

## A Dynamic Window Recognition Algorithm for SSVEP-Based Brain–Computer Interfaces Using a Spatio-Temporal Equalizer

Chen Yang and Xu Han

*Department of Biomedical Engineering, Tsinghua University  
Beijing, P. R. China*

Yijun Wang

*Institute of Semiconductors, Chinese Academy of Sciences  
Beijing, P. R. China*

Rami Saab, Shangkai Gao and Xiaorong Gao\*

*Department of Biomedical Engineering, Tsinghua University  
Beijing, P. R. China*  
\*gxr-dea@tsinghua.edu.cn

Accepted 11 June 2018

Published Online 13 August 2018

The past decade has witnessed rapid development in the field of brain–computer interfaces (BCIs). While the performance is no longer the biggest bottleneck in the BCI application, the tedious training process and the poor ease-of-use have become the most significant challenges. In this study, a spatio-temporal equalization dynamic window (STE-DW) recognition algorithm is proposed for steady-state visual evoked potential (SSVEP)-based BCIs. The algorithm can adaptively control the stimulus time while maintaining the recognition accuracy, which significantly improves the information transfer rate (ITR) and enhances the adaptability of the system to different subjects. Specifically, a spatio-temporal equalization algorithm is used to reduce the adverse effects of spatial and temporal correlation of background noise. Based on the theory of multiple hypotheses testing, a stimulus termination criterion is used to adaptively control the dynamic window. The offline analysis which used a benchmark dataset and an offline dataset collected from 16 subjects demonstrated that the STE-DW algorithm is superior to the filter bank canonical correlation analysis (FBCCA), canonical variates with autoregressive spectral analysis (CVARS), canonical correlation analysis (CCA) and CCA reducing variation (CCA-RV) algorithms in terms of accuracy and ITR. The results show that in the benchmark dataset, the STE-DW algorithm achieved an average ITR of 134 bits/min, which exceeds the FBCCA, CVARS, CCA and CCA-RV. In off-line experiments, the STE-DW algorithm also achieved an average ITR of 116 bits/min. In addition, the online experiment also showed that the STE-DW algorithm can effectively expand the number of applicable users of the SSVEP-based BCI system. We suggest that the STE-DW algorithm can be used as a reliable identification algorithm for training-free SSVEP-based BCIs, because of the good balance between ease of use, recognition accuracy, ITR and user applicability.

**Keywords:** Brain–computer interface; steady-state visual evoked potentials; spatio-temporal equalization; dynamic window.

### 1. Introduction

In recent years, neural signal processing has become an important area of research in the field of signal processing.<sup>1–5</sup> Brain–computer interfaces (BCIs), an application of neural signal processing, has received

widespread attention in neural engineering, neuroscience, and clinical rehabilitation communities.<sup>6–14</sup> Based on different signals produced by neurons in the brain, a variety of online BCI systems have been proposed and applied to various purposes and

\*Corresponding author.

target users.<sup>15–19</sup> Steady-state visual evoked potential (SSVEP)-based BCIs have become popular and one of the most promising types of BCIs because they are noninvasive, easy to train, relatively low-cost, and possess high information transfer rates (ITR). SSVEP is an inherent response in the human brain; when the eyes are focusing on a flashing stimulus at a fixed frequency, the brain's occipital region will produce a corresponding SSVEP with stable frequency characteristics, which can be used as the primary feature for target recognition.<sup>20</sup> In a typical SSVEP-based BCI system, each target flashes at a different frequency. When the subject gazes at one of the targets, the processing algorithm can identify the target that the subject is gazing on by analyzing the frequency components in the electroencephalogram (EEG) signal. Thus, SSVEP-based BCIs can be considered a special communication system using the human visual pathway as the channel.<sup>21</sup>

At present, research on SSVEP-based BCIs has predominantly focused on two aspects: stimulation paradigm design and recognition algorithms. The former is primarily interested in how to design a SSVEP-based BCIs stimulation system to expand the number of candidate targets and enhance the separability between different target-evoked EEG signals.<sup>22–24</sup> Alternately, recognition algorithm research is focused on techniques to improve the system's ITR.

In the study of traditional SSVEP recognition algorithms, spatial filtering is widely used as an important pre-processing method. Most traditional SSVEP recognition algorithms are based on spatial filtering approaches such as canonical correlation analysis (CCA)<sup>25</sup> and the multivariate synchronization index (MSI).<sup>26</sup> The hypothesis of these algorithms is that the background noise in the time domain is independent and identically distributed, and the power spectral density in the frequency domain is uniformly distributed. These algorithms extract features using a filter in the spatial domain, and neglect the temporal correlation of the background noise. However, it has been shown that the background noise present in EEG signals is not white noise, but rather a colored noise whose power spectrum exhibits a  $1/f$  distribution. The noise energy decreases with increasing frequency, and there is a high correlation between the noise samples in the time domain.<sup>27–29</sup> For this reason, there exists a

gap between the assumptions underpinning these algorithms and the reality of the EEG signal. In contrast, the minimum energy combination (MEC) algorithm simulates the projected background noise using the Auto Regression (AR) model which is estimated by real-time data and corrects for the statistical results of the different frequencies.<sup>30</sup> Using this approach, the adverse effect of background noise on the classification results can be reduced. The canonical variates with autoregressive spectral analysis (CVARS)<sup>31</sup> algorithm based on the MEC algorithm use a different decision function and statistics and is able to reduce the recognition error rates further. In addition, by introducing the deep networks to adaptively learn the nonlinear features of the SSVEP, deep canonical correlation analysis (DCCA) improves the signal-to-noise ratio (SNR) of the signal and hence reduces the impact of background noise.<sup>32</sup>

The equalization technology employed in communication theory can be used to solve the problem of colored noise in EEG signal processing. In wireless communication technology, equalization techniques are used to suppress frequency-specific interference. The propagation process of the wireless signal from the transmitter to the receiver may produce inter-symbol interference,<sup>33</sup> due to phase differences between the electromagnetic waves transmitting through different paths. This is expressed as colored frequency selectivity in the frequency domain. The equalizer is used to correct for the amplitude-frequency characteristics of the received signal, and suppress inter-symbol interference, which can achieve close to no-distortion transmission in all the channels. Feifei Qi *et al.* proposed the algorithm of regularized spatio-temporal filtering and classification (RSTFC) for motor imagery classification.<sup>34</sup> This algorithm utilized a spatio-temporal filter similar to the equalization techniques to filter the motor imagery EEG data. This method finds the optimal effective feature in the training template and effectively improves the recognition accuracy.

Traditionally, SSVEP recognition in BCI applications have been based on algorithms using a “fixed window length” in time domain, that is, the recognition time for each trial is constant such as CCA,<sup>25</sup> MSI,<sup>26</sup> MEC,<sup>30</sup> filter bank canonical correlation analysis (FBCCA),<sup>35</sup> etc. These algorithms generally estimate the optimal window length in

the offline dataset based on the BCI system performance and then apply that fixed window length to the online experiment. Due to the nonstationary nature of EEG background noise, the subject's attention and other factors, a unified window length may not be optimal across different subjects and trials. To solve this problem, some studies have proposed "dynamic window length" algorithms, by optimizing the stimulation time for each trial to maximize the ITR. The canonical correlation analysis reducing variation (CCA-RV) algorithm<sup>36</sup> based on the CCA algorithm can effectively reduce the system variance by checking whether a statistic "Score" exceeds the threshold to achieve dynamic optimization of time-window length. In another study, a kernel density estimation was used to estimate the subject-specific probability density function with maximum likelihood for target and nontarget feature values, and the dynamic window was achieved via hypothesis testing.<sup>37</sup> Both aforementioned algorithms include a complex learning and training process, which requires a certain period of time in order to collect a sufficient amount of calibration data. In another study, a system was designed to increase or decrease the stimulus window length during the experiment by dynamically optimizing the stimulus window based on the recognition results.<sup>38</sup> When an erroneous decision was made, the system would extend the stimulation window length of the follow-up trials, otherwise, it maintained or shortened the stimulation window. The system did not require training data but had a constraint on the experimental paradigm. Therefore, if we can find a "dynamic window length" recognition algorithm which does not require training and with fewer paradigm constraints, it may improve the practicality of the SSVEP-based BCI.

Based on the theory of adaptive equalization and hypothesis testing, this study designed a spatio-temporal equalization dynamic window (STE-DW) recognition algorithm which has low computational complexity and does not require training data. The algorithm uses the adaptive spatio-temporal equalizer to equalize the signal from both the spatial and temporal domain to reduce the adverse effects of colored noise. At the same time, according to the statistical test results, the algorithm can dynamically determine the stimulus window length in real time with the recognition process.

The structure of this paper is as follows. Section 2 presents the algorithm, in which Sec. 2.1 introduces the basic theory, Sec. 2.2 provides the algorithm implementation. Section 3 introduces experiment design and analysis method. Section 4 evaluates the real performance of the algorithm in a benchmark dataset<sup>39</sup> and the dataset collected in this study. Section 5 discusses the meaning of spatio-temporal equalization, parameter stability, practicality of the algorithm and so on.

## 2. Methods

In this paper, we first establish the SSVEP-EEG data model under colored background noise conditions using the communication channel model, which can take into account the spatial correlation and temporal correlation of background noise and is more consistent with real EEG data. Second, based on this model, a hypothesis testing method based on risk Bayesian estimation is used to establish the "dynamic window hypothesis". This hypothesis testing method can adaptively select either "continue to collect" or "give classification" according to different risk costs. Again, in order to build the model and apply the hypothesis test, we further introduced the estimation of all the parameters. Finally, for the purpose of system implementation, this section describes the implementation process of the entire algorithm in detail.

### 2.1. Theoretical model

#### 2.1.1. SSVEP-EEG data model

Let  $\mathbf{X} \in \mathbb{R}^{L \times N}$  denote  $L$  channels,  $N$  points EEG data,  $\mathbf{S}$  denote the SSVEP component and  $\mathbf{W}$  denote the background noise. The data model for the SSVEP-EEG is shown as

$$\mathbf{X} = \mathbf{S} + \mathbf{W}, \quad (1)$$

where  $\mathbf{W} \in \mathbb{R}^{L \times N}$  denotes  $L$ -channels,  $N$ -points colored background noise, including spontaneous EEG and system noise, etc. In this study, the background noise is treated as an integrated whole and the differences between components of the noise are ignored. It is generally considered that the SSVEP component  $\mathbf{S}$  is a linear mixture of multi-frequency sinusoidal signals, and is independent of background noise.<sup>20</sup>  $\mathbf{S}$  can be expressed as

$$\mathbf{S} = \mathbf{A}\Phi, \quad (2)$$

where  $\mathbf{A} \in \mathbb{R}^{L \times l}$  denotes the aliasing matrix of the sine template signal.  $\Phi \in \mathbb{R}^{l \times N}$  denotes  $l$ -dimensional,  $N$ -points sinusoidal template signal, which is the semi-orthogonal matrix. The  $i$ th row  $\dot{\phi}_i$  represents the  $i$ th complex frequency component in the template signal. It is generally assumed that the SSVEP component consists of the fundamental frequency of stimulus and multiple harmonics of stimulus. As such, if the SSVEP component includes a total of  $J$  harmonics, then  $l = 2J$ . The complex sine template  $\Phi$  can be defined as:

$$\Phi = \begin{bmatrix} \dot{\phi}_1^T & \dot{\phi}_2^T & \cdots & \dot{\phi}_l^T \end{bmatrix}^T, \\ \begin{cases} \dot{\phi}_{2k-1}(n) = \frac{1}{\sqrt{N}} e^{-jk\omega n} & k = 1, 2, \dots, \frac{l}{2} \\ \dot{\phi}_{2k}(n) = \frac{1}{\sqrt{N}} e^{jk\omega n} & n = 1, 2, \dots, N \end{cases}. \quad (3)$$

In (3),  $\omega$  denotes the angular frequency of stimulus, the coefficient  $1/\sqrt{N}$  ensured that each row in  $\Phi$  is close to orthogonal. So, the power spectrum of the EEG signal can be expressed as

$$P_{\mathbf{X}}(f) = P_{\mathbf{S}}(f) + P_{\mathbf{W}}(f), \quad (4)$$

where  $P_{\mathbf{X}}(f)$  represents the power spectrum of the EEG signal,  $P_{\mathbf{S}}(f)$  represents the power spectrum of the SSVEP and  $P_{\mathbf{W}}(f)$  represents the power spectrum of the background noise, as shown in Fig. 1(a).

**Assumption 1.** Variability between trials exists in the matrix  $\mathbf{A}$ .

In the EEG data, the amplitude of the SSVEP component between different trials usually varies,<sup>40</sup> and it can be modeled as the aliasing matrix  $\mathbf{A}$  changing over time. For training-free SSVEP-based BCIs, it is assumed that the aliasing matrix  $\mathbf{A}$  remains stable within a single trial but varies randomly between trials. As such, the maximum likelihood estimate of  $\mathbf{A}$  must be calculated based on the EEG data of each trial when solving the model.

**Assumption 2.** Background noise is a random signal, and there is a correlation in both the spatial and temporal domains.

In different trials, the background noise is a random signal.<sup>41</sup> Due to the volume conductor effect, the EEG signal represents a linear mixing of neuronal activities conducting from the cortex to the scalp.

According to the central limit theorem, the distribution of background EEG  $\mathbf{W}$  will be close to the Gaussian distribution<sup>42</sup> and will produce significant spatial and temporal correlations. That is, each sample point obeys a Gaussian distribution, and there are correlations between:

- (1) sample points between different channels at the same time,
- (2) adjacent sample points of the same channel
- (3) adjacent sample points of different channels.

**Assumption 3.**  $\mathbf{W}$  has both stationary and non-stationary features.

Background EEG signals are generated by the cumulative summation of millions of synaptic junctions and distributed networks of collaborating and synchronous neurons.<sup>43</sup> The background noise signals generated by these sources are conducted from the cortex to the scalp through the skull, where the aliasing process can be regarded as nonstationary and the conductive process as stationary. Thus, the background noise in the EEG signal contains both stationary<sup>44–46</sup> and nonstationary features.<sup>6</sup>

In this study, a channel system based on FIR-MIMO (Finite Impulse Response-Multiple Input Multiple Output) is proposed to simulate the background EEG generation and conduction process according to the background EEG characteristics. The process is described as follows:

- (i) First,  $L$  independent identically distributed Gaussian sources produce white noise  $\varepsilon_1, \dots, \varepsilon_L$ .
- (ii) Then,  $\varepsilon_1, \dots, \varepsilon_L$  pass through the nonstationary zero-order channel  $\mathbf{B}$ , and form nonstationary noise with spatial correlation.
- (iii) Finally, the noise passes through the stationary high-order channel  $\mathcal{L}(z)$  again and forms noise that has both spatial and temporal correlation.

