Frequency Domain Filtering Method for SSVEP-EEG Preprocessing

Wenqiang Yan, Bo He, Jin Zhao, Yongcheng Wu, Chenghang Du, and Guanghua Xu, *Member, IEEE*

***Abstract*— Steady-state visual evoked potential (SSVEP) signal collected from the scalp typically contains other types of electric signals, and it is important to remove these noise components from the actual signal by application of a pre-processing step for accurate analysis. High-pass or bandpass filtering of the SSVEP signal in the time domain is the most common pre-processing method. Because frequency is the most important feature information contained in the SSVEP signal, a technique for frequency-domain filtering of SSVEP was proposed here. In this method, the time-domain signal is extended to multi-dimensional signal by empirical mode decomposition (EMD), where each dimension represents a intrinsic mode function (IMF). The multi-dimensional signal is transformed to a frequency-domain signal by 2-D Fourier transform, the Gaussian high-pass filter function is constructed to perform high-pass filtering, and then the filtered signal is transformed to time domain by 2-D inverse Fourier transform. Finally, the filtered multi-dimensional intrinsic mode function is superimposed and averaged as the frequency-domain filtered signal. Compared with the time-domain filtering method, the experimental results revealed that frequency-domain filtering method effectively removed the baseline drift in signal and effectively suppressed the low-frequency interference component. This method was tested using data from public datasets and the results show that the proposed frequency-domain filtering method can significantly improve the feature recognition performance of canonical correlation analysis (CCA), filter bank canonical correlation analysis (FBCCA), and task-related component analysis (TRCA) methods. Thus, the results suggest that the application of frequency-domain filtering in the pre-processing stage allows improved noise removal. The proposed method extends SSVEP signal filtering from time-domain to frequency-domain, and the results suggest that this analysis tool significantly promotes the practical application of SSVEP systems.**

Manuscript received 28 September 2022; revised 4 April 2023; accepted 7 April 2023. Date of publication 17 April 2023; date of current version 26 April 2023. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 52105308, in part by the Fundamental Research Funds for the Central Universities under Grant xhj032021010-02, and in part by the Scientific and Technological Innovation 2030 Key Project of Ministry of Science and Technology of China under Grant 2022ZD0209800. *(Corresponding author: Wenqiang Yan.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Research Ethics Committee of Tsinghua University.

Wenqiang Yan and Guanghua Xu are with the State Key Laboratory for Manufacturing Systems Engineering, School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: yanwenq@xjtu.edu.cn).

Bo He, Jin Zhao, Yongcheng Wu, and Chenghang Du are with the School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China.

Digital Object Identifier 10.1109/TNSRE.2023.3266488

***Index Terms*— Brain–computer interface (BCI), steady-state visual evoked potential (SSVEP), pre-processing, empirical mode decomposition (EMD), frequency-domain filtering.**

I. Introduction

BRAIN-COMPUTER interface (BCI) technology constructs a direct information transmission pathway between the human brain and the external world by decoding information about the neural activity of the brain during human thought processes [1], [2], [3], [4], [5]. The steady-state visual evoked potential (SSVEP) is one of the main neural signals used in BCI systems because of its high information transmission rate and large number of encodable targets. When the human eye is exposed to a flickering or changing pattern of visual stimuli at a fixed frequency, the cortical potential activity is modulated, resulting in a continuous stimulus frequency-related response with a periodic rhythm that is similar to that of the visual stimulus, known as SSVEP [6]. SSVEP signal generation is thought to arise from inherent resonance frequency of the various neural networks distributed in the brain, as under normal conditions, these neural networks are not synchronized with each other and generate spontaneous electroencephalogram (EEG). With application of an external visual stimulus of constant frequency, the neural network, which coincides with the stimulus frequency or harmonic frequency, resonates, resulting in a significant change in the brain potential activity at the stimulus and harmonic frequencies and resulting in the SSVEP signal [7]. The SSVEP signal in power spectra exhibits a significant spectral peak at the stimulus and harmonic frequencies. Detection of the frequency at this spectral peak enables the identification of the stimulus source of the subject's visual gaze and inference of the subject's intention [8], [9], [10]. However, scalp-acquired SSVEP signals are susceptible to artifacts such as electrical signals from muscles and resting electrical potential that exists between the front and the back of the eye. Thus, it is important to develop an efficient focused target classification algorithm for the practical implementation of the SSVEP-BCI system.

