldsymbol{\chi}_{i,j}, \boldsymbol{\chi}_{i,h}) \in \mathbb{R}^{N_c \times N_c} \tag{5.3}$$

C_{12} and C_{21} refer to the covariance between the SSVEP training trials and the individual template, which can be represented as:

$$\boldsymbol{C}_{12} = \boldsymbol{C}_{21}^T = \sum_{j=1}^{N_t} cov(\boldsymbol{\chi}_{i,j}, \overline{\boldsymbol{\chi}}_i) \in \mathbb{R}^{N_c \times N_c} \tag{5.4}$$

The similarity between the training trials and artificially constructed reference is also incorporated, which can be denoted as :

$$\boldsymbol{C}_{13} = \boldsymbol{C}_{31}^T = \sum_{j=1}^{N_t} cov(\boldsymbol{\chi}_{i,j}, \boldsymbol{Y}_i) \in \mathbb{R}^{N_c \times 2N_h} \tag{5.5}$$

C_{23} and C_{32} are the covariance between the individual template and reference signal:

$$\boldsymbol{C}_{23} = \boldsymbol{C}_{32}^T = cov(\overline{\boldsymbol{\chi}}_i, \boldsymbol{Y}_i) \in \mathbb{R}^{N_c \times 2N_h} \tag{5.6}$$

In addition, C_{22} and C_{33} are denoted as:

$$\boldsymbol{C}_{22} = cov(\overline{\boldsymbol{\chi}}_i, \overline{\boldsymbol{\chi}}_i) \in \mathbb{R}^{N_c \times N_c} \tag{5.7}$$

$$\boldsymbol{C}_{33} = cov(\boldsymbol{Y}_i, \boldsymbol{Y}_i) \in \mathbb{R}^{2N_h \times 2N_h} \tag{5.8}$$

Therefore, the objective function is represented as $\boldsymbol{w}_i^T \boldsymbol{C} \boldsymbol{w}_i$.

The constraint is incorporated in above optimization problem, i.e., $\boldsymbol{w}_i^T \boldsymbol{Q} \boldsymbol{w}_i = 1$, where covariance matrix $\boldsymbol{Q}$ is denoted as follows:

$$\boldsymbol{Q} = blkdiag(\boldsymbol{Q}_1, \boldsymbol{Q}_2, \boldsymbol{Q}_3) \in \mathbb{R}^{2N_h \times 2N_h} \tag{5.9}$$

where

$$\boldsymbol{Q}_1 = cov(\hat{\boldsymbol{\chi}}_i, \hat{\boldsymbol{\chi}}_i) \in \mathbb{R}^{N_c \times N_c} \tag{5.10}$$

$$\boldsymbol{Q}_2 = cov(\overline{\boldsymbol{\chi}}_i, \overline{\boldsymbol{\chi}}_i) \in \mathbb{R}^{N_c \times N_c} \tag{5.11}$$

$$\boldsymbol{Q}_3 = cov(\boldsymbol{Y}_i, \boldsymbol{Y}_i) \in \mathbb{R}^{2N_h \times 2N_h} \tag{5.12}$$

Therefore, the constrained optimization problem can be formulated as:

$$\hat{\boldsymbol{w}}_i = \arg\max_{\boldsymbol{w}_i} \frac{\boldsymbol{w}_i^T \boldsymbol{C} \boldsymbol{w}_i}{\boldsymbol{w}_i^T \boldsymbol{Q} \boldsymbol{w}_i} \tag{5.13}$$

The spatial filter $\hat{\boldsymbol{w}}_i = [\hat{\boldsymbol{u}}_i^T, \hat{\boldsymbol{v}}_i^T, \hat{\boldsymbol{z}}_i^T]^T$ is obtained as the eigenvector of the matrix $\boldsymbol{Q}^{-1}\boldsymbol{C}$ corresponding to the largest eigenvalue. The N_f spatial filters are concatenated to make spatial filters $\widetilde{\boldsymbol{v}} = [\hat{\boldsymbol{v}}_1, \hat{\boldsymbol{v}}_2, ..., \hat{\boldsymbol{v}}_{N_f}]^T \in \mathbb{R}^{N_f \times N_c}$ and $\widetilde{\boldsymbol{z}} = [\hat{\boldsymbol{z}}_1, \hat{\boldsymbol{z}}_2, ..., \hat{\boldsymbol{z}}_{N_f}]^T \in \mathbb{R}^{N_f \times 2N_h}$. Hereafter, the variable with a right superscript n, $(n = 1, 2, ..., N_{sub})$ refers to the fact that it is provided by the n-th source subject. N_{sub} is the number of transferred source subjects. Therefore, the two kinds of transferred templates, i.e., transferred individual template $\boldsymbol{I}_i^n \in \mathbb{R}^{N_f \times N_s}$ and transferred reference template $\boldsymbol{R}_i^n \in \mathbb{R}^{N_f \times N_s}$, provided by n-th source subject can be represented as:

$$\boldsymbol{I}_i^n = \widetilde{\boldsymbol{v}} \times \overline{\boldsymbol{\chi}}_i \tag{5.14}$$

$$\boldsymbol{R}_i^n = \widetilde{\boldsymbol{z}} \times \boldsymbol{Y}_i \tag{5.15}$$

Let the variable with the double-dot superscript denote that it is provided by the target subject. The transferred spatial filters $\hat{\boldsymbol{s}}_{ij}$ and $\hat{\boldsymbol{t}}_{ij}$ for the i-th stimulus and the j-th training trial corresponding to the two kinds of transferred templates can be calculated by solving the following formula:

$$\hat{\boldsymbol{s}}_{ij}^n = \arg\min_{\boldsymbol{s}_{ij}} \|\boldsymbol{I}_i^n - \boldsymbol{s}_{ij}^T \ddot{\boldsymbol{\chi}}_{ij}\|_2^2 \tag{5.16}$$

$$\hat{\boldsymbol{t}}_{ij}^n = \arg\min_{\boldsymbol{t}_{ij}} \|\boldsymbol{R}_i^n - \boldsymbol{t}_{ij}^T \ddot{\boldsymbol{\chi}}_{ij}\|_2^2 \tag{5.17}$$

$\hat{\boldsymbol{s}}_{ij}^n$ and $\hat{\boldsymbol{t}}_{ij}^n$ can be estimated via least-squares regression [119]:

$$\hat{\boldsymbol{s}}_{ij}^n = (\ddot{\boldsymbol{\chi}}_{ij} \ddot{\boldsymbol{\chi}}_{ij}^T)^{-1} \ddot{\boldsymbol{\chi}}_{ij} \boldsymbol{I}_i^{nT} \tag{5.18}$$

$$\hat{\boldsymbol{t}}_{ij}^n = (\ddot{\boldsymbol{\chi}}_{ij} \ddot{\boldsymbol{\chi}}_{ij}^T)^{-1} \ddot{\boldsymbol{\chi}}_{ij} \boldsymbol{R}_i^{nT} \tag{5.19}$$

The final transferred spatial filters $\hat{\boldsymbol{S}}_i^n$ and $\hat{\boldsymbol{T}}_i^n$ provided by n-th source subject can be obtained by averaging all training trials. Suppose that $\boldsymbol{X}$ refers to the test data from the target subject, the two correlation coefficients can be calculated as:

$$r_i^{n,1} = corr((\hat{\boldsymbol{S}}_i^n)^T \boldsymbol{X}, \boldsymbol{I}_i^n) \tag{5.20}$$

$$r_i^{n,2} = corr((\hat{\boldsymbol{T}}_i^n)^T \boldsymbol{X}, \boldsymbol{R}_i^n) \tag{5.21}$$

According to the distance between the source subject and the target subject, weights will be assigned to correlation coefficients corresponding to different source subjects. For i-th stimulus and n-th source subject, the distance is measured by the correlation coefficient between the spatially filtered training trials of the target subject and the corresponding transferred template:

$$d_i^{n,1} = \sum_{j=1}^{N_t} corr((\hat{\boldsymbol{S}}_i^n)^T \ddot{\boldsymbol{\chi}}_{ij}, \boldsymbol{I}_i^n) \tag{5.22}$$

$$d_i^{n,2} = \sum_{j=1}^{N_t} corr((\hat{\boldsymbol{T}}_i^n)^T \ddot{\boldsymbol{\chi}}_{ij}, \boldsymbol{R}_i^n) \tag{5.23}$$

Therefore, the weights also called contribution scores are represented as:

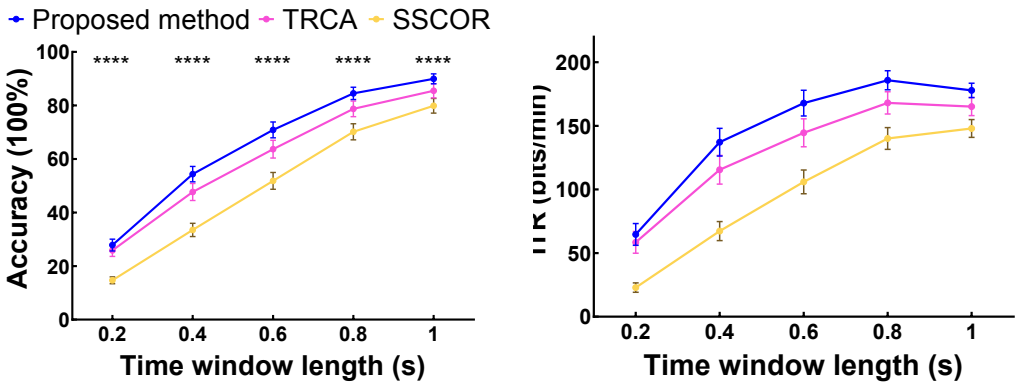
$$p_i^{n,1} = \frac{d_i^{n,1}}{\sum_{h=1}^{N_{sub}} d_i^{h,1}} \tag{5.24}$$

$$p_i^{n,2} = \frac{d_i^{n,2}}{\sum_{h=1}^{N_{sub}} d_i^{h,2}} \tag{5.25}$$

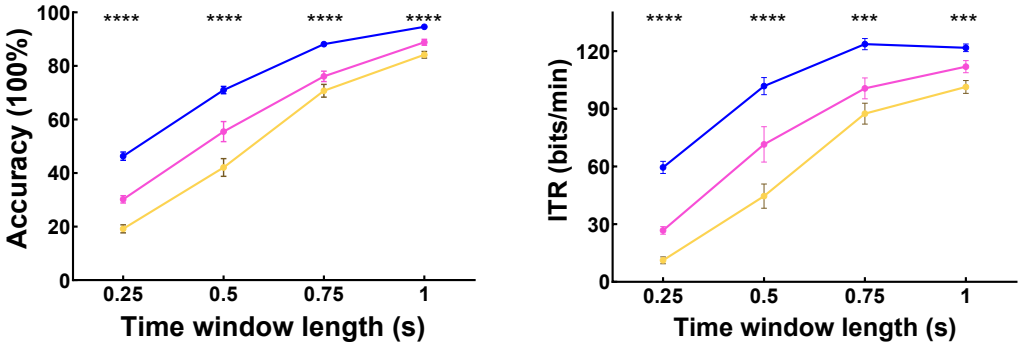
Therefore, for i-th stimulus frequency, the correlation vector $\boldsymbol{\rho}_i$ is denoted as follows:

$$\rho_i = \begin{bmatrix} \rho_i^{(1)} \\ \rho_i^{(2)} \\ \rho_i^{(3)} \\ \rho_i^{(4)} \end{bmatrix} = \begin{bmatrix} \sum_{n=1}^{N_{sub}} p_i^{n,1} corr((\hat{\boldsymbol{S}}_i^n)^T \boldsymbol{X}, \boldsymbol{I}_i^n) \\ \sum_{n=1}^{N_{sub}} p_i^{n,2} corr((\hat{\boldsymbol{T}}_i^n)^T \boldsymbol{X}, \boldsymbol{R}_i^n) \\ corr(\ddot{\boldsymbol{u}}_i^T \boldsymbol{X}, \ddot{\boldsymbol{v}}_i^T \ddot{\boldsymbol{\chi}}_i) \\ corr(\ddot{\boldsymbol{u}}_i^T \boldsymbol{X}, \ddot{\boldsymbol{z}}_i^T \boldsymbol{Y}_i) \end{bmatrix} \tag{5.26}$$

The above correlation coefficients are employed to construct the final feature for target



(a) benchmark dataset



(b) Dataset I

Figure 5.2: The average accuracy and ITR obtained by SSCOR, TRCA, and the proposed method at different time windows on (a) benchmark datasetand (b) Dataset I. The error bars represent SEM. The asterisks indicate significant difference between the three methods obtained by one-way repeated-measures ANOVA ($*$: p<0.05, $**$: p<0.01, $***$: p<0.001, $****$: p<0.0001).

recognition:

$$\beta_i = \sum_{l=1}^{4} \text{sign}(\rho_i^{(l)})(\rho_i^{(l)})^2 \tag{5.27}$$

Therefore, the frequency of test trial can be determined by the following formula:

$$f = \arg\max_{f_i} \beta_i, i = 1, 2, ..., N_f \tag{5.28}$$

The framework of the proposed cross-subject transfer learning method was shown in Fig. 5.1.

5.4 Results

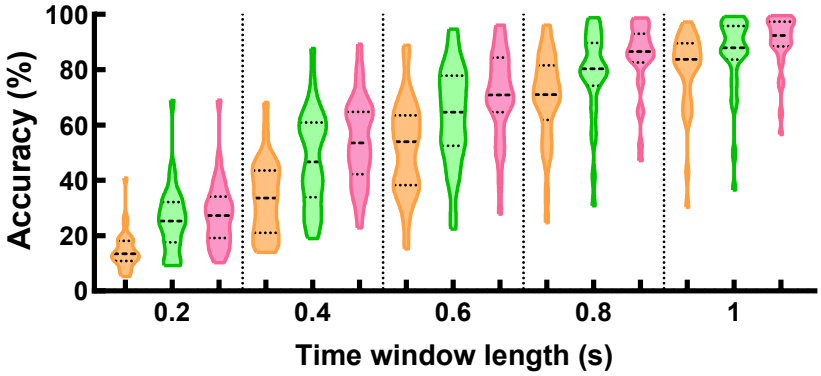
5.4.1 Performance Evaluation

Average classification accuracy and ITR are two widely used indicators to evaluate the performance of SSVEP-based BCIs. ITR (bits/min) can be calculated as follows:

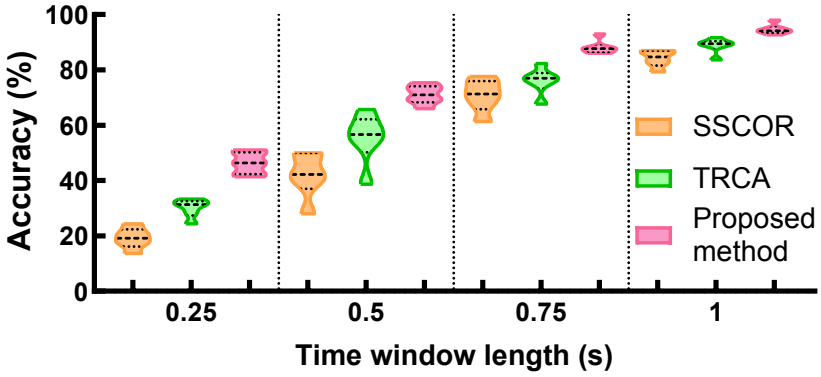
$$ITR = \left(\log_2 N_f + P\log_2 P + (1-P)\log_2\left[\frac{1-P}{N_f-1}\right]\right) \times \frac{60}{T} \tag{5.29}$$

where P is the accuracy of target identification, and T is the average time for a selection, including gaze shifting time (0.5 s) and gaze time. Fig. 5.2 shows the average accuracy and ITR for the proposed method and comparing methods. The sampling rates are different in the two datasets, thus different data lengths were used to keep the number of samples without decimals. The data lengths ranged from 0.2 s to 1 s with an interval of 0.2 s for benchmark datasetand 0.25 s to 1 s with an interval of 0.25 s for Dataset I. The accuracy and ITR were obtained via a leave-one-out cross-validation, where five or four blocks were used for training and a left-one block was used for testing on benchmark datasetand II. For the proposed method, source subjects are selected randomly for transfer. In order to get a general performance of the proposed method, each process was conducted ten times for benchmark datasetand five times for Dataset I. The different numbers of repeat times depend on the size of the two datasets being different. The averaged results were shown for performance evaluation. The number of source subjects is five for both datasets. The reason was clarified in Section 5.3.2. It is obvious that the proposed method can achieve higher accuracy and ITR than TRCA/SSCOR with different time windows (TWs) on two datasets. One-way repeated-measures ANOVA was conducted to explore the similarity of classification performance among the methods on two datasets. The statistical analysis results show that there are significant differences among these methods in accuracy and ITR with each data length.

Fig. 5.3 shows the probability density of classification accuracy for three methods on (a) benchmark datasetand (b) Dataset I via violin plots. The plots analyzed SSVEP signals with different data lengths. The violin plot focuses on illustrating the distribution of quantitative



(a) benchmark dataset



(b) Dataset I

Figure 5.3: Violin plots represent the distributions of classification accuracy of subjects achieved by the three methods with various TWs on (a) benchmark datasetand (b) Dataset I. The thick black line on each violin indicates the median and two black lines on each side represent interquartile ranges (25% and 75% percentiles).

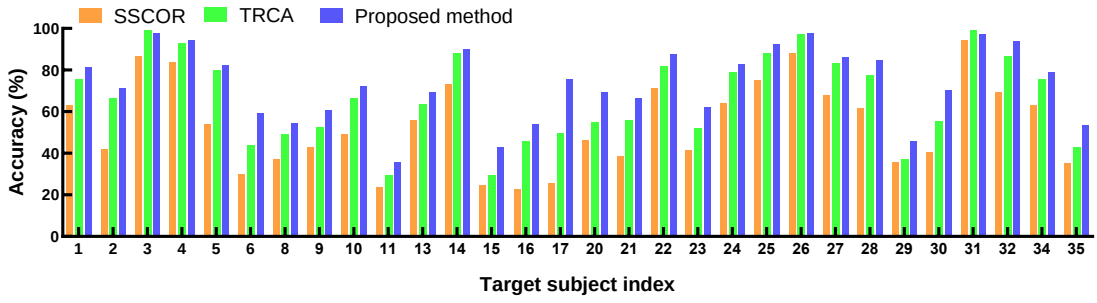


Figure 5.4: Comparison between the accuracy of the proposed method, TRCA and SSCOR for different target subjects with 0.6-long data length. The source subjects were selected randomly. In this case, the source subjects are [7 12 18 19 33], and the rest are target subjects.

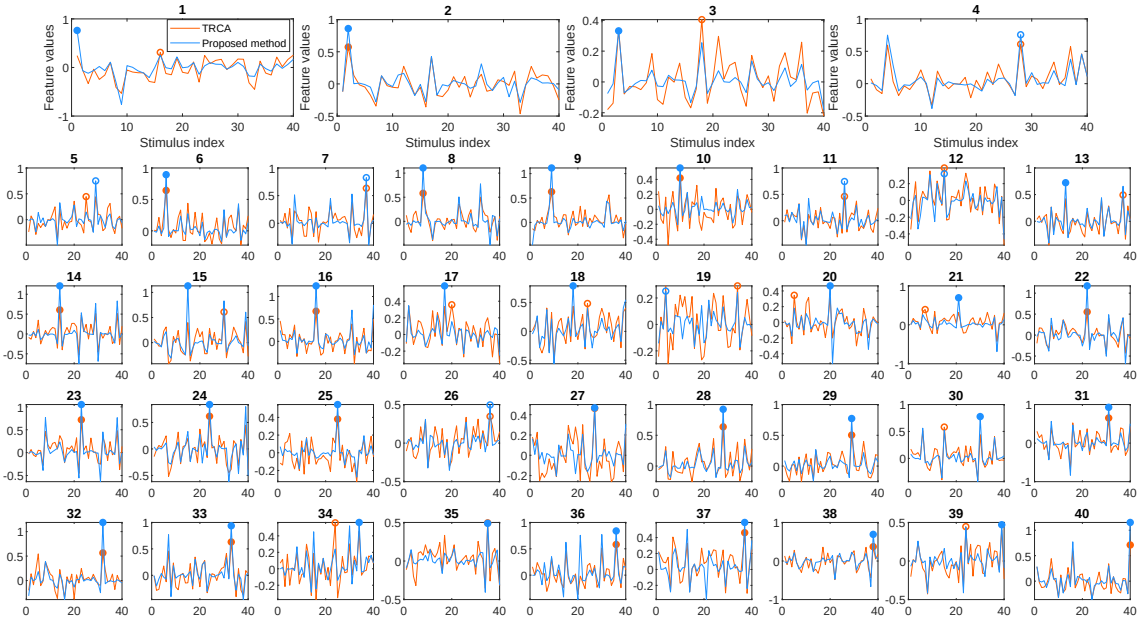


Figure 5.5: Feature values of the forty stimuli obtained by the proposed method and TRCA using a 0.6 s time window for an example subject (S17). The source subjects were selected randomly. The blue and orange circles represent the recognition results of the proposed method and TRCA. The hollow circles turned to solid ones as the results were accurate.

data in a visually intuitive way. The thick black line in the middle represents the median value, and the black lines on either side represent the interquartile range (25% and 75% percentiles). The wider regions of the violin plot denote values that appear more frequently. As shown in Fig. 5.3(a) and Fig. 5.3(b), the violin plots provided by the proposed method (i.e., the pink) generally present higher median values and more concentrated distributions. Therefore, the experimental results indicate that the proposed method can achieve a more stable and superb classification performance on various subjects compared with TRCA and SSCOR.

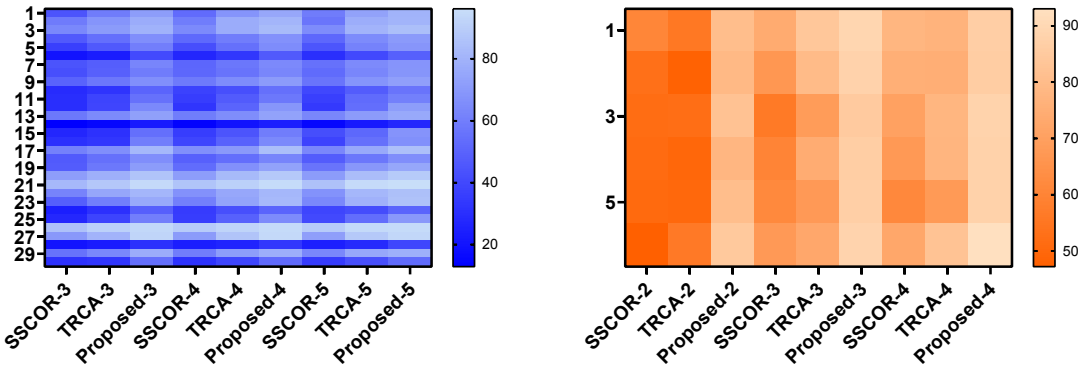
Fig. 5.4 as an example, shows the accuracy comparison between the proposed method, TRCA, and SSCOR for different target subjects. The source subjects were randomly selected, and the indexes are [7 12 18 19 33] in this case. The remaining thirty subjects were used as target subjects for performance comparisons. The experiment result shows that the proposed method achieves higher SSVEP classification accuracy for almost all target subjects.

Fig. 5.5 illustrates the feature values of forty stimuli provided by the proposed method and TRCA for an example target subject (S17) with 0.6 s data length. In comparison to SSCOR, the difference in performance between TRCA and the proposed method is narrower, prompting a more detailed comparison between these two methods. The feature values of the proposed method were calculated via (5.27). Each sub-figure represents a test trial, and the sub-title represents the accurate recognition result. The first four trials were selected and amplified for better viewing details. For each test trial, forty feature values were calculated, and the stimulus corresponding to the largest value was determined as the target via (5.28). The blue and orange circles represent the decisions of the proposed method and TRCA. The hollow circles turned to solid circles as the decisions were accurate. Obviously, the proposed method provided more accurate recognition results. Besides, for those trials where both methods provide correct results, the proposed method shows more distinctive and apparent feature values, such as 2, 6, 8, 14, and 16-th stimuli. It indicates the effectiveness of the proposed feature vector construction strategy in (5.26).