Here,  $\mathbf{B}$  is nonstationary and is used to simulate the source signal aliasing process;  $\mathcal{L}(z)$  is stationary within a period of time and simulates the conductive process of the background EEG transmitting from the cortex to the scalp. Specifically,  $\mathbf{B}$  is defined as a  $L \times L$  full rank square matrix, and  $\mathcal{L}(z)$  is a equalizable linear system, which satisfies  $\text{rank}(\mathcal{L}(z)) = L$  for all nonzero  $z \in \mathbf{C}$ .<sup>47</sup> Thus, the entire transmission

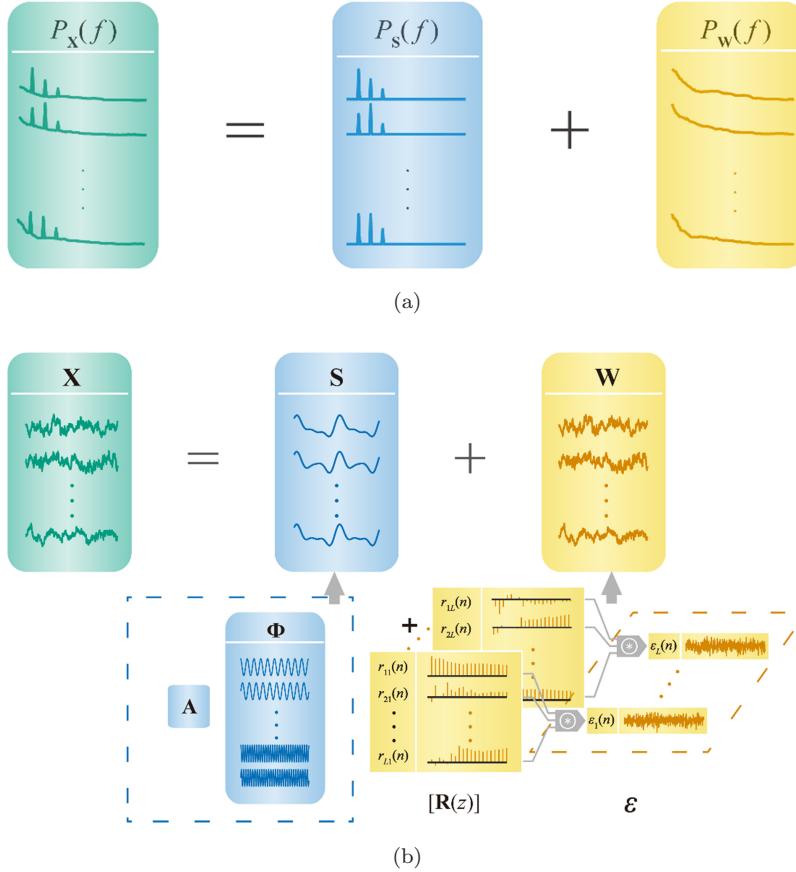


Fig. 1. (a) Schematic diagram of power spectrum components of SSVEP-EEG. Background noise **W** and SSVEP response **S** are independent of each other, and the EEG signal power spectrum of the EEG signal is equal to their addition. (b) SSVEP-EEG component model was defined in (1). According to Assumption 3, the background noise **W** is treated as the output of a FIR-MIMO channel  $\mathcal{R}(z)$  driven by a multidimensional independent Gaussian white noise  $\epsilon$ , so it can be expressed by (8). In this sub-figure, the  $L$ -dimensional independent identically distributed Gaussian white noise  $\epsilon_1, \dots, \epsilon_L$  is aliased into Gaussian colored noise **W** through sub-channels  $r_{11}(n), \dots, r_{LL}(n)$ . The aliasing process can be described as a linear convolution process described by “ $*$ ”, as in (7), where the  $r_{ij}(n)$  represents the impulse response of the  $j$ th input and the  $i$ th output.

process can be described by the high-order FIR-MIMO channel  $\mathcal{R}(z)$ ,

$$\mathcal{R}(z) = \mathcal{L}(z)\mathbf{B}. \quad (5)$$

$\gamma_{ij}(z)$  represents the transfer channel of the  $j$ th noise source to the  $i$ th receiver in  $\mathcal{R}(z)$ ,

$$\mathcal{R}(z) = \begin{bmatrix} \gamma_{11}(z) & \cdots & \gamma_{1L}(z) \\ \vdots & \ddots & \vdots \\ \gamma_{L1}(z) & \cdots & \gamma_{LL}(z) \end{bmatrix}, \quad (6)$$

and  $r_{ij}(n)$  represents the impulse response of the channel  $\gamma_{ij}(z)$ . It can be expressed as  $\gamma_{ij}(z) = \sum_{n=0}^K r_{ij}(n)z^{-n}$ , where  $K$  represents the highest order of  $\mathcal{R}(z)$ . So, the noise recorded by the  $i$ th

channel of EEG receiver can be expressed as

$$\begin{aligned} w_i(n) &= \sum_{j=1}^L r_{ij}(n) * \epsilon_j(n) \\ &= \sum_{j=1}^L \sum_{\tau=0}^T r_{ij}(\tau) \epsilon_j(n-\tau), \end{aligned} \quad (7)$$

where  $*$  denotes a linear convolution, and  $T$  denotes the number of data samples. The signal transmission process can be expressed as

$$\mathbf{W} = [\mathcal{R}(z)]\mathcal{E} = [\mathcal{L}(z)]\mathbf{B}\mathcal{E}, \quad (8)$$

where  $\mathcal{E}$  is the Gaussian white noise matrix,  $\mathcal{E} = [\epsilon_1^T, \dots, \epsilon_L^T]^T$ ,  $[\mathcal{R}(z)]$  represents that the multidimensional signal passing through the linear system

described by  $\mathcal{R}(z)$ , the same with  $[\mathcal{L}(z)]$ . Thus, the background noise  $\mathbf{W}$  will satisfy the features described by Remarks 2 and 3. In summary, the data model is shown in Fig. 1(b).

### 2.1.2. Dynamic window hypothesis

According to (1) and (2), for each SSVEP-BCI system with multiple targets, detection of each target can be abstracted as a hypothesis as

$$\mathcal{H}^{\{q\}} : \mathbf{X} = \mathbf{S}^{\{q\}} + \mathbf{W}, \quad (9)$$

where  $\mathcal{H}^{\{q\}}$  represents that the EEG data,  $\mathbf{X}$  contains the SSVEP component  $\mathbf{S}^{\{q\}}$  which is evoked by the  $q$ th stimulus frequency.

Suppose that the SSVEP-BCI system contains  $Q$  targets (the value range of  $q$  can be defined as  $1, \dots, Q$ ), and the EEG data are continuously received, an additional “erasure decision”<sup>48</sup>  $\mathcal{H}^{\{0\}}$  can be added to hypothesize. When  $\mathcal{H}^{\{0\}}$  is decided, it indicates that the current data are not enough to make a reasonable decision, so the system needs to wait for more new data.<sup>49</sup> It can be expressed as

$$\begin{cases} \mathcal{H}^{\{0\}} : \text{reject} \\ \mathcal{H}^{\{1\}} : \mathbf{X} = \mathbf{S}^{\{1\}} + \mathbf{W} \\ \vdots \\ \mathcal{H}^{\{Q\}} : \mathbf{X} = \mathbf{S}^{\{Q\}} + \mathbf{W}. \end{cases} \quad (10)$$

Under the condition of  $\mathcal{H}^{\{1\}}, \dots, \mathcal{H}^{\{Q\}}$ , the probability of each target is  $1/Q$ .  $\mathcal{G}^{\{q\}}$  represents the case where the real stimulus is the  $q$ th stimulus frequency, and the probability  $p(\mathcal{G}^{\{q\}}) = 1/Q$ .  $c(\mathcal{H}^{\{q_{\mathcal{H}}\}}, \mathcal{G}^{\{q\}})$  indicates the cost of the situation where the real case is  $\mathcal{G}^{\{q\}}$ , but the system makes the decision of  $\mathcal{H}^{\{q_{\mathcal{H}}\}}$ .

$$\begin{cases} c(\mathcal{H}^{\{q_{\mathcal{H}}\}}, \mathcal{G}^{\{q\}}) = 0 & q = q_{\mathcal{H}}, q_{\mathcal{H}} \neq 0 \\ c(\mathcal{H}^{\{q_{\mathcal{H}}\}}, \mathcal{G}^{\{q\}}) = 1 & q \neq q_{\mathcal{H}}, q_{\mathcal{H}} \neq 0 \\ c(\mathcal{H}^{\{q_{\mathcal{H}}\}}, \mathcal{G}^{\{q\}}) = \epsilon & q_{\mathcal{H}} = 0. \end{cases}$$

The statistics  $\gamma$  can be defined as (11), (see Appendix A for derivation)

$$\gamma = \min_{q_{\mathcal{H}}} \left[ 1 - \frac{p(\mathbf{X} | \mathcal{G}^{\{q_{\mathcal{H}}\}})}{\sum_{q=1}^Q p(\mathbf{X} | \mathcal{G}^{\{q\}})} \right], \quad (11)$$

where  $p(\mathbf{X} | \mathcal{G}^{\{q\}})$  represents the conditional probability of  $\mathbf{X}$  in the case of  $\mathcal{G}^{\{q\}}$ . Usually, as the stimulation time increases, the statistic  $\gamma$  will gradually decrease.

According to Appendix A, when  $\gamma \geq \epsilon$ , the cost of “erasure decisions”  $\mathcal{H}^{\{0\}}$  is considered as the smallest of all hypotheses, which indicates that the system should continue to receive the data. When  $\gamma < \epsilon$ , it suggests that at least one cost of the hypothesizes  $\mathcal{H}^{\{1\}}, \dots, \mathcal{H}^{\{Q\}}$  is lower than the cost of the  $\mathcal{H}^{\{0\}}$ , so the system can stop data acquisition. The system can then proceed to give the recognition result by the specific decision criteria which is shown below.

It should be noted that in this study, the dynamic window hypothesis is only used for dynamic window length selection to determine if the termination acquisition condition is met. Specific decision criteria are presented in Sec. 2.2.

### 2.1.3. Model parameter estimation

From (11), the calculation of  $\gamma$  needs to estimate the probability density of  $\mathbf{X}$  under various conditions. The traditional EEG probability model, where each of the channels is considered as one-dimensional, can only be used to describe the background noise as white noise. In this study, each sample in each channel is regarded as one-dimensional, so the dimension of the model is much higher than the traditional model. This can be used to characterize the nonwhite properties of background noise.

According to (9), the probability density function under the hypothesis  $\mathcal{H}^{\{q\}}$  can be expressed as<sup>50</sup>

$$p(\mathbf{X} | \mathcal{H}^{\{q\}}) = \frac{1}{(2\pi)^{\frac{LN}{2}} \det(\Sigma_{\text{vec}(\mathbf{W})})^{1/2}} \times e^{-\frac{1}{2} \bar{\theta}^{\{q\}} \bar{\theta}^{\{q\}}}, \quad (12)$$

$$\begin{aligned} \bar{\theta}^{\{q\}} &= \Sigma_{\text{vec}(\mathbf{W})}^{-1/2} \text{vec}(\mathbf{X}) - \Sigma_{\text{vec}(\mathbf{W})}^{-1/2} \\ &\quad \times (\Phi^{\{q\}}^H \otimes \mathbf{I}_L) \text{vec}(\mathbf{A}^{\{q\}}). \end{aligned} \quad (13)$$

In (12) and (13),  $\text{vec}(\cdot)$  denotes a vectorization,  $\otimes$  denotes a kronecker product, and  $\Sigma_{\text{vec}(\mathbf{W})} \in \mathbb{R}^{LN \times LN}$  denotes the covariance matrix of vectored noise  $\text{vec}(\mathbf{W})$ , which reflects the spatio-temporal correlation of background noise  $\mathbf{W}$ . Obviously, if (12) is substituted into (11),  $\det(\Sigma_{\text{vec}(\mathbf{W})})$  in both the numerator and denominator in (11) can be eliminated, so  $\det(\Sigma_{\text{vec}(\mathbf{W})})$  does not need to be explicitly calculated. So, only the variables  $\Sigma_{\text{vec}(\mathbf{W})}^{-1/2}$  and  $\mathbf{A}^{\{q\}}$  need to be estimated in (12) and (13).

Note that  $\Sigma_{\text{vec}(\mathbf{W})}^{-1/2}$  can be regarded as an equalizer for  $\text{vec}(\mathbf{W})$ , and any  $\Psi \in \mathbb{R}^{LN \times LN}$

satisfying (14) can be used as an estimate of  $\Sigma_{\text{vec}(\mathbf{W})}^{-1/2}$

$$\mathbb{E}([\Psi \text{vec}(\mathbf{W})][\Psi \text{vec}(\mathbf{W})]^H) = \mathbf{I}_{LN}, \quad (14)$$

where  $\mathbb{E}(\cdot)$  represents mathematical expectation.

According to the model assumptions, the background noise transmission process can be described using FIR-MIMO channel model  $\mathcal{R}(z)$ . Let the channel system  $\mathcal{R}(z)$  of  $L \times L$  dimension satisfying  $\deg(\mathcal{R}(z)) = K$  be equalizable, therefore there is always an equalizer  $\mathcal{D}(z)$  of  $L \times L$  dimension satisfying  $\deg(\mathcal{D}(z)) = \rho$  and  $\leq 2LK - 1$ , which enables  $\mathcal{D}(z)\mathcal{R}(z) = z^{-K}\mathbf{I}_L$ . The background noise  $\mathbf{W}$  passing through the spatio-temporal equalizer  $\mathcal{D}(z)$  will become the equalization noise  $\mathcal{E}'$  without the spatial and temporal correlation. The background noise conduction and equalization model is shown in Fig. 2.

According to (5), the matrix product of  $\mathcal{R}(z)$  is concatenated by the nonstationary channel  $\mathbf{B}$  and stationary channel  $\mathcal{L}(z)$ . Thus, the nonstationary equalizer  $\mathbf{C}$  and the stationary equalizer  $\mathcal{M}(z)$  are used to equalize the nonstationary channel  $\mathbf{B}$  and the stationary channel  $\mathcal{L}(z)$ , respectively.  $\mathcal{D}(z)$  can be expressed as

$$\mathcal{D}(z) = \mathbf{C}\mathcal{M}(z), \quad (15)$$

where  $\mathcal{M}(z)$  is  $\rho$ -order stationary FIR spatio-temporal equalizer and  $\mathbf{C}$  is zero-order nonstationary

spatial equalizer. And  $\mathcal{M}(z)$  and  $\mathbf{C}$  satisfy that

$$\mathcal{M}(z)\mathcal{L}(z)\mathcal{L}(1/z^*)^H\mathcal{M}(1/z^*)^H = \mathbf{I}_L, \quad (16)$$

$$\mathbf{C}\mathbf{B}\mathbf{B}^H\mathbf{C}^H = \mathbf{I}_L. \quad (17)$$

That is, the product of  $\mathcal{M}(z)$  and  $\mathcal{L}(z)$  is a paraunitary matrix, and the product of  $\mathbf{C}$  and  $\mathbf{B}$  is a unitary matrix.

Let  $\mathcal{F}_N(\mathcal{D})$  represent the transfer matrix of  $\mathcal{D}(z)$ , and  $\mathcal{D}_k$  represents the  $k$ th order transfer coefficient of  $\mathcal{D}(z)$ . That is,

$$\mathcal{F}_N(\mathcal{D})\text{vec}(\mathbf{W}) \sim N(\mathbf{0}, \mathbf{I}_{LN})$$

$$\mathcal{F}_N(\mathcal{D})$$

$$= \begin{bmatrix} \mathcal{D}_0 & & & \\ \vdots & \mathcal{D}_0 & & \\ \mathcal{D}_\rho & \vdots & \ddots & \\ & \mathcal{D}_\rho & & \mathcal{D}_0 \\ & & \ddots & \vdots & \ddots \\ & & & \mathcal{D}_\rho & \cdots & \mathcal{D}_0 \end{bmatrix} \left. \right\} N \text{ blocks}, \quad (18)$$

$$\mathcal{D}(z) = \sum_{k=0}^{\rho} \mathcal{D}_k z^{-k}. \quad (19)$$

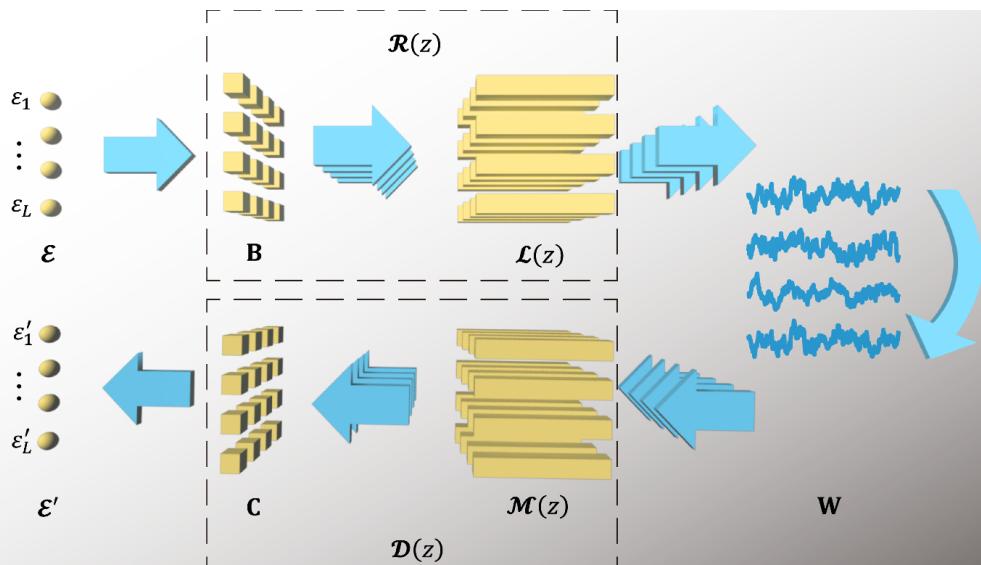


Fig. 2. The background noise conduction and equalization model. Where  $\mathbf{B}$  and  $\mathcal{L}(z)$  form the noise conduction channels  $\mathcal{R}(z)$ ,  $\mathbf{C}$  and  $\mathcal{M}(z)$  form the equalization channel  $\mathcal{D}(z)$  and the arrows represent the conduction direction. The noise source  $\mathcal{E}$  produced independent and identically distributed Gaussian noise, formed the spatial mixing noise after  $\mathbf{B}$ , and resulted in the spatial and temporal correlated background noise  $\mathbf{W}$  after  $\mathcal{L}(z)$ . At last, the background noise  $\mathbf{W}$  passed through the stationary equalizer  $\mathcal{M}(z)$  and the nonstationary equalizer  $\mathbf{C}$  and formed the equalized noise  $\mathcal{E}'$ .

