TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 11/15 pp505–522
DOI: 10.26599/TST.2020.9010015
Volume 26, Number 4, August 2021

# Cross-Target Transfer Algorithm Based on the Volterra Model of SSVEP-BCI

Jiajun Lin, Liyan Liang, Xu Han, Chen Yang, Xiaogang Chen, and Xiaorong Gao*

**Abstract:** In general, a large amount of training data can effectively improve the classification performance of the Steady-State Visually Evoked Potential (SSVEP)-based Brain-Computer Interface (BCI) system. However, it will prolong the training time and considerably restrict the practicality of the system. This study proposed a SSVEP nonlinear signal model based on the Volterra filter, which could reconstruct stable reference signals using relatively small number of training targets by transfer learning, thereby reducing the training cost of SSVEP-BCI. Moreover, this study designed a transfer-extended Canonical Correlation Analysis (t-eCCA) method based on the model to achieve cross-target transfer. As a result, in a single-target SSVEP experiment with 16 stimulus frequencies, t-eCCA obtained an average accuracy of 86.96%±12.87% across 12 subjects using only half of the calibration time, which exhibited no significant difference from the representative training classification algorithms, namely, extended canonical correlation analysis (88.32%±13.97%) and task-related component analysis (88.92%±14.44%), and was significantly higher than that of the classic non-training algorithms, namely, Canonical Correlation Analysis (CCA) as well as filter-bank CCA. Results showed that the proposed cross-target transfer algorithm t-eCCA could fully utilize the information about the targets and its stimulus frequencies and effectively reduce the training time of SSVEP-BCI.

**Key words:** Steady-State Visually Evoked Potential (SSVEP); Brain-Computer Interface (BCI); Volterra filter; cross-target information; transfer learning

## 1 Introduction

Steady-State Visually Evoked Potential (SSVEP), one of the dominant paradigms of the visual Brain Computer Interface (BCI), has received considerable attention because of its noninvasiveness, short training time, and relatively high performance[1–3]. In general, SSVEP is a physiological response that mainly occurs in the visual cortex of the brain's occipital lobe[4]. Specifically, when gazing at a periodic visual flicker, an electroencephalogram (EEG) response at certain fixed frequencies related to stimulus frequencies will be elicited in the visual cortex, such as harmonic frequencies under single-frequency stimulation and intermodulation frequencies under dual-frequency stimulation[5]. Therefore, signal detection can be completed in SSVEP-BCI to identify the visual stimulus that users gaze at using the stable nonlinear characteristics of SSVEP response. In recent years, signal decoding has been one of the research hotspots in SSVEP-BCI, which is classified as training decoding algorithm or non-training algorithm according to whether additional calibration data are needed[6–8].

• Jiajun Lin, Liyan Liang, Xu Han, and Xiaorong Gao are with the Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, China. E-mail: lin-jj17@mails.tsinghua.edu.cn; 18618488256@163.com; hanxu29@ucla.edu; gxr-dea@tsinghua.edu.cn.
• Chen Yang is with the School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China. E-mail: yuanchouyc@gmail.com.
• Xiaogang Chen is with the Institute of Biomedical Engineering, Chinese Academy of Medical Sciences and Peking Union Medical College, Tianjin 300192, China. E-mail: chenxg@bme.cams.cn.
* To whom correspondence should be addressed.
Manuscript received: 2020-04-06; revised: 2020-05-14; accepted: 2020-05-20

Among different template-based classification methods, the non-training algorithm primarily utilizes sine-cosine reference signals to obtain a spatial filter, including Canonical Correlation Analysis (CCA)[9], Filter-Bank Canonical Correlation Analysis (FBCCA)[10], and spatiotemporal equalization dynamic window[11]. By contrast, the training decoding algorithm achieves a high performance using subject-specific template from additional calibration data, with novel representative methods, such as extended Canonical Correlation Analysis (eCCA)[12] and Task-Related Component Analysis (TRCA)[13]. However, training classification algorithm generally needs to average multiple trials data of the same target because of the relatively week EEG signals[14], resulting in the performance of the training algorithm having a strong dependence on the number of training trials. When the number of calibration trials for each target is insufficient, the performance of the BCI system will be degraded significantly[13]. The main reason is that a small amount of training data leads to inaccurate covariance matrix estimation in CCA-based methods[15].

With the in-depth understanding of SSVEP features and the improvement of neural signal detection, the number of encoding targets in SSVEP-BCI increases rapidly. Cheng and Gao[16] proposed the first practical four-target cursor control system in 1999. In 2014, Nakanishi et al.[17] designed a high-speed spelling system with 32 characters. Chen et al.[12] realized the classic 40-target SSVEP-BCI system using the joint frequency-phase modulation method. Recently, Xu et al.[18] reported a 108-character hybrid BCI system using concurrent P300 and SSVEP features, the largest number of targets ever known. However, with the rapid increase in the number of targets, the total training time also increases considerably, resulting in the subject's visual fatigue and further degradation of the performance of the BCI system[19]. Thus, how to reduce the calibration time while maintaining the high decoding capability of the BCI system has become an urgent problem for researchers. A feasible approach to solve this issue is cross-subject transfer learning, which hypothesizes that the EEG data of different subjects have some shared information that can be used to expand the training data[20–22]. However, the lack of in-depth investigation on the specific difference between subjects has considerably limited the performance of the cross-subject transfer learning algorithm.

During the training stage, users are required to gaze at $N_{\text{train}}$ random visual stimuli, with each target repeated $N_{\text{trial}}$ times, and the stimulus time plus rest time of each trial equals $T_{\text{trial}}$. When the rest time between blocks is ignored, the total training time can be expressed as $T_{\text{train}} = N_{\text{train}} \times N_{\text{trial}} \times T_{\text{trial}}$[22]. Evidently, if the duration of each trial $T_{\text{trial}}$ is constant, the training time $T_{\text{train}}$ can be decreased by reducing the number of training targets $N_{\text{train}}$ or the number of training trials of each target $N_{\text{trial}}$. Wong et al.[23] used EEG data from both target stimulus and other adjacent stimuli to learn the reference signal, thereby effectively reducing the number of training trials of each target $N_{\text{trial}}$; however, the number of training targets $N_{\text{train}}$ was not decreased. Zhang et al.[24] and Han et al.[25] proposed a dynamic model that could effectively fit the SSVEP envelope, and made a preliminary attempt to reduce the training target $N_{\text{train}}$; however, the maximum average accuracy was only 70% with 1 s data length, which could not meet the requirements in actual application.

This study aims to realize effective cross-target transfer learning in SSVEP-BCI, thereby reducing the training time. The key concept of cross-target transfer learning is to explore and exploit the correlation of the characteristics of SSVEP signals between different targets (and its stimulus frequencies). This study mainly utilized the continuity of the frequency domain characteristics of SSVEP response, with the phase of SSVEP signals approximately linear[4, 26]. Based on the nonlinear transfer function theory[27], the transfer function was used to interpret the SSVEP response for the first time. Specifically, the brain's response under the SSVEP stimulation paradigm could be modeled as a Volterra filter. Based on the Volterra model, this study proposed a novel cross-target transfer algorithm, that is transfer-extended Canonical Correlation Analysis (t-eCCA). The proposed transfer algorithm primarily consisted of four steps: first, sampling all of the stimulus frequencies with equal intervals to obtain the training data; second, using Least Squares (LS) to identify the parameters of the Volterra model; third, utilizing polynomial interpolation to transfer the frequency response of the Volterra model from the training frequencies to all stimulus frequencies and further reconstruct the SSVEP templates; and fourth, using the reconstructed SSVEP reference signals as the template of the eCCA method. The results showed that the proposed cross-target transfer algorithm t-eCCA, which saved half of the training time, achieved similar classification performance to the representative training

algorithms eCCA and TRCA, and significantly higher classification performance than the classic non-training algorithms CCA and FBCCA.

The remainder of this paper is organized as follows: Section 2 describes the SSVEP data model, the specific flow of the proposed transfer algorithm, and the single-target experimental design. Section 3 shows the identification results of the Volterra model, the evaluation of the transferred SSVEP template, and the classification performance of the proposed transfer algorithm. Section 4 discusses the Volterra model in terms of the comparison of different signal models, physical interpretations, model complexities, multi-target experiments, and percentage of training stimulus frequencies and identifies several directions for further investigation. Section 5 concludes the study.

## 2 Material and Method

### 2.1 Data model

In general, the EEG data model of any channel $c$ is represented as an additive model, which is a single-input/single-output system,

$$x^c(n) = s^c(n) + w^c(n), \quad n = 0, 1, \ldots, L-1 \quad (1)$$

where $x^c(n)$ denotes the EEG data of channel $c$ at time $n$, $s^c(n)$ represents the SSVEP data of channel $c$ at time $n$, $w^c(n)$ is the background noise, and $L$ is the data length.

Given that the frequency response characteristics of the Volterra filter[28] are consistent with the nonlinear characteristics of SSVEP signals, the Volterra system was selected as the SSVEP signal model. The Voltterra model is defined as follows:

$$s^c(n) = \sum_{p=0}^{P} s_p^c(n), \quad n = 0, 1, \ldots, L-1 \quad (2)$$

where $P$ denotes the order of the Volterra model and $s_p^c(n)$ is the $p$-th-order response, which can be expressed as follows:

$$s_p^c(n) = \sum_{m_1=0}^{N_p-1} \sum_{m_2=0}^{N_p-1} \cdots \sum_{m_p=0}^{N_p-1} h_p^c(m_1, m_2, \ldots, m_p) \times u(n-m_1)u(n-m_2)\cdots u(n-m_p) \quad (3)$$

where $N_p$ denotes the memory span of the $p$-th-order response, $h_p^c(m_1, m_2, \ldots, m_p)$ represents the $p$-th-order Volterra kernel, and $u(n)$ is the stimulus signal. Equation (3) is a $p$-dimensional discrete convolution.

### 2.2 Framework of the transfer algorithm

Different from other transfer algorithms[20, 23], the proposed cross-target transfer algorithm aimed to transfer the information of one subject during an experiment, from signals at some stimulus frequencies to signals at other stimulus frequencies, which would reduce the training cost per subject. To transfer information, the Volterra model described in Section 2.1, which represents the characteristics of SSVEP in the frequency domain, was employed. Figure 1 illustrates the framework of the proposed cross-target transfer algorithm.