5.4.2 The Effect of Parameters

The Number of Training Blocks

An important purpose of the proposed method is to reduce the need for individual training data. The proposed method should classify SSVEP responses with sufficient accuracy even with a reduced number of individual training data blocks. Fig. 5.6 uses heat maps to show the SSVEP classification accuracy comparison between TRCA, SSCOR, and the proposed method with various numbers of training blocks on (a) benchmark datasetand (b) Dataset I. The heat map acts as a graphical representation of data, displaying values by color in two dimensions. It provides a more visual path to describe numeric values. In the heat map, the x-axis refers to the classification method with a corresponding number of training blocks, and the y-axis indicates the subject index. The accuracy of the target subjects is provided here. The number range of training blocks is [3, 5] for benchmark datasetand [2, 4] for Dataset I. The heat maps visualize the highest classification accuracy and lowest accuracy using colors on a scale from light to dark. As shown in Fig. 5.6(a) and Fig. 5.6(b),



(a) benchmark dataset (b) Dataset I

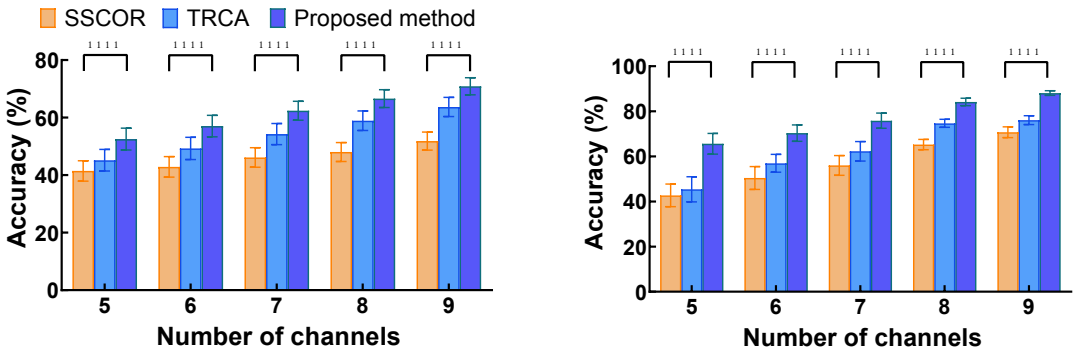
Figure 5.6: Heatmaps of the classification accuracy of three methods with different numbers of training blocks on (a) benchmark datasetand (b) Dataset I.

Table 5.1: Accuracy comparison among three methods with different numbers of training blocks

Methods	Accuracy with different number of training blocks					
	benchmark dataset			Dataset I		
	3	4	5	2	3	4
SSCOR	45.15	50.37	51.83	52.31	64.31	70.72
TRCA	52.23	59.52	63.69	51.85	74.31	76.11
Proposed	65.81	68.38	70.89	80.74	87.22	88.11
P-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

the proposed method generally provides the squares with the lightest color regardless of the number of training blocks. Besides, with the increasing training data scale, the squares generally turn lighter.

Table. 5.1 shows the numerical classification accuracy of three methods and corresponding one-way repeated-measures ANOVA analysis results. The results revealed that there was a statistically significant difference (i.e., $P < 0.0001$) between the methods with all numbers of training blocks for benchmark datasetand Dataset I. In conclusion, this table further demonstrates the effectiveness of the proposed method by providing more quantitative evidence.



(a) benchmark dataset

(b) Dataset I

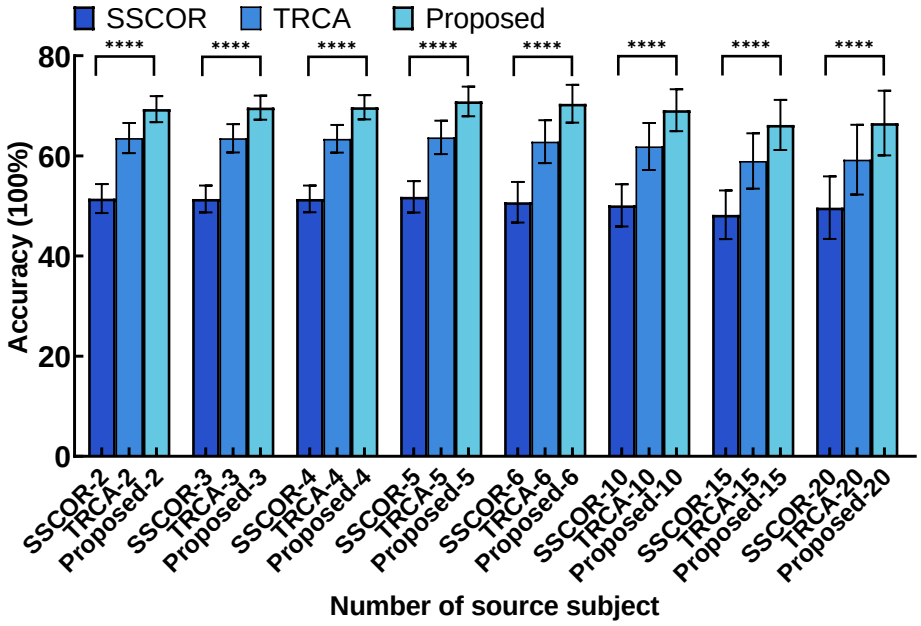
Figure 5.7: Barchart of the classification accuracy of three methods with different numbers of electrodes on (a) benchmark datasetand (b) Dataset I. The error bars represent SEM. The asterisks indicate significant differences between the three methods obtained by one-way repeated-measures ANOVA (∗: p<0.05, ∗∗: p<0.01, ∗ ∗ ∗: p<0.001, ∗ ∗ ∗ ∗: p<0.0001).

The Number of Channels

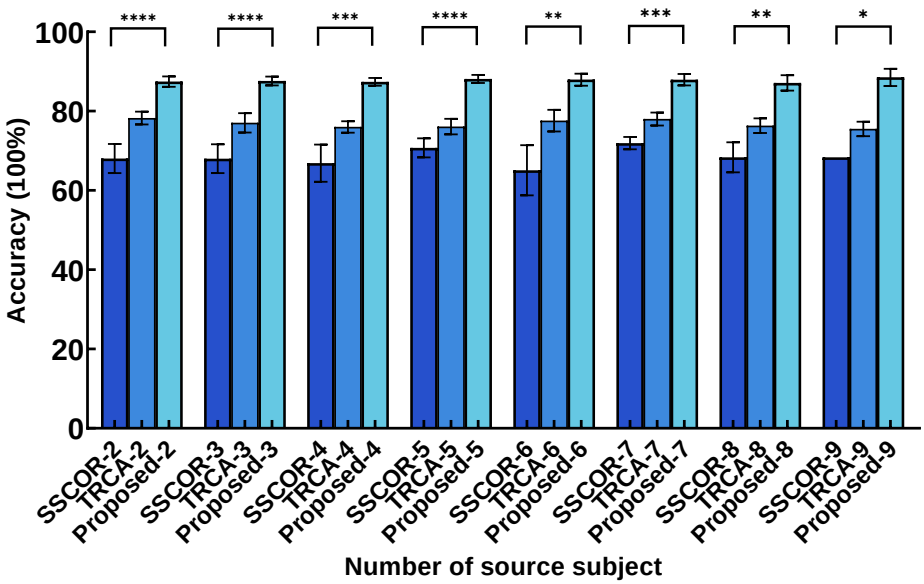
An exploration was conducted to assess how the number of electrodes affects the performance of all methods. Fig. 5.7 shows the classification accuracy results for (a) benchmark datasetand (b) Dataset I. As the number of channels increases, the recognition accuracy generally increases for all methods. As indicated in Fig. 5.7(a) and Fig. 5.7(b), the proposed method always provides the highest classification accuracy with a different number of channels ranging from five to nine for each dataset. Besides, the statistical analysis results show that there is a significant difference between the three methods.

The Number of Source Subject

Fig. 5.8 shows how the number of source subjects affects the performance of methods on (a) benchmark datasetand (b) Dataset I. The classification accuracy in Fig. 5.8(a) and Fig. 5.8(b) is calculated by the target subjects, which does not include source subjects. Therefore, to make the comparison more reasonable, TRCA and SSCOR also show various accuracy values for a different number of source subjects. As the number of source targets increases, the recognition performance of the proposed method generally improves slightly and then decreases. The highest value typically occurs at five, so the number of source targets is set at that in the analysis. The figure also shows that the number of source targets

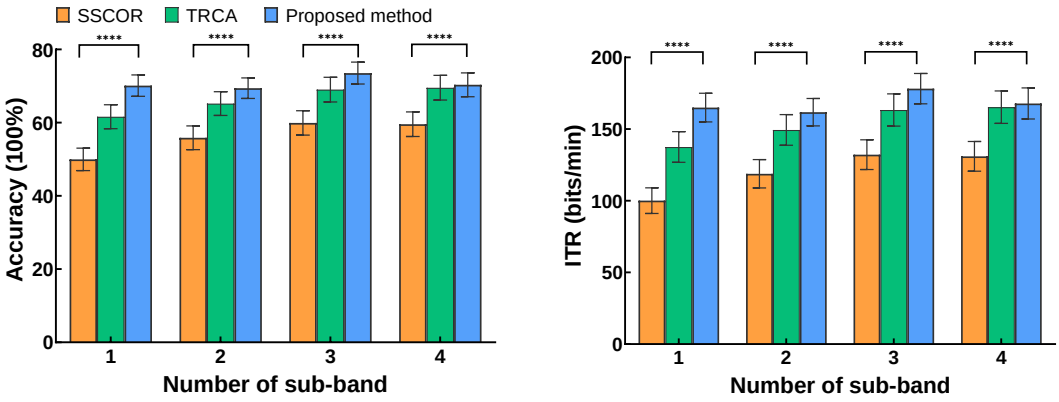


(a) benchmark dataset

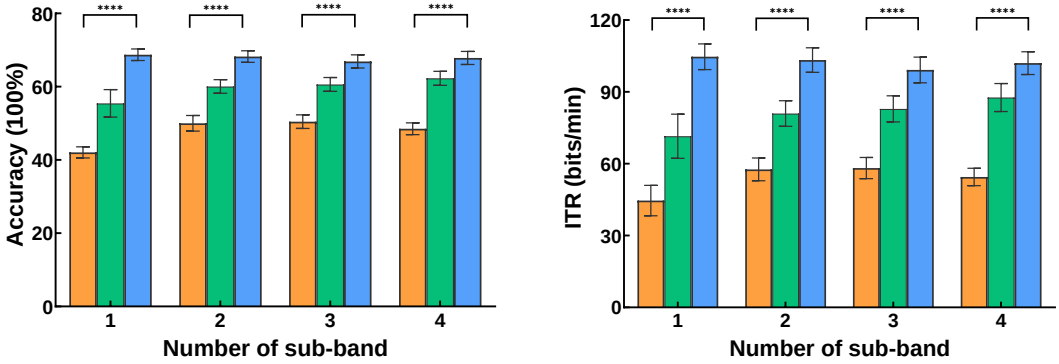


(b) Dataset I

Figure 5.8: Bar chart of the classification accuracy with different numbers of source subjects on (a) benchmark datasetand (b) Dataset I. The error bars represent SEM. The asterisks indicate a significant difference between the there methods obtained by one-way repeated-measures ANOVA ($*$: $p<0.05$, $**$: $p<0.01$, $***$: $p<0.001$, $****$: $p<0.0001$).



(a) benchmark dataset



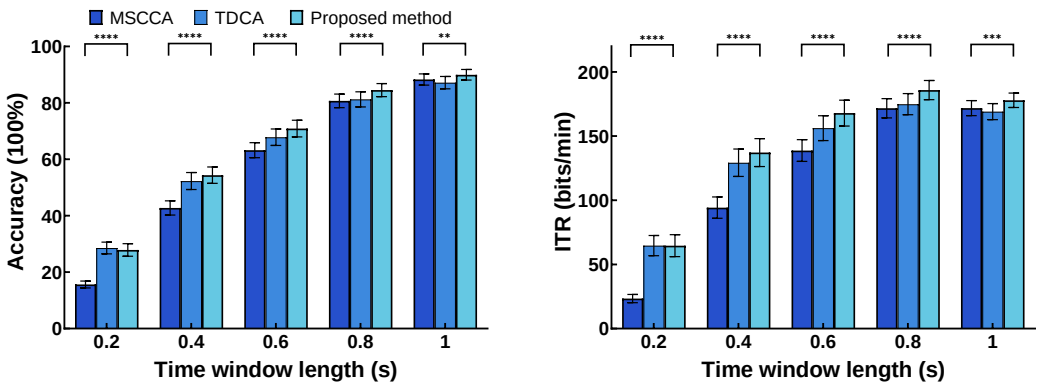
(b) Dataset I

Figure 5.9: Barchart of the classification accuracy and ITR of three methods with a different number of sub-band. The error bars represent SEM. The asterisks indicate significant differences between the three methods obtained by one-way repeated-measures ANOVA (∗: p<0.05, ∗∗: p<0.01, ∗ ∗ ∗: p<0.001, ∗ ∗ ∗∗: p<0.0001).

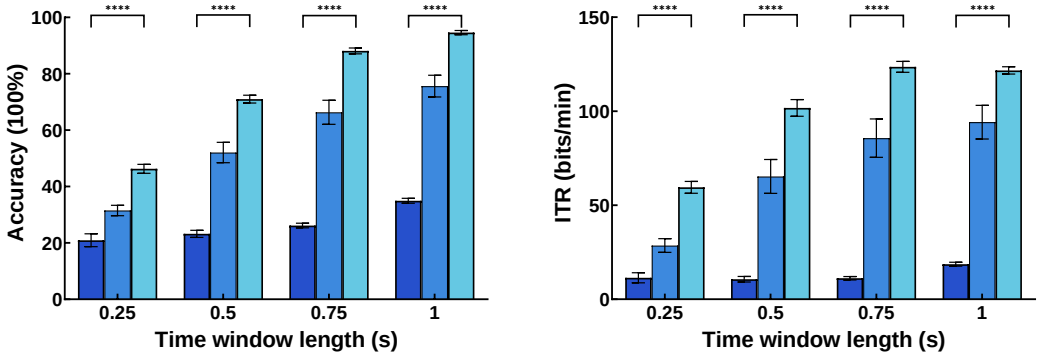
does not have a significant effect on the performance of the proposed method, making this parameter choice representative and reasonable. The [126] also has the same setting for the same dataset.

5.4.3 Filter-bank Analysis

Filter-bank analysis was used to further compare the recognition performance of the proposed method and other methods in this study. The filter-bank technology decomposes the SSVEP signals into N_b sub-band to investigate the information embedded in the harmonic components [69]. The cut-off frequency range was set between $b \times 8$ Hz and 90 Hz for the b-th sub-band, where $b = 1, 2, ..., N_b$ refers to the sub-band component number. The



(a) benchmark dataset



(b) Dataset I

Figure 5.10: The average accuracy and ITR obtained by MSCCA, TDCA, and the proposed method at different time windows. The error bars represent SEM. The asterisks indicate significant differences between the three methods obtained by one-way repeated-measures ANOVA ($*$: p<0.05, $**$: p<0.01, $***$: p<0.001, $****$: p<0.0001).

feature β_i^b was extracted from b-th sub-band signals and then a weighted summation was obtained from all sub-bands as $\Lambda_i = \sum_{b=1}^{N_b}(b^{-1.25}+0.25)\cdot\beta_i^b$ [30]. The target frequency can be recognized by the formula:

$$f = \arg\max_{f_i} \Lambda_i, i = 1, 2, ..., N_f \tag{5.30}$$

Fig. 5.9 shows the classification performance comparison of the proposed method and other methods with different numbers of sub-band on (a) benchmark datasetand (b) Dataset I. The proposed method provided the highest accuracy and ITRs for all data lengths. One-way repeated-measures ANOVA was conducted to further compare these methods. The statistical analysis results indicate that there are significant differences among the three

methods in terms of accuracy and ITRs in each dataset.

5.4.4 Performance Comparison with Data Augmentation Methods

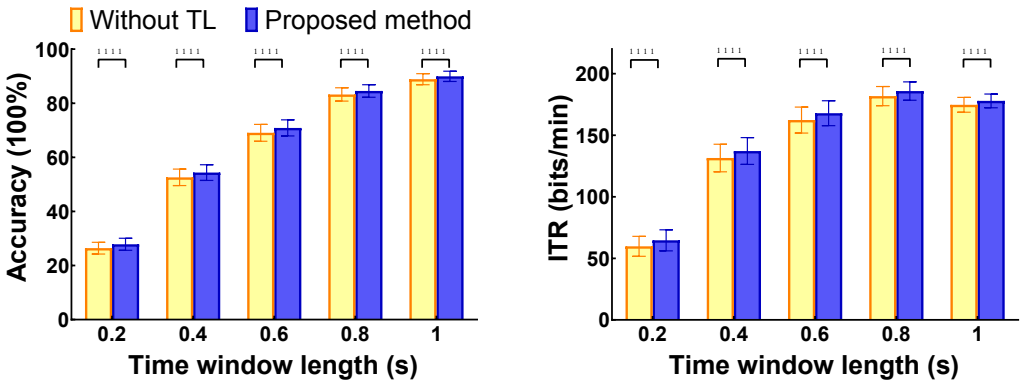
In this study, the proposed method incorporates SSVEP data from the source subject to effectively improve the recognition performance of the target subject. In other words, the data in the target domain was augmented via auxiliary data from the source domain. In this subsection, the proposed method was further compared with two data augmentation methods, including multi-stimulus eCCA (MSCCA) [130] and task-discriminant component analysis (TDCA) [91].

The number of channels and training blocks is set to nine and five for all methods. For TDCA, the number of subspaces and the number of delayed points are eight and one, respectively. In accordance with the comparison shown in Fig. 5.10, the proposed method achieved the highest accuracy and ITRs among all compared methods with almost data lengths. A one-way repeated-measures ANOVA revealed that there was a statistically significant difference between the compared methods. The evaluation results further demonstrated the effectiveness and feasibility of the proposed method for SSVEP recognition of the BCI system.

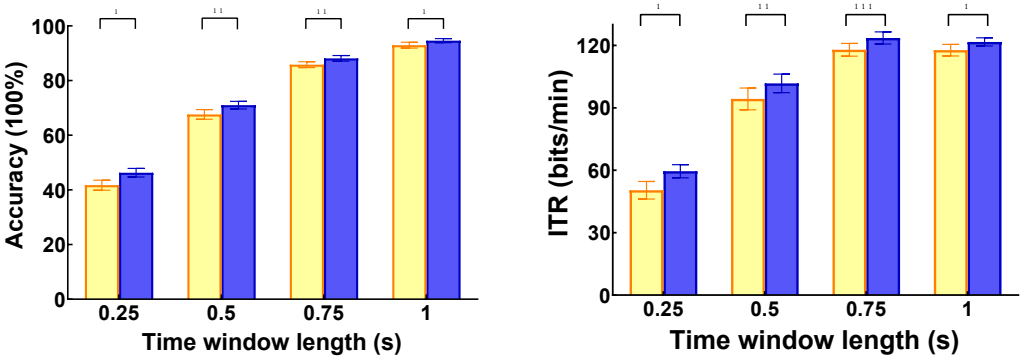
5.5 Discussion

5.5.1 Model's Performance

Almost recognition methods in SSVEP-based BCI fields built spatial filters by considering the relationship between the EEG signal and the artificial reference or the individual template, e.g., CCA and IT-CCA [68] or the relation across training trials, e.g., TRCA and SSCOR [1, 71]. In this study, the spatial filter was trained with multiple similarity constraints. Specifically, maximizing the reproducibility across trials could extract task-related components [1], but it may also bring task-related noise [223]. It is reasonable to remove noise and extract more SSVEP-related features by incorporating the covariance maximization between the training trial and the individual template, between the training trial and



(a) benchmark dataset



(b) Dataset I

Figure 5.11: Performance comparison between the proposed method and the method without transfer learning at different time windows on (a) benchmark datasetand (b) Dataset I. The error bars represent SEM. The asterisks indicate significant differences between the two methods obtained by paired t-test. (∗: $p<0.05$, ∗∗: $p<0.01$, ∗ ∗ ∗: $p<0.001$, ∗ ∗ ∗∗: $p<0.0001$).

the artificial reference, as well as between the two templates. As a cross-subject scheme, the transferred template and transferred spatial filter are used to boost the SSVEP detection performance for the target subject. As shown in Fig. 5.2, the accuracy of the proposed method is 7.19% higher than that of TRCA and 19.05% higher than that of SSCOR on benchmark datasetwith 0.6 s long data length. Besides, the proposed inter-subject transfer learning scheme does not require massive amounts of training data from the target subject and still achieves superior SSVEP classification performance. As shown in Fig. 5.6(a), the accuracy of TRCA with five training trials (i.e., 63.69%) is close to the accuracy of the proposed method with only three training trials (i.e., 65.81%) on Dataset I.

5.5.2 Feature Vector Construction

In this study, the feature vector (5.26) includes four types of correlation coefficients, two of which come from the source subjects and the other two from the target subject. An exploration was conducted into the difference in classification accuracy and ITR between this design and feature vector information only provided by the target subject. Fig. 5.11 shows the comparison results on (a) benchmark datasetand (b) Dataset I. The proposed method shows better SSVEP recognition performance compared with the method without transfer learning. Paired t-test was used to measure the similarity of these methods. The statistical results show that there are significant differences in accuracy or ITR between the two methods at each data length on two datasets. It means that transferred information from source subjects is beneficial to improving the SSVEP recognition performance of the target subject.

5.6 Conclusion

In this study, a cross-subject transfer learning scheme was proposed for enhancing SSVEP classification performance. The spatial filter was first trained via multiple covariance maximization. The relationships between training trials, the individual template, and artificial reference were properly considered in the spatial filter training process. The spatial filters were then applied to the aforementioned templates to construct two new transferred templates, on which the transferred spatial filter can be obtained accordingly. The contribution scores of different source subjects to the feature vector were calculated by their distances from the target subject. Finally, a four-dimensional feature vector was constructed for each stimulus to achieve SSVEP recognition. The effectiveness and feasibility of the proposed method were demonstrated via experimental evaluation on a publicly available dataset and a self-collected dataset.

Chapter 6

SSVEP-based Brain-Computer Interface Controlled Robotic Platform with Velocity Modulation

SSVEP-based BCIs have been extensively studied due to many benefits, such as non-invasiveness, high information transfer rate, and ease of use. SSVEP-based BCI has been investigated in various applications by projecting brain signals to robot control commands. However, the movement direction and speed are generally fixed and prescribed, neglecting the user's requirement for velocity changes during practical implementations. In this study, a velocity modulation method based on stimulus brightness was proposed for controlling the robotic arm in the SSVEP-based BCI system. A stimulation interface was designed, incorporating flickers, a target, and a cursor workspace. The synchronization of the cursor and robotic arm does not require the subject's eye to switch between the stimuli and the robot. The feature vector consists of the characteristics of the signal and the classification result. Subsequently, the GMM and Bayesian inference were used to calculate the posterior probabilities that the signal came from a high or low brightness flicker. A brain-actuated speed function was designed by incorporating the posterior probability difference. Finally, the historical velocity was considered to determine the final velocity. To demonstrate the

effectiveness of the proposed method, online experiments, including single- and multi-target reaching tasks, were conducted. The extensive experimental results validated the feasibility of the proposed method in reducing reaching time and achieving proximity to the target.