The background noise can be equalized by  $\mathcal{F}_N(\mathcal{D})$ , so  $\mathcal{F}_N(\mathcal{D})$  can be regarded as an estimate of  $\Sigma_{\text{vec}(\mathbf{W})}^{-1/2}$ . That is

$$\Sigma_{\text{vec}(\mathbf{W})}^{-1/2} = \mathcal{F}_N(\mathcal{D}). \quad (20)$$

Thus, the maximum likelihood estimate of the matrix  $\mathbf{A}^{\{q\}}$  can be obtained for each possible assumption of  $\mathcal{H}^{\{q\}}$ .

$$\hat{\mathbf{A}}^{\{q\}} = \underset{\mathbf{A}^{\{q\}}}{\operatorname{argmax}} p(\mathbf{X} | \mathbf{A}^{\{q\}}; \mathcal{H}^{\{q\}}, \hat{\Sigma}_{\bar{w}}^{-1/2}). \quad (21)$$

As can be seen from (21),

$$\begin{aligned} \text{vec}(\hat{\mathbf{A}}^{\{q\}}) &= [(\Phi^{\{q\}}^H \otimes \mathbf{I}_L)^H \mathcal{F}_N(\mathcal{D})^H \mathcal{F}_N(\mathcal{D}) \\ &\quad \times (\Phi^{\{q\}}^H \otimes \mathbf{I}_L)]^{-1} (\Phi^{\{q\}}^H \otimes \mathbf{I}_L)^H \\ &\quad \times \mathcal{F}_N(\mathcal{D})^H \mathcal{F}_N(\mathcal{D}) \text{vec}(\mathbf{X}). \end{aligned} \quad (22)$$

Refer to Appendix B for the derivation process.

## 2.2. Algorithm implementation

### 2.2.1. Processing flow

Based on the theoretical analysis above, the STE-DW algorithm can be achieved by the chart presented in Fig. 3, which includes four processes: stationary equalizer estimation, nonstationary equalizer estimation, statistic decision and classify.

#### (i) Stationary equalizer estimation

The stationary equalizer  $\mathcal{M}(z)$  equilibrates the stationary channel  $\mathcal{L}(z)$ , which remains relatively stable for a certain period of time, so that multiple adjacent trials can be used for  $\mathcal{M}(z)$  estimation.

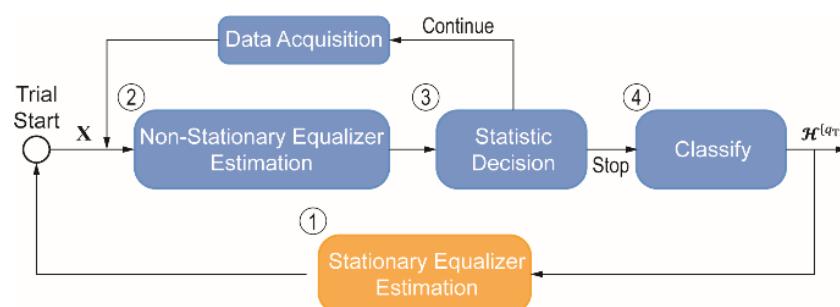


Fig. 3. (Color online) The STE-DW algorithm system flow chart. The yellow process is executed after the end of the last trial, and the blue processes are executed in real time during the trial.

In each trial, the EEG background noise estimation can be expressed as

$$\hat{\mathbf{W}} = \mathbf{X} - \hat{\mathbf{S}}, \quad (23)$$

and

$$\hat{\mathbf{S}} = \mathbf{X} \Phi^{\{q_T\}^H} \Phi^{\{q_T\}}, \quad (24)$$

where  $\Phi^{\{q_T\}}$  represents the complex sine template corresponding to the recognition result. Note that since odd-row elements and even-row elements are conjugated to each other in  $\Phi^{\{q_T\}}$ ,  $\hat{\mathbf{S}}$  is a real matrix.

In the case where the multiple adjacent trials' recognition processes have been completed, the stationary noise estimate  $\hat{\mathbf{W}}$  can be expressed as

$$\tilde{\mathbf{W}} = [\hat{\mathbf{W}}[1] \quad \hat{\mathbf{W}}[2], \dots, \hat{\mathbf{W}}[M],] \quad (25)$$

where  $\hat{\mathbf{W}}[i]$  represents the estimated noise of the  $i$ th trial. If the trials completed are less than  $M$ , all of them are used to estimate the background noise.  $\tilde{\mathbf{W}}$  reflects the stationary component of the background noise so that it can be used to estimate the stationary equalizer  $\mathcal{M}(z)$ .

$\mathcal{M}(z)$  contains  $L^2\rho$  unknown variables. Practically, it is not only difficult to calculate in real time, but also the total amount of sample data required makes it difficult to reliably estimate the associated parameters. In order to reduce the computational complexity, improve the numerical stability of the system and achieve spatio-temporal equalization, a zero-order stationary spatial equalizer  $\mathbf{P}$  and a high-order stationary time-domain equalizer  $\mathcal{T}(z)$  are used to approximate  $\mathcal{M}(z)$  as

$$\mathcal{M}(z) \approx \mathcal{T}(z)\mathbf{P}, \quad (26)$$

where  $\mathbf{P}$  is the  $L \times L$ -dimensional lower triangular real matrix such that

$$\mathbf{P} = \begin{bmatrix} p_{11} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ p_{L1} & \cdots & p_{LL} \end{bmatrix}. \quad (27)$$

$\mathcal{T}(z)$  is a polynomial matrix with only the main diagonal element, and  $\deg(\mathcal{T}(z)) = \rho$ , such that

$$\mathcal{T}(z) = \begin{bmatrix} T_1(z) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & T_L(z) \end{bmatrix}, \quad (28)$$

where the main diagonal elements are that

$$T_i(z) = 1 + \sum_{k=1}^{\rho} t_i(p) z^{-k}, \quad (29)$$

where  $t_i(p)$  represents the coefficients of the  $i$ th equalizer channel polynomial, and the specific solution method will be presented in Appendix C. The approximation of  $\mathbf{P}$  and  $\mathcal{T}(z)$  contains  $L(L+1)/2 + L\rho$  unknown variables, so the computational complexity is significantly reduced. Experiments show that the simplified equalizers can be applied well to real-time analysis.

$\mathbf{P}$  can be estimated from  $\tilde{\mathbf{W}}$  using Cholesky decomposition, which can be expressed as

$$\mathbf{P} = \text{cholesky}[\tilde{\mathbf{W}}\tilde{\mathbf{W}}^T]^{-1}. \quad (30)$$

$\tilde{\mathbf{U}}$  represents the output of the  $\tilde{\mathbf{W}}$  through  $\mathbf{P}$ , such that

$$\tilde{\mathbf{U}} = \mathbf{P}\tilde{\mathbf{W}}. \quad (31)$$

$\dot{\tilde{\mathbf{u}}}_i$  represents the  $i$ th row data in  $\tilde{\mathbf{U}}$ , which can be expressed as

$$\tilde{\mathbf{U}} = [\dot{\tilde{\mathbf{u}}}_1^T, \dots, \dot{\tilde{\mathbf{u}}}_L^T]^T.$$

Each diagonal element  $T_i(z)$  in  $\mathcal{T}(z)$  can be regarded as a one-dimensional stationary whitening filter for  $\dot{\tilde{\mathbf{u}}}_i$ . Therefore, the AR model can be established for  $\dot{\tilde{\mathbf{u}}}_i$  and the AR coefficients  $t_i(1), \dots, t_i(\rho_i)$  can be used as the filter coefficients, such that

$$\dot{\tilde{\mathbf{u}}}_i(n) = - \sum_{k=1}^{\rho_i} t_i(k) \dot{\tilde{\mathbf{u}}}_i(n-k) + e(n),$$

where  $\dot{\tilde{\mathbf{u}}}_i(n)$  represents the  $n$ th element in  $\dot{\tilde{\mathbf{u}}}_i$ , and  $e(n)$  represents the random error. The AR model coefficients can be estimated by the Burg algorithm,<sup>51</sup> and the order  $\rho_i$  is determined by the AIC

(Akaike information criterion). Within the specified range, the minimum order of the AIC criterion is chosen as the order  $\rho_i$  of  $T_i(z)$ . The computed results  $\rho_1, \dots, \rho_L$  are generally not the same, and  $\rho$  should be selected based on  $\rho = \max(\rho_i)$ . Those diagonal elements with order less than  $\rho$  will be appended with zeroes.

### (ii) Nonstationary equalizer estimation

Any matrix  $\mathbf{C}$  satisfying the constraint of (16) can be used as the nonstationary equalizer. In order to guarantee uniqueness, it is also possible to restrain  $\mathbf{C}$  as a lower triangular matrix, which is

$$\mathbf{C} = \begin{bmatrix} c_{11} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ c_{L1} & \cdots & c_{LL} \end{bmatrix}. \quad (32)$$

The nonstationary equalizer  $\mathbf{C}$  is used to equalize the nonstationary channel  $\mathbf{B}$ , which itself is related to the nonstationary characteristic of the background noise. Therefore, the nonstationary equalizer  $\mathbf{C}$  cannot depend on the priori data and must only rely on the EEG data of the current trial.

As evidenced in the experimental results in Sec. 4.2, the signal-to-noise ratio between the SSVEP component and the background noise component is extremely low: less than  $-10$  dB in most trials. Thus, it is possible to directly estimate the nonstationary equalizer  $\mathbf{C}$  using the current trial EEG data instead of the background noise data.

Let  $\mathbf{V}$  indicate the output of the current trial EEG data through the stationary equalizer, that is

$$\mathbf{V} = [\mathcal{M}(z)]\mathbf{X}. \quad (33)$$

Thus, the nonstationary equalizer  $\mathbf{C}$  can be expressed as

$$\mathbf{C} = \text{cholesky}[\mathbf{V}\mathbf{V}^T]^{-1}. \quad (34)$$

In summary,  $\mathcal{D}(z)$  can be expressed as

$$\mathcal{D}(z) = \mathbf{C}\mathcal{T}(z)\mathbf{P}. \quad (35)$$

So,

$$\mathcal{F}_{\mathbf{N}}(\mathcal{D}) = (\mathbf{I}_{\mathbf{N}} \otimes \mathbf{C})\mathcal{F}_{\mathbf{N}}(\mathcal{T})(\mathbf{I}_{\mathbf{N}} \otimes \mathbf{P}). \quad (36)$$

### (iii) Statistic decision

The conditional probability  $p(\mathbf{X} | \mathcal{H}^{\{q\}})$  under different conditions  $\mathcal{G}^{\{q_{\mathcal{H}}\}}$  can be calculated by the following steps.

First, the  $\text{vec}(\hat{\mathbf{A}}^{\{q\}})$  can be determined by substituting (18) into (22).

Then, in (13), replace  $\hat{\Sigma}_{\text{vec}(\mathbf{W})}^{-1/2}$  with  $\mathcal{F}_N(\mathcal{D})$  according to (20) and replace  $\text{vec}(\mathbf{A}^{\{q\}})$  with  $\text{vec}(\hat{\mathbf{A}}^{\{q\}})$  to obtain  $\bar{\boldsymbol{\theta}}^{\{q\}}$ .

Finally, substitute  $\bar{\boldsymbol{\theta}}^{\{q\}}$  into (12) to calculate  $p(\mathbf{X} | \mathcal{H}^{\{q\}})$ . So the statistic  $\gamma$  can also be obtained from (11). If  $\gamma$  is not less than the cost  $\epsilon$  of the “erasure decisions”  $\mathcal{H}^{\{0\}}$ , the data collection should be continued; otherwise, the classify stage should be entered.

#### (iv) Classify

Other studies have shown that the number of independent components of SSVEP is less than the number of the channels in most cases.<sup>25,26,30</sup> By optimizing the projection direction, it can effectively improve the signal-to-noise ratio of the received data and improve the decision accuracy.

Therefore, this study utilizes the MEC algorithm<sup>30</sup> and the minimum energy combination direction of the noise is selected as the projection direction of the equalized data, which can be expressed by the optimization function as

$$\begin{aligned}\lambda^{\{q\}} &= \min_{\boldsymbol{\beta}^{\{q\}}} \|\boldsymbol{\beta}^{\{q\}}^T \hat{\boldsymbol{\Theta}}^{\{q\}}\|_2 \\ &= \min_{\boldsymbol{\beta}^{\{q\}}} \boldsymbol{\beta}^{\{q\}}^T \hat{\boldsymbol{\Theta}}^{\{q\}} \hat{\boldsymbol{\Theta}}^{\{q\}}^T \boldsymbol{\beta}^{\{q\}},\end{aligned}\quad (37)$$

where

$$\begin{aligned}\hat{\boldsymbol{\Theta}}^{\{q\}} &= \text{unvec}_{L,N}(\bar{\boldsymbol{\theta}}^{\{q\}}) \\ &= \text{unvec}_{L,N}(\mathcal{F}_N(\mathcal{D}) \text{vec}(\mathbf{X} - \hat{\mathbf{A}}^{\{q\}} \boldsymbol{\Phi}^{\{q\}})),\end{aligned}\quad (38)$$

where  $\text{unvec}_{L,N}()$  indicates that the  $LN \times 1$  dimensional column vector is converted to  $L \times N$  dimensional matrix.

It is easy to note that  $\lambda^{\{q\}}$  is the smallest eigenvalue of  $\hat{\boldsymbol{\Theta}}^{\{q\}} \hat{\boldsymbol{\Theta}}^{\{q\}}^T$ .  $\boldsymbol{\beta}^{\{q\}}$  is the minimum energy projection direction of the equalized noise, which is also called the minimum noise energy diversity reception direction.

For  $q = 1, \dots, Q$ ,

$$q_T = \operatorname{argmin}_q \lambda^{\{q\}}.$$

The condition where the projected noise energy  $\lambda^{\{q\}}$  reaches the minimum is chosen as the final result of

the minimum noise energy diversity reception of the  $M$ th trial, that is, the  $\mathcal{H}^{\{q_T\}}$  assumption is accepted.

#### 2.2.2. Main steps

---

##### Steps

Initialize  $\mathbf{P} = \mathbf{I}_L$ ,  $\mathcal{T}(\mathbf{z}) = \mathbf{I}_L$ , and calculate  $\mathcal{M}(\mathbf{z})$  according to (26)

**For**  $m = 1, 2, 3, \dots$  all trials

**Do**

    Receive new data and update  $\mathbf{X}$

    Use  $\mathcal{M}(z)$  to filter  $\mathbf{X}$  and calculate  $\mathbf{V}$  according to (33)

    Calculate  $\mathbf{C}$ , according to (34)

    Calculate  $\mathcal{F}_N(\mathcal{D})$ , according to (36)

    Calculate  $\hat{\mathbf{A}}^{\{q\}}$  for each  $\mathcal{H}^{\{q\}}$ , according to (20) and (21)

    Calculate  $p(\mathbf{X} | \mathcal{H}^{\{q\}})$  for each  $\mathcal{H}^{\{q\}}$ , according to (12) and (13)

    Calculate  $\gamma$ , according to (11)

**While**  $\gamma \geq \epsilon$  **and** the stimulus time is less than a finite upper limit

        Calculate the minimum energy combination coefficient  $\lambda^{\{q\}}$  for each  $\mathcal{H}^{\{q\}}$ , according to (37)

        Choose the  $\mathcal{H}^{\{q\}}$  corresponding to the smallest  $\lambda^{\{q\}}$  as the result of the judgment. Constructs  $\tilde{\mathbf{W}}$  using the estimated noise of the last  $M$  trials, according to (25)

        Update  $\mathbf{P}$  according to (30)

        Update  $\mathcal{T}(z)$  using the Burg algorithm and AIC. See Appendix C for specific solution.