The cross-target transfer algorithm included four steps, of which the first three steps were classified as the training stage and the last step was classified as the testing stage. First, a sample training strategy was used to acquire the SSVEP training data. Second, the LS Volterra model was applied to fit the SSVEP signals quantitatively and the frequency response of the Volterra model at the training frequencies was determined. Third, polynomial interpolation was employed to obtain information about the frequency domain at all stimulus frequencies, so that the SSVEP template signals could be completely reconstructed. Fourth, the efficacy of the proposed transfer algorithm was tested by comparing the

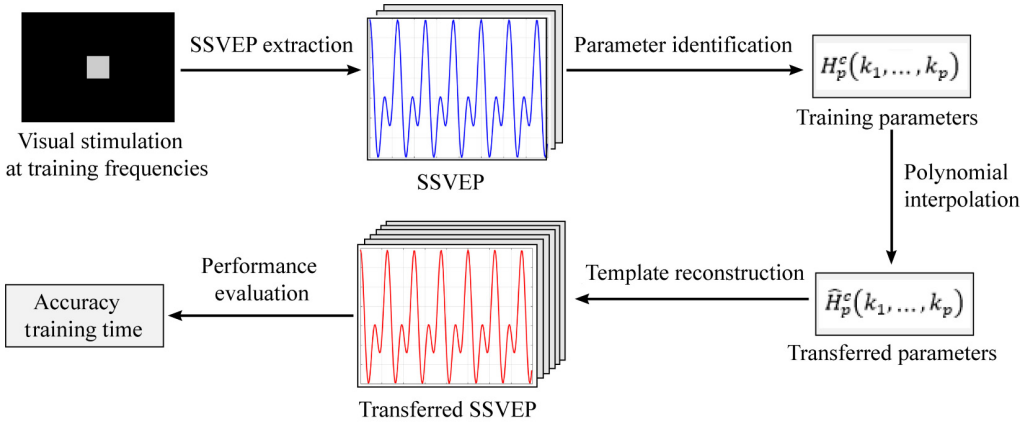


**Fig. 1 Flowchart of the proposed transfer algorithm.**

classification performance of different template-based SSVEP-BCI methods.

#### 2.2.1 Step 1: Sample training

For an SSVEP-BCI system with $N$ targets flickering at different frequencies, $N_{\text{train}}$ stimulus frequencies (targets) are equidistantly sampled to train, with the remaining $N - N_{\text{train}}$ stimulus frequencies not being presented during the training stage. To obtain the SSVEP component $s^c(n)$, first, the EEG signals $x^c(n)$ of different trials with the same stimulus frequency $f$ are averaged, with $\bar{x}^c(n)$ derived as follows:

$$\bar{x}^c(n) = \frac{1}{N_{\text{trial}}} \sum x^c(n) \tag{4}$$

where $N_{\text{trial}}$ denotes the number of one target's training trials.

After calculating the Fourier Transform (FT) of the averaged EEG data $\bar{x}^c(n)$, we finally derive the SSVEP signal $s^c(n)$ as follows:

$$s^c(n) = \sum_{k=1}^{K} |\bar{X}^c(\mathrm{e}^{\mathrm{j}w})| \cos(kwn + \arg(\bar{X}^c(\mathrm{e}^{\mathrm{j}w}))) \tag{5}$$

where $K$ denotes the number of harmonics of the SSVEP signals, $w = 2\pi f$ represents the digital angular frequency, $|\cdot|$ is the amplitude of a complex number, $\arg(\cdot)$ is the phase of a complex number, and $\bar{X}^c(\mathrm{e}^{\mathrm{j}w})$ is the FT of $\bar{x}^c(n)$, which is derived as follows:

$$\bar{X}^c(\mathrm{e}^{\mathrm{j}w}) = \sum_{n=0}^{L-1} \bar{x}^c(n)\mathrm{e}^{-\mathrm{j}wn} \tag{6}$$

In the single-target offline experiment, the total number of BCI targets $N$ was 16 ([8 : 0.5 : 15.5] Hz) and the number of training targets $N_{\text{train}}$ was 8 ([8 : 15] Hz).

#### 2.2.2 Step 2: Parameter identification

After obtaining the training SSVEP signal $s^c(n)$ (see Eq. (5)), the next step is to identify the parameters of the Volterra model, as follows:

(1) To establish the observed signal vector $\boldsymbol{s}^c$: The cascading SSVEP data of channel $c$ in ascending order of stimulus frequencies are derived as follows:

$$\boldsymbol{s}^c = [(\boldsymbol{s}_1^c)^{\mathrm{T}}, \ldots, (\boldsymbol{s}_{n_{\text{train}}}^c)^{\mathrm{T}}, \ldots, (\boldsymbol{s}_{N_{\text{train}}}^c)^{\mathrm{T}}]^{\mathrm{T}} \tag{7}$$

where $\boldsymbol{s}_{n_{\text{train}}}^c$ denotes the SSVEP signal of target $n_{\text{train}}$ at channel $c$, which is defined as follows:

$$\boldsymbol{s}_{n_{\text{train}}}^c = [s^c(0), s^c(1), \ldots, s^c(L-1)], \quad n_{\text{train}} = 1, 2, \ldots, N_{\text{train}} \tag{8}$$

(2) To establish data matrix $\boldsymbol{A}$:

$$\boldsymbol{A} = [\boldsymbol{A}_1^{\mathrm{T}}, \boldsymbol{A}_2^{\mathrm{T}}, \ldots, \boldsymbol{A}_{N_{\text{train}}}^{\mathrm{T}}]^{\mathrm{T}} \tag{9}$$

where the $n$-th row of $\boldsymbol{A}_{n_{\text{train}}}$ is expressed as follows:

$$\boldsymbol{A}_{n_{\text{train}}}[n,:] = [1, u(n), \ldots, u(n-N_1+1), \ldots, \underbrace{u(n)\cdots u(n)}_{P \text{ times}}, \ldots, \underbrace{u(n-N_p+1)\cdots u(n-N_p+1)}_{P \text{ times}}], \quad n_{\text{train}} = 1, 2, \ldots, N_{\text{train}} \tag{10}$$

(3) To establish parameter vector $\boldsymbol{\theta}^c$:

$$\boldsymbol{\theta}^c = [h_0, h_1(0), \ldots, h_1(N_1-1), \ldots, h_p \underbrace{(0, \ldots, 0)}_{P \text{ times}}, \ldots, h_p \underbrace{(N_p-1, \ldots, N_p-1)}_{P \text{ times}}]^{\mathrm{T}} \tag{11}$$

(4) To construct the linear equation:

$$\boldsymbol{s}^c = \boldsymbol{A}\boldsymbol{\theta}^c \tag{12}$$

(5) To get the LS solution $\hat{\boldsymbol{\theta}}^c$:

$$\hat{\boldsymbol{\theta}}^c = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{s}^c \tag{13}$$

After obtaining the LS solution $\hat{\boldsymbol{\theta}}^c$ of the parameter $\boldsymbol{\theta}^c$, each order impulse response of the Volterra model $h_p^c(m_1, \ldots, m_p)$ is determined. To intuitively understand the physical meaning of the Volterra model, each order frequency response of the system $H_p^c(k_1, \ldots, k_p)$ is calculated using the $N_p$ points Fast Fourier Transform (FFT) as follows:

$$H_p^c(k_1, \ldots, k_p) = \sum_{m_1=0}^{N_p-1} \cdots \sum_{m_p=0}^{N_p-1} h_p^c(m_1, \ldots, m_p) \times \mathrm{e}^{-\mathrm{j}\frac{2\pi}{N_p}m_1k_1} \cdots \mathrm{e}^{-\mathrm{j}\frac{2\pi}{N_p}m_pk_p}, \quad k_i = 0, \ldots, N_p-1, i = 1, \ldots, p \tag{14}$$

#### 2.2.3 Step 3: Template reconstruction

Given that the responses of adjacent stimulus frequencies have similar spatial distribution[29], this study assumed that the amplitude of each harmonic of the SSVEP signals changed continuously at adjacent stimulus frequencies, which will be verified in Section 3.2. Therefore, the amplitude of each harmonic of the SSVEP signals at the stimulus frequencies can be transferred to all stimulus frequencies by polynomial interpolation. Furthermore, with the same initial phase of the stimulus signals, each harmonic of the SSVEP responses has an approximately linear phase. The phase of each harmonic of the SSVEP data at the stimulus frequencies can thereby be transferred to all stimulus frequencies by linear interpolation. In Appendix, the frequency response of the Volterra model is capable of characterizing the amplitude and phase of the SSVEP responses. As a result, we can perform linear interpolation on the each-order frequency response of the Volterra model. The energy of the SSVEP response is mainly concentrated at the fundamental and second

harmonics of the stimulus frequency, which play an important role in SSVEP detection[9]. Therefore, the first- and second-order frequency responses of the Volterra model, which correspond to the fundamental and second harmonics of the SSVEP signals, respectively, are taken as examples to explain the transfer process in detail.

The first-order frequency response of the model $H_1^c(k_1)$ obtained in Step 2 has meaningful values at the training frequencies. For the sake of simplicity, the amplitude $|\hat{H}_1^c(k_1)|$ and the phase $\arg(\hat{H}_1^c(k_1))$ at other frequencies are obtained using linear interpolation, resulting in the frequency response after interpolation derived as follows:

$$\hat{H}_1^c(k_1) = |\hat{H}_1^c(k_1)|e^{j\arg(\hat{H}_1^c(k_1))} \tag{15}$$

Next, the first-order impulse response $\hat{h}_1^c(m_1)$ can be obtained by employing the $N_1$-point inverse FFT as follows:

$$\hat{h}_1^c(m_1) = \frac{1}{N_1}\sum_{k_1=0}^{N_1-1}\hat{H}_1^c(k_1)e^{j\frac{2\pi}{N_1}m_1k_1},$$
$$m_1 = 0,\ldots,N_1-1 \tag{16}$$

Finally, the reconstructed fundamental wave of the SSVEP response $\hat{s}_1^c(n)$ can be obtained by employing the following convolution operation:

$$\hat{s}_1^c(n) = \sum_{m_1=0}^{N_1-1}\hat{h}_1^c(m_1)u(n-m_1) \tag{17}$$

Similarly, the second-order frequency response of the model $H_2^c(k_1,k_2)$ obtained in Step 2 has meaningful values at the training frequencies. The amplitude $|\hat{H}_2^c(k_1,k_2)|$ and the phase $\arg(\hat{H}_2^c(k_1,k_2))$ at other frequencies are obtained using linear interpolation, resulting in the frequency response after interpolation derived as follows:

$$\hat{H}_2^c(k_1,k_2) = |\hat{H}_2^c(k_1,k_2)|e^{j\arg(\hat{H}_2^c(k_1,k_2))} \tag{18}$$

Next, the second-order impulse response $\hat{h}_2^c(m_1,m_2)$ can be obtained by employing the $N_2$-point two-dimensional inverse FFT as follows:

$$\hat{h}_2^c(m_1,m_2) = \frac{1}{(N_2)^2}\sum_{k_1=0}^{N_2-1}\sum_{k_2=0}^{N_2-1}\hat{H}_2^c(k_1,k_2)e^{j\frac{2\pi}{N_2}m_1k_1}\times$$
$$e^{j\frac{2\pi}{N_2}m_2k_2}, m_i = 0,\ldots,N_2-1, i = 1,2 \tag{19}$$

Finally, the reconstructed second harmonic of the SSVEP response $\hat{s}_2^c(n)$ can be obtained by employing the following two-dimensional convolution operation:

$$\hat{s}_2^c(n) = \sum_{m_1=0}^{N_2-1}\sum_{m_2=0}^{N_2-1}\hat{h}_2^c(m_1,m_2)u(n-m_1)u(n-m_2) \tag{20}$$

After obtaining the reconstructed fundamental wave and second harmonic of the SSVEP response, the reconstructed SSVEP signals are also determined, that is

$$\hat{s}^c(n) = \hat{s}_1^c(n) + \hat{s}_2^c(n) \tag{21}$$

#### 2.2.4 Step 4: Performance evaluation

To evaluate the performance of the proposed transfer algorithm, the transferred template signals were verified by computing the correlation between reconstructed $\hat{s}^c(n)$ and observed $s^c(n)$ SSVEP signals and comparing their frequency spectrum.