6.1 Introduction

Individuals who experience severe motor impairments encounter challenges in performing their daily tasks [224, 225]. Many studies attempted to help people achieve mobility with assistive robotic devices [226]. However, several traditional forms of assistance still require manual control ability from users. BCIs can establish a direct connection between brain signals and external devices, eliminating reliance on peripheral nerves and muscles [227]. In recent years, BCIs have been extensively used in various assistance and rehabilitation applications, such as wheelchair [228], speller [8], and robot arm [229, 230].

EEG is broadly employed in BCI research due to many advantages, such as non-invasiveness, high temporal resolution, and simple operation [231]. SSVEP as one of the EEG paradigms has attracted significant attention because of its minimal training requirements and high SNR [86,211,218]. The SSVEP-based BCI system maps the brain signals to robot commands and then transmits them to the corresponding manipulation [185]. Therefore, SSVEP-based BCIs have been widely explored in the field of assistive and rehabilitative applications in recent decades. For example, Guo *et.al* [154] designed an SSVEP-based BCI-controlled soft robotic glove for post-stroke hand function rehabilitation. Sakkalis *et.al* [162] implemented efficient electric wheelchair navigation by utilizing an SSVEP-controlled system. In a separate study, Wang *et.al* [148] built a portable SSVEP system specifically designed for rehabilitation exoskeletons. Chen *et.al* [147] integrated SSVEP with computer vision to fulfill robot pick-and-place tasks. Subsequently, Chen *et.al* [150] further employed the AR technique that eliminates the need for subjects to switch attention between the stimulation and the robotic arm. These studies effectively established the connection between SSVEP signals and external robots with prescribed velocity control. However, modulating the velocity of the robots in response to the user's intentions can bring more benefits in human motion assistance and rehabilitative scenarios.

In recent years, several studies have considered velocity control in SSVEP-based BCI systems. For instance, Zhao *et.al* [155] mapped three flashing squares to the three motion modes, where different frequencies represent fast, medium, and slow speeds. However, by replacing the stimulus-command pair with the stimulus-velocity pair, the BCI achieves velocity modulation in a discrete manner. To realize continuous velocity mapping, Zhang *et.al* [164] introduced stimuli with different frequencies to control a navigation robot. A self-defined mapping function on the basis of correlation values was proposed to output continuous control commands. Additionally, Sharma *et.al* [232] demonstrated that the amplitude of brain signals can be modulated by varying the distances between the subject and the stimulus. However, only one stimulus was designed for unidirectional robot manipulation since orthoptic eye accommodation may bring difficulties in multi-class interface design. In addition to stimulus frequency and distance, voluntary attention increased stimulus-driven EEG activity [233]. Integrating the attention factor, Molina-Cantero *et.al* [234] utilized SSVEP responses obtained from a single-channel EEG headset known as NeuroSky Mindwave (NM). The NM was able to provide attention levels, which were then used to modulate the speed of a controlled cursor. However, the requirement for a specific headset limits its generalizability and scalability. On the other hand, flicker brightness is another important component that influences the characteristics of the SSVEP signal; thus, some studies have investigated the brain response to the modulation of stimulus brightness [235]. For instance, the findings in [236] revealed that SSVEP amplitude generally increases with higher brightness levels. Zhang *el.al* found that the PSD of the stimulus maintained a certain positive correlation with brightness [237]. Nevertheless, studies investigating the feasibility and effectiveness of employing brightness in robot velocity modulation are still limited to SSVEP-based BCIs.

This study pursued the exploration of a velocity modulation approach for the control of a robotic arm in the context of SSVEP-based BCI. The key contributions encompass 1) Development of a stimulation interface including flickers, a target, and a cursor workspace. It enables synchronized movement of the cursor and robotic arm. Subjects are not required to switch their eyes between the stimulus and the robot. 2) Proposal of a stimulus brightness-

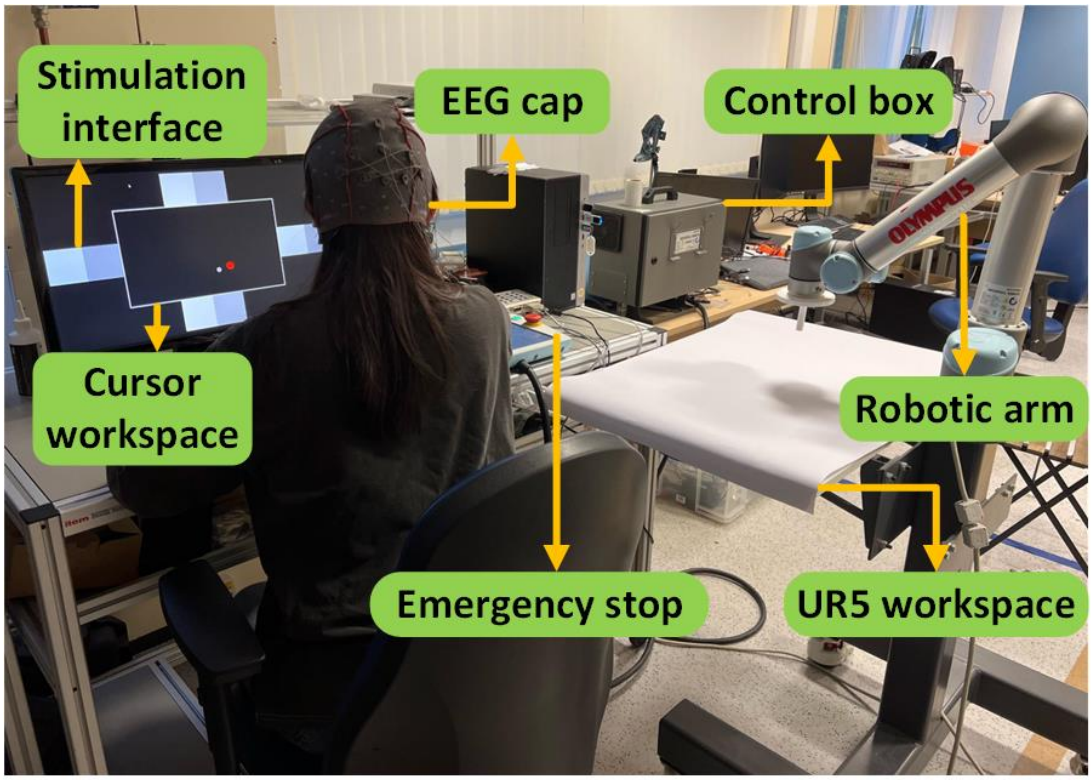


Figure 6.1: The experimental environment of the proposed SSVEP-based BCI system for robotic arm velocity control.

based velocity modulation model. The feature vector is constructed using the correlation coefficient and PSD. 3) Devising a speed function that leverages the difference in posterior probabilities of being a high- or low-brightness flicker. The historical velocity was incorporated in the velocity determination to avoid occasional misclassification. To evaluate the performance of the proposed method, online experiments, including single- and multi-target reaching tasks, were conducted. Two additional control methods were used for comparison. Extensive evaluations showed that the proposed method enabled the cursor or robotic arm to complete tasks in a shorter time and get closer to the target.

This paper is organized as follows: Section 6.2 introduces the SSVEP experiment and the proposed method. Section 6.3 presents the experimental results. The discussion and conclusion are shown in Sections 6.4 and 6.5, respectively.

6.2 Experimental Protocol

6.2.1 Experimental Environment

Fig. 6.1 showed the experimental setup of the SSVEP-based BCI system for robotic arm control. The computer screen displayed the stimulation interface, which includes flickers, the target, and the cursor workplace. The robotic arm and its workspace were located at the right hand of the subject. The experimental environment was detailed as follows:

Participants

Ten healthy subjects (five females and five males, mean age: twenty-eight years) volunteered for the experiment in this study. All subjects had a normal or corrected-to-normal vision. The experiment has been approved by the Faculty Research Ethics Committee of the University of Leeds. Each participant read and signed an informed consent form.

Simulation Interface

The stimulation interface was shown in Fig. 6.2. There are eight stimuli on a 23.6-inch LCD monitor with a resolution of 1920×1080 pixels and a refresh rate of 60 Hz. Two adjacent stimuli have the same frequency but different brightnesses. For instance, the top two stimuli flashed at 7 Hz, corresponding to upward movements. The maximum brightnesses were set at 255 and 180, respectively. The right, bottom, and left stimuli flickered at 8, 9, and 12 Hz, respectively. The size of the stimulus is 210×210 pixels. The rectangle with white edges is 1060×620 pixels in size. The red circle (radius: 25 pixels) and the white one (radius: 15 pixels) are the target and the cursor. The interface was developed in MATLAB using the Psychophysics Toolbox Version 3 [238].

Signal Acquisition

The data was recorded using equipment from g.tec medical engineering GmbH, and it was sampled at a rate of 512 Hz using the g.USB amplifier. Nine electrodes, i.e., Pz, PO3, POz, PO4, PO7, O1, Oz, O2, and PO8, in the parietal and occipital areas were selected. The ground electrode was placed over FPz, and the reference electrode was positioned on the

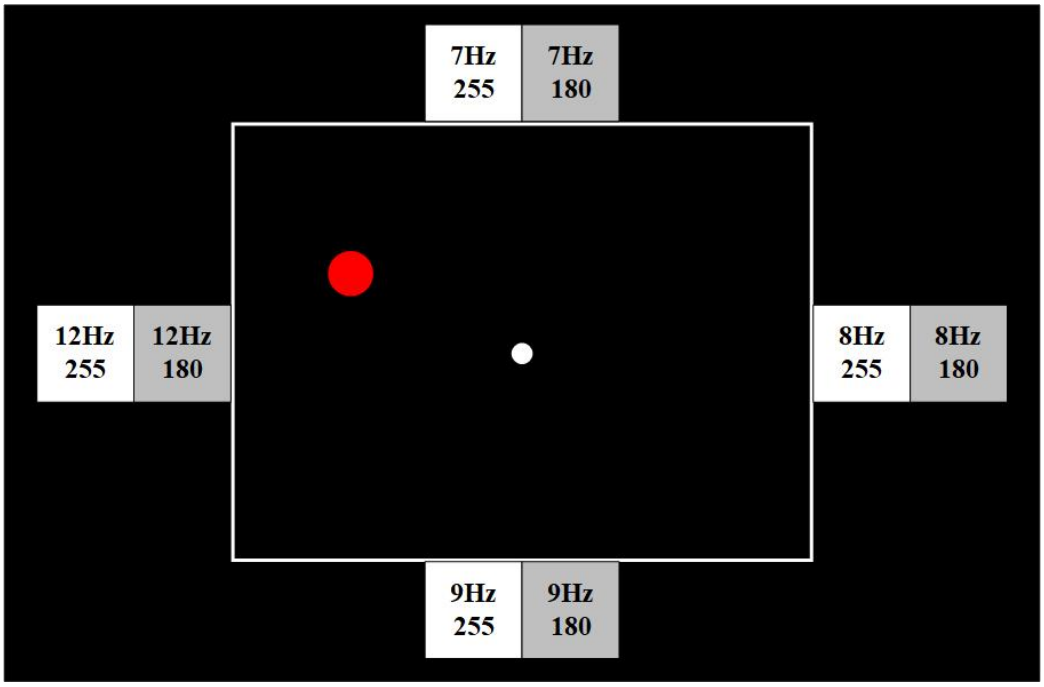


Figure 6.2: The simulation interface of the eight-target SSVEP-based BCI system. The frequency and maximal brightness were displayed for each visual stimulus. The workspace for the random target and the cursor was represented by a rectangle with white edges. The red circle is the target and the white one is the cursor.

right earlobe.

Robotic Arm

The UR5 robotic arm (Universal Robots) was employed in this study. Its maximum reaching radius is 850 mm. The workspace for the UR5 robotic arm was set up with a whiteboard positioned directly in front of the robot. It provides a designated horizontal area (803 mm $\times$ 470 mm) for the arm's actions and operations. The participants were asked to control the robotic arm to perform reaching tasks.

6.2.2 Experimental Protocol

Offline

For each subject, the experiment included four blocks, and each block contained eight trials corresponding to eight stimuli. Each trial began with a 2 s target cue (a red square). After the cue, all targets flickered for 7 s simultaneously. During the experiment, the subject

sat in a comfortable chair in a dimly lit and quiet room. The viewing distance from the computer screen was set at 60 cm. The subject was asked to focus on the stimulus and avoid eye movement. The subject can take a rest between two blocks.

Online

The online experiments include single- and multi-target reaching tasks. In the single-target reaching experiment, each subject should complete eight trials. In each trial, the target was generated randomly within the cursor's workspace. There is a one-to-one mapping between the target position in the cursor's workspace and a specific location in the robotic arm's workspace. The subject should move the cursor to reach the target by focusing on different visual stimuli. Meanwhile, the cursor velocity was transformed and then sent to the robotic arm, so the robotic arm can map the cursor's movement in an enlarged workspace. Therefore, subjects were not required to switch their eyes between the stimulus and the robotic arm. Fig. 6.3 showed the diagram of the proposed robotic arm velocity modulation method in the SSVEP-based BCI.

The basic rule for governing reaching tasks is that the cursor/robotic arm should move quickly when it is far from the target and slow down as it approaches. It ensures that the cursor/robotic arm approaches the target with greater precision within a shorter duration. Hence, the subject should focus on the lighter flicker to move the cursor/robotic arm faster while focusing on the darker flicker to move it slower. When the distance between the centers of the cursor and the target is consistently below a certain value for four consecutive times, and the trial time is within 40 s, the target is considered to be hit successfully. Otherwise, the trial fails. The distance was set to 30 pixels. Two other control methods were employed for performance comparison. The comparing methods would be described in the next subsection. In each comparative experiment, the target position is the same. Therefore, each subject should finish 24 trials in the single-target reaching task.

In the multi-target reaching task, the subject should reach three randomly generated targets successively within 200 s. After successfully hitting a target, the next target occurs automatically. Each participant was required to complete three reaching tasks using the

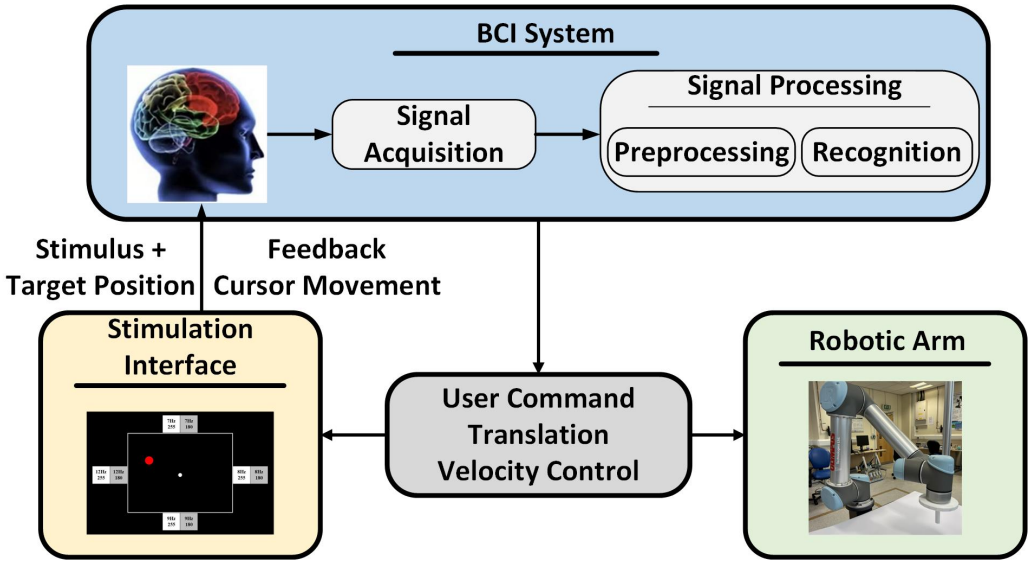


Figure 6.3: Diagram of the proposed SSVEP-based BCI system for robotic arm velocity control.

proposed method along with two comparative control methods, for a total of nine trials.

6.2.3 Data Pre-processing

Since the effect of visual latency in the human visual system, the data was extracted in $[0.14\ (0.14 + d)]s$, where d was the data length for performance analysis. The data were filtered by a Butterworth band-pass filter (5–60 Hz) and a notch filter (50 Hz). The band-pass range is determined by the visual stimulus frequency and the number of harmonics in CCA.

6.2.4 SSVEP-based BCI Controlled Robotic Platform with Velocity Modulation

A stimulus brightness-based velocity modulation method was presented for robotic arm control in the SSVEP-based BCI. The framework, as shown in Fig. 6.4, included two parts: offline and online experiments. The offline experiment includes data acquisition, feature extraction, and stimulus brightness-based speed modulation. In the online experiment, the velocity of the cursor/robotic arm was finally determined by the speed function trained in the offline and historical velocity. The data acquisition was described in previous sections. Subsequent content will explain the other modules in detail.

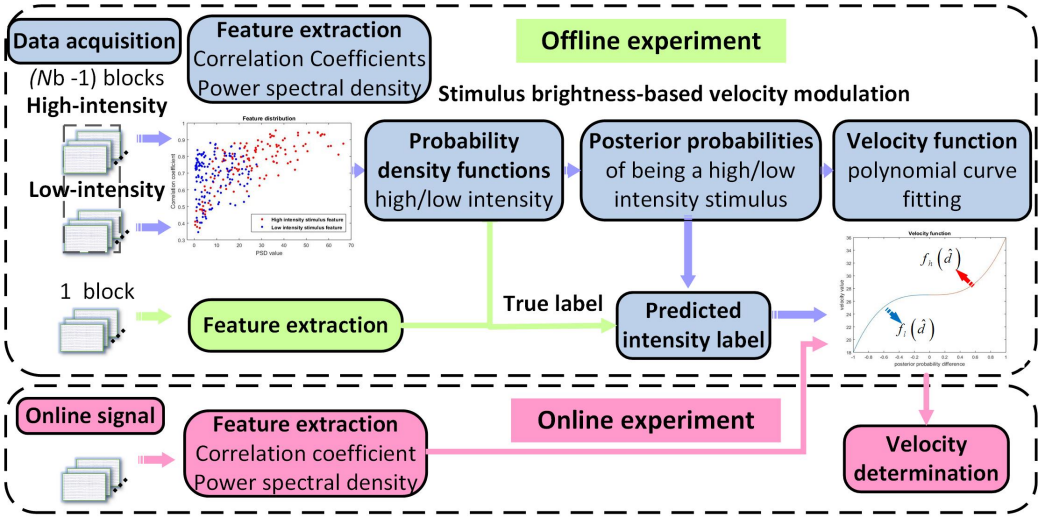


Figure 6.4: The detailed framework of the proposed velocity modulation process for robotic arm control in the SSVEP-based BCI.

Feature Extraction

Suppose that $\boldsymbol{X} \in \mathbb{R}^{N_c \times N_s}$ represents the two-dimensional SSVEP signal from the offline experiment. N_c, and N_s are the number of channels and samples, respectively. SSVEP signals can also be described as sine-cosine waves, so the reference signal $\boldsymbol{Y}_i \in \mathbb{R}^{2N_h \times N_s}$ for the i-th stimulus can be artificially constructed as follows:

$$\boldsymbol{Y}_i = \begin{bmatrix} \sin(2\pi f_i t) \\ \cos(2\pi f_i t) \\ \vdots \\ \sin(2\pi N_h f_i t) \\ \cos(2\pi N_h f_i t) \end{bmatrix}, \; t = [1/F_s, 2/F_s, ..., N_s/F_s] \tag{6.1}$$

where f_i, N_h, and F_s refer to the frequency of the stimulus, the number of harmonics, and the sampling rate, respectively. In this study, the number of harmonics in CCA was set to five.

This study used canonical correlation analysis (CCA) for feature construction and recognition. CCA finds two spatial filters $\boldsymbol{w}_i \in \mathbb{R}^{N_c \times 1}$ and $\boldsymbol{v}_i \in \mathbb{R}^{2N_h \times 1}$ to maximize the correlation

between linear projections $\boldsymbol{w}_i^T\boldsymbol{X}$ and $\boldsymbol{v}_i^T\boldsymbol{Y}_i$:

$$\rho_i = \max_{\boldsymbol{w}_i,\boldsymbol{v}_i} \frac{E[\boldsymbol{w}_i^T\boldsymbol{X}\boldsymbol{Y}_i^T\boldsymbol{v}_i]}{\sqrt{E[\boldsymbol{w}_i^T\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{w}_i]}\sqrt{E[\boldsymbol{v}_i^T\boldsymbol{Y}_i\boldsymbol{Y}_i^T\boldsymbol{v}_i]}} \tag{6.2}$$

The correlation coefficient ρ_i can be calculated between $\boldsymbol{X}$ and each reference signal $\boldsymbol{Y}_i, i = 1, 2, ..., N_f$. The frequency of the reference signal with maximal correlation coefficient is determined as the frequency of $\boldsymbol{X}$. If the classification is correct, the maximal correlation coefficient ρ was recorded for the feature construction. Additionally, the PSD of $\boldsymbol{X}_j \in \mathbb{R}^{1\times N_s}, j = 1, 2, ..., N_c$ was estimated by periodogram() in MATLAB. The PSD at the true frequency f for j-th channel was presented as $\text{PSD}_j(f)$. The average PSD value α across all channels was calculated as:

$$\alpha = \frac{1}{N_c}\sum_{j=1}^{N_c}\text{PSD}_j(f) \tag{6.3}$$

The signals from N_t blocks of the offline experiment were processed using the aforementioned procedure.

Stimulus Brightness-based Speed Function

As illustrated in Fig. 6.4, the offline signals were divided into two sets. Specifically, $(N_t - 1)$ data blocks were used to train conditional probability density functions, and one block served as an evaluation dataset to fine-tune the model's hyperparameters. A sliding time window with a step of 0.5 s was employed to divide the signals into multiple epochs. In the $(N_t - 1)$ blocks, ρ, and α were further divided into two sets based on whether they were obtained from high- or low-brightness flickers:

$$\begin{aligned}
\boldsymbol{\Phi}_h &= \begin{bmatrix} \rho_{h,1},\ \rho_{h,2},...,\ \rho_{h,N_h} \\ \alpha_{h,1},\ \alpha_{h,2},...,\ \alpha_{h,N_h} \end{bmatrix}^T \in \mathbb{R}^{N_h\times 2} \\
\boldsymbol{\Phi}_l &= \begin{bmatrix} \rho_{l,1},\ \rho_{l,2},...,\ \rho_{l,N_l} \\ \alpha_{l,1},\ \alpha_{l,2},...,\ \alpha_{l,N_l} \end{bmatrix}^T \in \mathbb{R}^{N_l\times 2}
\end{aligned} \tag{6.4}$$

where N_h and N_l represent the number of trials where the subject directed their attention toward high- and low-brightness visual stimuli, respectively. The feature matrices $\mathbf{\Phi}_h$ and $\mathbf{\Phi}_l$ are derived from stimuli with high and low brightness. Suppose that when the subject observed the high and low brightness stimuli, the corresponding trials were labeled as T_h and T_l. The probability density functions of the feature matrices for high and low brightness stimuli are denoted as $p(\mathbf{\Phi}|T_h)$ and $p(\mathbf{\Phi}|T_l)$. GMM was used to fit $\mathbf{\Phi}_h$ and $\mathbf{\Phi}_l$. The GMM is an efficient probabilistic model capable of building complex probability distribution functions [212]. Therefore, two probability distribution functions are expressed as follows:

$$\begin{aligned} p(\mathbf{\Phi}|T_h) = p(\mathbf{\Phi}_h) = \sum_{k=1}^{K} \lambda_k \mathcal{N}(\mathbf{\Phi}_h|\boldsymbol{\theta}_k) \\ p(\mathbf{\Phi}|T_l) = p(\mathbf{\Phi}_l) = \sum_{k=1}^{K} \varphi_k \mathcal{N}(\mathbf{\Phi}_l|\boldsymbol{\xi}_k) \end{aligned} \tag{6.5}$$

where K is the number of mixture components, λ_k and φ_k are the mixture wights subject to the constraints $\sum_{k=1}^{K} \lambda_k = 1$ and $\sum_{k=1}^{K} \varphi_k = 1$. The Gaussian density functions $\mathcal{N}(\mathbf{\Phi}_h)$ and $\mathcal{N}(\mathbf{\Phi}_l)$ are defined by the parameters $\boldsymbol{\theta}_k = (\boldsymbol{\mu}_k, \mathbf{\Sigma}_k)$ and $\boldsymbol{\xi}_k = (\boldsymbol{\nu}_k, \mathbf{\Gamma}_k)$, where $\boldsymbol{\mu}_k$ and $\boldsymbol{\nu}_k$ are the mean, and $\mathbf{\Sigma}_k$ and $\mathbf{\Gamma}_k$ are the covariance matrices. The GMM parameters, i.e., $\lambda_k, \eta_k, \boldsymbol{\mu}_k, \mathbf{\Sigma}_k, \boldsymbol{\nu}_k$, and $\mathbf{\Gamma}_k$, were estimated by the EM algorithm in MATLAB.