        Update  $\mathcal{M}(z)$ , according to (26)

**End for**

---

### 3. Experiment and Analysis

#### 3.1. Experimental design

##### 3.1.1. System configuration and experimental paradigm

In this experiment, a real-time online processing system was designed for SSVEP-based BCIs. The processing system consisted of a stimulator, an EEG recorder and an operator, as shown in Fig. 4(a). An ASUS VG278HE 27-inch display with 1920 \* 1080 resolution and 60 Hz refresh rate was used as the stimulator, a SynAmps2 system (Neuroscan, Inc.)

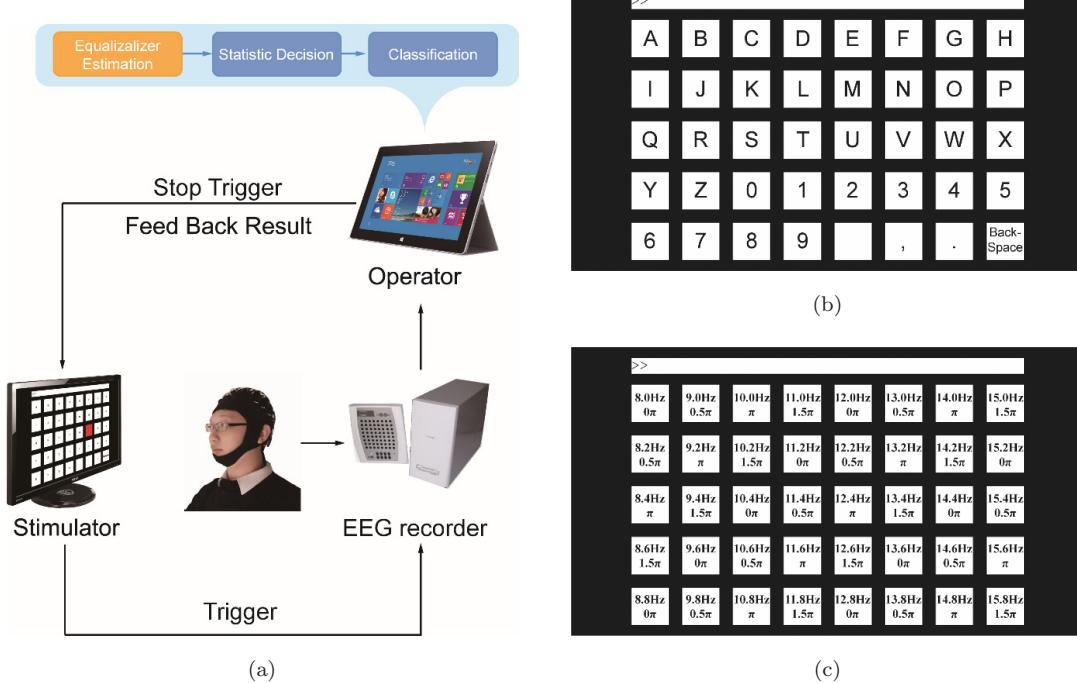


Fig. 4. (Color online) (a) Experimental system configuration, in which the red box in the display indicates the target cue. (b) Experimental stimulation interface. (c) The frequency and the initial phase of stimulus targets (the first line indicates the target frequency and the second line indicates the initial phase in each rectangle).

was used as the EEG recorder, and a Surface Pro 4 tablet PC with dual-core i5-6300U@2.4GHz CPU and DDR3 4GB memory was used as the operator. All of the programs were developed under MATLAB 2015b and the Psychophysics Toolbox Version 3<sup>52</sup> was also used in developing the stimulus program.

During the experiment, the stimulator sent a start trigger and an end trigger to the collector at the beginning and the end of each trial. The processing device sent a stop trigger to the stimulator when the stimulus stop condition was met and sent the feedback to the stimulus device after completing the recognition. In order to facilitate the comparison of algorithm performance and eliminate the interference of experimental designs, this experiment used the same experimental design as of Ref. 39. A 5×8 stimulation matrix containing 40 characters (A–Z total of 26 English letters, 0–9 total of 10 digits, 3 other symbols, and the backspace) was presented on the stimulator, as shown in Fig. 4(b). The frequency, initial phase, position and the size of each target block were consistent with the Ref. 39, as shown in Fig. 4(c). In the experiments, the sampled sinusoidal stimulation was used for each stimulation.<sup>24,53</sup>

### 3.1.2. Data acquisition

A 64 channels EEG collection cap based on 10–10 system was used for EEG data recording. Only nine channels of data (Pz, PO5, PO3, POz, PO4, PO6, O1, Oz and O2) were recorded during the experiment. The reference electrode was located at the vertex, and the impedance of each electrode was kept under 10 KΩ. The sampling rate was 1000 Hz and all data were down-sampled to 250 Hz during the pre-processing process. The raw data were preprocessed by a 1–100 Hz band-pass filter and a 50 Hz notch filter.

A total of 21 subjects participated in the experiment, including 16 subjects aged 17–30 years old (10 males and 6 females) and 5 subjects aged 59–68 years old (3 males and 2 females). One male subject aged 66 years old was unable to complete the experiment due to physical reasons, so his data were not included in the dataset. Therefore, the EEG data of 20 subjects were applied to analysis. Each subject was asked to read and sign an informed consent form approved by the Ethical Review Board of Tsinghua University before the experiment and

received a nominal compensation for their time at the end of the experiment.

S1–S16 subjects participated in an offline experiment, an STE-DW algorithm online experiment and a FBCCA algorithm online experiment, their information (age, gender) have been shown in Table 2. S17 (24 years, female), S18 (25 years, male), S19 (24 years, male), S20 (24 years, male) participated in an STE-DW algorithm online experiment and a FBCCA algorithm online experiment. Each experiment included four blocks, so S1–S16 subject was asked to participate in a total of 12 blocks, and S17–S20 subjects was asked to participate in eight blocks. In order to balance the effects between different experiments and the effects of fatigue, the experiments were performed randomly and alternately. Due to the uncertainty intervals of blocks, which may affect the estimation of stationary spatio-temporal filter, only the data in the same block was used for the stationary spatio-temporal filter adaptive learning.

### 3.1.3. Offline experiment

The experiment comprised two parts: offline and online. The offline experiment consisted of four blocks, each containing 80 trials. Each target was triggered once in both the first 40 trials and the last 40 trials with a random order. Each trial lasted 5.5 s. At the beginning 0.5 s of each trial, termed the “cue stage”, a red block appeared on the screen to indicate the target position (as shown in stimulator in Fig. 4(a)). The subject was asked to quickly shift their gaze to the target within this 0.5 s duration. During the next 5 s, called the “stimulus stage”, 40 stimuli flickered simultaneously at their predetermined frequencies, and their brightness followed the associated sine function. Within this 5 s, the subject was asked to focus on the target of the prompt. Meanwhile, a red triangle was shown below the target to facilitate visual fixation. After the end of this trial, the next trial would start. There was no feedback during the offline experiment.

### 3.1.4. Online experiment

The online experiment was used to verify the practical performance of the STE-DW algorithm and the FBCCA algorithm. The online experiment paradigm is the same as the offline experiment paradigm,

but used a different stimulus duration. During the online experiment, the EEG recorder sent a packet to the operator every 40 ms. The STE-DW algorithm dynamically determined the stimulus duration according to the hypothesis test result, and the stimulus duration of the FBCCA algorithm was set to 1.25 s which was the optimization used in Ref. 39. After the stimulus, the recognition result would be displayed at the top of the screen within 0.5 s.

## 3.2. Data processing

### 3.2.1. Data preprocessing

Due to the hysteresis effect of the visual system,<sup>54,55</sup> the data epochs were extracted after a certain period of time had elapsed from the start trigger generated by the stimulus program. In a typical cue-stimulus paradigm, the SSVEP component does not reach its peak amplitude at the same time as the stimulus begins but has a certain hysteresis which causes it to increase to a peak gradually. When the stimulus is stopped, the SSVEP component is not interrupted instantaneously, but remains for a short period of time and diminishes gradually, as shown in Fig. 5. Therefore, data with lower SNR at the initial segment of the trials could be avoided by delaying data extraction, and the amount of effective data can be increased by extending the duration of the data acquisition.<sup>35,56</sup> According to Ref. 35, EEG data epochs were extracted after a 140 ms delay. For example, if the duration of the stimulus was 1.25 s, the data from 0.14–1.39 s was extracted for analysis. For the dynamic window algorithm, after the termination condition was reached, the EEG Recorder received an additional 0.14 s data for the classification.

### 3.2.2. STE-DW processing flow

For the online experiment, the STE-DW algorithm processing flow is shown in Fig. 5.

- (i) At the beginning of each trial, the stimulator gave a cue which lasted 0.5 s.
- (ii) The stimulator sent a “Start trigger” marker to the operator via the recorder and entered the stimulus stage.
- (iii) The operator delayed processing data for 0.14 s from the “Start trigger”, and calculated the statistic  $\gamma$  via the iterative algorithm in real

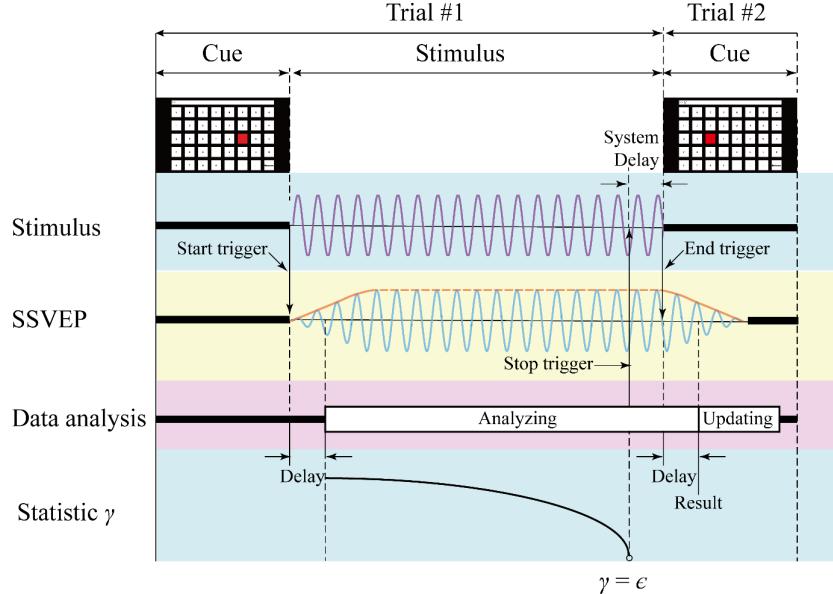


Fig. 5. STE-DW algorithm online processing flow diagram.

time. When the statistic  $\gamma$  was less than  $\varepsilon$  (the cost of “erasure decision”  $\mathcal{H}^{\{0\}}$ ) the operator sent a “Stop trigger” to the stimulator.

- (iv) The stimulus would stop after the stimulator received the “Stop trigger”, however, it may not stop immediately due to system latency caused by network, device response, etc.
- (v) After the stimulus had ceased, the stimulator sent the “End Trigger” to the operator via the EEG recorder and gave the next target cue immediately. The operator would receive an additional 0.14 s of data from the “End Trigger” and provide feedback of the recognition result within 0.5 s.
- (vi) During the cue stage of the next trial, the operator completed the stationary spatio-temporal filter update.

### 3.2.3. Parameter selection

The raw data of the Benchmark Dataset<sup>39</sup> was used for parameter optimization in this study. The dataset contained 35 subjects where 27 subjects had no experience and 8 subjects had previously participated in SSVEP experiments. The experimental paradigm of this dataset is basically the same as the offline experimental paradigm of this study, but each subject in the benchmark dataset was asked to perform six blocks, each of which included 40 trials. At the

beginning of each trial, there was a 0.5 s cue time, followed by a 5 s flicking stimulus, then the subject was given 0.5 s of rest time after the stimulus.<sup>39</sup>

By comparing the performance of different parameters on the benchmark raw dataset, the optimal parameters of each algorithm were selected and applied to the offline and online experiments. For optimizing the fixed window length algorithms, the stimulus time window was set to the range of 0.5 s–3.5 s with an interval 0.25 s. For optimizing the STE-DW algorithm, the cost coefficient was set to  $[10^0, 10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}, 10^{-10}, 10^{-12}, 10^{-14}, 10^{-16}]$ , the adaptive order range of stationary equalizer was set to  $20 \leq P \leq 40$ , and the maximum number of adjacent trials used in noise estimation was set to  $M = 10$ .

### 3.2.4. Signal-noise-ratio estimation

As mentioned above, the SSVEP-based BCI recognition problem was considered as a statistical hypothesis test problem, with the recognition accuracy being closely related to signal-to-noise ratio. It is generally believed that the SSVEP components are primarily composed of multiple sinusoidal signals, and can be defined as the projection of the SSVEP-EEG data in the subspace of the stimulus frequency and its harmonics. This can be expressed as

$$\text{Signal} = \mathbf{X}\Phi^H(\Phi\Phi^H)^{-1}\Phi. \quad (39)$$

And the noise is defined as

$$\text{Noise} = \mathbf{X} - \mathbf{X}\Phi^H(\Phi\Phi^H)^{-b}\Phi. \quad (40)$$

Under the condition of constraint  $\Phi\Phi^H = \mathbf{I}_l$ , the SNR is obtained by

$$\text{SNR} = \frac{\|\text{Signal}\|_F}{\|\text{Noise}\|_F} = \frac{\text{tr}(\mathbf{X}\Phi^H\Phi\mathbf{X}^H)}{\text{tr}[\mathbf{X}(\mathbf{I}_N - \Phi^H\Phi)\mathbf{X}^H]}. \quad (41)$$

In this experiment,  $\Phi$  was defined as the complex sine template composed of five harmonics of the stimulus frequency, that is, the template dimension  $l = 10$ . The EEG data from 0.14 s to 5.14 s after “Start trigger” in each trial was used to calculate the average SNR for the trial.

### 3.2.5. Performance evaluation

Accuracy is the basic measure of the identification accuracy of a system. It is defined as the proportion of all test trials that are correctly classified. That is, the total number of correctly classified trials divided by the total number of test trials.

Information transfer rate (ITR) is often considered to be an ideal criterion for measuring the efficiency of BCI systems,<sup>57</sup> and is one of the main optimization goals for BCI design. It is obtained by

$$\begin{aligned} \text{ITR} = & \frac{60}{T} \left[ \log_2 Q + P \log_2 P \right. \\ & \left. + (1 - P) \log_2 \left( \frac{1 - P}{Q - 1} \right) \right], \end{aligned} \quad (42)$$

where  $Q$  denotes the number of targets,  $P$  denotes the accuracy and  $T$  denoted the average trial time.

In order to simulate the results of the online experiment,  $T$  was defined as the average stimulus duration plus 0.5 s of simulated rest time in the offline experiment. And in the online experiment itself, it was defined as the average length of time from the beginning of a stimulus to the next (including the 0.5 s cue stage).

### 3.2.6. Algorithms for comparison

In order to evaluate the performance of STE-DW, the algorithms of CCA, CCA, FBCCA, CCA-RV, CVARS, and spatio-temporal equalization fixed window (STE-FW) were used for comparison.

According to Ref. 36, the CCA-RV algorithm requires adequate calibration to train the correlation

template of the target frequency and nontarget frequency. Therefore, leave-one-block cross-validation was used in CCA-RV algorithm analysis in this experiment. That is,  $N$  groups of offline analysis were calculated in the dataset containing  $N$  blocks, and each group of analysis retained one block as a test sample, and the remaining  $N - 1$  blocks were used as training samples. The recognition accuracy and ITR of each group were calculated, and the results were averaged to estimate the final accuracy and ITR.

In order to compare the effect of a dynamic window on performance, the STE-FW algorithm was also used for comparison. The STE-FW algorithm is akin to the STE-DW algorithm but when the constraint  $\gamma \geq \epsilon$  is removed, the stimulus window is set to a fixed value. The STE-FW algorithm only includes spatio-temporal equalization, not the dynamic time window, so the comparison between the STE-FW algorithm and the STE-DW algorithm can clearly reflect the significance of the dynamic time window.