The performance of the proposed transfer algorithm was further evaluated in terms of its classification accuracy by comparing the accuracy of different template-based classification methods, including the classic non-training algorithms CCA and FBCCA, and the novel representative training algorithms eCCA and TRCA. Table 1 lists the details of these methods.

To our knowledge, the M1 (eCCA) method uses both the standard sine-cosine template and training EEG template $\bar{x}^c(n)$ (see Eq. (4)). Compared with the M1 method, the M2 method, a variant of eCCA, replaced the training EEG reference signal with the training SSVEP template $s^c(n)$ (see Eq. (5)). Similarly, the proposed cross-target transfer algorithm M3 (t-eCCA) replaced the training EEG reference signal with the transferred SSVEP template $\hat{s}^c(n)$ (see Eq. (21)). The number of harmonics $N_h$ of the sine-cosine template was set to 2 in all methods. To verify the feasibility of the SSVEP template, the M2 method was compared with the M1, M4, and M5 methods on the basis of classification accuracy. Moreover, to verify the efficacy of the proposed transfer algorithm, the M3 method was compared with the M1, M4, M5, and M6 methods on the basis of classification accuracy.

The following method was used to calculate the average classification accuracy: For the training algorithm, one block of data was for training and five blocks of data were for testing at a certain time, with

**Table 1 Different templates used in different algorithms.**

| Method | Algorithm | Template |
|---|---|---|
| M1 | eCCA | Sine-cosine template + training EEG template |
| M2 | Variant of eCCA | Sine-cosine template + training SSVEP template |
| M3 | t-eCCA | Sine-cosine template + transferred SSVEP template |
| M4 | CCA | Sine-cosine template |
| M5 | FBCCA | Sine-cosine template |
| M6 | TRCA | Training EEG template |

the final classification accuracy calculated by averaging the accuracies of the six blocks. For the non-training algorithm, all six blocks of data were for testing, with the final classification accuracy calculated by averaging the accuracies of the six blocks. All of the methods described in Table 1 used all of the EEG data recorded by nine electrodes in the experiment.

### 2.3 Experiment

To verify the efficacy of the proposed cross-target transfer algorithm, the prerequisite was to design an appropriate experiment. Given that Han et al.[25] preliminarily explored the feasibility of reducing the number of training targets in SSVEP-BCI, part of the experimental design of this study used the same setup. The specific experimental setup is described in the subsequent section.

#### 2.3.1 Experimental procedure

The single-target offline experiment consisted of six blocks, with each block containing 80 trials (16 targets × 5 repetitions). All trials were conducted in random order. Each trial lasted for 3 s. Specifically, each trial started with a blank for 1 s and the subjects were only allowed to blink for a break during this period. After the blank offset, a sinusoidal stimulus appeared on the screen which lasted 2 s, and the subjects were asked to pay attention and gaze at the visual flash during this period. The flowchart of the experiment is shown in Fig. 2.

#### 2.3.2 Subject

A total of 12 healthy subjects (10 males and 2 females, with the mean age of 25 years) with normal or corrected-to-normal eyesight, participated in the experiment. Each subject signed the informed consent before the experiment and was paid for his/her participation. This study was approved by the Research Ethics Committee of Tsinghua University.

#### 2.3.3 Data acquisition

All of the EEG data were acquired using a SynAmps2 system (Neuroscan, Inc.) at a sampling rate of 1000 Hz. The usable bandwidth was 1–100 Hz. Nine electrodes (Pz, PO5, PO3, POz, PO4, PO6, O1, Oz, and O2) under the international 10-10 system were used in the experiment, with the reference electrode at Cz and the ground electrode at Fz. Electrode impedances were kept below 10 kΩ for all electrodes. During the EEG recording, the subjects were seated on a comfortable chair approximately 70 cm away from the screen in a shielded room.

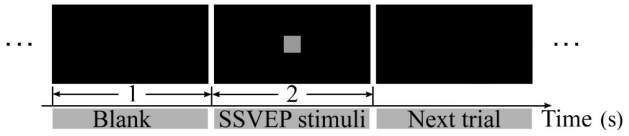


**Fig. 2 Flowchart of the single-target offline experiment.**

#### 2.3.4 Stimulation presentation

Referring to part of the experimental design of a previous study[25], the frequencies of 16 visual stimuli were selected from 8 to 15.5 Hz with an interval of 0.5 Hz, and the initial phase was 0 rad. The stimulus was rendered within a 250 pixel × 250 pixel square at the center of the screen. For each trial, Eq. (22) was used to generate the luminance $u(n)$ of the stimulus with frequency $f$ and initial phase $\phi$, where $n$ indicates the frame index in the stimulus sequence.

$$u(n)=\frac{1}{2}\left[1+\cos\left(2\pi f\left(\frac{n}{\text{RefreshRate}}\right)+\phi\right)\right] \tag{22}$$

The visual stimulus was presented on a 23.6-in screen with a resolution of 1920 pixel × 1080 pixel and a refresh rate of 60 Hz. The stimulation program was developed using the Psychophysics Toolbox[30] in the MATLAB software.

#### 2.3.5 Data preprocessing

Data epochs were extracted on the basis of event triggers marked by the stimulation program. On the basis of the apparent latency of the visual system[4], the data with a time window of [0.14, 2.14] s (time 0 represented the SSVEP stimulus onset) was intercepted, downsampled to 250 Hz, and band-pass filtered from 7 to 32 Hz using the Parks-McClellan optimal Finite Impulse Response (FIR) filter. The filtfilt() method was employed to implement the zero-phase FIR filter in MATLAB.

## 3 Result

### 3.1 Response of the Volterra model

To obtain an in-depth understanding of the Volterra model, the SSVEP data of all targets (and its stimulus frequencies = [8 : 0.5 : 15.5] Hz) were selected as the response signals of different signal models, with 2 s data length of each target (corresponding to 0.5 Hz interval at adjacent stimulus frequencies).

Figure 3 shows the fitting results of Subject S4 to all targets at the Oz channel in the time domain waveform using the Volterra model. Obviously, the response of the Volterra model could fit the SSVEP response of each target well, with the fitting amplitude and phase close to

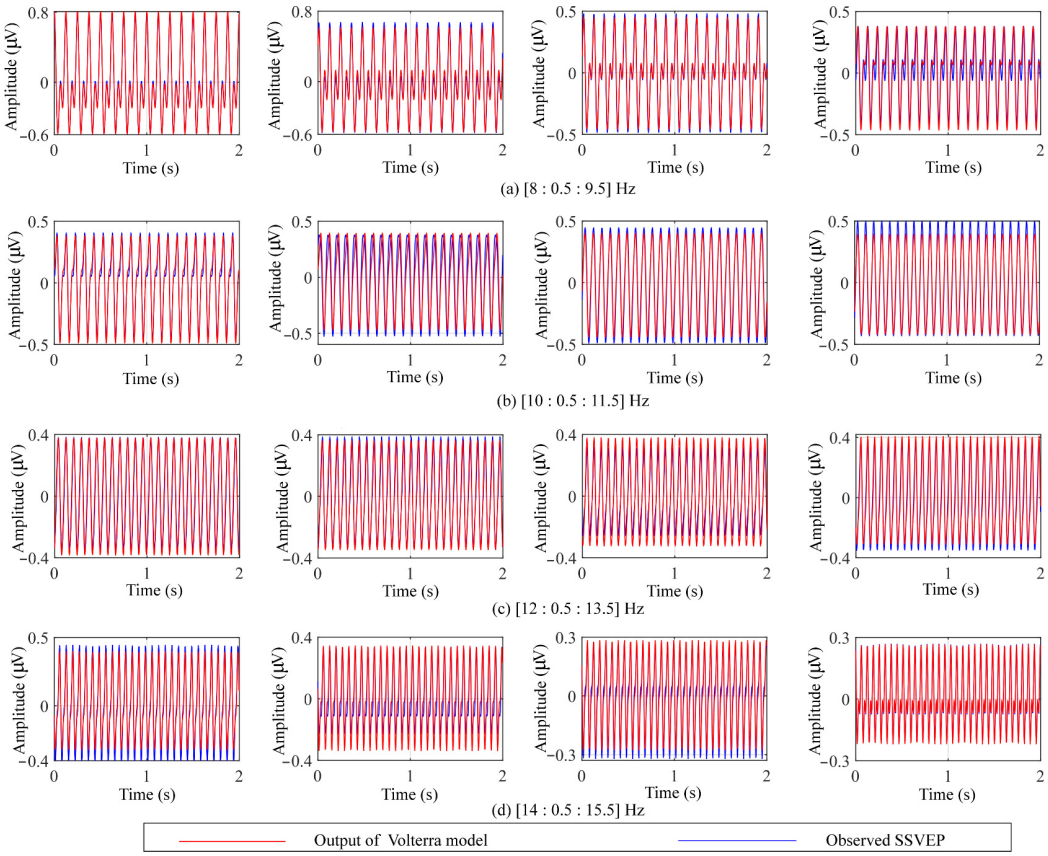


**Fig. 3 Fitting results of Subject S4 at the Oz channel in the time domain waveform for all visual stimuli at stimulus frequencies of [8 : 0.5 : 15.5] Hz using the Volterra model.**

the actual results.