The left-one block was used to determine hyperparameters. Specifically, each trial from the left block was analyzed to form the feature vector $\hat{\boldsymbol{\phi}} = [\hat{\rho}\ \hat{\alpha}]$. Bayesian inference was used to compute the posterior probabilities, indicating the trial was obtained by observing a high and low brightness stimulus:

$$\begin{aligned} P(T_h|\hat{\boldsymbol{\phi}}) = \frac{p(\hat{\boldsymbol{\phi}}|T_h)P(T_h)}{p(\hat{\boldsymbol{\phi}}|T_h)P(T_h) + p(\hat{\boldsymbol{\phi}}|T_l)P(T_l)} \\ P(T_l|\hat{\boldsymbol{\phi}}) = \frac{p(\hat{\boldsymbol{\phi}}|T_l)P(T_l)}{p(\hat{\boldsymbol{\phi}}|T_h)P(T_h) + p(\hat{\boldsymbol{\phi}}|T_l)P(T_l)} \end{aligned} \tag{6.6}$$

where $P(T_h)$ and $P(T_l)$ are prior probabilities of high and low brightness stimulus. They are both set to 0.5. The difference between two posterior probabilities is calculated as follows:

$$d = P(T_h|\hat{\boldsymbol{\phi}}) - P(T_l|\hat{\boldsymbol{\phi}}) \tag{6.7}$$

Subsequently, d is compared with a threshold σ:

$$L(\hat{\boldsymbol{\phi}}) = \begin{cases} T_h, \ d > \sigma \\ T_l, \ \text{otherwise} \end{cases} \tag{6.8}$$

where $L(\hat{\boldsymbol{\phi}})$ is the predicted high/low brightness label of $\hat{\boldsymbol{\phi}}$. $L(\hat{\boldsymbol{\phi}})$ would be compared with the true brightness label of $\hat{\boldsymbol{\phi}}$. If they are equal, it indicates that the above process successfully recognized the brightness label of the SSVEP trial. The ratio of successfully predicted trials to total trials in the left block is defined as *Acc*. The grid-search method was used to determine optimal values for K and σ via calculation of *Acc*. The K and σ ranges are specified as [1, 4] and [-0.5, 0.5], respectively. An exhaustive search was conducted for the K with an interval of 1 and for σ with an interval of 0.1. The values that yielded the highest *Acc* were chosen as optimal values for K and σ.

The speed range of the cursor is set to $[v_l, v_h]$, so the middle speed is $v_m = \frac{v_l + v_h}{2}$. The brain-actuated speed function based on the brightness label is represented as follows:

$$v = \begin{cases} f_l(d), & d < \sigma \\ v_m, & d = \sigma \\ f_h(d), & d > \sigma \end{cases} \tag{6.9}$$

where $f_l(\cdot)$ and $f_h(\cdot)$ were obtained by polyfit() in MATLAB. For $f_l(\cdot)$, polyfit fits a polynomial of degree three to $[-1, v_l]$ and $[\sigma, v_m]$. For $f_h(\cdot)$, polyfit fits a same degree polynomial to $[\sigma, v_m]$ and $[1, v_h]$. $v_l = 18$, $v_h = 36$ in this study.

Velocity Determination

In the online experiment, the feature of signal $\boldsymbol{\chi} \in \mathbb{R}^{N_c \times N_s}$ was extracted as $[\hat{\rho}, \hat{\alpha}]$. Following the above process, $\boldsymbol{\chi}$ can be classified as the high or low brightness stimulus based on the posterior probability difference $\hat{d}$ in (6.8). Subsequently, the brain-actuated speed v is obtained using (6.9). So the two-dimensional velocity vector $\boldsymbol{v}_d$ can be determined by the

speed and recognition result:

$$\boldsymbol{v}_d = [v_x \ v_y] = \begin{cases} [0, \ -v], & f_d = 1 \\ [v, \quad 0], & f_d = 2 \\ [0, \quad v], & f_d = 3 \\ [-v, \ 0], & f_d = 4 \end{cases} \tag{6.10}$$

where f_d is the predicted class of $\boldsymbol{\chi}$. f_d would be compared with the recognition result of the last trial f'_d. The final velocity vector of cursor $\hat{\boldsymbol{v}}_c$ can be determined as follows:

$$\hat{\boldsymbol{v}}_c = \begin{cases} \boldsymbol{v_d}, & f_d = f'_d \\ \hat{\boldsymbol{v}}'_c - \frac{\hat{\boldsymbol{v}}'_c}{4} + \boldsymbol{v_d}, & f_d \neq f'_d \end{cases} \tag{6.11}$$

where $\hat{\boldsymbol{v}}'_c$ is the cursor's velocity at the last moment. The cursor moves at a velocity of $\hat{\boldsymbol{v}}_c$ for a duration of 0.5 s. Subsequently, the robotic arm's velocity $\hat{\boldsymbol{v}}_a$ can be expressed proportionally:

$$\begin{aligned} v_{ax} &= \frac{-W_h \times \hat{\boldsymbol{v}}_c(2)}{w_h} \\ v_{ay} &= \frac{-W_w \times \hat{\boldsymbol{v}}_c(1)}{w_w} \end{aligned} \tag{6.12}$$

where W_h and W_w are the height and weight of the robotic arm's workspace, respectively. w_h and w_w are the height and weight of the cursor's workspace.

Based on the cursor's and target's positions, the subject adjusted the speed and direction via the stimulus's brightness and frequencies, respectively. The distance between their centers was calculated for each time duration. If the distance is less than 30 pixels for four consecutive instances within 40 s, it is considered a successful hit.

6.2.5 Velocity Control Strategies for Comparison

Two additional methods were compared with the proposed method. The first method is called discrete velocity (DV) control, which is commonly utilized in SSVEP-based BCIs for

controlling robots [154, 162]. The speed is a constant value v_{cons}, and its direction depends on the recognition result. The second method, named discrete attenuated velocity (DAV) control, incorporates velocity attenuation in (6.11) based on the configuration of the DV method. It shares the same setting as the proposed method. The only difference is that it does not include the proposed SSVEP-based BCI-actuated velocity modulation. Given that DV and DAV methods do not consider the brightness factor, only high-brightness stimuli were used for a fair comparison. Therefore, v_{cons} is derived by the d obtained from high-brightness stimuli. Specifically, the fitdist() function was used to fit these values, and consequently, the mean value d_m can be obtained. Therefore, v_{cons} can be expressed as follows:

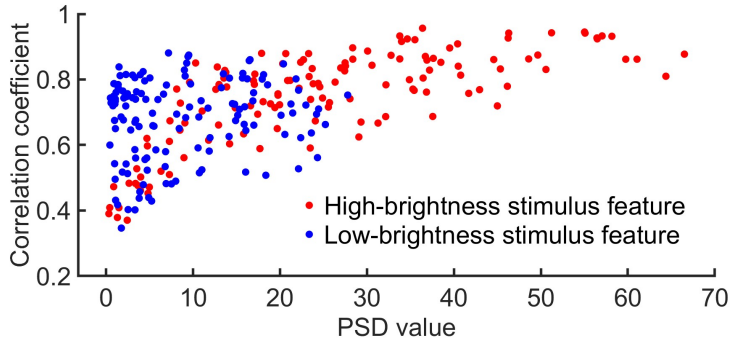
$$v_{cons} = f_h(\hat{d}_m) \tag{6.13}$$

To ensure a fair comparison, v_{cons} is not merely equivalent to v_m, but rather obtained through the above formula. Besides, the two comparison methods also employ the sliding window, the same as the proposed method.

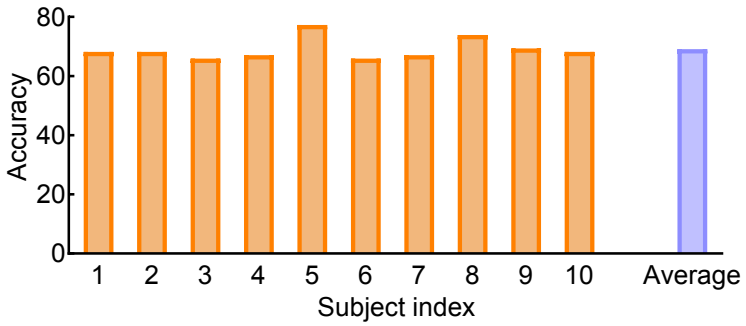
6.3 Results

6.3.1 Offline Experiment Analysis

Fig. 6.5(a) displays the feature distribution of high- and low-brightness visual stimuli for the fifth subject. The x- and y-axes represent the PSD value and correlation coefficients, respectively. The scatter plots reveal distinct feature distributions for the two kinds of stimuli. Specifically, high-brightness stimuli generally exhibit higher PSD values and correlation coefficients. Fig. 6.5(b) presents the classification accuracy of high- and low-brightness stimuli. The graph displays the accuracy for each subject, along with the average value (i.e., 69.1%). Notably, the classification accuracy for each subject is within an acceptable range in the offline experiment.



(a) Feature distribution



(b) Brightness classification accuracy

Figure 6.5: (a) Feature distribution of high- and low- brightness stimuli, and (b) brightness classification accuracy of each subject.

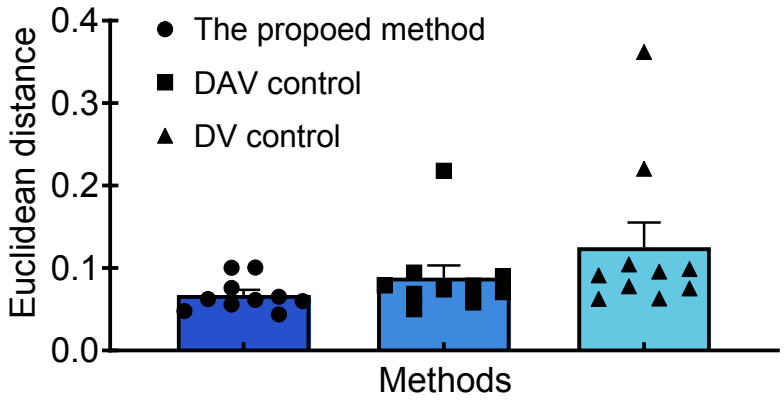


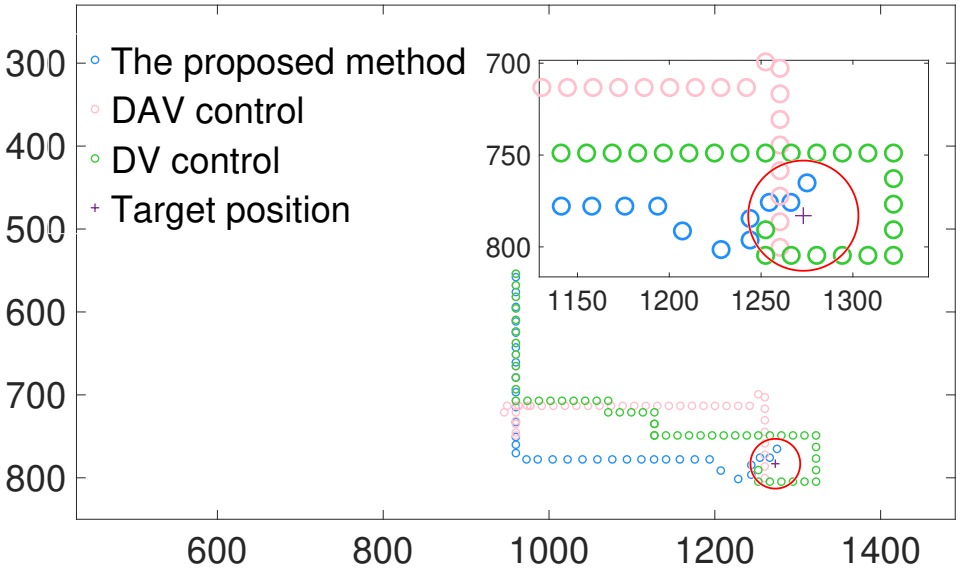
Figure 6.6: The Euclidean distance between the robotic arms' positions projected by the cursor's movements and its actual arrival positions in single-target reaching tasks. The scatter points refer to the Euclidean distances provided by different subjects.

6.3.2 Single-target Reaching Task Performance Evaluation

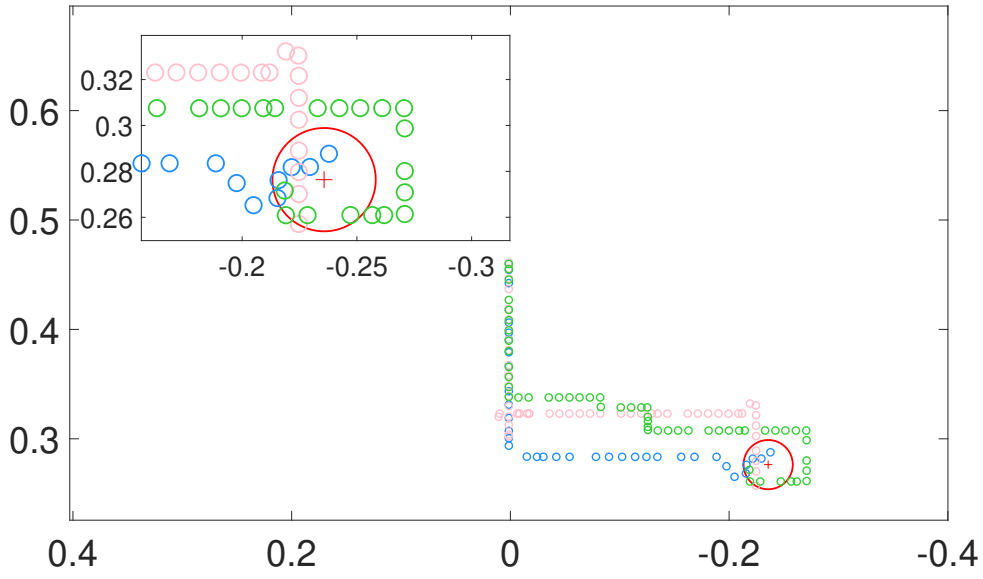
Fig. 6.6 showed the Euclidean distance between the robotic arm's positions projected by the cursor and its actual arrival positions using the high-brightness of the three methods. The scatters indicated the average distance across tasks for each subject. The Euclidean distance can also be interpreted as the tracking error of the robotic arm in following the cursor's movements. The numeric values were within an acceptable range, i.e., [0.044 0.10] for the proposed method, [0.050 0.22] for DAV control, and [0.063 0.36] for DV control. It indicated that the robotic arm largely followed the cursor's movements. Additionally, Fig. 6.6 demonstrates that the proposed method resulted in improved accuracy of the robotic arm in following the projected positions for most subjects. The average Euclidean distances across subjects were 0.067, 0.088, and 0.13 for the three methods, respectively.

Before the experiment, subjects were informed that when the cursor is far away from the target, they should concentrate on the high-brightness stimulus to accelerate the cursor/robotic arm. Therefore, it helps decrease the reaching time. When the cursor is close to the target, the user should focus on a low-brightness stimulus to initiate cursor/robotic arm deceleration. This ensures that the cursor gets closer to the target center or does not deviate too much from its intended position. Consequently, the shorter distance between the center of the target position and the center of the cursor's final position validates the efficiency of the deceleration process.

Fig. 6.7 shows the cursor and robotic arm movements in a single-target reaching task performed by subject 8 using the three methods. Each circle represents a movement of the cursor/robotic arm. The interval between two circles represents the distance covered by two consecutive movements. Therefore, the density distribution of the dots also reflects velocity changes. Fig. 6.7(b) illustrates the actual position of the robotic arm, which was obtained through the robot sensor feedback. Fig. 6.8(a) showed the brain-actuated speeds of the three control methods. For the DAV and DV control, the constant value was determined by (6.13). The proposed method enabled the cursor to accelerate when it was distant from the target. When the cursor was tuning, there were a few speed reduction steps. However, users can achieve acceleration again by focusing on high-brightness stimuli. Fig. 6.8(b) and

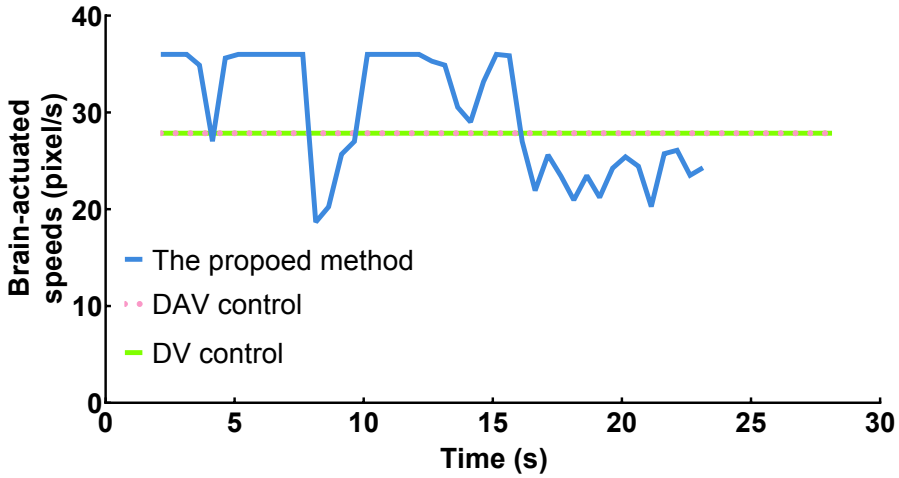


(a) Cursor movement

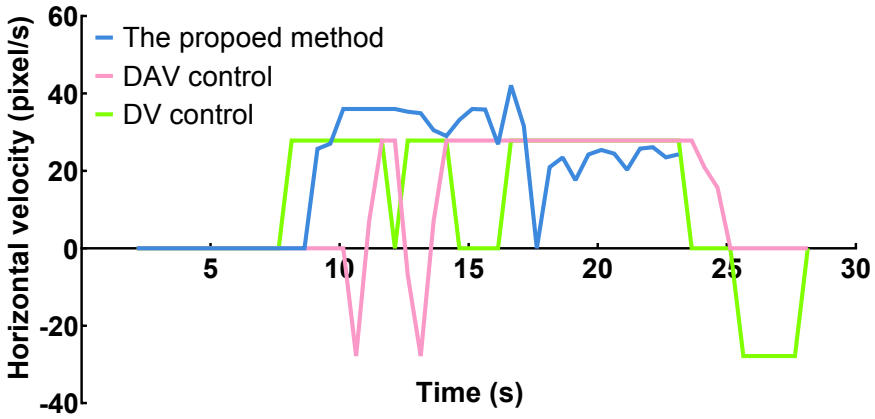


(b) Robotic arm movement

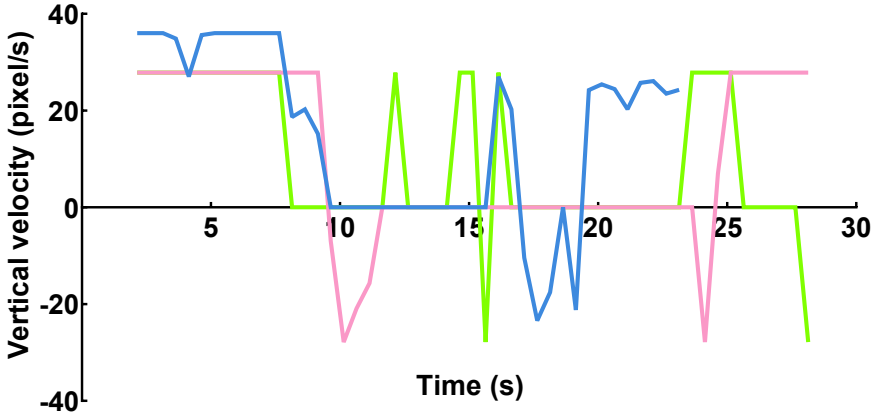
Figure 6.7: (a) Cursor and (b) robotic arm movements generated by the proposed method and two compared methods (i.e., DV and DAV) in a single-target reaching task. Each circle represents a movement. The red cross refers to the target. The interval between two circles represents the distance covered by two consecutive movements.



(a) The brain-actuated speeds



(b) The horizontal velocity



(c) The vertical velocity

Figure 6.8: (a) The brain-actuated speeds, (b) the horizontal velocity, and (c) the vertical velocity generated by the proposed method, DAV control, and DV control in a single-target reaching task.

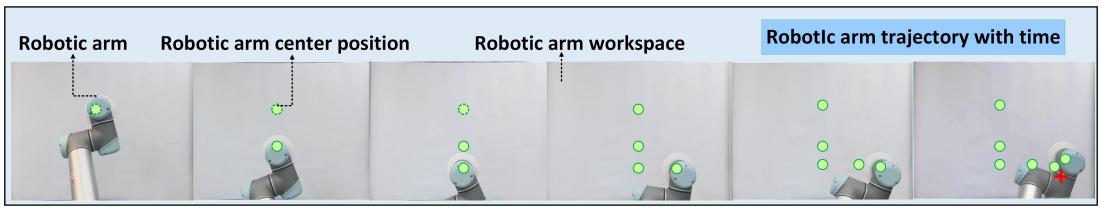
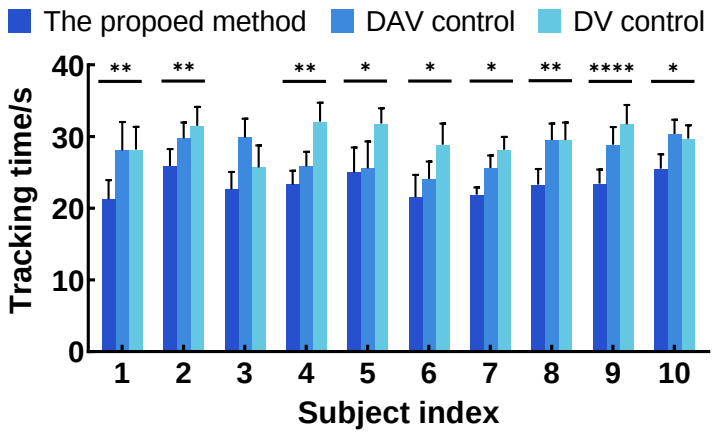
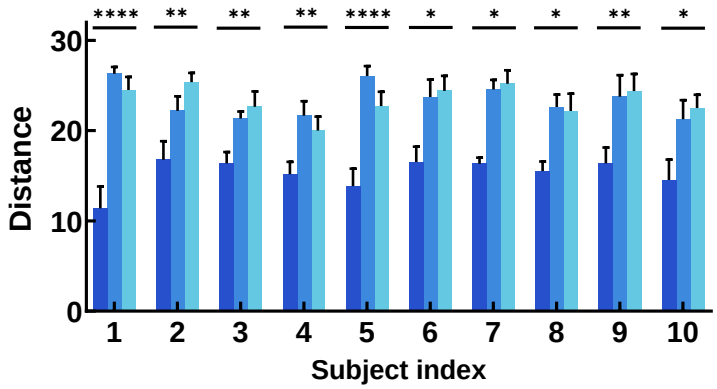


Figure 6.9: Robotic arm movements in a single-target reaching task generated by the proposed method. The green circle represents the center of the robotic arm, and the red cross denotes the center of the target.