## 4. Result

### 4.1. Parameter sweep

Figure 6 shows the results of parameters sweep for the six algorithms (STE-DW, STE-FW, FBCCA, CVARS, CCA and CCA-RV) using the Benchmark Dataset. The optimal parameters of each algorithm with the highest ITR using the benchmark dataset were applied to the offline and online experiment.

According to the analysis results, for recognizing the SSVEP data of 35 subjects using this dataset, the optimal average ITR for the STE-DW algorithm was significantly higher than that of the other algorithms. The Lilliefors test was applied on the ITR performance difference between the STE-DW algorithm and the other five algorithms. The results showed that the normality hypothesis could not be rejected in the pairwise pairing results under the significance threshold of 0.05. That is, the paired  $t$ -test can be used for comparative analysis. In order to avoid false-positive results due to multiple replicates, Bonferroni correction was used to correct the paired  $t$ -test results. The result of the paired  $t$ -tests with Bonferroni correction, versus STE-FW:  $p = 6.12 \times 10^{-11}$ , versus FBCCA:  $p = 8.93 \times 10^{-4}$ , versus CVARS:  $p = 2.57 \times 10^{-14}$ , versus CCA:  $p = 1.21 \times 10^{-15}$ , versus CCA-RV:  $p = 3.92 \times 10^{-11}$ ). The results are shown in Table 1.

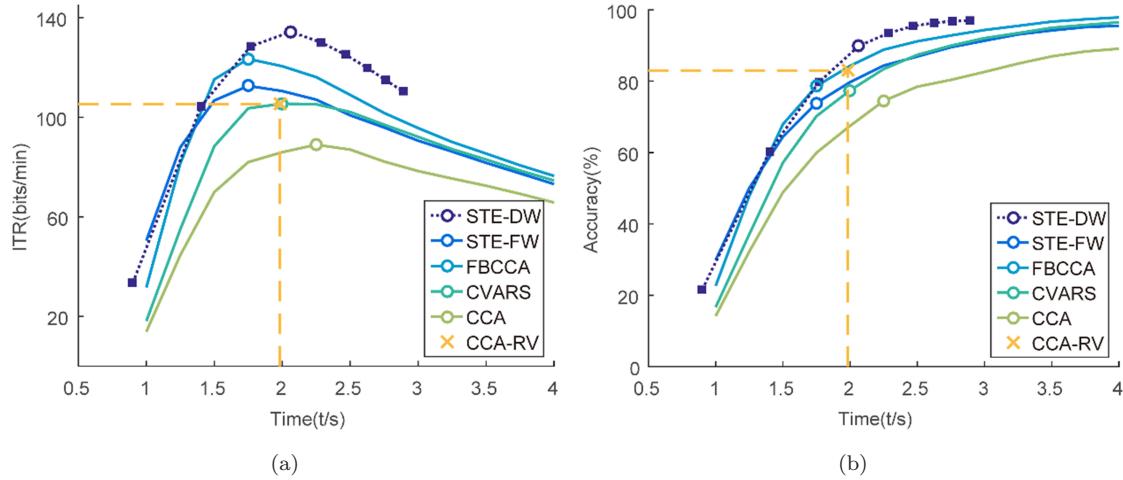


Fig. 6. The average results of 35 subjects in the benchmark dataset (a) Average ITR for six algorithms. (b) Average accuracy for 6 algorithms. On the STE-DW algorithm performance curve, ‘■’ represents the results for the different cost coefficients  $\epsilon$ . On other curves, ‘o’ represents the optimal parameter points for the various algorithms in the benchmark dataset.

Table 1. The optimal result of algorithms in benchmark dataset.

	STE-DW	STE-FW	FBCCA	CVARS	CCA	CCA-RV
Avg ITR(bits/min)	$134.4 \pm 39.2$	$112.6 \pm 45.6$	$123.4 \pm 44.1$	$105.5 \pm 38.5$	$89.0 \pm 37.0$	$105.3 \pm 31.2$
Avg Accuracy(%)	$89.8 \pm 8.0$	$73.9 \pm 21.4$	$78.6 \pm 19.4$	$77.4 \pm 19.8$	$74.5 \pm 21.9$	$83.0 \pm 10.5$
Optimal Trial length(s)*	$2.06 \pm 0.43$	1.75	1.75	2.0	2.25	$1.98 \pm 0.64$

Notes: \*The optimal trial length for STE-DW algorithm represents the average trials time with  $\epsilon = 10^{-6}$ , and for CCA-RV algorithm represents the average trials time.

## 4.2. Experiment data

### 4.2.1. Offline experiment result

The performance of each algorithm in offline experiments is shown in Fig. 7. The optimal average ITR of the STE-DW algorithm (116.5 bits/min) was higher than the other algorithms, 29.1% better than that of the STE-FW algorithm (90.2 bits/min), 32.8% better than that of the FBCCA algorithm (87.7 bits/min), 42.6% better than that of the CVARS algorithm (81.7 bits/min), 51.1% better than that of the CCA algorithm (77.1 bits/min), and 20.35% better than the CCA-RV algorithm (96.8 bits/min). The Lilliefors test was applied on the ITR performance difference between the STE-DW algorithm and the other five algorithms. The results showed that the normality hypothesis could not be rejected in the pairwise pairing results under the significance threshold 0.05. That is, the paired  $t$ -test can be used for comparative analysis. The paired  $t$ -test with

Bonferroni correction showed that there were significant differences between the ITR of the STE-DW algorithm and those of the other algorithms, the  $p$ -values, respectively were  $3.90 \times 10^{-5}$  (versus STE-FW),  $5.53 \times 10^{-5}$  (versus FBCCA),  $1.08 \times 10^{-5}$  (versus CVARS),  $2.06 \times 10^{-6}$  (versus CCA), and  $5.26 \times 10^{-4}$  (versus CCA-RV).

As can be seen from Fig. 7(a), compared with the Benchmark Dataset, the trial time window corresponding to the optimal ITR of the STE-FW and FBCCA algorithms changed in the offline experiment. In the offline experiment, the optimal trial time for these two algorithms was 2.25 s. Using this trial time, the STE-FW algorithm and the FBCCA algorithms could achieve optimal average ITRs of 97.0 bits/min and 96.5 bits/min, respectively, which were significantly lower than those in the Benchmark Dataset. However, the average performance gap between the dynamic window algorithms STE-DW and CCA-RV was small between the benchmark

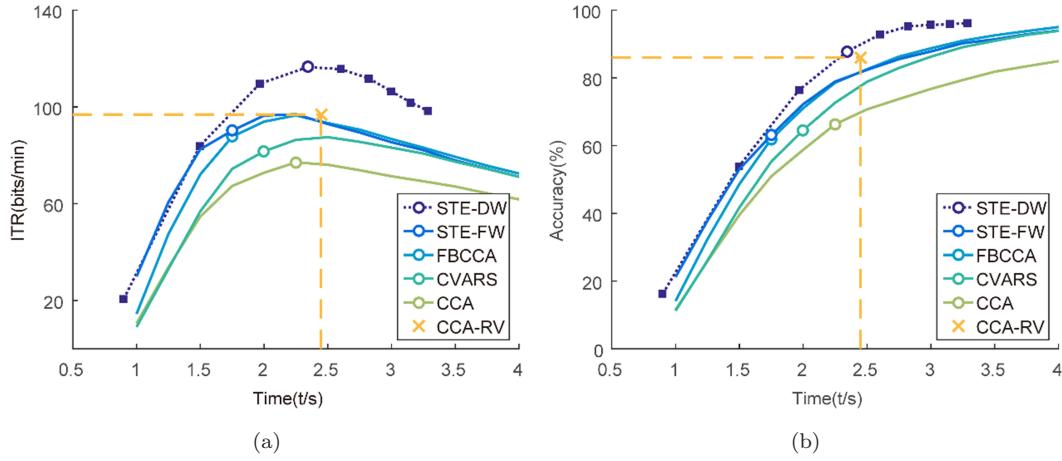


Fig. 7. The average results of 16 subjects in offline experiment (a) Average ITR for six algorithms. (b) Average accuracy for six algorithms. On the STE-DW algorithm performance curve, ‘■’ represents the results for the different cost coefficients  $\epsilon$ . On other curves, ‘○’ represents the optimal parameter points from the benchmark dataset.

dataset and the offline experiment, and the optimal value of the cost coefficient  $\epsilon$  in the STE-DW algorithm was the same as that in the benchmark dataset.

This situation further demonstrates that the fixed window algorithms lack stability between different datasets, and the performance of these algorithms is heavily dependent on the SSVEP response intensity of the subject. Therefore, the fixed window algorithms are very likely to be difficult to apply in the subjects with low SNR.

The STE-DW algorithm can adaptively adjust the time window according to the SNR, so as to

optimize the ITR. In the conducted BCI experiments, there were large differences in the SNR between different subjects. As shown in Figs. 8(a) and 8(b), the trial SNRs of S7 were concentrated at about  $-10\text{ dB}$ , and most of the trial SNRs of S14 were concentrated at  $-16\text{ dB}$ . From Table 2, the accuracies of other fixed window algorithms for S14 were about 20%, but the accuracy of the STE-DW algorithm could reach 70%. As illustrated in Figs. 8(c) and 8(d), the STE-DW algorithm dynamically shortened the time window for S7 so that the recognition speed could be increased as much as possible while maintaining a high classification accuracy.

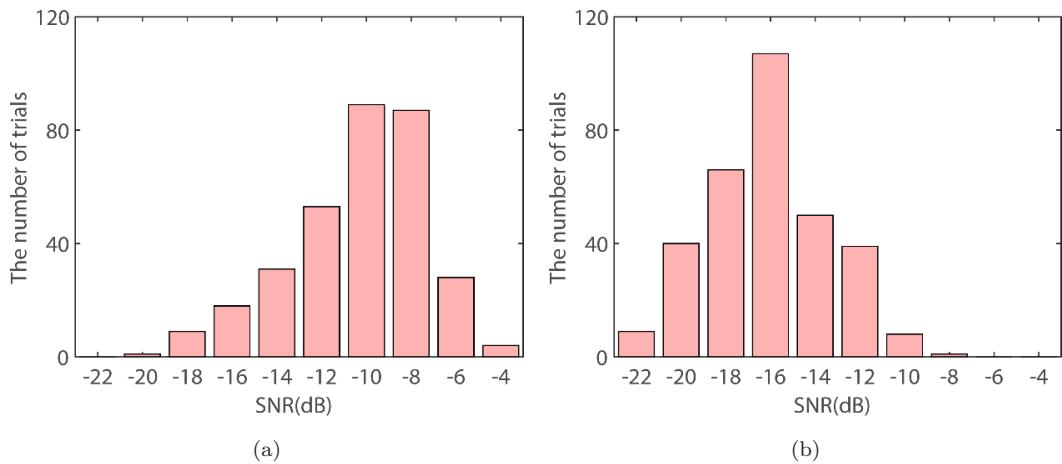


Fig. 8. The distribution of SNR and trial time of the STE-DW algorithm with typical subjects. (a) The distribution of SNR with subject S7. (b) The distribution of SNR with subject S14. (c) The distribution of trial time of the STE-DW algorithm with subject S7. (d) The distribution of trial time of the STE-DW algorithm with subject S14.

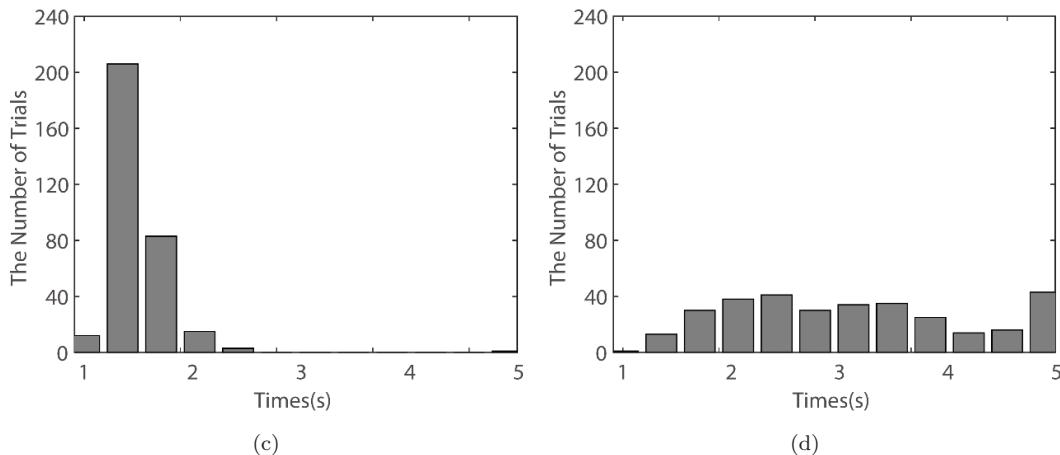


Fig. 8. (*Continued*)

Table 2. The offline experiment results (ITR: bits/min, Acc: %, Time: s).

Sub	Gen*	Age	STE-DW			STE-FW		FBCCA		CVARS		CCA		CCA-RV		
			ITR	Acc*	Time	ITR	Acc	ITR	Acc	ITR	Acc	ITR	Acc	ITR	Acc	Time
1	M	22	156.1	93.8%	1.79	136.9	85.9%	135.2	85.3%	115.4	84.1%	92.1	78.8%	109.9	88.1%	1.92
2	F	24	69.8	83.1%	3.25	31.2	34.1%	35.0	36.6%	33.6	38.8%	34.2	42.2%	63.4	84.7%	3.34
3	M	22	135.6	94.7%	2.10	97.2	69.7%	94.4	68.4%	110.4	81.9%	109.1	87.2%	118.5	93.1%	1.98
4	F	68	88.4	89.4%	2.90	43.0	41.6%	39.9	39.7%	36.7	40.9%	70.3	66.6%	90.0	88.4%	2.46
5	F	62	164.2	98.4%	1.87	156.7	92.8%	144.7	88.8%	141.4	94.4%	137.6	98.8%	153.7	93.1%	1.44
6	M	23	86.9	79.4%	2.42	45.6	43.1%	36.5	37.5%	21.5	29.4%	14.0	24.4%	64.6	94.1%	4.00
7	M	24	198.1	98.1%	1.54	172.2	97.5%	173.3	97.8%	155.9	99.1%	140.7	99.7%	149.4	95.3%	1.58
8	F	28	95.5	88.8%	2.65	44.0	42.2%	25.0	29.7%	48.9	49.1%	71.9	67.5%	90.8	89.4%	2.48
9	F	24	135.7	90.6%	1.94	132.7	84.4%	116.8	78.1%	126.6	88.8%	132.2	96.9%	134.0	96.3%	1.84
10	M	59	68.9	76.9%	2.89	31.2	34.1%	34.6	36.3%	37.2	41.3%	31.5	40.0%	60.6	80.9%	3.36
11	F	18	173.6	96.3%	1.69	164.6	95.3%	162.6	94.7%	146.8	96.3%	137.6	98.8%	151.4	96.9%	1.61
12	M	28	93.4	85.6%	2.55	88.3	65.6%	93.1	67.8%	81.4	67.8%	65.1	63.4%	75.1	60.9%	1.51
13	M	30	132.5	95.3%	2.17	105.0	73.1%	107.2	74.1%	87.5	70.9%	72.9	68.1%	96.6	91.6%	2.42
14	M	62	50.5	70.0%	3.39	11.7	19.1%	19.1	25.3%	16.4	25.0%	9.1	19.1%	32.9	55.0%	3.19
15	F	17	142.8	95.6%	2.03	130.2	83.4%	136.9	85.9%	91.9	73.1%	96.9	81.3%	110.1	89.4%	2.00
16	M	18	71.4	68.8%	2.33	52.6	47.2%	48.8	45.0%	54.9	52.8%	18.1	28.4%	48.5	79.7%	4.03
Avg			116.5	87.8%	2.34	90.2	63.1%	87.7	61.9%	81.7	64.6%	77.1	66.3%	96.8	86.1%	2.45
Std			43.9	9.7%	0.55	53.8	25.6%	53.4	25.5%	47.2	25.1%	46.4	27.9%	38.0	12.1%	0.88

*Notes:* \*Acc is short for accuracy. Gen is short for gender, and 'M' means male, 'F' means female.

On the other hand, the STE-DW algorithm was able to maintain a relatively stable recognition accuracy for S14 by extending the test window length.