Figure 4 illustrates the comparison results of the frequency response of the Volterra model and the frequency spectrum of the SSVEP signals of Subject S4 at the Oz channel. Figure 4a shows the comparison between the first-order amplitude frequency characteristic curve of the Volterra system and the amplitude of the fundamental wave of the SSVEP. As shown in Fig. 4a, while maintaining a high fitting accuracy, the Volterra model had smooth amplitude frequency characteristic curve, which meant that the model had good generalization capability and partly reduced the effect of spontaneous EEG activities. Similarly, the second-order amplitude frequency characteristic of the model yielded analogical results (see Fig. 4c). Figures 4b and 4d show the first-order and second-order phase frequency characteristics of the Volterra model, respectively. Notably, the phase response of SSVEP was approximately linear, but exhibited shifts in the phase response at 10.5 Hz (see Fig. 4b) and 23 Hz (see Fig. 4d) because averaging the EEG data could only reduce, not eliminate, the effect of spontaneous EEG activities. Thus, $s^c(n)$ (see Eq. (5)) could only be an estimate of the true evoked components of SSVEP. Obviously, the model could describe the approximate linear phase characteristics of each harmonic of the SSVEP response. The results indicated that the system under the SSVEP paradigm could be modeled as the Volterra model, so that the nonlinear transfer function theory could be used to analyze the SSVEP response.

The amplitude of the SSVEP response at adjacent stimulus frequencies was continuously changing and the phase was approximately linear (see Fig. 4), which showed that the features of SSVEP response in the frequency domain had a relatively strong correlation at adjacent stimulus frequencies. The key concept of the cross-target transfer algorithm is to exploit the correlation of SSVEP signal characteristics between different targets. Therefore, the SSVEP information at other stimulus frequencies could be obtained using interpolation by exploiting the SSVEP information at the training frequencies, which achieved the reconstructed SSVEP reference signals. In other words, it was feasible

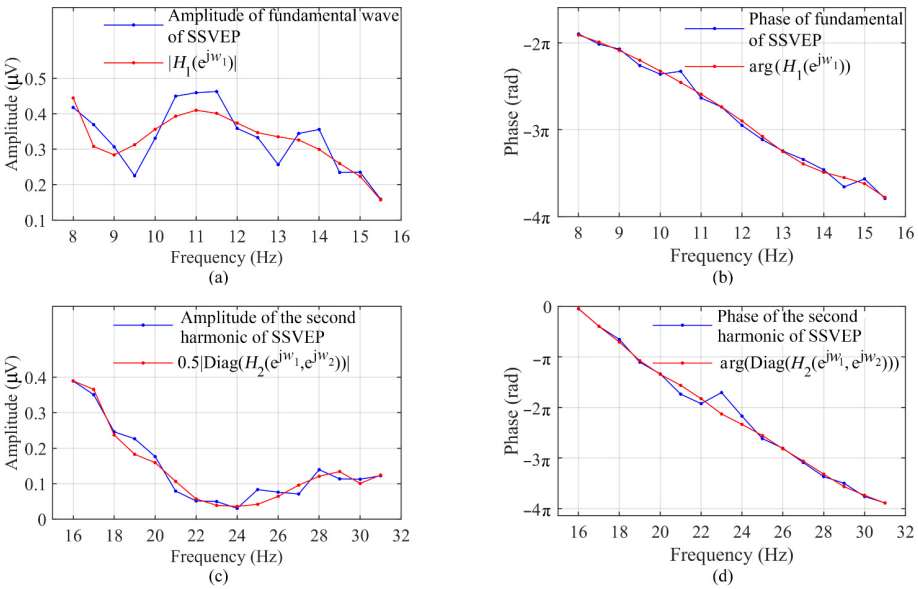


**Fig. 4 Identification results of Subject S4 at the Oz channel using the Volterra model. (a) First-order amplitude frequency response of the Volterra model. (b) First-order phase frequency response of the Volterra model. (c) Second-order amplitude frequency response of the Volterra model, which only showed the amplitude of the diagonal component of $H_2^c(e^{jw_1}, e^{jw_2})$ (see Eq. (A8)). (d) Second-order phase frequency response of the Volterra model, which only showed the phase of the diagonal component of $H_2^c(e^{jw_1}, e^{jw_2})$ (see Eq. (A8)).**

to transfer between different targets through only a part of the stimulus frequencies (training targets). We further analyzed the experimental data to test the efficacy of transferred SSVEP signals and the identification performance of the transfer algorithm.

### 3.2 Evaluation of the transferred SSVEP signal

To verify the efficacy of the transferred SSVEP signal, the SSVEP data of eight targets (and its stimulus frequencies = [8 : 15] Hz) were selected, with 1 s data length of each target (corresponding to 1 Hz interval at adjacent stimulus frequencies). The transferred SSVEP template was obtained by following the steps presented in Section 2.2. Figure 5 shows the comparison results of the frequency domain characteristic of the transferred and observed SSVEP signals of Subject S4 at the Oz channel. As shown in Fig. 5, the reconstructed SSVEP signal was highly similar to the original SSVEP response in the frequency domain. The evoked component $s^c(n)$ (see Eq. (5)) obtained by averaging was only an approximation of the true evoked EEG, which also included part of spontaneous EEG activities. To ensure that the Volterra model has good generalization capability and reduce the influence of spontaneous EEG activities, completely fitting the observed data was not our purpose, which explained why the fitting result was not exactly the same as the observation result (see Fig. 5a (left)). If necessary, to ensure the fitting accuracy of the model, we increased the memory span of the Volterra model as the data length increased.

We further calculated the average correlation coefficient matrix between transferred and original SSVEP signals across all subjects at the Oz channel (see Fig. 6). Evidently, the matrix showed the highest correlation between transferred and original SSVEP signals at the same stimulus frequency (correlation coefficients >0.9), but transferred and original SSVEP signals were nearly uncorrelated at different stimulus frequencies. Furthermore, the average correlation coefficients at the training frequencies of [8 : 15] Hz were significantly higher than that at the transferred frequencies of [8.5 : 15.5] Hz ($0.996 \pm 0.004$ vs. $0.968 \pm 0.020$, $p < 0.05$). In conclusion, the proposed transfer algorithm was able to effectively transfer SSVEP information at different stimulus frequencies.

### 3.3 Evaluation of the classification performance

#### 3.3.1 Efficacy testing of the training SSVEP template

Figure 7 shows the average classification accuracy across 12 subjects. We calculated the classification results for two kinds of training template-based eCCA methods using different data lengths from 0.5 to 2 s with an interval of 0.25 s. Section 2.2.4 described the three kinds of template-based eCCA methods in detail. Notably, with the increase in time window length,

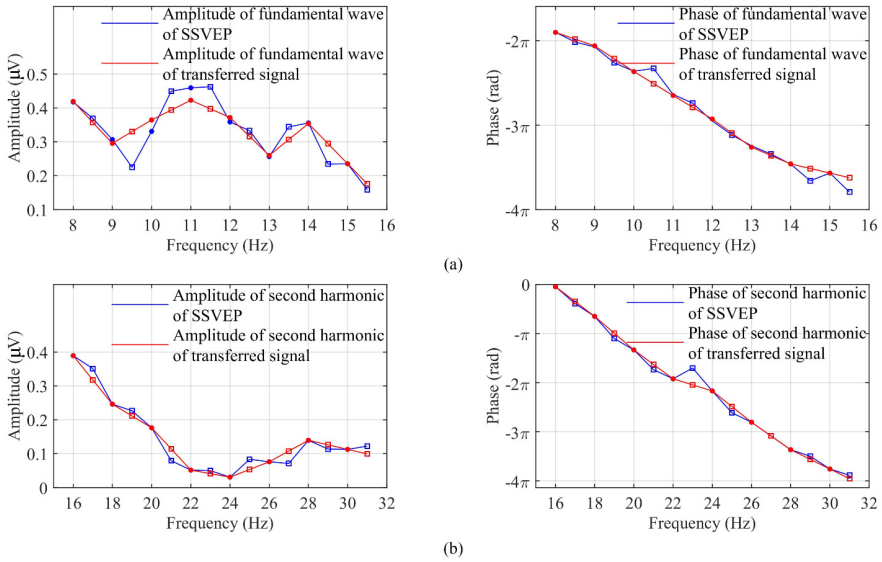


**Fig. 5 Comparison results of (a) amplitude of the fundamental wave (left), phase of the fundamental wave (right) and (b) amplitude of the second harmonic (left), phase of the second harmonic (right) of the transferred and observed SSVEP signals of Subject S4 at the Oz channel using the Volterra model. The dot marker indicates the training frequencies ([8 : 15] Hz) and the double frequencies ([16 : 2 : 30] Hz); the square marker indicates the transferred frequencies ([8.5 : 15.5] Hz) and the double frequencies ([17 : 2 : 31] Hz)**

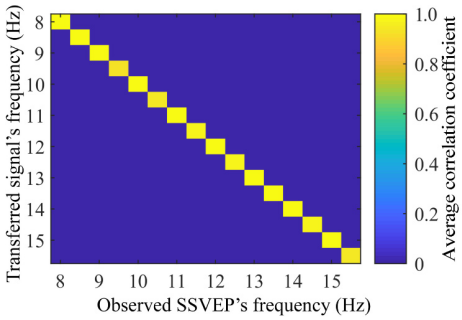


**Fig. 6 Average correlation matrix across all subjects at the Oz channel between transferred and observed SSVEP signals at different stimulus frequencies.**

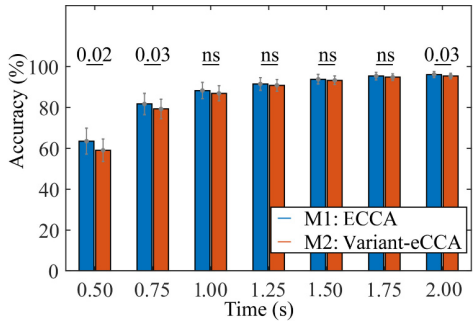


**Fig. 7 Target identification accuracy calculated by eCCA and the variant of eCCA using different data lengths from 0.5 to 1.0 s with an interval of 0.25 s. The numbers above the horizontal bars represent the $p$ values between different methods, ns denotes not significant, and the error bars indicate the standard errors.**

the difference in classification accuracies between M2 (variant of eCCA) and M1 (standard eCCA) decreased continually from 4.46% (M2 vs. M1 : 59.05% ± 19.26% vs. 63.51% ± 22.08%) with 0.5 s data length to 0.75% (M2 vs. M1: 95.40%±4.62% vs. 96.15%±5.16%) with 2 s data length. With the decrease in time window length (less than 1 s), the classification performance of M2 was slightly lower than that of M1 ($p$ >0.01 but $p$ <0.05). In fact, the steady-state response of SSVEP mainly appears after 0.5 s stimulus duration, before which the transient response dominates[31, 32].