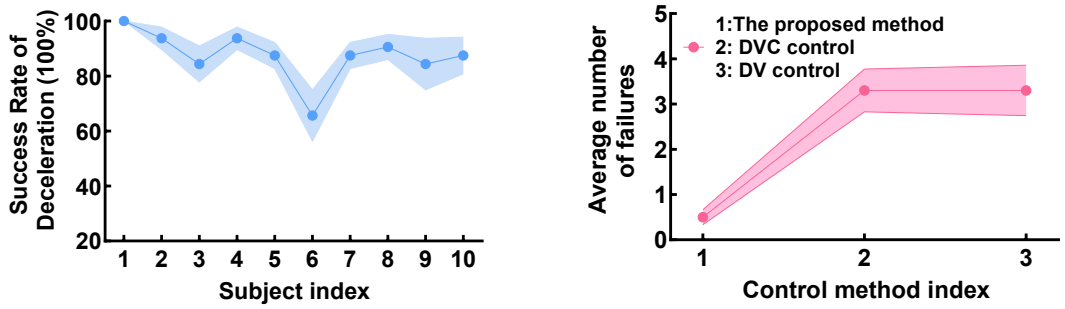


(a) Reaching time comparison in single-target tasks



(b) Distance comparison in single-target tasks

Figure 6.10: Performance comparison in single-target reaching tasks. The (a) average reaching time and (b) average distance were compared between the three methods. The distance was calculated between the centers of the target and the last cursor. The error bars are the standard error of the mean (SEM). The asterisks indicate a significant difference between the three methods obtained by one-way repeated-measures ANOVA ($*$: $p<0.05$, $**$: $p<0.01$, $***$: $p<0.001$, $****$: $p<0.0001$).



(a) Rate of successful deceleration

(b) Average number of failures

Figure 6.11: The (a) average rate of successful deceleration across tasks and (b) the average number of failures across subjects with different control methods. The SEM was shown as error envelopes.

Fig. 6.8(c) show the vertical and horizontal velocity comparisons. The proposed method (blue line) controlled the cursor to initially move downward, leading to higher vertical velocities. As it turned to the horizontal, the velocity in this direction also exceeded that of other methods. Finally, it achieved deceleration in the last few moments. Overall, this method exhibited a shorter reaching time compared to other methods. Fig. 6.9 showed the robotic arm's actual movements in the proposed method.

The reaching time of the three methods with various subjects were shown in Fig. 6.10(a). The values were averaged across tasks. It illustrated that the proposed method consistently achieved the shortest time for each subject. The reaching time of the proposed method improved that of DAV control by 0.56 s-7.25 s and DV control by 3.05 s-8.76 s. The average reaching time across subjects of the three methods was 23.37 s, 27.75 s, and 29.72 s, respectively. A one-way repeated-measures ANOVA was conducted to explore the similarity of reaching time across methods. The statistical results revealed significant differences in reaching time among the methods for most subjects. Additionally, Fig. 6.10(b) illustrated the distance between the centers of the target and the last cursor position for the three methods. The results indicated that the cursor controlled by the proposed method demonstrated greater proximity to the target across tasks. For example, the distance of the proposed method increased DAV control by 4.84 to 13.10 pixels and DV control by 4.96 to 14.91 pixels. The results of the statistical analysis indicated significant differences among

Table 6.1: Cursor's average distance of the last four positions in single-target reaching tasks

Subject index	Cursor's average distance across tasks (pixels)		
	DV control	DAV control	The proposed method
Subject 1	22.10	22.42	**15.52**
Subject 2	18.55	20.48	**15.95**
Subject 3	17.46	19.30	**15.77**
Subject 4	16.80	19.40	**16.36**
Subject 5	18.78	21.79	**16.72**
Subject 6	20.16	20.48	**16.36**
Subject 7	19.53	19.50	**16.33**
Subject 8	16.37	16.10	**14.42**
Subject 9	18.91	19.28	**16.93**
Subject 10	18.39	16.96	**14.33**
Average	18.70	19.57	**15.87**

the methods for each subject. The difference is particularly pronounced for subject 1 and subject 5, with a highly significant p-value ($P<0.0001$).

The rate of successful deceleration from the proposed method was shown in Fig. 6.11(a). It was calculated based on whether its speeds in the last four steps were lower than those in the DV/DAV methods. The values were averaged across tasks. The results revealed that as the cursor/robotic arm approached the target, most subjects completed successful decelerations. For example, subjects 1, 2, and 8 achieved success rates of 100%, 93.75%, and 90.62%, respectively. The average rate across subjects was 87.5%. It further demonstrated the efficacy of stimulus brightness-based velocity control. TABLE.6.1 and TABLE. 6.2 displayed the cursor's and robotic arm's average distances across tasks for the last four positions. The cursor's average distances across subjects of the proposed method, DAV control, and DV control method were 15.87, 19.57, and 18.70 pixels, respectively. Meanwhile, the average distances across subjects of the three methods provided by the robotic arm were 1.670, 2.163, and 2.552 cm, respectively. The compared methods achieved similar performance, while the proposed method generally provided the shortest distances. The results indicated that when utilizing the proposed method, the cursor and robotic arm tended to get closer or avoid excessive misses in the last four steps. The number of failures before the subject successfully conducted each task was counted in Fig. 6.11(b). The values were summed

Table 6.2: Robotic arm's average distance of the last four positions in single-target reaching tasks

Subject index	Robotic arm's average distance across tasks (cm)		
	DV control	DAV control	The proposed method
Subject 1	2.277	2.726	**1.960**
Subject 2	**1.701**	2.279	1.914
Subject 3	2.175	2.079	**1.501**
Subject 4	5.641	1.982	**1.520**
Subject 5	2.217	1.729	**1.639**
Subject 6	2.293	2.111	**1.549**
Subject 7	2.193	1.749	**1.608**
Subject 8	2.039	1.696	**1.472**
Subject 9	1.920	1.960	**1.488**
Subject 10	3.064	3.317	**2.048**
Average	2.552	2.163	**1.670**

across tasks and averaged across subjects. The results demonstrated that the proposed method exhibits fewer failures, primarily due to its velocity control, which allows it to reach the target within a limited time.

6.3.3 Multi-target Reaching Task Performance Evaluation

Fig. 6.12 shows the Euclidean distance between the projected positions of the robotic arm based on cursor movements and its actual arrival positions for the three methods. It can represent the cumulative tracking error when the robotic arm moves with the cursor. The error ranges exhibited by different subjects are [0.11 0.35] for the proposed method, [0.25 0.60] for the DAV control, and [0.26 0.59] for the DV control, respectively. The proposed method provides relatively small tracking errors for most subjects. It can be attributed to the fact that the velocity direction of the robotic arm does not change suddenly like in DV control, resulting in a smoother trajectory. Besides, it provides a shorter tracking time compared to DAV control.

Fig. 6.13 shows the cursor and robotic arm movements generated by the three methods in a multi-target reaching task of subject 6. Three targets were generated randomly. The targets remain consistent across different methods for fair comparisons. Fig. 6.14(a) displayed the brain-actuated speeds of the methods. DAV and DV methods controlled the cursor or

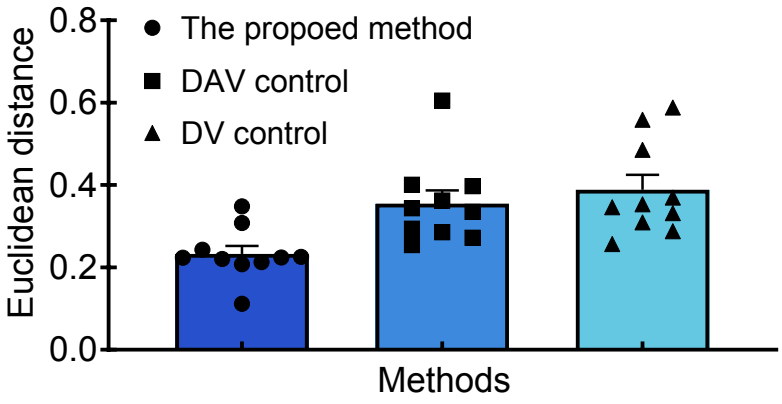
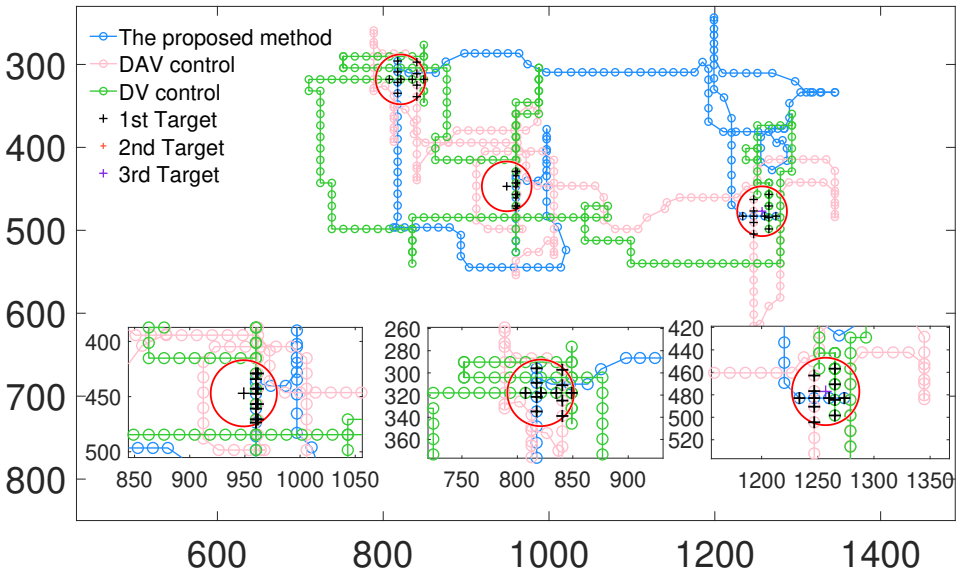


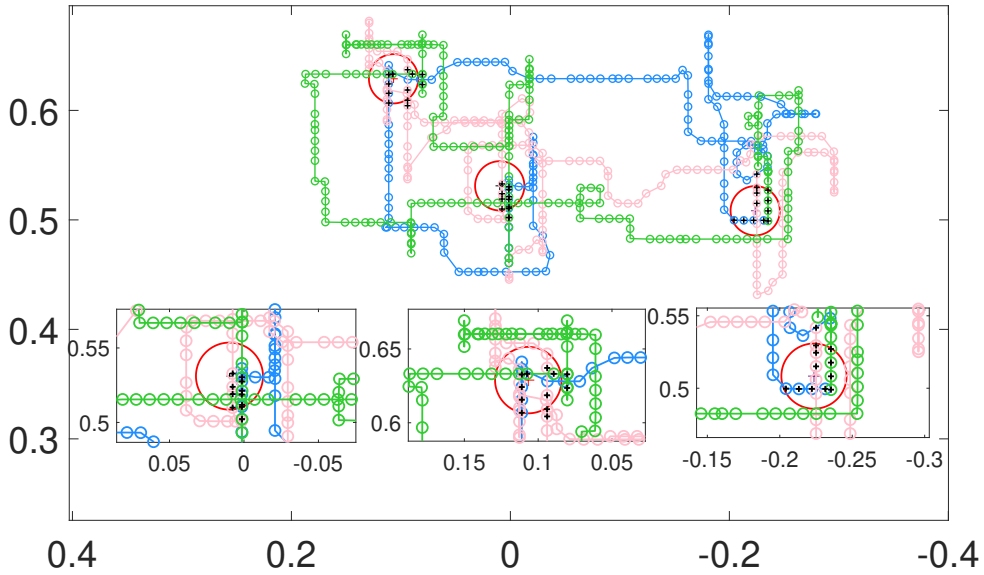
Figure 6.12: The Euclidean distance between the robotic arms' positions projected by the cursor's movements and its actual arrival positions in multi-target reaching tasks. The scatter points refer to the Euclidean distances provided by different subjects.

robotic arm with a constant speed regardless of distance from the target. The speed of the proposed method was subject-driven and relied on distance. For example, since the first target was near the cursor's origin, the cursor accelerated briefly and then approached the target at a slower speed. The speeds of the last four steps were marked in red. The third target was far away from the second target, so the cursor experienced a long acceleration period and then decelerated as it approached the target. The reaching time of the proposed method (74.82 s) was shorter than others, i.e., DAV control: 88.71 s; DV control: 99.45 s. Fig. 6.14(b) and Fig. 6.14(c) showed the horizontal and vertical velocities, which offer a more comprehensive view of velocity changes for each method.

The reaching time comparisons of multi-target reaching tasks among these methods were shown in TABLE. 6.3. The proposed method consistently achieved the shortest reaching time. For each multi-target trial, three single tasks were included. To further evaluate its performance, Fig. 6.15 presented the reaching time and distance comparisons for each single task. The values were averaged across tasks. The results show that the proposed method provided the shortest reaching time and distance compared with other methods. The average reaching time across different subjects of the three methods were 28.34 s, 35.81 s, and 47.07 s, respectively. In addition, the average distances between the final position of the cursor and the target were 16.78, 21.64, and 23.60 pixels for the three methods,

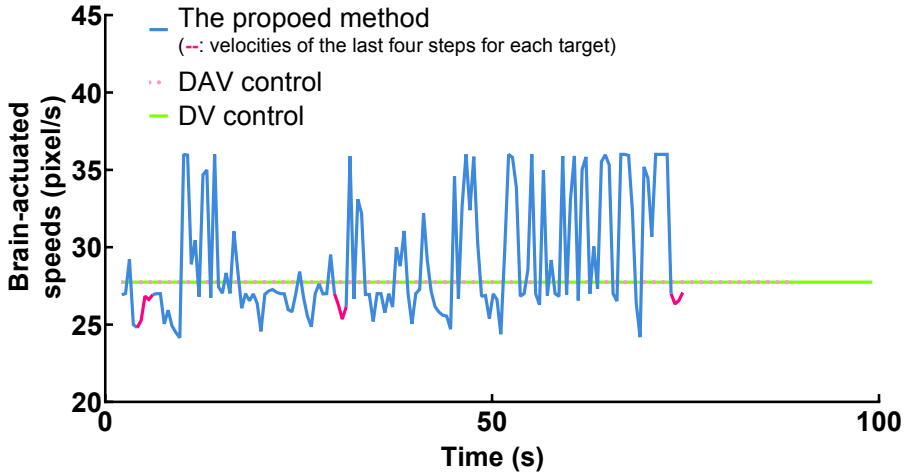


(a) Cursor movement

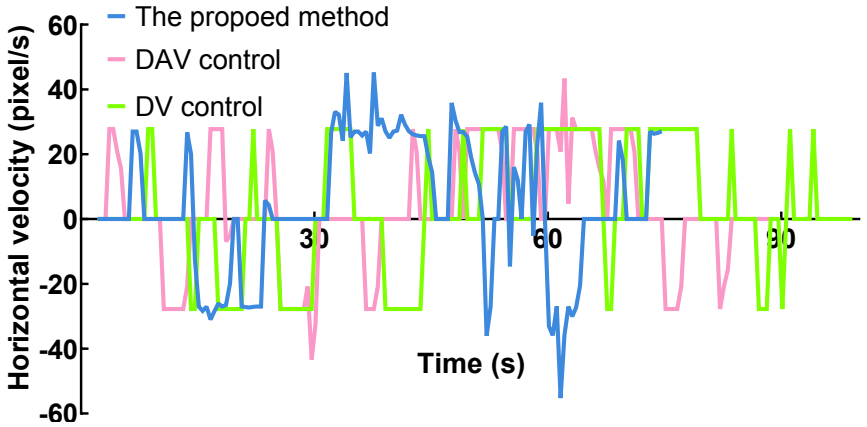


(b) Robotic arm movement

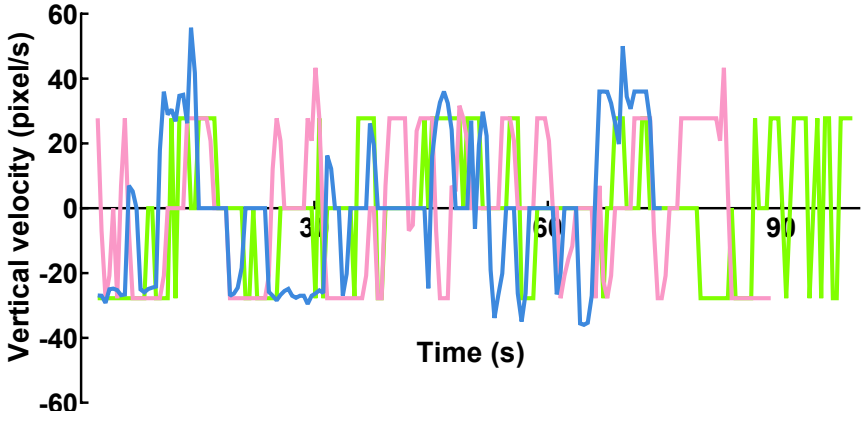
Figure 6.13: (a) Cursor and (b) robotic arm movements generated by the proposed method, DAV control, and DV control in a multi-target reaching task. Each circle represents a movement. Three targets are represented as crosses in different colors. The interval between two circles represents the distance covered by two consecutive movements.



(a) The brain-actuated speeds



(b) The horizontal velocity



(c) The vertical velocity

Figure 6.14: (a) The brain-actuated speeds (b) the horizontal velocity, and (c) the vertical velocity generated by the proposed method, DAV control, and DV control in a multi-target reaching task. The red short lines in (a) refer to the speed of the last four steps for each target.

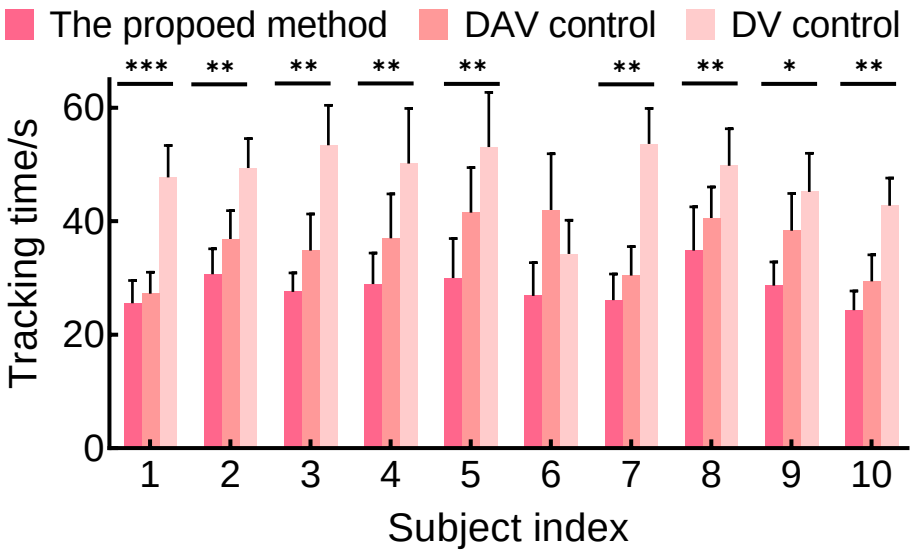
Table 6.3: Average reaching time of three methods in multi-target reaching tasks

Subject index	Average reaching time across trials (s)		
	DV control	DAV control	The proposed method
Subject 1	143.34	81.69	**76.93**
Subject 2	148.33	110.98	**92.17**
Subject 3	160.49	104.78	**82.85**
Subject 4	150.80	111.42	**87.08**
Subject 5	159.27	124.75	**90.28**
Subject 6	102.94	126.07	**80.65**
Subject 7	135.90	91.68	**78.53**
Subject 8	149.32	121.92	**104.61**
Subject 9	135.78	115.29	**86.20**
Subject 10	128.51	88.35	**73.35**
Average	141.47	107.70	**85.27**

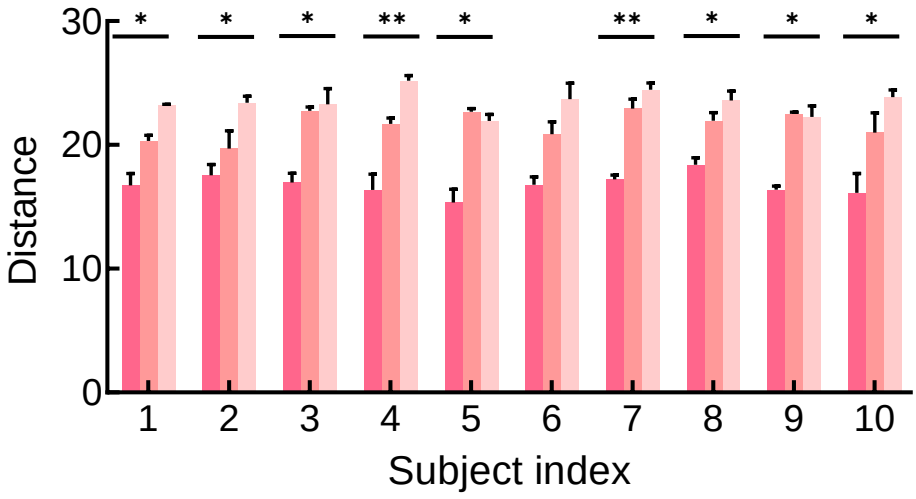
respectively. To investigate the similarity of the average reaching time and distance across different methods, a one-way repeated-measures ANOVA was conducted. The statistical analysis revealed significant differences in both performance indicators among these methods for most subjects.

The cursor's and robotic arm's average distances across tasks for the last four positions were given in the TABLE.6.4 and TABLE.6.5. The results indicated that the proposed method generally achieved the shortest distance for both the cursor and the robotic arm. Specifically, the cursor's average distances across subjects of DV, DAV, and the proposed method were 19.20, 17.95, and 16.62 pixels, respectively. Meanwhile, the robotic arm's average distances across subjects were 3.484, 3.450, and 2.740 cm, respectively. The outcomes demonstrated that the proposed method achieves a higher degree of precision in reaching the target position.

The average success rate of deceleration across tasks was shown in Fig. 6.16(a). Most subjects achieved a high rate of successful deceleration. The average rate across subjects was 76.39%. The numbers of failed trials of the three methods were shown in Fig. 6.16(b). The figure indicated that subjects encountered fewer failures when utilizing the proposed method to control the cursor/robotic arm.



(a) Reaching time comparison in multi-target reaching tasks



(b) Distance comparison in multi-target reaching tasks

Figure 6.15: Performance comparison in multi-target reaching tasks. The (a) average reaching time and (b) average distance were compared between the three methods. The distance was calculated between the centers of the target and the last cursor position. The error bars represent SEM. The asterisks indicate a significant difference between the three methods obtained by one-way repeated-measures ANOVA ($*$: $p<0.05$, $**$: $p<0.01$, $***$: $p<0.001$).

Table 6.4: Cursor's average distance of the last four positions in multi-target reaching tasks

Subject index	Cursor's average distance across tasks (pixels)		
	DV control	DAV control	The proposed method
Subject 1	19.71	**17.04**	18.64
Subject 2	20.15	17.63	**16.13**
Subject 3	20.17	19.20	**16.15**
Subject 4	19.64	16.78	**15.62**
Subject 5	18.31	17.80	**16.48**
Subject 6	18.38	19.36	**16.86**
Subject 7	20.19	17.32	**16.57**
Subject 8	18.07	**16.20**	18.25
Subject 9	18.42	20.05	**16.47**
Subject 10	18.91	18.15	**14.98**
Average	19.20	17.95	**16.62**

6.4 Discussion

6.4.1 Model's Performance

Currently, most brain-controlled robotic platforms use discrete movement commands and constant velocity [144]. Studies involving velocity modulation explored the effects of factors, such as attention, distance, and frequency [164, 232, 234]. Differently, this study focused on stimulus brightness-based velocity modulation. The proposed method was used to control the cursor and robotic arm simultaneously to track the randomly generated target. The velocities in Fig. 6.8 and Fig. 6.14 both depicted that the proposed method achieved acceleration as the cursor was far away from the target and deceleration when their distance was small. The acceleration process helped reach the target more quickly. For example, the reaching time of the proposed method, DAV control, and DV control were 74.82 s, 88.71 s, and 99.45 s in Fig. 6.14. The deceleration process assisted the robotic arm in getting closer to the target or decreasing overshooting. For instance, the DV method exhibited a shorter distance (21.84 versus 24.10 pixels) between the cursor and the target when it initially entered the third hit area (i.e., the red circle) in Fig. 6.13. However, the cursor controlled by the DV method gradually moved away from the target due to the fixed speed, increasing the distance over time (22.90 versus 17.65 pixels).