Due to the nonwhite characteristics of background EEG noise, there were also large differences among the SNR of the SSVEP under different stimulus frequencies. The STE-DW algorithm could achieve the dynamic optimization of the time window to balance the recognition accuracies among different stimuli. The red curve in Fig. 9(a) shows the

average SNR of the SSVEP under different stimulus frequencies. As the stimulus frequency increases, the average signal-to-noise ratio of the SSVEP decreases gradually, resulting in a decrease in recognition accuracy. According to Fig. 9(b), the dynamic time window of the STE-DW algorithm adaptively extended the trial window length of the high-frequency stimuli, thereby effectively accumulating the SSVEP signal energy and balancing the accuracies among all frequencies. Figures 9(c) and 9(d) show the recognition

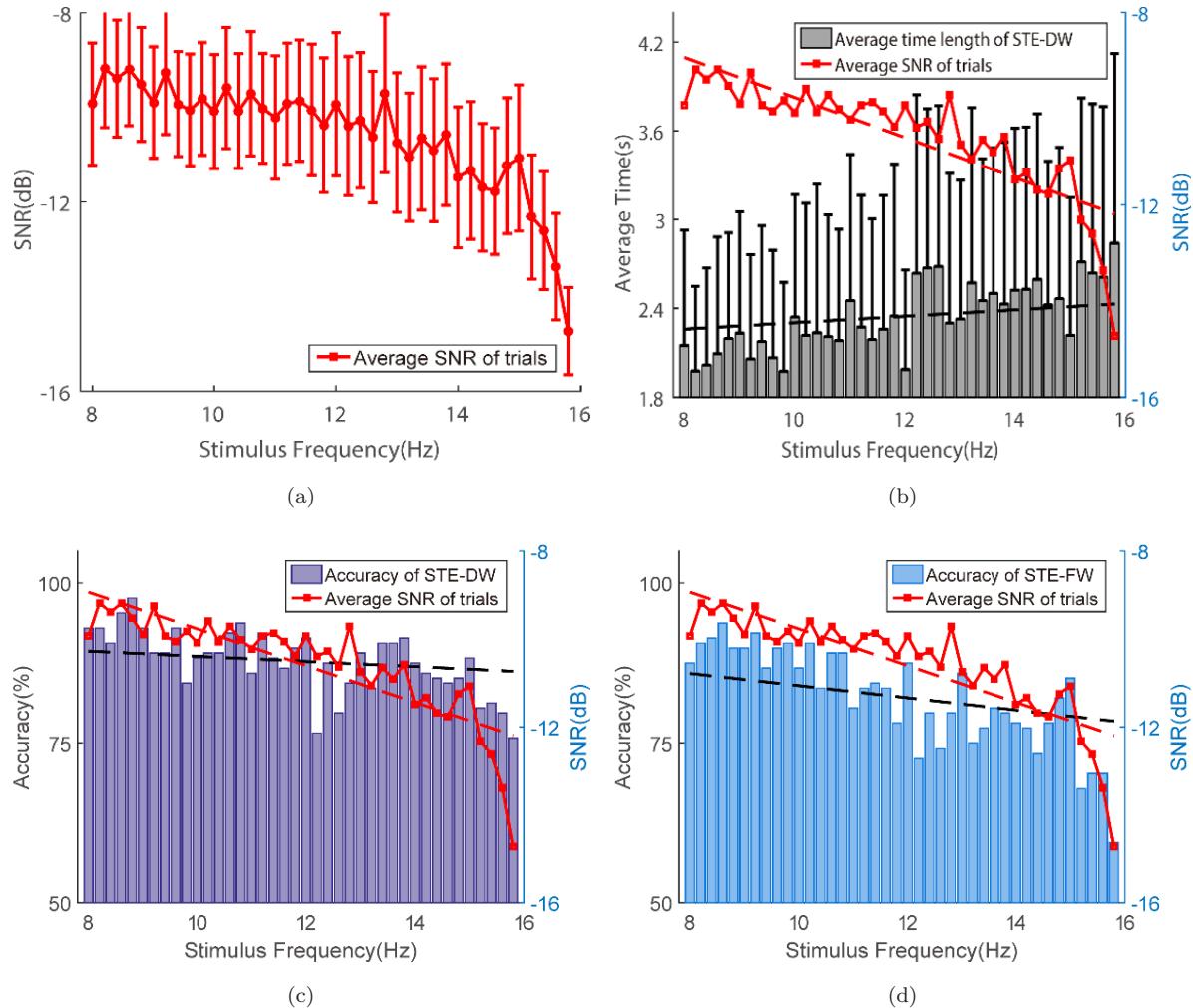


Fig. 9. (a) Average SNR of SSVEP induced by each stimulus frequency for all trials. The error bar represents the standard deviation. (b) Average time window of the STE-DW algorithm under different stimulus frequencies. The histogram shows the average time window of the STE-DW algorithm in each stimulus frequency, and the black dotted line indicates the linear fitting result of the average time window. The slope of the fitting result is  $k = 0.022 \text{ s/Hz}$ . (c) Average accuracy of the STE-DW algorithm under different stimulus frequencies. The histogram shows the recognition accuracy of the STE-DW algorithm at each stimulus frequency, and the black dotted line indicates the linear fitting result of the recognition accuracies. The slope of the fitting result  $k = -0.40\%/\text{Hz}$  (d) average accuracy of the STE-FW algorithm under different stimulus frequencies. The slope of the fitting result  $k = -0.96\%/\text{Hz}$ . The linear fitting result of the average SNR of SSVEP in (b), (c) and (d) is denoted by the red dotted line with a slope  $k = -0.42 \text{ dB/Hz}$ . The STE-FW algorithm time window was set to 2.34 s which was the average trial time in the offline experiment of the STE-DW algorithm with  $\epsilon = 10^{-6}$ .

accuracies of the STE-DW algorithm and STE-FW algorithm at different stimulus frequencies using a trial time of 2.34 s for the STE-FW algorithm. The recognition accuracy of the STE-DW algorithm was 92.97% and the recognition accuracy of the STE-FW algorithm was 90.6% using a stimulus frequency of 8.2 Hz. Using a stimulus frequency of 15.8 Hz, the recognition accuracy of the STE-DW algorithm was 75.8% but the STE-FW algorithm accuracy was only 59.4%. These results suggest that the

STE-DW algorithm is superior to the STE-FW algorithm in terms of recognition accuracy and frequency stability.

#### 4.2.2. Online experiment data

As presented in Fig. 6, the algorithm FBCCA performs better than the other training-free algorithms using the Benchmark Dataset, therefore, FBCCA has been used as the reference algorithm for online

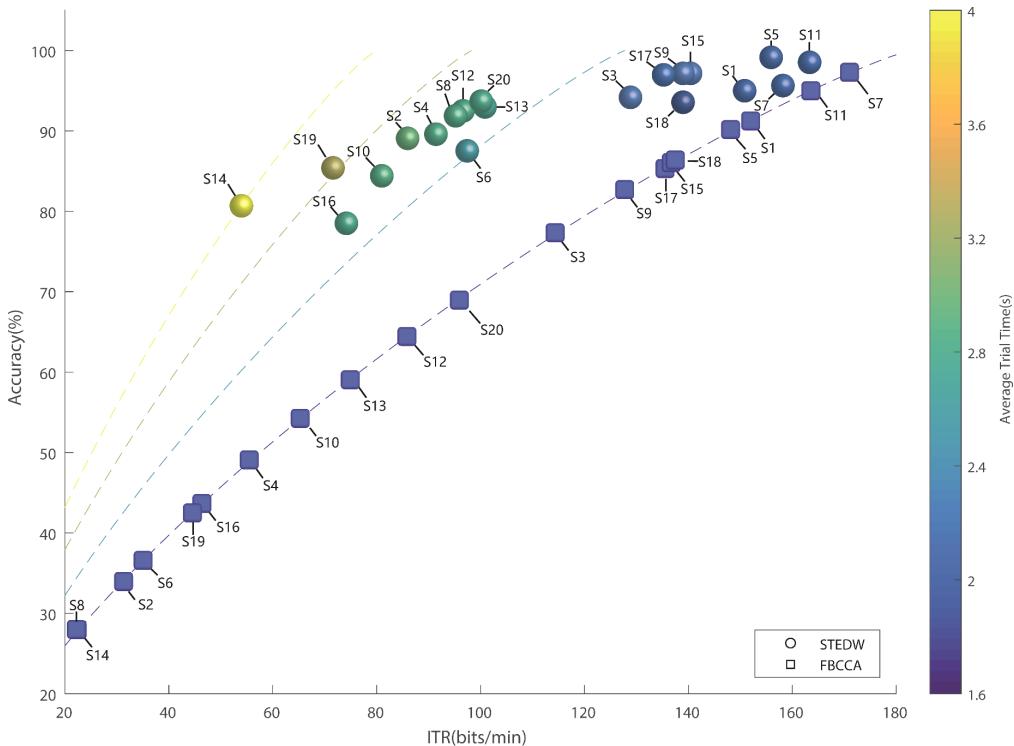


Fig. 10. The average ITR, accuracy and trial duration of each subject using the STE-DW algorithm and the FBCCA algorithm in the online experiment. The dotted lines show the fitting relation between accuracy and ITR at different trial times ( $T = 1.75, 2.25, 3.25, 4$  s).

experiments. Figure 10 shows the results from these experiments for each subject comparing the performance of STE-DW to FBCCA.

Due to the use of a fixed stimulus time 1.25 s (trial time 1.75 s) for the FBCCA algorithm, the experimental results of each subject are strictly distributed along the dotted line  $T = 1.75$  s. For the FBCCA algorithm, there were large variances in the ITR and accuracy of different subjects. The typical subject S7 in the FBCCA online experiment achieved an accuracy of 97.2% and an ITR of 171.2 bits/min, but the typical subject S14 achieved an accuracy of only 27.8% and ITR of only 22.4 bits/min. The lowest accuracies of the FBCCA algorithm were 27.8% (S8, S14) in the online experiment. For the STE-DW algorithm, the typical subject S7 achieved an accuracy of 95.6% and ITR of 158.3 bits/min, which was slightly lower than that of the FBCCA. However, the other typical subject S14 achieved an accuracy of 80.6% and ITR of 54.3 bits/min in the online experiment of the STE-DW algorithm. Across all subjects, the lowest accuracy of the STE-DW algorithm was 78.4% (S16) in the online experiment.

In order to compare the differences between the STE-DW algorithm and the FBCCA algorithm in different groups, all subjects were divided into two groups according to their FBCCA performance in the online experiment. S1, S3, S5, S7, S9, S11, S15, S17, S18 and S20 were placed into the high performance group, because their average ITR was higher than 90 bits/min in the FBCCA online experiment. S2, S4, S6, S8, S10, S12, S13, S14, S16 and S19 were placed into the low-performance group, because of their lower ITR in the FBCCA online experiment. The results demonstrated that, in the FBCCA online experiment, the average ITR of the high-performance group was  $138.3 \pm 22.3$  bits/min and the average classification accuracy rate was  $85.9\% \pm 8.3\%$ . In the STE-DW online experiment, the average ITR of the high-performance group was  $141.2 \pm 18.1$  bits/min and the average classification accuracy rate was  $96.0\% \pm 2.0\%$ . In the high-performance group, the performance advantage of the STE-DW algorithm was not significant when compared with the FBCCA algorithm, but the STE-DW algorithm showed a great advantage in the low-performance group. In the

FBCCA online experiment, the average ITR of the low-performance group was  $48.2 \pm 21.9$  bits/min and the average accuracy rate was  $43.8\% \pm 12.7\%$ , and in the STE-DW online experiment, the average ITR of low-performance group was  $84.9 \pm 146$  bits/min and the average accuracy rate was  $87.2\% \pm 5.0\%$ . This result suggests that the STE-DW algorithm may have a more pronounced optimization effect on poor performing subjects.

Although there is variation in ITR between subjects, since the STE-DW algorithm adopts a dynamic time window, the recognition accuracy of each subject can be consistently maintained at a high level which may enhance the adaptability of the system. To enhance the efficiency of information transmission while still maintaining a certain accuracy, the STE-DW algorithm is likely a more suitable choice over FBCCA.

## 5. Discussion

### 5.1. Proportion of effective population

Because of the differences among different users' brain signals, it is often difficult to create a BCI system that is usable by everyone. The proportion of the population that can effectively use a BCI (termed the "effective population") is also regarded as an important performance index of BCI systems. References 58 and 59 argued that a BCI was effective for a subject if the accuracy was over 70%. Reference 60 further explored how many (and what kinds of) people could use SSVEP-based BCIs. However, due to the small number of stimulus targets in the study, most of the subjects under sufficient stimulation time could achieve high recognition accuracy. Therefore, this section examines the effective population proportion of the different algorithms in the benchmark dataset and offline experiment with 70% accuracy as the validity criterion.

Figure 11 shows the effective proportions of different algorithms in the two datasets. The accuracy of the STE-DW algorithm in the benchmark dataset for all subjects was greater than or equal to 70%. It can be seen in the offline experiment results (see Table 2), that only the recognition accuracy rate of S16 (68.8%) was less than 70%.

Dynamic window algorithms such as the STE-DW and CCA-RV algorithms have a relatively large effective population compared to the fixed

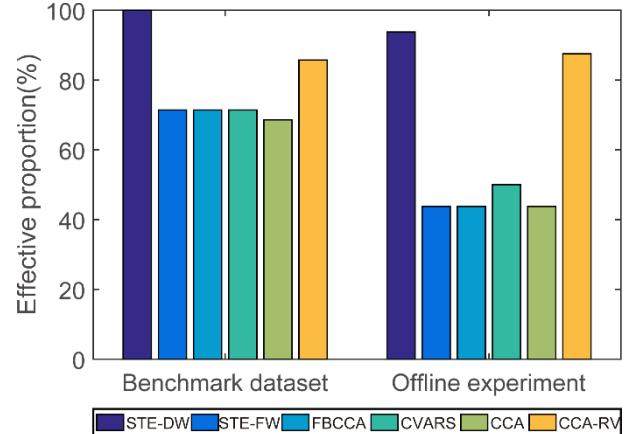


Fig. 11. The effective population proportion of different algorithms in benchmark dataset and offline experiment.

window algorithms. This is because the recognition accuracies of fixed windows algorithms largely depend on the selection of the time window, so that it is difficult to take both the transmission efficiency and effectiveness into account in the design of the fixed window algorithm. However, the dynamic window algorithm can not only optimize the time window for each subject, but also for each trial, taking into account both efficiency and effectiveness. This increases the number of people able to effectively use the BCI while maintaining a high ITR.

It should be noted that the accuracy of the "fixed window length" algorithms will increase accordingly with increasing stimulation time. Therefore, the proportion of people who exceed the 70% accuracy threshold will correspondingly increase. However, longer stimulation times often result in a reduction in the ITR, which in turn reduces overall system performance.

In this part of the study, the "effective population"<sup>58,59</sup> is based on the parameters corresponding to the optimal average ITR in the benchmark dataset of each algorithm. The benchmark dataset contains a large number of subjects (35), so the parameters optimized on the dataset can be extended to practical applications. In the absence of individual calibration data, this optimization parameter has a certain reference value.

Another way to increase the effective population is to use a subject's own calibration data for parameter optimization. Studies have shown that system performance can be significantly improved

by using individual-optimized parameters.<sup>61</sup> However, in practice, this method usually has some limitations. Firstly, the time needed to optimize the parameters using the calibration dataset is generally long, and it needs to be recalibrated before each use. Secondly, to ensure that the parameter optimization is sufficiently reliable, enough data must be prepared as calibration data, so it is also necessary to extend the experiment. Finally, if the calibration data satisfies the requirements, some better performing training algorithms can also be used, such as the Task-Related Component Analysis.<sup>62</sup> Therefore, it is generally believed that the training-free methods that do not require calibration data are more conducive to practical applications.

Subject selection is an important problem in BCI experiments. A total of 21 subjects were invited to participate in offline and/or online experiments, which included 16 persons under the age group of 17 to 30, and five persons under the age group of 59–68. A 66-year old male subject failed to complete all the experiments successfully due to physical reasons. His data were not involved in the statistics. Before conducting offline and online experiments, we hope to verify whether the dynamic window algorithm can help the weaker subjects to increase their recognition efficiency. It is assumed that the signal quality of older people might be lower than that of younger subjects, so five older subjects were invited to participate in the experiment. However, the results were not as expected. In the STE-DW off-line experiment results, although the best performer was a 24-year-old male (S7), the worst performer was a 62-year-old male (S14), but overall, the difference in performance between older subjects (S4, S5, S10, S14) and younger subjects was not significant. Therefore, we did not group the results by age. Instead, we considered all subjects as a whole and divided them into two groups according to their performance in the FBCCA online experiment.