The M1 method used the EEG template, which contains the transient response of the first 0.5 s, whereas the M2 used the SSVEP template which does not consider the transient process, explaining the finding that the classification performance of M2 was slightly lower than that of M1 with a short time window. To make the BCI system suitable for most participants, the data length used for target identification should not be short (generally at least 1 s), so that the system can maintain a relatively high accuracy (average accuracy higher than 90%). With 1 s data length, no significant difference between the accuracies of M2 and M1 was observed (M2 vs. M1 : 86.94% ± 12.91% vs. 88.32% ± 13.97%, paired t-test: $p$ >0.1). Similarly, with other data lengths longer than 1 s ([1.25 : 0.25 : 1.75] s), no significant difference between the accuracies of M2 and M1 was

observed ($p$ >0.1). Although the accuracies of these two methods exhibited a significant difference ($p$ = 0.03) with 2 s data length, the eCCA based on the training SSVEP template was still effective, which will be explained subsequently. Table 2 lists the classification accuracies of each subject calculated by different template-based eCCA methods with 2 s data length. With 2 s data length, the difference in average accuracies between M2 and M1 was only 0.75% (M2 vs. M1: 95.40% ± 4.62% vs. 96.15% ± 5.16%) and was merely 2.04% (for S2, M2 vs. M1 : 93.50% vs. 95.54%) in the worst case. More importantly, compared with the non-training algorithms M4 and M5 (for more details about M4 and M5, see Fig. 8), the classification performance of M2 had been considerably improved. The classification accuracy of M1 was slightly higher than that of M2 and M3

**Table 2 Detailed target identification accuracy of 12 subjects with 2 s data length calculated by different methods.** (%)

| Subject | M1 | M2 | M3 | M6 |
|---|---|---|---|---|
| S1 | 81.13 | 82.33 | 80.83 | 83.50 |
| S2 | 95.54 | 93.50 | 93.63 | 96.67 |
| S3 | 99.38 | 98.63 | 97.79 | 98.71 |
| S4 | 94.83 | 93.50 | 94.08 | 93.71 |
| S5 | 98.04 | 97.04 | 97.63 | 98.50 |
| S6 | 99.50 | 97.88 | 97.92 | 98.88 |
| S7 | 92.17 | 91.29 | 91.42 | 93.96 |
| S8 | 99.83 | 98.33 | 98.88 | 99.29 |
| S9 | 99.79 | 99.17 | 99.33 | 99.67 |
| S10 | 99.75 | 98.50 | 98.67 | 98.96 |
| S11 | 99.17 | 98.50 | 98.63 | 98.75 |
| S12 | 94.71 | 96.08 | 94.92 | 93.88 |
| Mean | 96.15 | 95.40 | 95.31 | 96.20 |
| Std | 5.16 | 4.62 | 4.99 | 4.40 |

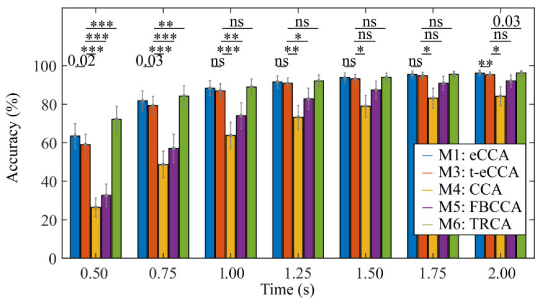


**Fig. 8 Target identification accuracy as functions of data length (from 0.5 s to 2 s with an interval of 0.25 s) calculated by different methods. The asterisks and numbers above the horizontal bars represent the $p$ values between M3 and other methods (∗: $p$<0.05; ∗∗: $p$<0.005; ∗∗∗: $p$<0.0005), ns denotes not significant, and the error bars indicate the standard errors.**

because M1 used the EEG template signal $\bar{x}^c(n)$ (see Eq. (4)), which contains not only evoked components but also spontaneous EEG components, whereas M2 and M3 only used the evoked components. This meant that spontaneous EEG components contributed to the actual EEG signal detection, but the evoked components dominated in the classification process. The results showed that the SSVEP signal can be feasibly used as template for eCCA in practical applications.

### 3.3.2 Efficacy testing of the transfer algorithm

Figure 8 illustrates the average classification accuracy across all subjects. The classification results were computed for five kinds of template-based methods with different data lengths ([0.5 : 0.25 : 2] s). First, the performance difference between M3 (t-eCCA) and M1 (standard eCCA) was analyzed. Evidently, with the increase in data length, the difference in classification performance between M3 and M1 decreased continually from 4.47% (M3 vs. M1 : 59.04% ± 18.52% vs. 63.51% ± 22.08%) with 0.5 s data length to 0.84% (M3 vs. M1 : 95.31% ± 4.99% vs. 96.15% ± 5.16%) with 2 s data length. With a short data length (less than 1 s), the classification accuracy of M3 was slightly lower than that of M1 ($p$ >0.01 but $p$ <0.05).

The classification accuracy of M1 was slightly higher than that of M3 with a short data length because the template signal of M1 includes the transient response, whereas that of M3 only considers the steady-state response. With 1 s data length, no significant difference between the classification performance of M2 and M1 was observed (M3 vs. M1 : 86.96% ± 12.87% vs. 88.32% ± 13.97%, $p$>0.1). Similarly, with other data lengths longer than 1 s ([1.25 : 0.25 : 1.75] s), no significant difference between the classification performance of M3 and M1 was observed ($p$ >0.05). Notably, although a significant difference between M3 and M1 with 2 s data length was observed ($p$ <0.005), the following interpretation showed that the proposed cross-target transfer algorithm M3 was still effective. As shown in Table 2, with 2 s data length, the difference in average accuracies between M3 and M1 was merely 0.84% (M3 vs. M1 : 95.31% ± 4.99% vs. 96.15% ± 5.6%) and was 1.91% (for S2, M3 vs. M1 : 93.63% vs. 95.54%) in the worst case. In other words, with 2 s data length, the actual performance of M3 was slightly different from that of M1. Moreover, the method used to analyze the classification performance of M3 and M6 (TRCA) at different data lengths was basically the same

as that used to analyze the classification performance of M1. The detailed classification results of M3 and M6 are shown in Fig. 8 and Table 2. The results proved that, while saving half of the training time, M3 exhibited a similar performance to M1 and M6. Thus, the proposed cross-target transfer algorithm t-eCCA was effective.

Next, the performance difference between M3 and the classic non-training algorithms M4 (CCA) and M5 (FBCCA) was analyzed. As shown in Fig. 8, with all data lengths, the classification performance of M3 was significantly higher than that of M4 (i.e., with 1 s data length, M3 vs. M4 : 86.96% ± 12.87% vs. 63.82% ± 23.42%, $p < 0.0005$). With different data lengths ([0.5 : 0.25 : 1.25] s), the performance of M3 was significantly higher than that of M5 (i.e., with 1 s data length, M3 vs. M5: 86.96% ± 12.87% vs. 74.05% ± 23.08%, $p < 0.005$). The following would explain why the overall performance of M3 was still better than that of M5 although no significant difference between them with data lengths from 1.5 to 2 s was observed ($p > 0.05$). Table 3 lists in detail the classification accuracy of each subject calculated by M3 and M5 with data lengths from 1.5 to 2 s. The data shown in Table 3 could be divided into two categories: the first category of data with classification accuracy of M5 higher than 98.33% (data in bold in Table 3) and the second category of data with classification accuracy of M5 lower than 98.33%. For the first category of data, although the performance of M3 was slightly lower than that of M5, the classification accuracies are both close to 100% (for 1.5 s data length, M3 vs. M5: 97.93% ± 0.87% vs. 99.79% ± 0.17%; for 1.75 s data length, M3 vs. M5: 98.47% ± 0.67% vs. 99.50% ± 0.57%; and for 2 s data length, M3 vs. M5: 98.37% ± 0.63% vs. 99.44% ± 0.59%). For the second category of data, with all of the three data lengths ([1.5 : 0.25 : 2] s), the performance of M3 was always higher than that of M5 (for 1.5 s data length, M3 vs. M5: 91.59% ± 8.25% vs. 83.54% ± 16.47%; for 1.75 s data length, M3 vs. M5: 92.20% ± 6.54% vs. 84.73% ± 13.11%; and for 2 s data length, M3 vs. M5: 92.25% ± 5.54% vs. 84.69% ± 11.40%, $p > 0.1$). This meant that, for several participants, when the classification performance still had much room for improvement, the proposed algorithm t-eCCA had a good improvement effect. The results showed that the performance of the proposed cross-target transfer algorithm was higher than that of the classic non-training algorithms CCA and FBCCA.

**Table 3 Detailed target identification accuracy of 12 subjects as functions of data length (from 1.5 to 2 s with an interval of 0.25 s) calculated by the M3 and M5 methods.**

(%)

| | 1.5 s | | 1.75 s | | 2 s | |
|---|---|---|---|---|---|---|
| Subject | M3 | M5 | M3 | M5 | M3 | M5 |
| S1 | 70.58 | 45.42 | 77.46 | 57.08 | 80.83 | 62.08 |
| S2 | 90.92 | 71.88 | 93.33 | 82.08 | 93.63 | 84.79 |
| S3 | **96.75** | **99.58** | **97.25** | **99.79** | **97.79** | **99.79** |
| S4 | 92.63 | 89.58 | 94.33 | 93.13 | 94.08 | 92.29 |
| S5 | 97.54 | 96.04 | 97.75 | 96.88 | **97.63** | **98.54** |
| S6 | 96.71 | 96.25 | 97.67 | 98.33 | **97.92** | **98.75** |
| S7 | 87.29 | 73.75 | 89.67 | 79.58 | 91.42 | 81.88 |
| S8 | 98.38 | 98.13 | **98.75** | **99.17** | **98.88** | **99.58** |
| S9 | **98.83** | **100.00** | **99.29** | **100.00** | **99.33** | **100.00** |
| S10 | **98.21** | **99.79** | **98.54** | **100.00** | **98.67** | **100.00** |
| S11 | 98.13 | 97.08 | **98.50** | **98.54** | 98.63 | 98.33 |
| S12 | 92.17 | 81.04 | 95.17 | 86.04 | 94.92 | 88.75 |
| Mean | 93.18 | 87.38 | 94.81 | 90.89 | 95.31 | 92.07 |
| Std | 7.66 | 15.96 | 5.89 | 12.38 | 4.99 | 10.94 |

On the basis of the analysis results, we determined that when using only half of the training time, the transfer algorithm t-eCCA exhibited a classification performance that is close to that of the representative training algorithms eCCA and TRCA, and significantly higher than that of the classic non-training algorithms CCA and FBCCA.