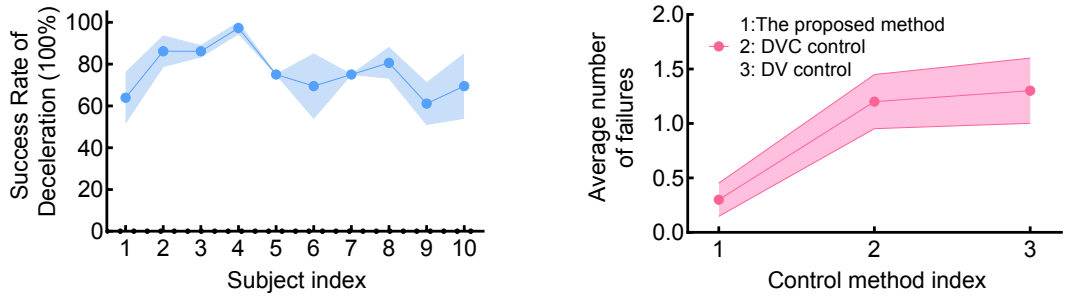
Table 6.5: Robotic arm's average distance of the last four positions in multi-target reaching tasks

Subject index	Robotic arm's average distance across tasks (cm)		
	DV control	DAV control	The proposed method
Subject 1	**1.967**	4.187	2.966
Subject 2	**1.957**	3.897	4.544
Subject 3	3.392	3.820	**3.080**
Subject 4	3.116	2.277	**2.189**
Subject 5	4.442	4.516	**2.198**
Subject 6	**2.157**	2.460	2.581
Subject 7	4.999	2.285	**1.479**
Subject 8	3.220	2.674	**2.263**
Subject 9	4.396	5.698	**3.486**
Subject 10	5.190	2.691	**2.608**
Average	3.484	3.450	**2.740**

In the single-target task, the cursor's maximum distance to the target is 840 pixels. The slowest DV method would reach this position in about 32 s if there were no misclassifications. However, considering the occurrence of misclassifications, the task duration is set to 40 s to allow for extra correction time. For each multi-target reaching task, there are three single-reaching tasks. The first task is the same as the single-target reaching task, while in the other two tasks, the cursor starts away from the center, doubling the furthest distance. Consequently, the total duration for the multi-target reaching task is calculated to be 200 s.

6.4.2 Velocity Modulation Design

As described in (6.11), the velocity determination in the proposed method considered previous moment information when the classification result changed. The reasons are as follows. Firstly, occasional changes in velocity direction may be misclassified for various reasons, such as user fatigue and background noise. The proposed method preserved a part of the speed in the previous direction and then combined it with the speed in the new direction to determine the final movement of the cursor/robotic arm. If the classification was wrong, the proposed method can move toward the target while adjusting for the error. If the same direction was obtained again, it was considered the subject's intentional behavior, and the



(a) Success rate of deceleration

(b) Average number of failures

Figure 6.16: The (a) average success rate of deceleration across tasks and (b) average number of failures across various subjects with different control methods. The SEM was shown as error envelopes.

previous velocity would be overwritten without keeping the previous information. The DV method lengthens the experiment time since it is limited to adjusting erroneous directions individually and cannot simultaneously move toward the target. Secondly, discrete changes in direction have adverse effects on practical implementation. The robotic arm will experience mechanical stress and exhibit jerky movements. As shown in Fig. 6.6 and Fig. 6.12, the DV control method produced a larger tracking error in two kinds of reaching tasks with average values across subjects of 0.13 and 0.39. In comparison, the proposed method achieved lower tracking errors with average values of 0.067 and 0.23. Rapid, discrete direction changes cause discomfort when users conduct reaching tasks assisted by the robotic arm, especially for individuals with mobility limitations or during rehabilitation exercises. The proposed method allows smooth movements of the robotic arm.

Compared to the DAV method, the proposed stimulus brightness-based velocity modulation saves time in reaching tasks. The protocol also reflects the subjects' practical behaviors. While observing a distant object, individuals tend to increase their speed, whereas when nearing the object, they tend to slow down. It can also improve subjects' motivation and participation in conducting various activities, such as assistance and rehabilitation. Intentional participation is important to enhance brain plasticity, thereby increasing the chances of motor recovery [231].

In this study, a third-order polynomial was used to fit the speed function. The curve

demonstrates an S-shaped behavior as the posterior probability difference changes. It is characterized by an initial increase, followed by a smooth transitional stage, and finally an increase. This indicates that the model outputs at higher speeds when the probability of the SSVEP signal being excited by a high-brightness stimulus significantly exceeds that of a low-brightness stimulus. Otherwise, the model outputs at a slower speed. As the probability difference approaches the training threshold, it becomes challenging to determine the signal evoked by which brightness stimulus. To address this uncertainty, a middle speed should be assigned, corresponding to a third-order polynomial that features a smooth transitional stage in the middle. Meanwhile, this setting aligns with practical scenarios. In Fig. 6.5(b), the average brightness classification accuracy was 69.1% across subjects. Therefore, assigning speeds around the median aims to strike a balance in cases where the model's uncertainty arises due to the proximity of probabilities to the threshold.

6.4.3 Feature Extraction

In this study, the correlation coefficient and PSD value were integrated into the feature vector. The correlation coefficient was provided by the CCA recognition method. One reason for choosing CCA is due to its simple implementation and low computational complexity [69]. Additionally, it offered sufficient features to train probability density functions and fine-tune hyperparameters in the offline experiment via the sliding window. It also showed satisfactory performance in the offline experiment's four-class problem. The average recognition accuracy across subjects is 90.5%. Thus, it can be a valid choice for the online experiment.

6.5 Conclusion

In this study, a stimulus brightness-based velocity modulation method was proposed for robotic arm control in the SSVEP-based BCI system. The flickers with different frequencies and brightnesses were employed to achieve velocity modulation. The feature vector was constructed from the correlation coefficient and PSD. The GMM model and Bayesian inference were then used to calculate the posterior probabilities that the signal came from a

high- and low-brightness flicker. The speed function was designed using the posterior probability difference, and the velocity from the previous moment was incorporated to derive the final direction and speed. For performance comparison, two velocity control methods were included. The effectiveness and feasibility of the proposed method were demonstrated via online experiments involving single- and multi-target reaching tasks.

Chapter 7

Conclusion and Future Work

This chapter summarises the research works and contributions of this thesis. Besides, the limitations of this work and potential future development in the SSVEP-based BCI are also included.

7.1 Conclusion

This thesis commences with an introductory chapter that offers readers an overview of the research's significance of the SSVEP-based BCI and provides an outline of the subsequent chapters. To better understand the development of the SSVEP-based BCI, studies about SSVEP signal pre-processing methods, recognition methods, classification confidence analysis, transfer learning methods, and SSVEP-based BCI-controlled robots are reviewed in Chapter 2. Following the objectives mentioned in Chapter 1, this thesis presented outcomes in the SSVEP-based BCI, mainly focusing on the improvement of recognition accuracy and reliability, the building of an effective cross-subject transfer learning scheme, and the velocity modulation method in the robotic arm control. Technical achievements of this thesis can be concluded as follows:

1. Multi-objective optimisation-based high-pass spatial filtering method for improving the SSVEP detection accuracy: Recognition accuracy is a critical aspect of SSVEP-based BCI systems. Higher accuracy means more precise control of external devices

through the BCI. It can enhance the user experience and system efficiency. To this end, Chapter 3 proposed a multi-objective optimisation-based high-pass spatial filtering method for SSVEP recognition performance enhancement. Specifically, the derivation of the filter involved maximizing the correlation between the training signal and the individual template corresponding to the same target, while minimizing the correlation between signals from other targets and the template. Furthermore, a constraint was introduced requiring the sum of spatial filter elements to be zero. This setup allows high spatial-frequency SSVEP signals to pass through while attenuating low spatial-frequency signals. Consequently, the proposed approach demonstrated the potential to extract features relevant to the target, reject irrelevant information, and mitigate the effects of volume conduction simultaneously.

2. Bayesian-based classification confidence estimation method to enhance recognition reliability: By analyzing the classification confidence of SSVEP signals, the BCI system can avoid making incorrect decisions when the confidence level is low. This leads to a significant reduction in errors, which is critical for improving the reliability of the BCI and ensuring the safety of the subject. Chapter 4 proposed a Bayesian-based classification confidence estimation method to solve this issue. In this study, the feature vector was defined based on the differences between the largest correlation coefficient and other values. The probability density functions of feature vectors for correct and wrong classifications were estimated using the GMM model. During the test process, Bayesian inference was applied to calculate the posterior probabilities of accurate and wrong recognition using the newly obtained feature vector. The CCValue, representing the difference between these two probabilities, served as a measure of classification confidence. The decision-making process utilized the CCValue to determine whether to accept trustworthy results or reject uncertain ones. By rejecting uncertain results, the proposed method may enhance the safety of various practical applications, such as assistive and rehabilitative devices.

3. Cross-subject transfer learning for boosting recognition performance of the target subject: One of the major limitations of SSVEP-based BCI systems is the long cali-

bration time required. It necessitates collecting a substantial volume of training data to train the system's parameters for each individual user. Direct data transfer between users is not possible due to the presence of non-stationarity. Chapter 5 proposed a cross-subject scheme to transfer SSVEP knowledge from source subjects to effectively strengthen the recognition performance for the target subject. Under this scheme, an informative multidimensional feature vector is created partly by the transferred spatial filter and the transferred SSVEP template from the source subject, and partly by the spatial filter of the target subject through multiple covariance maximization. Additionally, a contribution score is incorporated for each source subject, based on the distance between the source subject and the target subject. This study helps reduce the calibration time required for each individual user. It can also lead to improved recognition performance of the target subject by transferring SSVEP data from the source subject. These advancements have the potential to enhance the efficiency and practicality of the BCI system, as well as user adaptability.

4. SSVEP-based brain-computer interface controlled robotic platform with velocity modulation: Robotic systems with velocity control play a pivotal role in SSVEP-based BCIs by facilitating more accurate and natural interactions with the environment. This is particularly important in tasks that require delicate and precise movements. Meanwhile, it allows the robotic system to adjust its movement speed according to the user's intention, making the interaction more intuitive and personalized. Chapter 6 introduced a velocity modulation method for controlling the robotic arm in the SSVEP-based BCI. In this chapter, a stimulation interface was created, comprising flickers, a target, and a cursor workspace, enabling synchronized movement of the cursor and robotic arm. This design eliminates the need for subjects to shift their gaze between the stimulus and the robot, thus enhancing the user experience. A stimulus brightness-based velocity modulation model was constructed using the correlation coefficient and power spectral density. Additionally, a speed function was devised, utilizing the difference in posterior probabilities between high- and low-brightness flickers. The velocity determination incorporates historical velocity data to avoid occasional

misclassification, ensuring smoother and more reliable robotic movement control.

7.2 Future Work

7.2.1 Hybrid BCIs

One key characteristic of SSVEP-based BCI is the requirement of constant attention to the light source, which may be challenging and bothersome for people. A promising approach to address this issue is through Hybrid BCIs, which enhance the quality of BCI systems by combining two or more BCI paradigms [239]. To be specific, in hybrid paradigms, the number of control commands can be increased through simultaneous decoding of multiple brain activities [240]. For example, in a Tetris game study conducted by Wang et al. [241], the rotating command requires a continuous gaze of visual stimulus to evoke SSVEP signals. Meanwhile, motor imagery (MI) is employed to generate two control commands, enabling the movement of bricks towards the left and right. This multi-modal approach effectively mitigates the discomfort associated with prolonged gazing stimuli. Additionally, hybrid BCIs have the potential to enhance the accuracy of system classification. For instance, Wang *et al.* [242] designed a novel hybrid paradigm (shape-changing and flickering-hybrid) based on P300 and SSVEP, which demonstrated improved performance for certain subjects. The research on hybrid BCIs has been increasing in recent years, but there is a need for further commercialization of portable, wearable, and low-cost products suitable for public use [240]. Additionally, many SSVEP-based hybrid systems currently utilize the standard CCA as the target detection method [241–243]. However, higher performance can be achieved by incorporating advanced signal analysis methods, as illustrated in Section 2.2.

7.2.2 Amplitude Modulation

In general, SSVEP-based BCI systems primarily emphasize frequency and phase features, but several studies have specifically explored amplitude modulation [244]. Sharma et al. [245] proposed a setup that utilizes amplitude modulation of SSVEP through eye accommodation, enabling the generation of multiple commands from a single flickering stimulus. Autthasan *et al.* [236] pointed out that the SSVEP amplitude changes as a function of

stimulus luminance contrast and then proposed an integrated architecture to predict the frequency and contrast-related amplitude modulations of the SSVEP signal simultaneously. Moreover, attention generally enhances rhythmic brain responses at the frequency of the stimulus. For example, Gulbinaite *et al.* [246] explored the effect of attention on the amplitude of SSVEPs in a wide range of temporal frequencies (3-80 Hz). The research results showed that such influence is frequency-dependent, namely different flicker frequency bands like theta, gamma, and alpha have various relationships with amplitudes. Considering the eye fatigue problem, Chang et al. [244] introduced an amplitude-modulated visual stimulation that combines the benefits of both low-frequency SSVEPs, such as high amplitude and low BCI illiteracy, and high-frequency SSVEPs, such as reduced eye fatigue and a lower risk of epileptic seizures. Despite achieving good performance in research settings, these approaches are seldom incorporated into practical scenarios. However, considering amplitude as a crucial feature of SSVEP signals, it has the potential to enhance the precise speed control of robots. The reduced eye fatigue and increased control options provided by amplitude modulation contribute to a more pleasant and comfortable user experience, making the SSVEP-based BCI system more practical and user-friendly.

7.2.3 Neural Plasticity

Patients with impaired motor function resulting from neuronal injuries require proper rehabilitation. Successful rehabilitation is often accompanied by neuronal reorganization, known as neural plasticity, which refers to the brain's inherent ability to reorganize its function and structure in response to training experiences [247]. For instance, several studies [154, 224] have demonstrated that BCI-based rehabilitation has a positive effect on post-stroke hand recovery, which is often attributed to the promotion of neural plasticity. The attentive participation of patients [248] stands as another crucial factor influencing rehabilitation programs. Many studies have highlighted the critical role of attention as a modulator of plasticity [249]. They emphasize that sustaining attention during motor exercises enhances neural plasticity and motor learning, directly influencing the overall outcome of patient rehabilitation [250]. In Chapter 6, a user-driven velocity modulation method based on SSVEP was introduced for controlling the robotic arm. By directing their attention to either high-

brightness or low-brightness stimuli, the user can adjust the speed of the robotic arm. Compared to conventional fixed-speed control methods in the SSVEP paradigm, this method provides users with higher levels of autonomy and engagement. This more comprehensive user-driven method may be conducive to neural plasticity. Therefore, future research should focus on integrating both the user's attention and intention into the SSVEP-based BCI to facilitate improved rehabilitation outcomes.

Bibliography

[1] 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.

[2] T. Tanaka and M. Arvaneh, *Signal processing and machine learning for brain-machine interfaces*. Institution of Engineering & Technology, 2018.

[3] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan, "Brain-computer interface technology: a review of the first international meeting," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 164–173, 2000.

[4] Y. Wang, X. Gao, B. Hong, and S. Gao, "Practical designs of brain–computer interfaces based on the modulation of eeg rhythms," *Brain-computer interfaces: revolutionizing human-computer interaction*, pp. 137–154, 2010.

[5] A. Nijholt, D. Tan, G. Pfurtscheller, C. Brunner, J. d. R. Millán, B. Allison, B. Graimann, F. Popescu, B. Blankertz, and K.-R. Müller, "Brain-computer interfacing for intelligent systems," *IEEE Intell. Syst.*, vol. 23, no. 3, pp. 72–79, 2008.

[6] U. Chaudhary, N. Birbaumer, and A. Ramos-Murguialday, "Brain–computer interfaces for communication and rehabilitation," *Nat. Rev. Neurol.*, vol. 12, no. 9, p. 513, 2016.

[7] X. Deng, Z. L. Yu, C. Lin, Z. Gu, and Y. Li, "A bayesian shared control approach for wheelchair robot with brain machine interface," *IEEE Trans. Neural Syst. Rehabil.*

Eng., 2019.

[8] J. J. Podmore, T. P. Breckon, N. K. 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, 2019.

[9] D. Tan and A. Nijholt, *Brain-computer interfaces and human-computer interaction.* Springer, 2010.

[10] N. Birbaumer, "Breaking the silence: brain–computer interfaces (bci) for communication and motor control," *Psychophysiology*, vol. 43, no. 6, pp. 517–532, 2006.

[11] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update," *J. Neural Eng.*, vol. 15, no. 3, p. 031005, 2018.

[12] N. Naseer and K.-S. Hong, "fNIRS-based brain-computer interfaces: a review," *Front. Hum. Neurosci.*, vol. 9, p. 3, 2015.

[13] M. Lührs, B. Sorger, R. Goebel, and F. Esposito, "Automated selection of brain regions for real-time fMRI brain–computer interfaces," *J. Neural Eng.*, vol. 14, no. 1, p. 016004, 2016.

[14] Y.-H. Chen, J. Saby, E. Kuschner, W. Gaetz, J. C. Edgar, and T. P. Roberts, "Magnetoencephalography and the infant brain," *NeuroImage*, 2019.

[15] F. Wallois, M. Mahmoudzadeh, A. Patil, and R. Grebe, "Usefulness of simultaneous eeg–nirs recording in language studies," *Brain and language*, vol. 121, no. 2, pp. 110–123, 2012.

[16] T. Liu, M. Pelowski, C. Pang, Y. Zhou, and J. Cai, "Near-infrared spectroscopy as a tool for driving research," *Ergonomics*, vol. 59, no. 3, pp. 368–379, 2016.

[17] T. J. Sejnowski, P. S. Churchland, and J. A. Movshon, "Putting big data to good use in neuroscience," *Nature neuroscience*, vol. 17, no. 11, pp. 1440–1441, 2014.

[18] P. Sawangjai, S. Hompoonsup, P. Leelaarporn, S. Kongwudhikunakorn, and T. Wilaiprasitporn, "Consumer grade EEG measuring sensors as research tools: A review," *IEEE Sensors J.*, vol. 20, no. 8, pp. 3996–4024, 2019.

[19] I. Lazarou, S. Nikolopoulos, P. C. Petrantonakis, I. Kompatsiaris, and M. Tsolaki, "EEG-based brain–computer interfaces for communication and rehabilitation of people with motor impairment: a novel approach of the 21st century," *Front. Hum. Neurosci.*, vol. 12, p. 14, 2018.

[20] 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," *J. Neural Eng.*, vol. 6, no. 4, p. 046002, 2009.

[21] R. Zerafa, T. Camilleri, O. Falzon, and K. P. Camilleri, "To train or not to train? A survey on training of feature extraction methods for SSVEP-based BCIs," *J. Neural Eng.*, vol. 15, no. 5, p. 051001, 2018.

[22] F.-B. Vialatte, M. Maurice, J. Dauwels, and A. Cichocki, "Steady-state visually evoked potentials: focus on essential paradigms and future perspectives," *Prog. Neurobiol.*, vol. 90, no. 4, pp. 418–438, 2010.

[23] 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.

[24] Y. Jiao, Y. Zhang, Y. Wang, B. Wang, J. Jin, and X. Wang, "A novel multilayer correlation maximization model for improving CCA-based frequency recognition in SSVEP brain–computer interface," *Int. J. Neural Syst.*, vol. 28, no. 04, p. 1750039, 2018.

[25] P. F. Diez, V. A. Mut, E. M. A. Perona, and E. L. Leber, "Asynchronous BCI control using high-frequency SSVEP," *J. Neuroeng. Rehabil.*, vol. 8, no. 1, p. 39, 2011.

[26] L. Niu, J. Bin, J. kong shuai Wang, G. Zhan, L. Zhang, Z. Gan, and X. Kang, "A dynamically optimized time-window length for SSVEP based hybrid BCI-VR system,"

Biomed. Signal. Proces., vol. 84, p. 104826, 2023.

[27] Y.-J. Chen, A. R. A. See, and S.-C. Chen, "SSVEP-based BCI classification using power cepstrum analysis," *Electron. Lett.*, vol. 50, no. 10, pp. 735–737, 2014.

[28] G. Hakvoort, B. Reuderink, and M. Obbink, "Comparison of PSDA and CCA detection methods in a SSVEP-based BCI-system," *Centre for Telematics & Information Technology University of Twente*, 2011.

[29] W. Yan, Y. Wu, G. Xu, and C. Du, "Cross-subject spatial filter transfer method for SSVEP-EEG feature recognition," *J. Neural Eng.*, vol. 19, no. 3, p. 036008, 2022.

[30] C. M. Wong, Z. Wang, B. Wang, K. F. Lao, A. Rosa, P. Xu, T.-P. Jung, C. P. Chen, and F. Wan, "Inter-and intra-subject transfer reduces calibration effort for high-speed SSVEP-based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 10, pp. 2123–2135, 2020.

[31] D. Wu, Y. Xu, and B.-L. Lu, "Transfer learning for EEG-based brain-computer interfaces: A review of progress made since 2016," *IEEE Trans. Cogn.*, vol. 14, no. 1, pp. 4–19, 2020.

[32] M. Kołodziej, A. Majkowski, Ł. Oskwarek, and R. J. Rak, "Comparison of EEG signal preprocessing methods for SSVEP recognition," in *Proc. 39th Int. Conf. Telecommun. Signal.* IEEE, 2016, pp. 340–345.

[33] Q. Liu, Y. Jiao, Y. Miao, C. Zuo, X. Wang, A. Cichocki, and J. Jin, "Efficient representations of EEG signals for SSVEP frequency recognition based on deep multiset CCA," *Neurocomputing*, vol. 378, pp. 36–44, 2020.

[34] Q. Liu, K. Chen, Q. Ai, and S. Q. Xie, "Recent development of signal processing algorithms for SSVEP-based brain computer interfaces," *J. Med. Biol. Eng.*, vol. 34, no. 4, pp. 299–309, 2014.

[35] Q. Ai, Q. Liu, W. Meng, and S. Q. Xie, *Advanced rehabilitative technology: neural interfaces and devices.* Academic Press, 2018.

[36] S. Leske and S. S. Dalal, "Reducing power line noise in EEG and MEG data via spectrum interpolation," *Neuroimage*, vol. 189, pp. 763–776, 2019.

[37] S.-C. Chen, Y.-J. Chen, I. A. Zaeni, and C.-M. Wu, "A single-channel SSVEP-based BCI with a fuzzy feature threshold algorithm in a maze game," *Int. J. Fuzzy. Syst.*, vol. 19, no. 2, pp. 553–565, 2017.

[38] 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, 2006.

[39] X. Gao, D. Xu, M. Cheng, and S. Gao, "A BCI-based environmental controller for the motion-disabled," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 137–140, 2003.

[40] Y. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "SSVEP recognition using common feature analysis in brain–computer interface," *J. Neurosci. Methods.*, vol. 244, pp. 8–15, 2015.

[41] Z. Wu and D. Yao, "Frequency detection with stability coefficient for steady-state visual evoked potential (SSVEP)-based BCIs," *J. Neural Eng.*, vol. 5, no. 1, p. 36, 2007.