### 5.2. Parameter selection and the stability of the STE-DW algorithm

The STE-DW algorithm is relatively stable for all subjects, after determining the cost factor  $\epsilon$ . Figure 12 shows that most subjects in the offline experiment can achieve or approach the maximum value of the ITR curve when the parameter  $\epsilon = 10^{-6}$ . It can also be seen that the optimal cost function  $\epsilon$

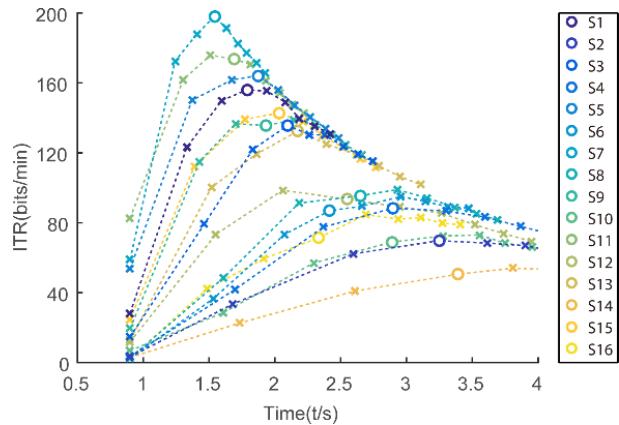


Fig. 12. The curve of ITR for each subject in the offline experiment with different cost coefficients  $\epsilon$ . Each dotted line represents one subject, ‘ $\times$ ’ represents the different cost coefficients of the STE-DW algorithm, from left to right, followed by  $\epsilon = 10^0, 10^{-2}, \dots, 10^{-16}$ , and ‘ $\circ$ ’ means  $\epsilon = 10^{-6}$ .

in the STE-DW algorithm is related to the experimental system, and less related to the subject. In the design of STE-DW system, the optimal cost coefficient  $\epsilon$  is closely related to the total number of stimuli in the experimental paradigm, the frequency distribution of the stimulus, the data sampling rate and the number of EEG channels. Therefore, it is advisable to optimize the cost factor coefficients  $\epsilon$  for each specific experimental paradigm by parameter sweeping. In addition to the cost coefficients, the influence of other parameters on the accuracy of recognition is relatively small. The only main concern is that using background noise which is too short for adaptive learning may lead to poor equalizer stability, while using excessive background noise for adaptive learning may cause the equalizer difficulties in dynamically tracking the background noise changes and thus increase the learning complexity of the stationary equalizer. According to this study, choosing 0.5–2 minutes of data of background noise as the stationary equalizer self-learning data is the best.

### 5.3. The significance of spatio-temporal equalizing and algorithm application

Since the transmission pathway of the SSVEP is the same as the background noise, the equalizer based on the background noise can, to a certain extent, offset

the difference caused by the volume conductor effect between the SSVEPs induced by different stimuli.

The spatio-temporal model also takes the spatial correlation and the temporal correlation of EEG background noise into consideration, while CCA, MEC, MSI and other space-based second-order statistics algorithms are theoretically equivalent to assuming that the background noise is independent and identically distributed at each time point. However, it is known that this is actually not true and this assumption is not fulfilled. Therefore, assuming that there is both spatial and temporal correlation in the background noise is more precise and better reflects the characteristics of the source signals. The spatio-temporal algorithm uses adaptive stationary spatial-temporal equalizers to remove the spatial and temporal correlation between the signals, and then uses a set of adaptive spatial equalizers to correct the stationary equalizers so that each point of the output signal is closer to the fully uncorrelated.

The spatio-temporal equalizers used in this paper are a simplification of the theory. The spatio-temporal correlation of background noise of EEG is inseparable, and the equalizer should be designed according to the spatio-temporal correlation. However, in practical applications, the computational complexity of spatio-temporal equalizer is high and is difficult to achieve in real-time processing. In this paper, the stationary equalizer  $\mathcal{M}(z)$  was decomposed into a stationary spatial equalizer  $\mathbf{P}$  and a stationary temporal equalizer  $\mathcal{T}(z)$ , so that the number of free variables could drop from  $L^2\rho$  to  $L(L+1)/2 + L\rho$ . Offline analysis showed that the use of  $\mathbf{P}$  and  $\mathcal{T}(z)$  or the use of  $\mathcal{M}(z)$  as equalizers had little effect on the accuracy of classification, and the former had better computational efficiency.

Although the spatio-temporal equalization used in this study is a type of spatio-temporal filtering algorithm, its purpose is significantly different from the other BCI spatio-temporal filtering methods.<sup>22,34</sup> Most of the traditional spatio-temporal filtering techniques are based on optimizing the characteristics of the target signal, whose features can be enhanced by spatio-temporal filtering. The spatio-temporal equalization algorithm used in this study is mainly used to suppress the background noise correlation. Therefore, this algorithm differs from other known spatio-temporal filtering algorithms in principle and parameter estimation. As a dynamic window algorithm, the

STE-DW algorithm needs to be calculated in real-time based on the received data. In real-time processing, the computational complexity of the process calculating  $\hat{\mathbf{A}}^{\{q\}}$  is still high according to (22), so it is preferable to use iterative updating to avoid the direct estimation of  $\hat{\mathbf{A}}^{\{q\}}$ . After optimization, the computational complexity of the STE-DW algorithm can be significantly reduced and can be much lower than that of FBCCA.

#### **5.4. Some details of the experiment design**

It is worthy to pay attention to the problem of the monitor refresh rate in visual response experiments. In this study, the refresh rate of the monitor was set to 60 Hz, and the stimulation frequency range was set to 8–15.8 Hz. All stimulation frequencies are less than 1/3 of the display refresh rate to ensure that the sinusoidal stimulus can be fully rendered. Therefore, the effect of monitor refresh rate on stimulus integrity can be largely ignored. The stimulus paradigm design and monitor refresh rate settings in this study are consistent with the benchmark dataset paper.<sup>39</sup> In this paper, the best-performed JFPM method<sup>56</sup> achieves an average ITR of 200 bits/min with training algorithms. Therefore, we believe that the refresh rate is sufficient to meet the demand of 8–15.8 Hz flickering stimulation.

In order to reduce the potential risk of inducing epileptic seizure on the experimentation, a certain degree of precautionary measures was set in the experiment. First, in terms of experimental design, we follow the paradigm design method provided by Ref. 39. The experimental paradigm has been validated on 35 subjects and no significant adverse effects have been reported. Second, we restricted the intensity of the stimulus by controlling the brightness of the monitor during the experiment to avoid discomfort. Finally, all the subjects strictly were required to have no neurological diseases and no history of epilepsy.

#### **5.5. BCI system performance comparison**

As a human-computer interaction system, a BCI system can be evaluated according to the following aspects.

- (i) Efficiency: The ITR, accuracy and trial duration are generally used as the criteria for evaluating the efficiency of a BCI system. High ITR, high average accuracy and short response time to stimulus, all contribute to an improved efficiency.
- (ii) Effectiveness: The proportion of “effective population” reflects the effectiveness of the BCI system. A higher proportion of effective population indicates the higher adaptability of the BCI, thus the system has higher application value.
- (iii) Economy: The core of the BCI system is the recognition algorithm. The computational complexity of the recognition algorithm determines the implementation cost of the whole system. An algorithm with lower complexity is easier to implement with low-cost equipment, and thus has higher economic efficiency.
- (iv) Ease-of-use: The training time required before a subject may effectively use a BCI online which

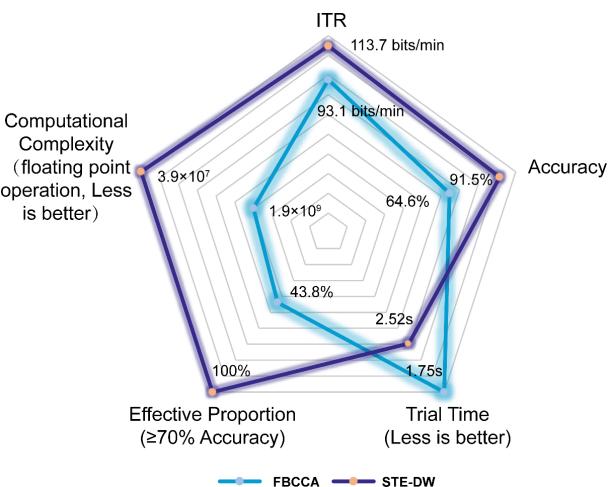


Fig. 13. Performance comparison of the STE-DW and FBCCA algorithm in the online experiment. The trial time includes the 0.5 s cue stage and the average stimulus stage which is 2.02 s for the STE-DW algorithm and 1.25 s for the FBCCA algorithm. The computational complexity is the average computational complexity of a single trial. The computational complexity of the STE-DW algorithm is estimated according to the optimized algorithm, which includes the real-time processing and the filter adaptive updating steps in the single trial calculation, as well as the computation complexity of the CCA step of the FBCCA algorithm estimated using the ‘canoncorr’ function in Matlab, 2015b.

affects the ease of use of the system. The shorter the training time required, the easier it is for the BCI to be widely used.

Based on the aspects above, the online experimental data were used to compare the performance of the STE-DW algorithm and the FBCCA algorithm in practical applications (since the STE-DW algorithm and the FBCCA algorithm are both training-free algorithms, the ease of use was not compared). Figure 13 illustrates that except for trial time, other performances of the STE-DW algorithm are better than those of the FBCCA algorithm. During the practical use, the STE-DW algorithm finishes the analysis step in 1 ms, and the adaptive updating of the stationary equalizer can be completed within 10 ms. In Ref. 35, it was mentioned that a single recognition of the FBCCA algorithm requires an average of about 40 ms.

## 6. Conclusion

In summary, by building up the mathematical model, the STE-DW algorithm achieved a dynamic time window detection system based on the hypothesis test. The system has simple design parameters, low computational complexity and no need for single subject training data. Online experimental results show that the STE-DW method can significantly increase ITR compared to the traditional FBCCA method (STE-DW: 113.7 bit/min versus FBCCA: 93.1 bits/min) and the average classification accuracy (STE-DW: 91.5% versus FBCCA: 64.6%). Meanwhile, due to the use of an adaptive window detection mechanism, the STE-DW algorithm can achieve high adaptability, and can be applied to the vast majority of subjects. Therefore, the BCI system based on the STE-DW algorithm has achieved a good balance in terms of efficiency, effectiveness, economy and ease of use.

## Acknowledgments

The authors would like to thank Yuting Liu, Zhiqiang Yuan for assisting in drawing, and Erwei Yin for the support in the implementation of the CCA-RV algorithm. This work was supported by the National Key R&D Program of China Under Grant No. 2017YFB1002505 and National Natural Science

Foundation of China Under Grant Nos. 61431007 and 91320202.

### Appendix A. Derivation of the Statistics $\gamma$

If the EEG data  $\mathbf{X}$  has been collected, the cost expectation of the hypothesis  $\mathcal{H}^{\{q_{\mathcal{H}}\}}$  can be expressed by

$$\begin{aligned} \text{Cost}_{\mathcal{H}^{\{q_{\mathcal{H}}\}}}(\mathbf{X}) &= \sum_{\mathcal{G}^{\{q\}}} c(\mathcal{H}^{\{q_H\}}, \mathcal{G}^{\{q\}}) p(\mathcal{G}^{\{q\}} | \mathbf{X}) \\ &= \frac{1}{Q} \sum_{q=1}^Q p(\mathbf{X} | \mathcal{G}^{\{q\}}) - \frac{1}{Q} p(\mathbf{X} | \mathcal{G}^{\{q_H\}}), \end{aligned} \quad (\text{A.1})$$

and the cost expectation of the hypothesis  $\mathcal{H}^{\{0\}}$  can be expressed by

$$\begin{aligned} \text{Cost}_{\mathcal{H}^{\{0\}}}(\mathbf{X}) &= \sum_{\mathcal{G}^{\{q\}}} c(\mathcal{H}^{\{0\}}, \mathcal{G}^{\{q\}}) p(\mathcal{G}^{\{q\}} | \mathbf{X}) \\ &= \frac{\epsilon}{Q} \sum_{q=1}^Q p(\mathbf{X} | \mathcal{G}^{\{q\}}). \end{aligned} \quad (\text{A.2})$$

When  $\text{Cost}_{\mathcal{H}^{\{0\}}}(\mathbf{X}) \leq \text{Cost}_{\mathcal{H}^{\{q_{\mathcal{H}}\}}}(\mathbf{X})$  for all  $q_{\mathcal{H}} \neq 0$ ,  $\mathcal{H}^{\{0\}}$  should be decided, that is, the system does not make a judgment, but continues to collect data.

In such condition, that is,  $\forall q_{\mathcal{H}} \neq 0$ , there is

$$\begin{aligned} \frac{\epsilon}{Q} \sum_{q=1}^Q p(\mathbf{X} | \mathcal{G}^{\{q\}}) &\leq \frac{1}{Q} \sum_{q=1}^Q p(\mathbf{X} | \mathcal{G}^{\{q\}}) - \frac{1}{Q} p(\mathbf{X} | \mathcal{G}^{\{q_{\mathcal{H}}\}}), \end{aligned} \quad (\text{A.3})$$

which can be simplified to

$$1 - \frac{p(\mathbf{X} | \mathcal{G}^{\{q_{\mathcal{H}}\}})}{\sum_{q=1}^Q p(\mathbf{X} | \mathcal{G}^{\{q\}})} \geq \epsilon. \quad (\text{A.4})$$

It can be seen that when  $\gamma \geq \epsilon$ , the cost of  $\mathcal{H}^{\{0\}}$  is smallest, and the system should continue to collect data. When  $\gamma < \epsilon$ , it is assumed that there is already a hypothesis with a cost that is lower than that of  $\mathcal{H}^{\{0\}}$ , then the system can stop data acquisition, and give the recognition results through a specific decision criteria.

### Appendix B. The Maximum Likelihood Estimation of the Matrix $\mathbf{A}$

By applying the logarithmic likelihood function, that is

$$\begin{aligned} L(\mathbf{X} | \mathbf{A}^{\{q\}}; \mathcal{H}^{\{q\}}, \hat{\Sigma}_{\bar{w}}^{-1/2}) &= -\ln p(\mathbf{X} | \mathbf{A}^{\{q\}}; \mathcal{H}^{\{q\}}, \hat{\Sigma}_{\bar{w}}^{-1/2}) \\ &= \frac{1}{2} \bar{\theta}^{\{q\}^H} \bar{\theta}^{\{q\}} + \frac{LN}{2} \ln(2\pi) + \frac{1}{2} \ln \det(\Sigma_{\bar{w}}). \end{aligned} \quad (\text{B.1})$$

Finding  $\hat{\mathbf{A}}^{\{q\}}$ , is equivalent to solving the following optimization problem,

$$\text{vec}(\hat{\mathbf{A}}^{\{q\}}) = \underset{\text{vec}(\mathbf{A}^{\{q\}})}{\text{argmax}} L(\mathbf{X} | \mathbf{A}^{\{q\}}; \mathcal{H}^{\{q\}}, \hat{\Sigma}_{\bar{w}}^{-1/2}). \quad (\text{B.2})$$

So,

$$\begin{aligned} \frac{\partial L(\mathbf{X} | \mathbf{A}^{\{q\}}; \mathcal{H}^{\{q\}}, \hat{\Sigma}_{\bar{w}}^{-1/2})}{\partial \text{vec}(\mathbf{A}^{\{q\}})} &= \hat{\Sigma}_{\bar{w}}^{-\frac{1}{2}} \text{vec}(\mathbf{X}) - \hat{\Sigma}_{\bar{w}}^{-\frac{1}{2}} (\Phi^{\{q\}^H} \otimes \mathbf{I}_L) \text{vec}(\mathbf{A}) = 0. \end{aligned} \quad (\text{B.3})$$

Take  $\hat{\Sigma}_{\bar{w}}^{-1/2} = \mathcal{F}_N(\mathcal{D})$  into (B.3), that is

$$\begin{aligned} \text{vec}(\hat{\mathbf{A}}^{\{q\}}) &= [(\Phi^{\{q\}^H} \otimes \mathbf{I}_L)^H \mathcal{F}_N(\mathcal{D})^H \mathcal{F}_N(\mathcal{D}) \\ &\quad \times (\Phi^{\{q\}^H} \otimes \mathbf{I}_L)]^{-1} (\Phi^{\{q\}^H} \otimes \mathbf{I}_L)^H \\ &\quad \times \mathcal{F}_N(\mathcal{D})^H \mathcal{F}_N(\mathcal{D}) \text{vec}(\mathbf{X}). \end{aligned} \quad (\text{B.4})$$

### Appendix C. Estimation of the Stationary Temporal Equalizer $\mathcal{T}(z)$

In this section, the line number  $i$  is ignored, and the subscript  $( )_{[k]}$  represents the parameters under the condition of order  $k$ , all coefficients are calculated by iteration.