## 4 Discussion

### 4.1 Further understanding of the Volterra model

#### 4.1.1 Comparison of different signal models

As discussed in Section 2.1, when the order of the Volterra model $P$ equals 1 ($P = 1$), it degenerates into the Moving Average (MA) model as follows:

$$s_c(n) = \sum_{m_1=0}^{N_1-1} h_1^c(m_1)u(n-m_1) \tag{23}$$

When the order of the Volterra model $P$ equals 2 ($P = 2$) and both memory spans equal 1 ($N_1 = N_2 = 1$), it degenerates into the Quadratic Polynomial (QP) model as follows:

$$s^c(n) = h_0^c + h_1^c u(n) + h_2^c u^2(n) \tag{24}$$

The experimental results illustrate the rationality of the Volterra model as the SSVEP signal model compared with the MA and QP models. The SSVEP data of eight targets (and its stimulus frequency range is 8–15 Hz) were selected as the response signals of different signal models, with 1 s data length of each target (corresponding to 1 Hz interval at adjacent stimulus frequencies).

Figure 9 shows the results of different signal models for a typical subject. The identification results in the time domain waveform (see Fig. 9a) and amplitude spectrum (see Fig. 9b) show that MA could only characterize the fundamental component of the SSVEP response, not the second harmonic. In addition, although the QP model could simultaneously characterize the fundamental and second harmonic of the SSVEP response, the fitting accuracy was low (see Figs. 9c and 9d). Figures 9e and 9f show that the Volterra model has relatively high modeling capability to accurately describe the fundamental and second harmonic of the SSVEP response. Table 4 shows the Normalized Mean Square Error (NMSE) results for the three different signal models.

As shown in Table 4, the NMSE of the QP model was the largest, that was 1; the NMSE of the MA model reached 0.21; and the NMSE of the Volterra model was the smallest, only 0.04. The results showed that the MA and QP models could not be used as the SSVEP signal model and that the Volterra model could be reasonably used as the SSVEP signal model because of its strong fitting capability.

**Table 4 NMSE of fitting results of different signal models.**

| Model | NMSE |
|---|---|
| MA | 0.21 |
| QP | 1.00 |
| Volterra | 0.04 |

### 4.1.2 Physical interpretation of the Volterra model

As indicated by the $p$-th-order response of the Volterra model $s_p^c(n)$ (see Eq. (3)), the system can effectively characterize the response of the human brain under the SSVEP paradigm as follows:

(1) Nonlinear: the output $s^c(n)$ is expressed as the sum of the products of the input signal $u(n)$ delays, which is a nonlinear process.

(2) Not memoryless: the $p$-th-order response $s^c(n)$ is an $(N_p-1)$-th-order Markov chain, that is, past stimuli also contribute to the current response.

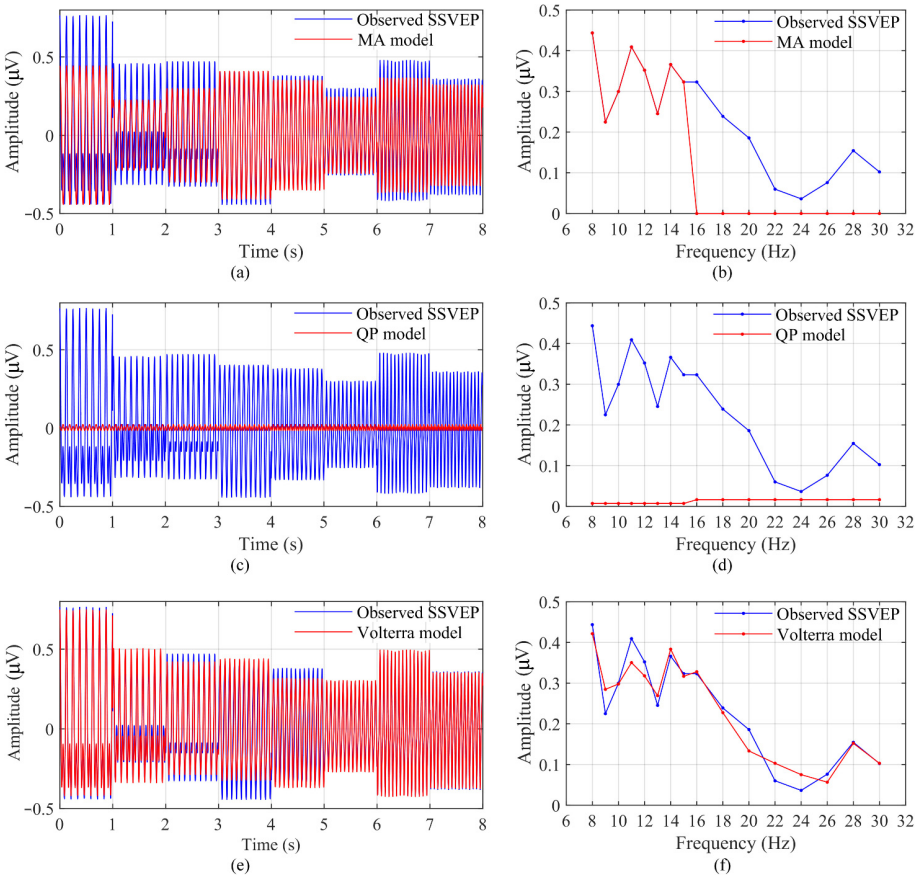


**Fig. 9 Comparison of the fitting results of different signal models in the time domain waveform and amplitude spectrum: (a) fitting result of the MA model in the time domain waveform; (b) fitting result of the MA model in the amplitude spectrum; (c) fitting result of the QP model in the time domain waveform; (d) fitting result of the QP model in the amplitude spectrum; (e) fitting result of the Volterra model in the time domain waveform; and (f) fitting result of the Volterra model in the amplitude spectrum.**

(3) Causal: the lower limit of convolution starts from 0, indicating that the model is a causal system.

To intuitively understand the physical meaning of the Volterra model, the steady-state response of the model was investigated in the frequency domain[28]. When the stimulus signal is the sum of two sinusoidal signals (stimulus frequencies: $f_1$ and $f_2$), the first-order response of the Volterra model $s_1^c(n)$ (see Eq. (A2)) contains only the $f_1$ and $f_2$ frequencies. The first-order response of a sinusoidal signal with frequency $f_i$ is also a sinusoidal signal of the same frequency, which has amplitude weighting (the value $|H_1^c(\mathrm{e}^{\mathrm{j}w_i})|$ of the first-order amplitude frequency response at $f_i$) and phase shift (the value $\arg(H_1^c(\mathrm{e}^{\mathrm{j}w_i}))$ of the first-order phase frequency response at $f_i$). The second-order response of the Volterra model $s_2^c(n)$ (see Eq. (A4)) contains not only the second-frequency components $2f_1$ and $2f_2$, but also the second-order intermodulation $f_1 \pm f_2$, which also have amplitude weighting and phase shift with the values of the second-order frequency response at corresponding frequencies. Obviously, with dual-frequency input, the response of the Volterra model was completely consistent with the nonlinear characteristic of SSVEP response, that is, harmonic and intermodulation.

Specifically, when the stimulus signal is a single sinusoidal signal (digital angular frequency: $w$), the first-order response of the Volterra model $s_1^c(n)$ (see Eq. (A7)) generates the fundamental component of SSVEP, which has amplitude weighting (the value $|H_1^c(\mathrm{e}^{jw})|$ of the first-order amplitude frequency response at $w$) and phase shift (the value $\arg(H_1^c(\mathrm{e}^{\mathrm{j}w}))$ of the first-order phase frequency response at $w$). The second-order response of the Volterra model $s_2^c(n)$ (see Eq. (A8)) produces the second harmonic component, which has amplitude weighting (the value $|H_2^c(\mathrm{e}^{\mathrm{j}w}, \mathrm{e}^{\mathrm{j}w})|/2$ of the second-order amplitude frequency response at $w$) and phase shift (the value $\arg(H_2^c(\mathrm{e}^{\mathrm{j}w}, \mathrm{e}^{\mathrm{j}w})$ of the second-order phase frequency response at $w$).

In general, the $p$-th-order response of the Volterra model includes all of the frequency components of $|af_1 \pm bf_2|, a + b = p, a, b \in \mathbf{N}$. Therefore, the $P$-th-order Volterra system can characterize the highest $P$-th-order harmonic and $P$-th-order intermodulation.

From the previous presented analysis, the Volterra model can characterize the SSVEP response, whose frequency response can describe the frequency features of SSVEP. Given that the Volterra model is a typical representative of the nonlinear transfer function, the transfer function can be used to explain the SSVEP response.

#### 4.1.3 Complexity of the Volterra model

In general, the complexity of the $p$-th-order kernel of the Volterra model is exponential ($O((N_p)^p)$), which means that the number of parameters in the Volterra model increases rapidly with the order of the nonlinearity $P$ and the length of the system memory $N_P$. The large number of coefficients of the Volterra model will limit its application in real-life environments. Specifically, when the order of nonlinearity is large, the model has high complexity and can easily overfit[27]. The energy of SSVEP response is mainly concentrated at the fundamental and second harmonics of the stimulus frequency, which play an important role in SSVEP detection[9]. For most subjects, the data length used for signal detection is short because of the increasing performance of SSVEP classification algorithms[13], which means that the data used to identify the parameters are limited. On the basis of these considerations, this study initially investigated a simple second-order system, that is, the order of the Volterra model is equal to 2 ($P = 2$) and considers only the fundamental and second harmonics of the SSVEP signals.

However, higher harmonic information in SSVEP response can improve the performance of BCI systems[10, 33]. To improve the performance of the proposed transfer algorithm, further study is required to increase the order of the Volterra system. Moreover, to avoid overlearning of the model, we can simplify the complexity of the model using the symmetry of Volterra series expansion[34] and several regularization techniques[35].

### 4.2 Multi-target experiment

This study further introduced a classic benchmark dataset[36] of SSVEP-BCI to test the universality of the proposed transfer algorithm in multi-target experiments. The dataset contained data from 35 subjects, with each subject having six blocks of data and each block including 40 trials (targets). Each trial comprised 6 s recording, including 0.5 s cue time in the beginning, 5 s visual stimulus, and 0.5 s rest time after the stimulation (for more details see Ref. [36]). A total of 20 targets (and its stimulus frequencies = [8 : 0.4 : 15.8] Hz) were selected for training, and the data length of each target was 2.5 s (corresponding to 0.4 Hz interval at adjacent stimulus frequencies) in multi-target experiments. This sixfold cross-validation method was used to calculate

the classification accuracy, that is, five blocks of data used for training and one block for testing.