[42] Y. Wang, R. Wang, X. Gao, B. Hong, and S. Gao, "A practical VEP-based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 234–240, 2006.

[43] V. Pukhova, E. Gorelova, G. Ferrini, and S. Burnasheva, "Time-frequency representation of signals by wavelet transform," in *Proc. IEEE Conf. Russian. Young. Researchers. Elect. Electron. Eng.*, 2017, pp. 715–718.

[44] Z. Zhang, X. Li, and Z. Deng, "A CWT-based SSVEP classification method for brain-computer interface system," in *Proc. Int. Conf. Intell. Control. Inf. Process.* IEEE, 2010, pp. 43–48.

[45] I. Rejer, "Wavelet transform in detection of the subject specific frequencies for SSVEP-based BCI," in *Proc. Int. Multi-Conf. Adv. Comput. Syst.*, 2016, pp. 146–155.

[46] E. Sayilgan, Y. Yuce, and Y. Isler, "Investigating the effect of flickering frequency pair and mother wavelet selection in steady-state visually-evoked potentials on two-command brain-computer interfaces," *IRBM*, vol. 43, no. 6, pp. 594–603, 2022.

[47] R. Chen, G. Xu, X. Zhang, C. Han, and S. Zhang, "Multi-scale noise transfer and feature frequency detection in SSVEP based on FitzHugh–Nagumo neuron system," *J. Neural Eng.*, vol. 18, no. 5, p. 056054, 2021.

[48] V. P. Oikonomou, G. Liaros, K. Georgiadis, E. Chatzilari, K. Adam, S. Nikolopoulos, and I. Kompatsiaris, "Comparative evaluation of state-of-the-art algorithms for SSVEP-based BCIs," *arXiv preprint arXiv:1602.00904*, 2016.

[49] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *P. Roy. Soc. Lond. A-Math. Phy. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, 1998.

[50] C.-H. Wu, H.-C. Chang, P.-L. Lee, K.-S. Li, J.-J. Sie, C.-W. Sun, C.-Y. Yang, P.-H. Li, H.-T. Deng, and K.-K. Shyu, "Frequency recognition in an SSVEP-based brain computer interface using empirical mode decomposition and refined generalized zero-crossing," *J. Neurosci. Methods.*, vol. 196, no. 1, pp. 170–181, 2011.

[51] W. Liu, L. Zhang, and C. Li, "A method for recognizing high-frequency steady-state visual evoked potential based on empirical modal decomposition and canonical correlation analysis," in *Proc. IEEE 3rd Inf. Technol. Netw. Electron. Automat. Control. Conf.*, 2019, pp. 774–778.

[52] L. Zhao, P. Yuan, L. Xiao, Q. Meng, D. Hu, and H. Shen, "Research on SSVEP feature extraction based on HHT," in *Proc. 7th Int. Conf. Fuzzy. Syst. Knowl. Discovery.*, vol. 5. IEEE, 2010, pp. 2220–2223.

[53] R. M. Tello, S. M. T. Müller, T. Bastos-Filho, and A. Ferreira, "Comparison of new techniques based on EMD for control of a SSVEP-BCI," in *Proc. IEEE 23rd Int. Symp. Ind. Electron.*, 2014, pp. 992–997.

[54] J. Xu, M. Grosse-Wentrup, and V. Jayaram, "Tangent space spatial filters for interpretable and efficient riemannian classification," *J. Neural Eng.*, vol. 17, no. 2, p. 026043, 2020.

[55] G. K. Kumar and M. R. Reddy, "Constructing an exactly periodic subspace for enhancing SSVEP based BCI," *Adv. Eng. Inform.*, vol. 44, p. 101046, 2020.

[56] E. Erkan and M. Akbaba, "A study on performance increasing in SSVEP based BCI application," *Eng. Sci. Technol. Int. J.*, vol. 21, no. 3, pp. 421–427, 2018.

[57] O. Trigui, W. Zouch, and M. B. Messaoud, "Anti-noise Capability Improvement of Minimum Energy Combination Method for SSVEP Detection," *Int. J. Adv. Comput. Sc.*, vol. 7, no. 9, 2016.

[58] I. Volosyak, T. Malechka, D. Valbuena, and A. Gräser, "A novel calibration method for SSVEP based brain-computer interfaces," in *Proc. Eur. Signal Process. Conf.* IEEE, 2010, pp. 939–943.

[59] O. Friman, I. Volosyak, and A. Graser, "Multiple channel detection of steady-state visual evoked potentials for brain-computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 4, pp. 742–750, 2007.

[60] S. Parini, L. Maggi, A. C. Turconi, and G. Andreoni, "A robust and self-paced BCI system based on a four class SSVEP paradigm: algorithms and protocols for a high-transfer-rate direct brain communication," *Comput. Intel. Neurosc.*, vol. 2009, 2009.

[61] O. Falzon, K. Camilleri, and J. Muscat, "Complex-valued spatial filters for SSVEP-based BCIs with phase coding," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 9, pp. 2486–2495, 2012.

[62] B. Thompson, "Canonical correlation analysis," *Encyclopedia of statistics in behavioral science*, 2005.

[63] A. Cichocki, R. Zdunek, A. H. Phan, and S.-i. Amari, *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*. John Wiley & Sons, 2009.

[64] T.-K. Kim and R. Cipolla, "Canonical correlation analysis of video volume tensors for action categorization and detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 8, pp. 1415–1428, 2008.

[65] 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 *Proc. Int. Conf. Neural. Inf. Process.*, 2011, pp. 287–295.

[66] 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.

[67] 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.

[68] 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.

[69] 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.

[70] Y. Zhang, D. Guo, F. Li, E. Yin, Y. Zhang, P. Li, Q. Zhao, T. Tanaka, D. Yao, and P. Xu, "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, 2018.

[71] K. K. GR and R. Reddy, "Designing a sum of squared correlations framework for enhancing SSVEP-based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27,

no. 10, pp. 2044–2050, 2019.

[72] G. Bin, X. Gao, Y. Wang, Y. Li, B. Hong, and S. Gao, "A high-speed BCI based on code modulation VEP," *J. Neural Eng.*, vol. 8, no. 2, p. 025015, 2011.

[73] 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.

[74] Y. Wang, M. Nakanishi, Y.-T. Wang, and T.-P. Jung, "Enhancing detection of steady-state visual evoked potentials using individual training data," in *Proc. 36th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2014, pp. 3037–3040.

[75] J. Huang, P. Yang, B. Xiong, B. Wan, K. Su, and Z.-Q. Zhang, "Latency aligning task-related component analysis using wave propagation for enhancing SSVEP-based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 851–859, 2022.

[76] 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.

[77] Z. M. Zhang and Z. D. Deng, "A kernel canonical correlation analysis based idle-state detection method for SSVEP-based brain-computer interfaces," in *Advanced Materials Research*, vol. 341, 2012, pp. 634–640.

[78] G. Andrew, R. Arora, J. Bilmes, and K. Livescu, "Deep canonical correlation analysis," in *Proc. Int. Conf. Mach. Learn.*, 2013, pp. 1247–1255.

[79] N.-S. Kwak, K.-R. Müller, and S.-W. Lee, "A convolutional neural network for steady state visual evoked potential classification under ambulatory environment," *PloS one*, vol. 12, no. 2, p. e0172578, 2017.

[80] O. B. Guney, M. Oblokulov, and H. Ozkan, "A deep neural network for SSVEP-based brain-computer interfaces," *IEEE Trans. Biomed. Eng.*, 2021.

[81] S. Zavala, J. López, K. Chicaiza, and S. G. Yoo, "Review of steady state visually evoked potential brain-computer interface applications: technological analysis and classification," *J. Eng. Appl. Sci.*, vol. 15, pp. 659–678, 2019.

[82] 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.

[83] C. Carmeli, M. G. Knyazeva, G. M. Innocenti, and O. De Feo, "Assessment of EEG synchronization based on state-space analysis," *Neuroimage*, vol. 25, no. 2, pp. 339–354, 2005.

[84] K. Qin, R. Wang, and Y. Zhang, "Filter bank-driven multivariate synchronization index for training-free SSVEP BCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 934–943, 2021.

[85] J. Rupnik, A. Muhic, G. Leban, P. Skraba, B. Fortuna, and M. Grobelnik, "News across languages-cross-lingual document similarity and event tracking," *J. Artif. Intell. Res.*, vol. 55, pp. 283–316, 2016.

[86] J. Huang, P. Yang, B. Xiong, Q. Wang, B. Wan, Z. Ruan, K. Yang, and Z.-Q. Zhang, "Incorporating neighboring stimuli data for enhanced SSVEP-Based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 71, pp. 1–9, 2022.

[87] A. Delorme, T. Sejnowski, and S. Makeig, "Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis," *NeuroImage*, vol. 34, no. 4, pp. 1443–1449, 2007.

[88] T.-P. Jung, S. Makeig, C. Humphries, T.-W. Lee, M. J. Mckeown, V. Iragui, and T. J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, no. 2, pp. 163–178, 2000.

[89] H. Tanaka, T. Katura, and H. Sato, "Task-related component analysis for functional neuroimaging and application to near-infrared spectroscopy data," *NeuroImage*, vol. 64, pp. 308–327, 2013.

[90] ——, "Task-related oxygenation and cerebral blood volume changes estimated from nirs signals in motor and cognitive tasks," *NeuroImage*, vol. 94, pp. 107–119, 2014.

[91] 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.

[92] Z. İşcan and V. V. Nikulin, "Steady state visual evoked potential (SSVEP) based brain-computer interface (BCI) performance under different perturbations," *PloS one*, vol. 13, no. 1, p. e0191673, 2018.

[93] S. F. Anindya, H. H. Rachmat, and E. Sutjiredjeki, "A prototype of SSVEP-based BCI for home appliances control," in *Proc. 1st Int. Conf. Biomed. Eng.* IEEE, 2016, pp. 1–6.

[94] S. Akaho, "A kernel method for canonical correlation analysis," *arXiv preprint cs/0609071*, 2006.

[95] S. N. Carvalho, T. B. Costa, L. F. Uribe, D. C. Soriano, G. F. Yared, L. C. Coradine, and R. Attux, "Comparative analysis of strategies for feature extraction and classification in SSVEP BCIs," *Biomed. Signal. Proces.*, vol. 21, pp. 34–42, 2015.

[96] X. Zheng, G. Xu, Y. Zhang, R. Liang, K. Zhang, Y. Du, J. Xie, and S. Zhang, "Anti-fatigue performance in SSVEP-based visual acuity assessment: A comparison of six stimulus paradigms," *Front. Hum. Neurosci.*, vol. 14, p. 301, 2020.

[97] Y. Peng, C. M. Wong, Z. Wang, A. C. Rosa, H. T. Wang, and F. Wan, "Fatigue detection in SSVEP-BCIs based on wavelet entropy of EEG," *IEEE Access*, vol. 9, pp. 114 905–114 913, 2021.

[98] Y. Zhao, J. Tang, Y. Cao, X. Jiao, M. Xu, P. Zhou, D. Ming, and H. Qi, "Effects of distracting task with different mental workload on steady-state visual evoked potential based brain computer interfaces—an offline study," *Front. Neurosci-Switz*, vol. 12, p. 79, 2018.

[99] J. J. Norton, J. Mullins, B. E. Alitz, and T. Bretl, "The performance of 9–11-year-old children using an SSVEP-based BCI for target selection," *J. Neural Eng.*, vol. 15, no. 5, p. 056012, 2018.

[100] H. A. Lamti, M. M. B. Khelifa, and V. Hugel, "Cerebral and gaze data fusion for wheelchair navigation enhancement: Case of distracted users," *Robotica*, vol. 37, no. 2, pp. 246–263, 2019.

[101] F. Iacopi and C.-T. Lin, "A perspective on electroencephalography sensors for brain-computer interfaces," *Prog. Biomed. Eng.*, vol. 4, no. 4, p. 043002, 2022.

[102] 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.

[103] H. Cecotti, "A self-paced and calibration-less SSVEP-based brain–computer interface speller," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 2, pp. 127–133, 2010.

[104] N. Mora, I. De Munari, and P. Ciampolini, "Subject-independent, ssvep-based bci: Trading off among accuracy, responsiveness and complexity," in *Proc. 7th Int. IEEE/EMBS Conf. Neural. Eng. (NER)*. IEEE, 2015, pp. 146–149.

[105] E. K. Kalunga, S. Chevallier, Q. Barthélemy, K. Djouani, E. Monacelli, and Y. Hamam, "Online SSVEP-based BCI using Riemannian geometry," *Neurocomputing*, vol. 191, pp. 55–68, 2016.

[106] H. Cecotti, "Adaptive time segment analysis for steady-state visual evoked potential based brain–computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 3, pp. 552–560, 2020.

[107] Y. Chen, C. Yang, X. Chen, Y. Wang, and X. Gao, "A novel training-free recognition method for SSVEP-based BCIs using dynamic window strategy," *J. Neural Eng.*, vol. 18, no. 3, p. 036007, 2021.

[108] J. Jiang, E. Yin, C. Wang, M. Xu, and D. Ming, "Incorporation of dynamic stopping strategy into the high-speed SSVEP-based BCIs," *J. Neural Eng.*, vol. 15, no. 4, p. 046025, 2018.

[109] S. Lemm, B. Blankertz, T. Dickhaus, and K.-R. Müller, "Introduction to machine learning for brain imaging," *Neuroimage*, vol. 56, no. 2, pp. 387–399, 2011.

[110] D. Garrett, D. A. Peterson, C. W. Anderson, and M. H. Thaut, "Comparison of linear, nonlinear, and feature selection methods for EEG signal classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 141–144, 2003.

[111] C. Q. Lai, H. Ibrahim, M. Z. Abdullah, J. M. Abdullah, S. A. Suandi, and A. Azman, "Artifacts and noise removal for electroencephalogram (eeg): A literature review," in *Proc. Conf. Symp. Comput. Appl. Ind. Electron.* IEEE, 2018, pp. 326–332.

[112] A. Kamrud, B. Borghetti, and C. Schubert Kabban, "The effects of individual differences, non-stationarity, and the importance of data partitioning decisions for training and testing of EEG cross-participant models," *Sensors*, vol. 21, no. 9, p. 3225, 2021.

[113] A. Kilicarslan, R. G. Grossman, and J. L. Contreras-Vidal, "A robust adaptive denoising framework for real-time artifact removal in scalp EEG measurements," *J. Neural Eng.*, vol. 13, no. 2, p. 026013, 2016.

[114] C. Hema and A. A. Osman, "Single trial analysis on EEG signatures to identify individuals," in *Proc. 6th Int. Conf. Colloquium. Signal. Process. Appl.* IEEE, 2010, pp. 1–3.

[115] P. Yuan, X. Chen, Y. Wang, X. Gao, and S. Gao, "Enhancing performances of SSVEP-based brain–computer interfaces via exploiting inter-subject information," *J. Neural Eng.*, vol. 12, no. 4, p. 046006, 2015.

[116] M. Nakanishi, Y. Wang, and T.-P. Jung, "Session-to-session transfer in detecting steady-state visual evoked potentials with individual training data," in *Foundations of Augmented Cognition: Neuroergonomics and Operational Neuroscience: 10th Int.*

Conf. AC 2016, Held as Part of HCI Int. 2016, Toronto, ON, Canada, July 17-22, 2016, Proceedings, Part I 10. Springer, 2016, pp. 253–260.

[117] N. R. Waytowich, J. Faller, J. O. Garcia, J. M. Vettel, and P. Sajda, "Unsupervised adaptive transfer learning for steady-state visual evoked potential brain-computer interfaces," in *Proc. IEEE Int. Conf. Syst. Man. Cybern.*, 2016, pp. 004 135–004 140.

[118] H. Wang, Y. Sun, Y. Li, S. Chen, and W. Zhou, "Inter-and intra-subject template-based multivariate synchronization index using an adaptive threshold for SSVEP-based BCIs," *Front. Neurosci-Switz*, vol. 14, p. 717, 2020.

[119] M. Nakanishi, Y.-T. Wang, C.-S. Wei, K.-J. Chiang, and T.-P. Jung, "Facilitating calibration in high-speed BCI spellers via leveraging cross-device shared latent responses," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 4, pp. 1105–1113, 2019.

[120] B. Liu, X. Chen, X. Li, Y. Wang, X. Gao, and S. Gao, "Align and pool for EEG headset domain adaptation (ALPHA) to facilitate dry electrode based SSVEP-BCI," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 2, pp. 795–806, 2021.

[121] X. Liu, B. Liu, G. Dong, X. Gao, and Y. Wang, "Facilitating Applications of SSVEP-Based BCIs by Within-Subject Information Transfer," *Front. Neurosci-Switz*, vol. 16, p. 863359, 2022.

[122] K.-J. Chiang, C.-S. Wei, M. Nakanishi, and T.-P. Jung, "Cross-subject transfer learning improves the practicality of real-world applications of brain-computer interfaces," in *Proc. 9th Int. IEEE/EMBS Conf. Neural. Eng.*, 2019, pp. 424–427.

[123] Chiang, Kuan-Jung and Wei, Chun-Shu and Nakanishi, Masaki and Jung, Tzyy-Ping, "Boosting template-based SSVEP decoding by cross-domain transfer learning," *J. Neural Eng.*, vol. 18, no. 1, p. 016002, 2021.

[124] W. Yan, Y. Wu, C. Du, and G. Xu, "An improved cross-subject spatial filter transfer method for SSVEP-based BCI," *J. Neural Eng.*, vol. 19, no. 4, p. 046028, 2022.

[125] H. Tanaka, "Group task-related component analysis (gTRCA): A multivariate method for inter-trial reproducibility and inter-subject similarity maximization for EEG data

analysis," *Scientific Reports*, vol. 10, no. 1, p. 84, 2020.

[126] H. Wang, Y. Sun, F. Wang, L. Cao, W. Zhou, Z. Wang, and S. Chen, "Cross-subject assistance: inter-and intra-subject maximal correlation for enhancing the performance of SSVEP-based BCIS," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 517–526, 2021.

[127] Q. Wei, Y. Zhang, Y. Wang, and X. Gao, "A Canonical Correlation Analysis-Based Transfer Learning Framework for Enhancing the Performance of SSVEP-Based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, 2023.

[128] 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.

[129] Z. Wang, C. M. Wong, A. Rosa, T. Qian, T.-P. Jung, and F. Wan, "Stimulus-stimulus transfer based on time-frequency-joint representation in SSVEP-based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 70, no. 2, pp. 603–615, 2022.

[130] C. M. Wong, F. Wan, B. Wang, Z. Wang, W. Nan, K. F. Lao, P. U. Mak, M. I. Vai, and A. Rosa, "Learning across multi-stimulus enhances target recognition methods in SSVEP-based BCIs," *J. Neural Eng.*, vol. 17, no. 1, p. 016026, 2020.

[131] C. M. Wong, Z. Wang, A. C. Rosa, C. P. Chen, T.-P. Jung, Y. Hu, and F. Wan, "Transferring subject-specific knowledge across stimulus frequencies in SSVEP-based BCIs," *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 2, pp. 552–563, 2021.

[132] A. Capilla, P. Pazo-Alvarez, A. Darriba, P. Campo, and J. Gross, "Steady-state visual evoked potentials can be explained by temporal superposition of transient event-related responses," *PloS one*, vol. 6, no. 1, p. e14543, 2011.

[133] N. Shi, X. Li, B. Liu, C. Yang, Y. Wang, and X. Gao, "Representative-based cold start for adaptive ssvep-bci," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 1521–1531, 2023.

[134] M. Li, D. He, C. Li, and S. Qi, "Brain–computer interface speller based on steady-state visual evoked potential: A review focusing on the stimulus paradigm and performance," *Brain. Sci.*, vol. 11, no. 4, p. 450, 2021.

[135] I. Volosyak, "SSVEP-based Bremen–BCI interface boosting information transfer rates," *J. Neural Eng.*, vol. 8, no. 3, p. 036020, 2011.

[136] E. Yin, Z. Zhou, J. Jiang, Y. Yu, and D. Hu, "A dynamically optimized SSVEP brain–computer interface (BCI) speller," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 6, pp. 1447–1456, 2014.

[137] S. Sadeghi and A. Maleki, "Character encoding based on occurrence probability enhances the performance of SSVEP-based BCI spellers," *Biomed. Signal. Proces.*, vol. 58, p. 101888, 2020.

[138] A. Saboor, M. Benda, F. Gembler, and I. Volosyak, "Word prediction support model for SSVEP-based BCI web speller," in *International Work-Conference on Artificial Neural Networks*. Springer, 2019, pp. 430–441.

[139] L. Cao, T. Liu, L. Hou, Z. Wang, C. Fan, J. Li, and H. Wang, "A novel real-time multi-phase BCI speller based on sliding control paradigm of SSVEP," *IEEE Access*, vol. 7, pp. 133 974–133 981, 2019.

[140] T.-H. Nguyen and W.-Y. Chung, "A single-channel ssvep-based bci speller using deep learning," *IEEE Access*, vol. 7, pp. 1752–1763, 2018.

[141] H.-J. Hwang, J.-H. Lim, Y.-J. Jung, H. Choi, S. W. Lee, and C.-H. Im, "Development of an SSVEP-based BCI spelling system adopting a QWERTY-style LED keyboard," *J. Neurosci. Methods.*, vol. 208, no. 1, pp. 59–65, 2012.

[142] X. Chen, Z. Chen, S. Gao, and X. Gao, "A high-ITR SSVEP-based BCI speller," *Brain-Computer Interfaces*, vol. 1, no. 3-4, pp. 181–191, 2014.

[143] X. Chen, Y. Wang, M. Nakanishi, T.-P. Jung, and X. Gao, "Hybrid frequency and phase coding for a high-speed SSVEP-based BCI speller," in *2014 36th Annual In-*

ternational Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2014, pp. 3993–3996.

[144] F. Peng, M. Li, S.-n. Zhao, Q. Xu, J. Xu, and H. Wu, "Control of a robotic arm With an optimized common template-based CCA method for SSVEP-Based BCI," *Front. Neurorobotics.*, vol. 16, 2022.

[145] X. Chen, B. Zhao, Y. Wang, S. Xu, and X. Gao, "Control of a 7-dof robotic arm system with an ssvep-based bci," *Int. J. Neural. Syst.*, vol. 28, no. 08, p. 1850018, 2018.

[146] X. Chen, B. Zhao, and X. Gao, "Noninvasive brain-computer interface based high-level control of a robotic arm for pick and place tasks," in *Proc. 14th Int. Conf. Natural. Comput, Fuzzy. Syst. Knowledge. Discovery (ICNC-FSKD)*. IEEE, 2018, pp. 1193–1197.

[147] X. Chen, B. Zhao, Y. Wang, and X. Gao, "Combination of high-frequency SSVEP-based BCI and computer vision for controlling a robotic arm," *J. Neural Eng.*, vol. 16, no. 2, p. 026012, 2019.

[148] F. Wang, Y. Wen, J. Bi, H. Li, and J. Sun, "A portable SSVEP-BCI system for rehabilitation exoskeleton in augmented reality environment," *Biomed. Signal. Proces.*, vol. 83, p. 104664, 2023.

[149] Y. Ke, P. Liu, X. An, X. Song, and D. Ming, "An online SSVEP-BCI system in an optical see-through augmented reality environment," *J. Neural Eng.*, vol. 17, no. 1, p. 016066, 2020.

[150] X. Chen, X. Huang, Y. Wang, and X. Gao, "Combination of augmented reality based brain-computer interface and computer vision for high-level control of a robotic arm," *IEEE Trans. Neural Syst. Rehabil. Eng.*, 2020.

[151] J. Ai, J. Meng, X. Mai, and X. Zhu, "BCI Control of a Robotic arm based on SSVEP with Moving Stimuli for Reach and grasp Tasks," *IEEE J Biomed Health Inform.*, 2023.

[152] G. R. Muller-Putz and G. Pfurtscheller, "Control of an electrical prosthesis with an SSVEP-based BCI," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 1, pp. 361–364, 2007.