The stationary temporal equalizer uses the Burg algorithm to calculate the AR model coefficients and then uses the AIC criterion for order optimization. The reflection coefficient  $\eta_{[k]}$  in the Burg algorithm can be iterated according to the data  $\mathbf{u}_i$ . Both of the  $k$ -order forward error  $e_{[k]}^f(n)$  and the inverse error

$e_{[k]}^b(n)$  can be obtained by iterating the reflection coefficient. At the same time, the reflection coefficient  $\eta_{[k]}$  can also be optimized according to the  $k - 1$  order forward error  $e_{[k-1]}^f(n)$  and the inverse error  $e_{[k-1]}^b(n)$ . The order  $\rho$  can be defined as the value at which the AIC criterion reaches the minimum between the lower bound  $\rho_{\min}$  and the upper bound  $\rho_{\max}$ .

After determining the optimal order  $\rho$ , the Levinson algorithm can be used to obtain all the AR coefficients  $t$  in the order range  $1, \dots, \rho$  as the stationary temporal equalizer coefficients.

The solution is as follows,<sup>51</sup>

Initial

$$e_{[0]}^f(n) = e_{[0]}^b(n) = u(n), t_{[0]}(0) = 1.$$

Do the iteration of  $\rho$ ,

$$\eta_{[k]} = \frac{-2 \sum_{n=k}^{N-1} e_{[k-1]}^f(n) e_{[k-1]}^{b*}(n-1)}{\sum_{n=k}^{N-1} |e_{[k-1]}^f(n)|^2 + \sum_{n=k}^{N-1} |e_{[k-1]}^b(n)|^2},$$

$$e_{[k]}^f(n) = e_{[k-1]}^f(n) + \eta_{[k]} e_{[k-1]}^f(n-1),$$

$$e_{[k]}^b(n) = e_{[k-1]}^b(n-1) + \eta_{[k]}^* e_{[k-1]}^f(n).$$

There is

$$E(k) = (1 - \|\eta_{[k]}\|^2) E(k-1),$$

and select

$$\begin{aligned} \rho &= \operatorname{argmin}_k \{N_L \ln[E(k)] + 2k\}, \\ \text{s.t. } \rho_{\min} &\leq \rho \leq \rho_{\max}. \end{aligned}$$

The  $\rho$ -order AR coefficients are obtained by Levinson recursion.

$$\begin{cases} t_{[k]}(i) = t_{[k-1]}(i) + \eta_{[k]} t_{[k-1]}(\rho-i) \\ \quad i = 1, 2, \dots, P-1 \\ t_{[\rho]}(\rho) = \eta_{[\rho]}. \end{cases}$$

## References

- A. Ortiz-Rosario, I. Berrios-Torres, H. Adeli and J. A. Buford, Combined corticospinal and reticulospinal effects on upper limb muscles, *Neurosci. Lett.* **561** (2014) 30–34.
- A. Ortiz-Rosario, H. Adeli and J. A. Buford, Wavelet methodology to improve single unit isolation in primary motor cortex cells, *J. Neurosci. Meth.* **246** (2015) 106–118.
- H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, A spatio-temporal wavelet-chaos methodology for EEG-based diagnosis of Alzheimer's disease, *Neurosci. Lett.* **444** (2008) 190–194.
- E. Giraldo-Suarez, J. D. Martinez-Vargas and G. Castellanos-Dominguez, Reconstruction of neural activity from EEG data using dynamic spatiotemporal constraints, *Int. J. Neural Syst.* **26** (2016) 1650026.
- M. F. Li, W. Li and H. H. Zhou, Increasing N200 potentials via visual stimulus depicting humanoid robot behavior, *Int. J. Neural Syst.* **26** (2016) 1550039.
- J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller and T. M. Vaughan, Brain-computer interfaces for communication and control, *Clin. Neurophysiol.* **113** (2002) 767–791.
- X. R. Gao, D. F. Xu, M. Cheng and S. K. Gao, A BCI-based environmental controller for the motion-disabled, *IEEE Trans. Neural Syst. Rehabil. Eng.* **11** (2003) 137–140.
- G. R. Muller-Putz, C. Pokorny, D. S. Klobassa and P. Horki, A single-switch bci based on passive and imagined movements: Toward restoring communication in minimally conscious patients, *Int. J. Neural Syst.* **23** (2013) 1250037.
- D. Zhang, B. S. Huang, W. Wu and S. L. Li, An idle-state detection algorithm for SSVEP-based brain-computer interfaces using a maximum evoked response spatial filter, *Int. J. Neural Syst.* **25** (2015) 1550030.
- A. Ortiz-Rosario and H. Adeli, Brain-computer interface technologies: From signal to action, *Rev. Neurosci.* **24** (2013) 537–552.
- A. Burns, H. Adeli and J. A. Buford, Brain-computer interface after nervous system injury, *Neuroscientist* **20** (2014) 639–651.
- I. Rembado, E. Castagnola, L. Turella, T. Ius, R. Budai, A. Ansaldi, G. N. Angotzi, F. Debertoldi, D. Ricci, M. Skrap and L. Fadiga, Independent component decomposition of human somatosensory evoked potentials recorded by micro-electrocorticography, *Int. J. Neural Syst.* **27** (2017) 1650052.
- D. G. Zhang, F. Xu, H. Xu, P. B. Shull and X. Y. Zhu, Quantifying different tactile sensations evoked by cutaneous electrical stimulation using electroencephalography features, *Int. J. Neural Syst.* **26** (2016) 1650006.
- S. W. Chuang, C. H. Chuang, Y. H. Yu, J. T. King and C. T. Lin, EEG alpha and gamma modulators mediate motion sickness-related spectral responses, *Int. J. Neural Syst.* **26** (2016) 1650007.
- E. W. Yin, T. Zeyl, R. Saab, D. W. Hu, Z. T. Zhou and T. Chau, An auditory-tactile visual saccade-independent P300 brain-computer interface, *Int. J. Neural Syst.* **26** (2016) 1650001.
- R. Liu, Y. X. Wang, G. I. Newman, N. V. Thakor and S. Ying, EEG classification with a sequential decision-making method in motor imagery BCI, *Int. J. Neural Syst.* **27** (2017) 16.

17. F. Z. Xu, W. D. Zhou, Y. L. Zhen, Q. Yuan and Q. Wu, Using fractal and local binary pattern features for classification of ECOG motor imagery tasks obtained from the right brain hemisphere, *Int. J. Neural Syst.* **26** (2016) 1650022.
18. A. R. Sereshkeh, R. Trott, A. Bricout and T. Chau, Online EEG classification of covert speech for brain-computer interfacing, *Int. J. Neural Syst.* **27** (2017) 1750033.
19. M. P. Xu, J. Liu, L. Chen, H. Z. Qi, F. He, P. Zhou, B. K. Wan and D. Ming, Incorporation of inter-subject information to improve the accuracy of subject-specific P300 classifiers, *Int. J. Neural Syst.* **26** (2016) 1650010.
20. A. M. Norcia, L. G. Appelbaum, J. M. Ales, B. R. Cottereau and B. Rossion, The steady-state visual evoked potential in vision research: A review, *J. Vis.* **15** (2015) 1–46.
21. S. Gao, Y. Wang, X. Gao and B. Hong, Visual and auditory brain-computer interfaces, *IEEE Trans. Biomed. Eng.* **61** (2014) 1436–1447.
22. B. Wittevrongel and M. M. Van Hulle, Frequency- and phase encoded SSVEP using spatiotemporal beamforming, *Plos One* **11** (2016) e159988.
23. 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, *Int. J. Neural Syst.* **24** (2014) 1450019.
24. N. V. Manyakov, N. Chumerin, A. Robben, A. Combaz, M. van Vliet and M. M. Van Hulle, Sampled sinusoidal stimulation profile and multichannel fuzzy logic classification for monitor-based phase-coded SSVEP brain-computer interfacing, *J. Neural Eng.* **10** (2013) 036011.
25. 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.* **6** (2009) 046002.
26. Y. Zhang, P. Xu, K. Cheng and D. Yao, Multivariate synchronization index for frequency recognition of SSVEP-based brain-computer interface, *J. Neurosci. Meth.* **221** (2014) 32–40.
27. A. Paris, A. Vosoughi and G. Atia, Whitening 1/f-type noise in electroencephalogram signals for steady-state visual evoked potential brain-computer interfaces, *2014 48th Asilomar Conf. Signals, Systems and Computers*, 2014, pp. 204–207.
28. B. J. He, Scale-free brain activity: Past, present, and future, *Trends. Cogn. Sci.* **18** (2014) 480–487.
29. G. Buzsáki, *Rhythms of the Brain* (Oxford University Press, Oxford, 2006), pp. 119–126.
30. O. Friman, I. Volosyak and A. Graser, Multiple channel detection of steady-state visual evoked potentials for brain-computer interfaces, *IEEE Trans. Biomed. Eng.* **54** (2007) 742–750.
31. M. Abu-Alqumsan and A. Peer, Advancing the detection of steady-state visual evoked potentials in brain-computer interfaces, *J. Neural Eng.* **13** (2016) 036005.
32. V. Hanh, B. Koo and S. Choi, Frequency detection for SSVEP-based BCI using deep canonical correlation analysis, in *2016 IEEE Int. Conf. Syst. Man Cybern. (SMC)*, 2016, pp. 001983–001987.
33. T. S. Rappaport, *Wireless Communications: Principles and Practice* (Prentice Hall, Upper Saddle River, N.J., 1996), pp. 299–324.
34. F. Qi, Y. Li and W. Wu, RSTFC: A novel algorithm for spatio-temporal filtering and classification of single-trial EEG, *IEEE Trans. Neural. Netw. Learn. Syst.* **26** (2015) 3070–3082.
35. X. Chen, Y. Wang, S. Gao, T. R. Jung and X. Gao, Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain-computer interface, *J. Neural Eng.* **12** (2015) 046008.
36. E. Yin, Z. Zhou, J. Jiang, Y. Yu and D. Hu, A dynamically optimized SSVEP brain-computer interface (BCI) speller, *IEEE Trans. Biomed. Eng.* **62** (2015) 1447–1456.
37. M. Nakanishi, W. Yijun, W. Yu-Te and J. Tzyy-Ping, A dynamic stopping method for improving performance of steady-state visual evoked potential based brain-computer interfaces, *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2015** (2015) 1057–1060.
38. J. N. da Cruz, F. Wan, C. M. Wong and T. Cao, Adaptive time-window length based on online performance measurement in SSVEP-based BCIs, *Neurocomput.* **149** (2015) 93–99.
39. Y. Wang, X. Chen, X. Gao and S. Gao, A benchmark dataset for SSVEP-based brain-computer interfaces, *IEEE Trans. Neural Syst. Rehabil. Eng.* **25** (2017) 1746–1752.
40. T. Limpiti, B. D. Van Veen, H. T. Attias and S. S. Nagarajan, A spatiotemporal framework for estimating trial-to-trial amplitude variation in event-related MEG/EEG, *IEEE Trans. Biomed. Eng.* **56** (2009) 633–645.
41. K. H. Britten, M. N. Shadlen, W. T. Newsome and J. A. Movshon, The analysis of visual-motion — A comparison of neuronal and psychophysical performance, *J. Neurosci.* **12** (1992) 4745–4765.
42. W. Wu, C. Wu, S. Gao, B. Liu, Y. Li and X. Gao, Bayesian estimation of ERP components from multicondition and multichannel EEG, *Neuroimage* **88** (2014) 319–339.
43. S. Baillet, J. C. Mosher and R. M. Leahy, Electromagnetic brain mapping, *IEEE Signal Process. Mag.* **18** (2001) 14–30.
44. S. Blanco, H. Garcia, R. Q. Quiroga, L. Romanelli and O. A. Rosso, Stationarity of the Eeg series, *IEEE Eng. Med. Biol. Mag.* **14** (1995) 395–399.
45. J. C. Jimenez, R. Biscay and O. Montoto, Modeling the electroencephalogram by means of spatial spline smoothing and temporal autoregression, *Biol. Cybern.* **72** (1995) 249–259.

46. C. W. Anderson, E. A. Stoltz and S. Shamsunder, Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks, *IEEE Trans. Biomed. Eng.* **45** (1998) 277–286.
47. Y. Inouye and R. W. Liu, A system-theoretic foundation for blind equalization of an FIR MIMO channel system, *IEEE Trans. Circuits and Syst. I Fundam. Theory Appl.* **49** (2002) 425–436.
48. T. A. Schonhoff and A. A. Giordano, *Detection and Estimation Theory and Its Applications* (Prentice Hall, 2006), pp. 229–234.
49. C. W. Baum and V. V. Veeravalli, A sequential procedure for multihypothesis testing, *IEEE Trans. Inf. Theory* **40** (1994) 1994–2007.
50. S. M. Kay, Fundamentals of statistical signal processing, in *Prentice Hall Signal Processing Series* (Prentice-Hall PTR, Englewood Cliffs, N.J., 1998), pp. 490–500.
51. J. P. Burg, Maximum entropy spectral analysis, Ph.D. thesis, Stanford University, Ann Arbor (1975).
52. D. H. Brainard, The psychophysics toolbox, *Spat Vis* **10** (1997) 433–436.
53. X. Chen, Z. Chen, S. Gao and X. Gao, A high-ITR SSVEP-based BCI speller, *Brain-Comput. Interf.* **1** (2014) 181–191.
54. F. Di Russo and D. Spinelli, Electrophysiological evidence for an early attentional mechanism in visual processing in humans, *Vision Res.* **39** (1999) 2975–2985.
55. B. Wittevrongel, E. Khachatryan, M. Fahimi Hnazaee, E. Carrette, L. De Taeye, A. Meurs, P. Boon, D. Van Roost and M. M. Van Hulle M. M., Representation of steady-state visual evoked potentials elicited by luminance flicker in human occipital cortex: An electrocorticography study, *NeuroImage* **175** (2018) 315–326.
56. X. Chen, Y. Wang, M. Nakanishi, X. Gao, T. P. Jung and S. Gao, High-speed spelling with a non-invasive brain-computer interface, *Proc. Natl. Acad. Sci. USA* **112** (2015) E6058–E6067.
57. P. Yuan, X. Gao, B. Allison, Y. Wang, G. Bin and S. Gao, A study of the existing problems of estimating the information transfer rate in online brain-computer interfaces, *J. Neural Eng.* **10** (2013) 026014.
58. A. Kubler, N. Neumann, J. Kaiser, B. Kotchoubey, T. Hinterberger and N. P. Birbaumer, Brain-computer communication: Self-regulation of slow cortical potentials for verbal communication, *Arch. Phys. Med. Rehabil.* **82** (2001) 1533–1539.
59. G. Townsend, B. K. LaPallo, C. B. Boulay, D. J. Krusienski, G. E. Frye, C. K. Hauser, N. E. Schwartz, T. M. Vaughan, J. R. Wolpaw and E. W. Sellers, A novel P300-based brain-computer interface stimulus presentation paradigm: Moving beyond rows and columns, *Clin. Neurophysiol.* **121** (2010) 1109–1120.
60. B. Allison, T. Luth, D. Valbuena, A. Teymourian, I. Volosyak and A. Graser, BCI demographics: How many (and what kinds of) people can use an SSVEP BCI? *IEEE Trans. Neural Syst. Rehabil. Eng.* **18** (2010) 107–116.
61. Y. Wang, R. P. Wang, X. R. Gao, B. Hong and S. K. Gao, A practical VEP-based brain-computer interface, *IEEE Trans. Neural Syst. Rehabil. Eng.* **14** (2006) 234–239.
62. 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 Trans. Biomed. Eng.* **65** (2018) 104–112.