Figure 10 shows the average classification accuracies across all subjects calculated using different template-based methods at different data lengths ([0.5 : 0.25 : 2.5] s). With all of the data lengths, the performance of M3 was significantly higher than that of M4 ($p < 0.01$). As the data length increased, the performance improvement of M3 relative to M5 continuously decreased from 18.86% (M3 vs. M5: 63.88% ± 17.43% vs. 45.02% ± 18.43%, $p < 0.0001$) with 1 s data length to 0.67% (M3 vs. M5: 86.04% ± 10.33% vs. 86.71% ± 15.96%, $p > 0.05$) with 2.5 s data length. Similarly, with the increase in data length, although the difference between M3 and M1 continuously decreased from 16.37% (M3 vs. M1 : 63.88% ± 17.43% vs. 80.25% ± 21.06%, $p < 0.0001$) with 1 s data length to 6.19% (M3 vs. M1 : 86.04% ± 10.33% vs. 92.23% ± 16.65%, $p < 0.05$) with 2.5 s data length, the performance of M3 was still significantly lower than that of M1. In multi-target experiments, different targets have different spatial positions, and other targets are usually located around the stimulus target. Furthermore, the difference in the spatial location of visual stimuli and the stimulation of the peripheral vision will affect the response of the central visual field[37–39]. However, the influence of spatial information was not included in the signal model, so that the transferred SSVEP template could not accurately contain the response under multi-target stimulation, which decreased the performance of the proposed transfer algorithm. Therefore, further investigations are required to explore the effect of peripheral visual stimuli on the SSVEP response. A feasible approach is to quantitatively compensate the effect of the peripheral stimuli of each target according to the target arrangement of the actual stimulus paradigm to improve the SSVEP signal model, which will further improve the classification performance of the proposed transfer algorithm. In addition, as discussed in Section 3.3, the performance of the proposed transfer algorithm with a short data length still had much room for improvement. Therefore, we consider adding the transient response to the SSVEP signal model to further improve the modeling process. Moreover, the use of the proposed transfer algorithm in practical systems requires further investigation to design an online multi-target experiment to test its efficacy.

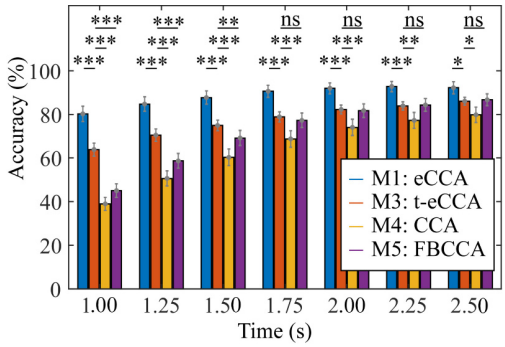


**Fig. 10 Target identification accuracy as a function of data length (from 1 to 2.5 s with an interval of 0.25 s) calculated by different methods. The asterisks above the horizontal bars represent the $p$ values between M3 and other methods (∗: $p < 0.05$; ∗∗: $p < 0.005$; ∗∗∗: $p < 0.0005$), ns indicates not significant, and the error bars indicate the standard errors.**

### 4.3 Percentage of training stimulus frequencies

In the single-target SSVEP experiment with 16 stimulus frequencies, eight stimulus frequencies were used to train the model to obtain the proposed transfer algorithm t-eCCA, with the percentage of training stimulus frequencies of 50%. On the basis of the results presented in Section 3.3.2, with half of the training time, the proposed transfer algorithm t-eCCA achieved a classification accuracy that is close to that of the representative training algorithms eCCA and TRCA. The sampling rate of stimulus frequencies played an important role in the transfer paradigm. Quantitatively evaluating the percentage of training stimulus frequencies could provide a clear guide for collecting training data.

Figure 11 illustrates the classification accuracy of t-eCCA at different sampling rates of stimulus frequencies. Specifically, four different sampling rates (i.e., 25%, 37.5%, 50%, and 62.5%) were compared, corresponding to the number of sampling frequencies 4, 6, 8, and

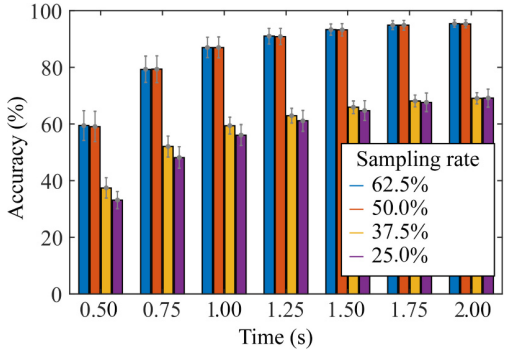


**Fig. 11 Target identification accuracy as functions of data length (from 0.5 to 2 s with an interval of 0.25 s) calculated by the proposed transfer algorithm t-eCCA at different sampling rates (i.e., 25%, 37.5%, 50%, and 62.5%). The error bars indicate the standard errors.**

10. Obviously, as the sampling rate decreased, the classification performance of the proposed transfer algorithm also decreased as a whole. As the sampling rate decreased, but not less than 50%, the classification accuracy of the transfer algorithm remained stable as a whole, close to that of the representative training algorithm eCCA. When the sampling rate was less than 50%, the classification performance of the transfer algorithm decreased sharply because as the sampling rate decreased, the available information also decreased, resulting in insufficient accuracy of the reconstructed template. On the basis of the assumption that the evoked components of SSVEP were relatively stable, a possible solution was to adopt subject-specific interpolation, which was expected to improve the accuracy of the transferred template with a small amount of training data. Considering the training time and classification accuracy, the best sampling rate of stimulus frequencies for the single-target experiment was 50%.

## 5 Conclusion

This study proposed an effective cross-target transfer algorithm for the SSVEP-BCI system. On the basis of the nonlinear transfer function theory, the SSVEP response was modeled as a Volterra filter by taking full advantage of the nonlinear characteristics of SSVEP signals. In addition, the cross-target transfer algorithm based on the Volterra model could effectively reduce the training time of the BCI system by utilizing the correlation of the frequency domain features of SSVEP signals. As a result, in a single-target SSVEP experiment with 16 stimulus frequencies, the proposed transfer algorithm t-eCCA obtained an average accuracy of 86.96% ± 12.87% across 12 subjects using only half of the calibration time, which exhibited no significant difference from the representative training classification algorithms, namely, eCCA (88.32% ± 13.97%) and TRCA (88.92% ± 14.44%), and was significantly higher than that of the classic non-training algorithms, namely, CCA and FBCCA. The results showed that the proposed cross-target transfer algorithm t-eCCA can fully utilize the information about the targets and its stimulus frequencies and effectively reduce the training time of SSVEP-BCI.

### Appendix

The stimulus signal $u(n)$ is assumed to be the sum of two sinusoidal signals (stimulus frequencies: $f_1$ and $f_2$), which can be expressed as follows:

$$u(n) = A_1\cos(w_1 n) + A_2\cos(w_2 n),\quad w_1 \geqslant w_2,\quad n = 0, 1, \ldots, L-1 \tag{A1}$$

where $A_i$ denotes the amplitude of the sinusoidal signal and $w_i = 2\pi f_i$ represents the digital angular frequency ($i = 1$ and 2).

The first-order response of the Volterra model $s_1^c(n)$ is defined as follows:

$$s_1^c(n) = A_1|H_1^c(\mathrm{e}^{\mathrm{j}w_1})|\cos(w_1 n + \arg(H_1^c(\mathrm{e}^{\mathrm{j}w_1}))) + A_2|H_1^c(\mathrm{e}^{\mathrm{j}w_2})|\cos(w_2 n + \arg(H_1^c(\mathrm{e}^{\mathrm{j}w_2}))) \tag{A2}$$

where $|*|$ denotes the amplitude of a complex number and arg() represents the phase of a complex number. $H_1^c(\mathrm{e}^{\mathrm{j}w_1})$ is the first-order frequency response of the Volterra system, which can be expressed as follows:

$$H_1^c(\mathrm{e}^{\mathrm{j}w_1}) = \sum_{m_1=-\infty}^{+\infty} h_1^c(m_1)\mathrm{e}^{-\mathrm{j}w_1 m_1} \tag{A3}$$

The second-order response of Volterra model $s_2^c(n)$ is defined as follows:

$$\begin{aligned} s_2^c(n) =& \frac{A_1^2}{2}|H_2^c(\mathrm{e}^{\mathrm{j}w_1}, \mathrm{e}^{\mathrm{j}w_1})|\times \\ & \cos(2w_1 n + \arg(H_2^c(\mathrm{e}^{\mathrm{j}w_1}, \mathrm{e}^{\mathrm{j}w_1}))) + \\ & \frac{A_2^2}{2}|H_2^c(\mathrm{e}^{\mathrm{j}w_2}, \mathrm{e}^{\mathrm{j}w_2})|\times \\ & \cos(2w_2 n + \arg(H_2^c(\mathrm{e}^{\mathrm{j}w_2}, \mathrm{e}^{\mathrm{j}w_2}))) + \\ & A_1 A_2|H_2^c(\mathrm{e}^{\mathrm{j}w_1}, \mathrm{e}^{\mathrm{j}w_2})|\times \\ & \cos((w_1 + w_2)n + \arg(H_2^c(\mathrm{e}^{\mathrm{j}w_1}, \mathrm{e}^{\mathrm{j}w_2}))) + \\ & A_1 A_2|H_2^c(\mathrm{e}^{\mathrm{j}w_1}, \mathrm{e}^{-\mathrm{j}w_2})|\times \\ & \cos((w_1 - w_2)n + \arg(H_2^c(\mathrm{e}^{\mathrm{j}w_1}, \mathrm{e}^{-\mathrm{j}w_2}))) \end{aligned} \tag{A4}$$

where $H_2^c(\mathrm{e}^{\mathrm{j}w_1}, \mathrm{e}^{\mathrm{j}w_2})$ denotes the second-order frequency response of the system,

$$H_2^c(\mathrm{e}^{\mathrm{j}w_1}, \mathrm{e}^{\mathrm{j}w_2}) = \sum_{m_1=-\infty}^{+\infty}\sum_{m_2=-\infty}^{+\infty} h_2^c(m_1, m_2)\times \mathrm{e}^{-\mathrm{j}w_1 m_1}\mathrm{e}^{-\mathrm{j}w_2 m_2} \tag{A5}$$

In particular, suppose $A_1 = A_2 = 1/2$, $w_1 = w_2 = w$, and $u(n)$ the stimulus signal becomes single-frequency,

$$u(n) = \cos(wn), n = 0, 1, \ldots, L-1 \tag{A6}$$

The first-order response of the single-frequency signal $s_1^c(n)$ is derived as follows:

$$s_1^c(n) = |H_1^c(\mathrm{e}^{\mathrm{j}w})|\cos(wn + \arg(H_1^c(\mathrm{e}^{\mathrm{j}w}))) \tag{A7}$$

The second-order response of the single-frequency signal $s_2^c(n)$ is derived as follows:

$$s_2^c(n) = \frac{1}{2}|H_2^c(\mathrm{e}^{\mathrm{j}w}, \mathrm{e}^{\mathrm{j}w})|\times \cos(2wn + \arg(H_2^c(\mathrm{e}^{\mathrm{j}w}, \mathrm{e}^{\mathrm{j}w}))) \tag{A8}$$

## Acknowledgment

This work was supported by the National Key Basic Research and Development Program of China (No. 2017YFB1002505), Key Research and Development Program of Guangdong Province (No. 2018B030339001), and the National Natural Science Foundation of China (No. 61431007).