[153] R. Ortner, B. Z. Allison, G. Korisek, H. Gaggl, and G. Pfurtscheller, "An SSVEP BCI to control a hand orthosis for persons with tetraplegia," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 19, no. 1, pp. 1–5, 2010.

[154] N. Guo, X. Wang, D. Duanmu, X. Huang, X. Li, Y. Fan, H. Li, Y. Liu, E. H. K. Yeung, M. K. T. To *et al.*, "SSVEP-based brain computer interface controlled soft robotic glove for post-stroke hand function rehabilitation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 1737–1744, 2022.

[155] X. Zhao, Y. Chu, J. Han, and Z. Zhang, "SSVEP-based brain–computer interface controlled functional electrical stimulation system for upper extremity rehabilitation," *IEEE T. Syst. Man.Cy-S.*, vol. 46, no. 7, pp. 947–956, 2016.

[156] T. Sakurada, T. Kawase, K. Takano, T. Komatsu, and K. Kansaku, "A BMI-based occupational therapy assist suit: asynchronous control by SSVEP," *Front. Neurosci-Switz*, vol. 7, p. 172, 2013.

[157] C. Wang, X. Wu, Z. Wang, and Y. Ma, "Implementation of a brain-computer interface on a lower-limb exoskeleton," *IEEE access*, vol. 6, pp. 38 524–38 534, 2018.

[158] N.-S. Kwak, K.-R. Müller, and S.-W. Lee, "A lower limb exoskeleton control system based on steady state visual evoked potentials," *J. Neural Eng.*, vol. 12, no. 5, p. 056009, 2015.

[159] Z. Qi, W. Chen, J. Wang, J. Zhang, and X. Wang, "Lower limb rehabilitation exoskeleton control based on ssvep-bci," in *Proc. 16th Conf. Ind. Electron. Appl (ICIEA)*. IEEE, 2021, pp. 1954–1959.

[160] X. Zeng, G. Zhu, L. Yue, M. Zhang, and S. Xie, "A feasibility study of SSVEP-based passive training on an ankle rehabilitation robot," *J. Healthc. Eng.*, vol. 2017, 2017.

[161] P. F. Diez, S. M. T. Müller, V. A. Mut, E. Laciar, E. Avila, T. F. Bastos-Filho, and M. Sarcinelli-Filho, "Commanding a robotic wheelchair with a high-frequency steady-

state visual evoked potential based brain–computer interface," *Med. Eng. Phys.*, vol. 35, no. 8, pp. 1155–1164, 2013.

[162] V. Sakkalis, M. Krana, C. Farmaki, C. Bourazanis, D. Gaitatzis, and M. Pediaditis, "Augmented reality driven steady-state visual evoked potentials for wheelchair navigation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 2960–2969, 2022.

[163] H. Rivera-Flor, D. Gurve, A. Floriano, D. Delisle-Rodriguez, R. Mello, and T. Bastos-Filho, "CCA-based compressive sensing for SSVEP-based brain-computer interfaces to command a robotic wheelchair," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–10, 2022.

[164] D. Zhang, S. Liu, J. Zhang, G. Li, D. Suo, T. Liu, J. Luo, Z. Ming, J. Wu, and T. Yan, "Brain-controlled 2d navigation robot based on a spatial gradient controller and predictive environmental coordinator," *IEEE J Biomed Health Inform.*, vol. 26, no. 12, pp. 6138–6149, 2022.

[165] Q. Gao, X. Zhao, X. Yu, Y. Song, and Z. Wang, "Controlling of smart home system based on brain-computer interface," *Technol. Health. Care.*, vol. 26, no. 5, pp. 769–783, 2018.

[166] A. Saboor, A. Rezeika, P. Stawicki, F. Gembler, M. Benda, T. Grunenberg, and I. Volosyak, "SSVEP-based BCI in a smart home scenario," in *Proc. Int. Work-Conf. Artif. Neural. Netw.*, 2017, pp. 474–485.

[167] M. Adams, M. Benda, A. Saboor, A. F. Krause, A. Rezeika, F. Gembler, P. Stawicki, M. Hesse, K. Essig, S. Ben-Salem *et al.*, "Towards an SSVEP-BCI controlled smart home," in *Proc. IEEE Int. Conf. Syst. Man. Cybern.*, 2019, pp. 2737–2742.

[168] L. Shao, L. Zhang, A. N. Belkacem, Y. Zhang, X. Chen, J. Li, and H. Liu, "EEG-controlled wall-crawling cleaning robot using SSVEP-based brain-computer interface," *J. Healthc. Eng.*, vol. 2020, 2020.

[169] J. Ha, S. Park, C.-H. Im, and L. Kim, "A hybrid brain–computer interface for real-life meal-assist robot control," *Sensors*, vol. 21, no. 13, p. 4578, 2021.

[170] C. J. Perera, I. Naotunna, C. Sadaruwan, R. A. R. C. Gopura, and T. D. Lalitharatne, "SSVEP based BMI for a meal assistance robot," in *Proc. IEEE Int. Conf. Syst. Man. Cybern.*, 2016, pp. 002 295–002 300.

[171] S. M. T. Müller, T. F. Bastos, and M. S. Filho, "Proposal of a SSVEP-BCI to command a robotic wheelchair," *J. Control. Autom. Elec.*, vol. 24, no. 1-2, pp. 97–105, 2013.

[172] A. Schielke and B. Krekelberg, "Steady state visual evoked potentials in schizophrenia: A review," *Front. Neurosci-Switz*, vol. 16, p. 988077, 2022.

[173] R. Kuś, A. Duszyk, P. Milanowski, M. Łabkecki, M. Bierzyńska, Z. Radzikowska, M. Michalska, J. Żygierewicz, P. Suffczyński, and P. J. Durka, "On the quantification of SSVEP frequency responses in human EEG in realistic BCI conditions," *PloS one*, vol. 8, no. 10, p. e77536, 2013.

[174] Y. Deng, Q. Sun, C. Wang, Y. Wang, and S. K. Zhou, "TRCA-Net: using TRCA filters to boost the SSVEP classification with convolutional neural network," *J. Neural Eng.*, vol. 20, no. 4, p. 046005, 2023.

[175] B. Liu, X. Huang, Y. Wang, X. Chen, and X. Gao, "BETA: A large benchmark database toward SSVEP-BCI application," *Front. Neurosci-Switz*, vol. 14, p. 627, 2020.

[176] R. Debnath, V. C. Salo, G. A. Buzzell, K. H. Yoo, and N. A. Fox, "Mu rhythm desynchronization is specific to action execution and observation: Evidence from time-frequency and connectivity analysis," *Neuroimage*, vol. 184, pp. 496–507, 2019.

[177] R. Luo, M. Xu, X. Zhou, X. Xiao, T.-P. Jung, and D. Ming, "Data augmentation of ssveps using source aliasing matrix estimation for brain-computer interfaces," *IEEE Trans. Biomed. Eng.*, 2022.

[178] X. Yu, M. Z. Aziz, M. T. Sadiq, Z. Fan, and G. Xiao, "A new framework for automatic detection of motor and mental imagery EEG signals for robust BCI systems," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021.

[179] S. Chevallier, E. K. Kalunga, Q. Barthélemy, and E. Monacelli, "Review of Riemannian distances and divergences, applied to SSVEP-based BCI," *Neuroinformatics*, vol. 19, no. 1, pp. 93–106, 2021.

[180] P. Stegman, C. S. Crawford, M. Andujar, A. Nijholt, and J. E. Gilbert, "Brain–computer interface software: A review and discussion," *IEEE T. Hum-Mach. Syst.*, vol. 50, no. 2, pp. 101–115, 2020.

[181] L. Angrisani, P. Arpaia, A. Esposito, and N. Moccaldi, "A wearable brain–computer interface instrument for augmented reality-based inspection in industry 4.0," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 4, pp. 1530–1539, 2019.

[182] P. Arpaia, L. Duraccio, N. Moccaldi, and S. Rossi, "Wearable brain–computer interface instrumentation for robot-based rehabilitation by augmented reality," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 9, pp. 6362–6371, 2020.

[183] L. Angrisani, P. Arpaia, D. Casinelli, and N. Moccaldi, "A single-channel SSVEP-based instrument with off-the-shelf components for trainingless brain-computer interfaces," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 10, pp. 3616–3625, 2018.

[184] A. Liu, Q. Liu, X. Zhang, X. Chen, and X. Chen, "Muscle artifact removal towards mobile SSVEP-based BCI: A comparative study," *IEEE Trans. Instrum. Meas.*, 2021.

[185] 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.

[186] P. Gaur, H. Gupta, A. Chowdhury, K. McCreadie, R. B. Pachori, and H. Wang, "A sliding window common spatial pattern for enhancing motor imagery classification in EEG-BCI," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–9, 2021.

[187] C. M. Wong, B. Wang, Z. Wang, K. F. Lao, A. Rosa, and F. Wan, "Spatial filtering in SSVEP-based BCIs: unified framework and new improvements," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 11, pp. 3057–3072, 2020.

[188] X. Jiang, G.-B. Bian, and Z. Tian, "Removal of artifacts from EEG signals: a review," *Sensors*, vol. 19, no. 5, p. 987, 2019.

[189] A. Khadem and G.-A. Hossein-Zadeh, "Quantification of the effects of volume conduction on the EEG/MEG connectivity estimates: an index of sensitivity to brain interactions," *Physiol Meas.*, vol. 35, no. 10, p. 2149, 2014.

[190] A. Anzolin, P. Presti, F. Van de Steen, L. Astolfi, S. Haufe, and D. Marinazzo, "Effect of head volume conduction on directed connectivity estimated between reconstructed EEG sources," *BioRxiv*, p. 251223, 2018.

[191] N. Caramia, F. Lotte, and S. Ramat, "Optimizing spatial filter pairs for EEG classification based on phase-synchronization," in *Proc. Int. Conf. Acoust. Speech. Signal. Process.* IEEE, 2014, pp. 2049–2053.

[192] J. Kayser and C. E. Tenke, "Issues and considerations for using the scalp surface Laplacian in EEG/ERP research: A tutorial review," *Int. J. Psychophysiol.*, vol. 97, no. 3, pp. 189–209, 2015.

[193] R. J. Bufacchi, C. Magri, G. Novembre, and G. D. Iannetti, "Local spatial analysis: an easy-to-use adaptive spatial EEG filter," *J. Neurophysiol.*, vol. 125, no. 2, pp. 509–521, 2021.

[194] F. P. Kalaganis, E. Chatzilari, S. Nikolopoulos, N. A. Laskaris, and Y. Kompatsiaris, "A collaborative representation approach to detecting error-related potentials in SSVEP-BCIs," in *Proc. Thematic. Workshop. ACM. Multimedia.*, 2017, pp. 262–270.

[195] J. Lu, D. J. McFarland, and J. R. Wolpaw, "Adaptive laplacian filtering for sensorimotor rhythm-based brain–computer interfaces," *J. Neural Eng.*, vol. 10, no. 1, p. 016002, 2012.

[196] S. E. Kober, C. Neuper, and G. Wood, "Differential effects of up-and down-regulation of SMR coherence on EEG activity and memory performance: A neurofeedback training study," *Front. Hum. Neurosci.*, vol. 14, 2020.

[197] R. Srinivasan, F. A. Bibi, and P. L. Nunez, "Steady-state visual evoked potentials: distributed local sources and wave-like dynamics are sensitive to flicker frequency," *Brain. Topogr.*, vol. 18, no. 3, pp. 167–187, 2006.

[198] S. Mane and M. N. Rao, "Many-objective optimization: Problems and evolutionary algorithms–a short review," *Int. J. Appl. Eng. Res.*, vol. 12, no. 20, pp. 9774–9793, 2017.

[199] A. Prajapati, "A comparative study of many-objective optimizers on large-scale many-objective software clustering problems," *Complex. Intell. Syst.*, vol. 7, no. 2, pp. 1061–1077, 2021.

[200] J. Jin, H. Sun, I. Daly, S. Li, C. Liu, X. Wang, and A. Cichocki, "A novel classification framework using the graph representations of electroencephalogram for motor imagery based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 20–29, 2021.

[201] Y. Zhou, Z. Xu, Y. Niu, P. Wang, X. Wen, X. Wu, and D. Zhang, "Cross-task cognitive workload recognition based on EEG and domain adaptation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, 2022.

[202] Y.-C. Jiang, R. Ma, S. Qi, S. Ge, Z. Sun, Y. Li, J. Song, and M. Zhang, "Characterization of bimanual cyclical tasks from single-trial EEG-fNIRS measurements," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 146–156, 2022.

[203] J. Jin, Z. Chen, R. Xu, Y. Miao, X. Wang, and T.-P. Jung, "Developing a novel tactile P300 brain-computer interface with a cheeks-stim paradigm," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 9, pp. 2585–2593, 2020.

[204] Y.-P. Lin, C.-H. Wang, T.-P. Jung, T.-L. Wu, S.-K. Jeng, J.-R. Duann, and J.-H. Chen, "EEG-based emotion recognition in music listening," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 7, pp. 1798–1806, 2010.

[205] P. P. Roy, P. Kumar, and V. Chang, "A hybrid classifier combination for home automation using EEG signals," *Neural Comput. Appl.*, vol. 32, no. 20, pp. 16 135–16 147,

2020.

[206] 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.

[207] J. Jin, Z. Wang, R. Xu, C. Liu, X. Wang, and A. Cichocki, "Robust similarity measurement based on a novel time filter for SSVEPs detection," *IEEE Trans. Neural. Netw. Learn. Syst.*, 2021.

[208] T. Bao, S. A. R. Zaidi, S. Q. Xie, P. Yang, and Z.-Q. Zhang, "CNN confidence estimation for rejection-based hand gesture classification in myoelectric control," *IEEE T. Hum-Mach. Syst.*, vol. 52, no. 1, pp. 99–109, 2021.

[209] J. Zhao, W. Li, and M. Li, "Comparative study of SSVEP-and P300-based models for the telepresence control of humanoid robots," *PLoS One*, vol. 10, no. 11, p. e0142168, 2015.

[210] M. Wang, R. Li, R. Zhang, G. Li, and D. Zhang, "A wearable SSVEP-based BCI system for quadcopter control using head-mounted device," *IEEE Access*, vol. 6, pp. 26 789–26 798, 2018.

[211] 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.

[212] Y. Yang, W. Wu, B. Wang, and M. Li, "Analytical reformulation for stochastic unit commitment considering wind power uncertainty with gaussian mixture model," *IEEE Trans. Power Syst.*, vol. 35, no. 4, pp. 2769–2782, 2019.

[213] G. Xuan, W. Zhang, and P. Chai, "EM algorithms of Gaussian mixture model and hidden Markov model," in *Proc. Int. Conf. Image. Process.*, vol. 1. IEEE, 2001, pp. 145–148.

[214] X.-Y. Zhou, J. S. Lim, I. Kwon *et al.*, "EM algorithm with GMM and naive Bayesian to implement missing values," *Adv. Sci. Technol. Lett.*, vol. 46, pp. 1–5, 2014.

[215] M. Hasnain, M. F. Pasha, I. Ghani, M. Imran, M. Y. Alzahrani, and R. Budiarto, "Evaluating trust prediction and confusion matrix measures for web services ranking," *IEEE Access*, vol. 8, pp. 90 847–90 861, 2020.

[216] J. Taylor, *Introduction to error analysis, the study of uncertainties in physical measurements*. University Science Books, 1997.

[217] A. Ravi, J. Lu, S. Pearce, and N. Jiang, "Enhanced system robustness of asynchronous BCI in augmented reality using steady-state motion visual evoked potential," *IEEE Trans. Neural Syst. Rehabil. Eng.*, 2022.

[218] Y. Li, J. Xiang, and T. Kesavadas, "Convolutional correlation analysis for enhancing the performance of SSVEP-based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, 2020.

[219] X. Chai, Z. Zhang, K. Guan, Y. Lu, G. Liu, T. Zhang, and H. Niu, "A hybrid BCI-controlled smart home system combining SSVEP and EMG for individuals with paralysis," *Biomed. Signal. Proces.*, vol. 56, p. 101687, 2020.

[220] V. Jayaram, M. Alamgir, Y. Altun, B. Scholkopf, and M. Grosse-Wentrup, "Transfer learning in brain-computer interfaces," *IEEE Comput. Intell. Mag.*, vol. 11, no. 1, pp. 20–31, 2016.

[221] W. Liu, Y. Ke, P. Liu, J. Du, L. Kong, S. Liu, X. An, and D. Ming, "A cross-subject SSVEP-BCI based on task related component analysis," in *Proc. 41st Annu.Int. Conf. Eng. Med. Biol. Soc (EMBS)*. IEEE, 2019, pp. 3022–3025.

[222] Q. Sun, M. Chen, L. Zhang, X. Yuan, and C. Li, "Improving SSVEP identification accuracy via generalized canonical correlation analysis," in *Proc. 10th Int. IEEE/EMBS Conf. Neural. Eng (NER)*. IEEE, 2021, pp. 61–64.

[223] Q. Sun, M. Chen, L. Zhang, C. Li, and W. Kang, "Similarity-constrained task-related component analysis for enhancing SSVEP detection," *J. Neural Eng.*, vol. 18, no. 4, p. 046080, 2021.

[224] N. Cheng, K. S. Phua, H. S. Lai, P. K. Tam, K. Y. Tang, K. K. Cheng, R. C.-H. Yeow, K. K. Ang, C. Guan, and J. H. Lim, "Brain-computer interface-based soft robotic glove rehabilitation for stroke," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 12, pp. 3339–3351, 2020.

[225] A. Kumar, L. Gao, E. Pirogova, and Q. Fang, "A review of error-related potential-based brain–computer interfaces for motor impaired people," *IEEE Access*, vol. 7, pp. 142 451–142 466, 2019.

[226] A. Kapsalyamov, P. K. Jamwal, S. Hussain, and M. H. Ghayesh, "State of the art lower limb robotic exoskeletons for elderly assistance," *IEEE Access*, vol. 7, pp. 95 075–95 086, 2019.

[227] Y. Zhang, S. Q. Xie, C. Shi, J. Li, and Z.-Q. Zhang, "Cross-subject transfer learning for boosting recognition performance in SSVEP-Based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 1574–1583, 2023.

[228] M. Y. Naser and S. Bhattacharya, "Towards practical BCI-driven wheelchairs: a systematic review study," *IEEE Trans. Neural Syst. Rehabil. Eng.*, 2023.

[229] Q. Ai, M. Zhao, K. Chen, X. Zhao, L. Ma, and Q. Liu, "Flexible coding scheme for robotic arm control driven by motor imagery decoding," *J. Neural Eng.*, vol. 19, no. 5, p. 056008, 2022.

[230] Y. Xu, H. Zhang, L. Cao, X. Shu, and D. Zhang, "A shared control strategy for reach and grasp of multiple objects using robot vision and noninvasive brain–computer interface," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 1, pp. 360–372, 2020.

[231] D. Liu, W. Chen, K. Lee, R. Chavarriaga, F. Iwane, M. Bouri, Z. Pei, and J. d. R. Millán, "Eeg-based lower-limb movement onset decoding: Continuous classification and asynchronous detection," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 8, pp. 1626–1635, 2018.

[232] K. Sharma and S. K. Maharaj, "Continuous and spontaneous speed control of a robotic arm using SSVEP," in *Proc. 9th Int. Conf. Brain. Comput. Interface (BCI).*

IEEE, 2021, pp. 1–5.

[233] Y. Joon Kim, M. Grabowecky, K. A. Paller, K. Muthu, and S. Suzuki, "Attention induces synchronization-based response gain in steady-state visual evoked potentials," *Nat. Neurosci*, vol. 10, no. 1, pp. 117–125, 2007.

[234] A. J. Molina-Cantero, J. A. Castro-García, F. Gómez-Bravo, R. López-Ahumada, R. Jiménez-Naharro, and S. Berrazueta-Alvarado, "Controlling a mouse pointer with a single-channel EEG sensor," *Sensors*, vol. 21, no. 16, p. 5481, 2021.

[235] X. Chen, Y. Wang, S. Zhang, S. Gao, Y. Hu, and X. Gao, "A novel stimulation method for multi-class SSVEP-BCI using intermodulation frequencies," *J. Neural Eng.*, vol. 14, no. 2, p. 026013, 2017.

[236] P. Autthasan, X. Du, J. Arnin, S. Lamyai, M. Perera, S. Itthipuripat, T. Yagi, P. Manoonpong, and T. Wilaiprasitporn, "A single-channel consumer-grade EEG device for brain–computer interface: enhancing detection of SSVEP and its amplitude modulation," *IEEE Sensors J.*, vol. 20, no. 6, pp. 3366–3378, 2019.

[237] X. Zhang, G. Xu, J. Xie, and X. Zhang, "Brain response to luminance-based and motion-based stimulation using inter-modulation frequencies," *PLoS One*, vol. 12, no. 11, p. e0188073, 2017.

[238] D. H. Brainard and S. Vision, "The psychophysics toolbox," *Spatial vision*, vol. 10, no. 4, pp. 433–436, 1997.

[239] H. Banville and T. Falk, "Recent advances and open challenges in hybrid brain-computer interfacing: a technological review of non-invasive human research," *Brain-Computer Interfaces*, vol. 3, no. 1, pp. 9–46, 2016.

[240] K.-S. Hong and M. J. Khan, "Hybrid brain–computer interface techniques for improved classification accuracy and increased number of commands: a review," *Front. Neurorobotics.*, vol. 11, p. 35, 2017.

[241] Z. Wang, Y. Yu, M. Xu, Y. Liu, E. Yin, and Z. Zhou, "Towards a hybrid BCI gaming paradigm based on motor imagery and SSVEP," *Int. J. Hum-Comput. Int.*, vol. 35,

no. 3, pp. 197–205, 2019.

[242] M. Wang, I. Daly, B. Z. Allison, J. Jin, Y. Zhang, L. Chen, and X. Wang, "A new hybrid BCI paradigm based on P300 and SSVEP," *J. Neurosci. Methods.*, vol. 244, pp. 16–25, 2015.

[243] S. Jalilpour, S. H. Sardouie, and A. Mijani, "A novel hybrid BCI speller based on RSVP and SSVEP paradigm," *Comput. Meth. Prog. Bio.*, vol. 187, p. 105326, 2020.

[244] M. H. Chang, H. J. Baek, S. M. Lee, and K. S. Park, "An amplitude-modulated visual stimulation for reducing eye fatigue in SSVEP-based brain–computer interfaces," *Clin. Neurophysiol.*, vol. 125, no. 7, pp. 1380–1391, 2014.

[245] K. Sharma and S. Kar, "Extracting multiple commands from a single SSVEP flicker using eye-accommodation," *Biocybern. Biomed. Eng.*, vol. 39, no. 3, pp. 914–922, 2019.

[246] R. Gulbinaite, D. H. Roozendaal, and R. VanRullen, "Attention differentially modulates the amplitude of resonance frequencies in the visual cortex," *NeuroImage*, vol. 203, p. 116146, 2019.

[247] H. Lim, S. Kim, and J. Ku, "Distraction classification during target tracking tasks involving target and cursor flickering using EEGnet," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 1113–1119, 2022.

[248] J.-H. Song, "The role of attention in motor control and learning," *Curr. Opin. Psychol.*, vol. 29, pp. 261–265, 2019.

[249] M. R. Kamke, A. E. Ryan, M. V. Sale, M. E. Campbell, S. Riek, T. J. Carroll, and J. B. Mattingley, "Visual spatial attention has opposite effects on bidirectional plasticity in the human motor cortex," *J. Neurosci. Methods.*, vol. 34, no. 4, pp. 1475–1480, 2014.

[250] M. Maier, B. R. Ballester, and P. F. Verschure, "Principles of neurorehabilitation after stroke based on motor learning and brain plasticity mechanisms," *Front. Syst. Neurosci.*, vol. 13, p. 74, 2019.