## References

[1] S. K. Gao, Y. J. Wang, X. R. Gao, and B. Hong, Visual and auditory brain-computer interfaces, *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 5, pp. 1436–1447, 2014.

[2] K. Lin, S. K. Gao, and X. R. Gao, Boosting the information transfer rate of an SSVEP-BCI system using maximal-phase-locking value and minimal-distance spatial filter banks, *Tsinghua Science and Technology*, vol. 24, no. 3, pp. 262–270, 2019.

[3] S. G. Zhang, X. Han, and X. R. Gao, Studying the effect of the pre-stimulation paradigm on steady-state visual evoked potentials with dynamic models based on the zero-pole analytical method, *Tsinghua Science and Technology*, vol. 25, no. 3, pp. 435–446, 2020.

[4] D. Regan, *Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine*. New York, NY, USA: Elsevier, 1989.

[5] A. M. Norcia, L. G. Appelbaum, J. M. Ales, B. R. Cottereau, and B. Rossion, The steady-state visual evoked potential in vision research: A review, *Journal of Vision*, vol. 15, no. 6, p. 4, 2015.

[6] O. Friman, I. Volosyak, and A. Graeser, Multiple channel detection of steady-state visual evoked potentials for brain-computer interfaces, *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 4, pp. 742–750, 2007.

[7] Y. S. Zhang, P. Xu, K. W. Cheng, and D. Z. Yao, Multivariate synchronization index for frequency recognition of SSVEP-based brain-computer interface, *Journal of Neuroscience Methods*, vol. 221, pp. 32–40, 2014.

[8] M. Nakanishi, Y. J. 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, p. e0140703, 2015.

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

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

[11] C. Yang, X. Han, Y. J. Wang, R. Saab, S. K. Gao, and X. R. Gao, A dynamic window recognition algorithm for SSVEP-based brain-computer interfaces using a spatio-temporal equalizer, *International Journal of Neural Systems*, vol. 28, no. 10, p. 1850028, 2018.

[12] X. G. Chen, Y. J. Wang, M. Nakanishi, X. R. Gao, T. P. Jung, and S. K. Gao, High-speed spelling with a noninvasive brain-computer interface, *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 44, pp. 6058–6067, 2015.

[13] M. Nakanishi, Y. J. Wang, X. G. Chen, Y. T. Wang, X. R. Gao, and T. P. Jung, Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis, *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 1, pp. 104–112, 2018.

[14] D. Regan, Some characteristics of average steady-state and transient responses evoked by modulated light, *Electroencephalography and Clinical Neurophysiology*, vol. 20, no. 3, pp. 238–248, 1966.

[15] C. O. Sakar, O. Kursun, and F. Gurgen, Ensemble canonical correlation analysis, *Applied Intelligence*, vol. 40, no. 2, pp. 291–304, 2014.

[16] M. Cheng and S. K. Gao, An EEG-based cursor control system, in *Proceedings of the First Joint BMES/EMBS Conference*, Atlanta, GA, USA, 1999, p. 669.

[17] M. Nakanishi, Y. J. Wang, Y. T. Wang, Y. Mitsukura, and T. P. Jung, A high-speed brain speller using steady-state visual evoked potentials, *International Journal of Neural Systems*, vol. 24, no. 6, p. 1450019, 2014.

[18] M. Xu, J. Han, Y. Wang, T. Jung, and D. Ming, Implementing over 100 command codes for a high-speed hybrid brain-computer interface using concurrent P300 and SSVEP features, doi:10.1109/TBME.2020.2975614.

[19] T. Cao, F. Wan, C. M. Wong, J. N. Da Cruz, and Y. Hu, Objective evaluation of fatigue by EEG spectral analysis in steady-state visual evoked potential-based brain-computer interfaces, doi: 10.1186/1475-925X-13-28.

[20] P. Yuan, X. G. Chen, Y. J. Wang, X. R. Gao, and S. K. Gao, Enhancing performances of SSVEP-based brain-computer interfaces via exploiting inter-subject information, *Journal of Neural Engineering*, vol. 12, no. 4, p. 046006, 2015.

[21] N. R. Waytowich, J. Faller, J. Garcia, J. M. Vettel, and P. Sajda, Unsupervised adaptive transfer learning for steady-state visual evoked potential brain-computer interfaces, in *Proceedings of 2016 IEEE International Conference on Systems, Man, and Cybernetics*, Budapest, Hungary, 2016, pp. 4135–4140.

[22] K. Suefuda and T. Tanaka, Reduced calibration by efficient transformation of templates for high speed hybrid coded SSVEP brain-computer interfaces, presented at the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 2017.

[23] C. M. Wong, F. Wan, B. Y. Wang, Z. Wang, W. Y. 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, *Journal of Neural Engineering*, vol. 17, no. 1, p. 016026, 2020.

[24] S. G. Zhang, X. Han, X. G. Chen, Y. J. Wang, S. K. Gao,

and X. R. Gao, A study on dynamic model of steady-state visual evoked potentials, *Journal of Neural Engineering*, vol. 15, no. 4, p. 046010, 2018.
[25] X. Han, S. G. Zhang, and X. R. Gao, A study on reducing training time of BCI system based on an SSVEP dynamic model, presented at the 2019 7th International Winter Conference on brain-computer interface (BCI), Gangwon, South Korea, 2019.
[26] J. Pan, X. R. Gao, F. Duan, Z. Yan, and S. K. Gao, Enhancing the classification accuracy of steady-state visual evoked potential-based brain-computer interfaces using phase constrained canonical correlation analysis, *Journal of Neural Engineering*, vol. 8, no. 3, p. 036027, 2011.
[27] V. J. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*. New York, NY, USA: Wiley, 2000.
[28] J. J. Bussgang, L. Ehrman, and J. W. Graham, Analysis of nonlinear systems with multiple inputs, *Proceeding of the IEEE*, vol. 62, no. 8, pp. 1088–1119, 1974.
[29] R. Kus, A. Duszyk, P. Milanowski, M. Labecki, M. Bierzynska, Z. Radzikowska, M. Michalska, J. Zygierewicz, P. Suffczynski, 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.
[30] D. H. Brainard, The psychophysics toolbox, *Spatial Vision*, vol. 10, no. 4, pp. 433–436, 1997.
[31] G. G. Celesia, Steady-state and transient visual evoked-potentials in clinical-practice, *Annals of the New York Academy of Sciences*, vol. 388, pp. 290–305, 1982.
[32] O. Sharon and Y. Nir, Attenuated fast steady-state visual evoked potentials during human sleep, *Cerebral Cortex*, vol. 28, no. 4, pp. 1297–1311, 2018.
[33] G. R. Mueller-Putz, R. Scherer, C. Brauneis, and G. Pfurtscheller, Steady-State Visual Evoked Potential (SSVEP)-based communication: Impact of harmonic frequency components, *Journal of Neural Engineering*, vol. 2, no. 4, pp. 123–130, 2015.
[34] R. D. Nowak and B. D. Vanveen, Random and pseudorandom inputs for Volterra filter identification, *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 2124–2135, 1994.
[35] G. Birpoutsoukis and J. Schoukens, Nonparametric Volterra kernel estimation using regularization, in *Proceedings of 2015 IEEE International Instrumentation and Measurement Technology Conference*, Pisa, Italy, 2015, pp. 222–227.
[36] Y. J. Wang, X. G. Chen, X. R. Gao, and S. K. Gao, A benchmark dataset for SSVEP-based brain-computer interfaces, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 10, pp. 1746–1752, 2017.
[37] F. D. Russo, S. Pitzalis, T. Aprile, G. Spitoni, F. Patria, A. Stella, D. Spinelli, and S. A. Hillyard, Spatiotemporal analysis of the cortical sources of the steady-state visual evoked potential, *Human Brain Mapping*, vol. 28, no. 4, pp. 323–334, 2007.
[38] J. J. Chen, D. Zhang, A. K. Engel, Q. Gong, and A. Maye, Application of a single-flicker online SSVEP BCI for spatial navigation, *Plos One*, vol. 12, no. 5, p. e0178385, 2017.
[39] J. J. Chen, A. Maye, A. K. Engel, Y. J. Wang, X. R. Gao, and D. Zhang, Simultaneous decoding of eccentricity and direction information for a single-flicker SSVEP BCI, *Electronics*, vol. 8, no. 12, p. 1554, 2019.

**Jiajun Lin** received the BS degree from Beijing University of Posts and Telecommunications in 2017. He is currently a master student at the Department of Biomedical Engineering, Tsinghua University. His research interests are biomedical signal processing, brain-computer interface, and machine learning.

**Liyan Liang** received the BE degree from Beijing University of Technology in 2015. He is currently working toward the PhD degree in Tsinghua University. His research interests include brain-computer interface, biomedical signal processing, and machine learning.

**Xu Han** received the BS degree from Huazhong University of Science and Technology in 2016. He is now a master student in Tsinghua University. His research interests focus on biomedical signal processing, brain-computer interface, and machine learning.

**Chen Yang** received the MS degree in biomedical engineering from Beijing University of Posts and Telecommunications in 2012, and the PhD degree in biomedical engineering from Tsinghua University in 2019. Since 2019, he has been doing postdoctoral research in Beijing University of Posts and Telecommunications. His research interests include brain-computer interface and signal processing.

**Xiaogang Chen** received the BEng degree in biomedical engineering from Xianning College, Xianning, China in 2008, the MEng degree in biomedical engineering from Hebei University of Technology, Tianjin, China in 2011, and the PhD degree in biomedical engineering from Tsinghua University, Beijing, China in 2015. He is currently working as an associate research fellow in Institute of Biomedical Engineering, Chinese Academy of Medical Sciences. His research interests include brain-computer interface and biomedical signal processing.

**Xiaorong Gao** received the BS degree from Zhejiang University in 1986, the MS degree from Peking Union Medical College in 1989, and the PhD degree from Tsinghua University in 1992. He is currently a professor at the Department of Biomedical Engineering, Tsinghua University. His current research interests include biomedical signal processing and medical instrumentation, especially the study of brain-computer interface.