Canonical correlation analysis (CCA) [11] is a classical method used for SSVEP feature recognition. In this method, the maximum correlation coefficient between the test signal and the reference signal is determined by solving a spatial filter. Filter bank canonical correlation analysis (FBCCA) [12] is an improved method based on CCA, in which the testing signal

is decomposed into several sub-band signals by using multiple filter banks. CCA analysis is then performed on each sub-band signal, and the obtained correlation coefficient ensembles are used as the final feature discriminant coefficients. There can be significant individual variations in EEG signals. Incorporating user training data can specifically characterize users' SSVEP responses for significantly improved detection performance. Extended canonical correlation analysis (eCCA) [13] and task-related component analysis (TRCA) [14] are the representatives of SSVEP training feature recognition algorithm. Based on the CCA algorithm, eCCA performs spatial filtering of multi-channel brain signals by solving the spatial filter between the training signal, testing signal, and the reference signal, composed of sine and cosine functions. Then, the correlation between signals after spatial filtering is determined to identify the user focused target. TRCA is a spatial filtering approach for time-locked EEG signals, which finds the spatial filter by maximizing the covariance of the time-locked signals across different trials. Recently, Riemannian geometry was also used for SSVEP feature recognition [15], [16], [17]. This classifier is not designed to estimate a spatial filter, but to map the data directly to a geometric space with a suitable metric. The adaptive classifier is another promising method for SSVEP classification [18], [19], [20]. In this method, the parameters of the classifier are re-estimated and updated as new EEG data are obtained, thus allowing the classifier to track the distribution of features that may change in brain signals.

Generally, methods that incorporate training can achieve higher SSVEP feature recognition accuracy than training-free methods. However, it can be very time-consuming to collect sufficient training data. Additionally, a user may experience visual fatigue caused by the lengthy data calibration process and this will seriously affect the signal-to-noise ratio of SSVEP responses and the overall performance of the BCI system. As an alternative, there is growing interest in the development of transfer learning techniques for SSVEP feature recognition. The transfer template-based canonical correlation analysis (tt-CCA) method assumes that different subjects can share one universal SSVEP template, and the grand average of all existing subjects' SSVEP data is assumed as the new subject's SSVEP template [21]. The least-squares transformation (LST)-based transfer learning method uses least-square transformation to reduce the difference between the source domain (existing subjects) and the target domain (a new subject), and then reliable spatial filters could be learned from all the SSVEP data [22]. The subject transfer-based canonical correlation analysis (stCCA) method assumes a common pattern in a subject's spatial filters across different stimulus frequencies, and a subject's spatially filtered SSVEP template could be approximated by the weighted summation of spatially filtered SSVEP templates from existing subjects [23]. The transfer learning canonical correlation analysis (tlCCA) method transfers a subject's SSVEP templates and spatial filters from the source frequency to the target frequency [24]. The cross-subject spatial filter transfer (CSSFT) method transfers the existing user model with good SSVEP response to the new user test data without collecting any training data from the new user [25], [26]. In addition to the above feature recognition methods based on a transfer learning framework, deep learning techniques have also been applied for EEG feature classification [27], [28], [29], [30]. The most studied deep learning networks in BCI are convolutional neural network (CNN) and deep belief network (DBN), and a combined CNN-DBN classification network has also been studied.

Techniques such as matrix classifier (spatial filter), adaptive classifier, transfer learning, and deep learning have greatly improved the classification recognition accuracy of SSVEP. These algorithms usually include steps of pre-processing, feature extraction, and feature classification. As the initial part of feature recognition, pre-processing has an important impact on the final classification accuracy. The stimulus frequency range of SSVEP is usually set above 7 Hz, therefore, filtering the low-frequency components of signals is a key step in SSVEP pre-processing. Most current studies use digital filters such as Butterworth and Chebyshev for time-domain filtering of brain signals. However, since the frequency spectrum of noise significantly overlaps with that of brain signals, these filters unfortunately suppress the brain signals of interest during the suppression of the noise. Principal component analysis (PCA) [31] and independent component analysis (ICA) [32] are two classical EEG denoising methods. PCA removes noise and redundancy by reducing the dimension of multidimensional data. The signals obtained after PCA processing are not correlated. It is worth noting that the SSVEP signal is not unrelated to noise, so PCA often fails to obtain good SSVEP feature recognition performance. ICA decomposes multi-channel observation signals into several independent components through an optimization algorithm that assumes the statistical independence of source signals. However, the independent components obtained by ICA have no specific arrangement, and the SSVEP component cannot be directly determined from the decomposed independent components. Wavelet threshold method is also applied as a EEG signal denoising method. In recent years, researchers have also proposed a hybrid strategy combining signal decomposition and blind source separation (BSS) to eliminate EEG noise [33], [34]. This strategy uses wavelet transform or ensemble empirical mode decomposition (EEMD) to decompose single-channel signals into multi-dimensional signal components, then BSS techniques (e.g. ICA) are used to further decompose those generated multi-dimensional signal components into meaningful sources.

In this study, we investigated the feasibility of frequency-domain filtering of SSVEP since the main characteristic of SSVEP signal is frequency. A frequency-domain filtering (FDF) method based on empirical mode decomposition (EMD) and 2-D Fourier transform (FFT) filtering is proposed. To do this, the time-domain signal is first expanded into a multi-dimensional signal through EMD and the multi-dimensional brain signal is transformed into frequency-domain by 2-D FFT. Next, the 2-D spectral signal is filtered by high-pass filter in the frequency-domain and then 2-D inverse FFT is used to obtain the frequency-domain filtered multi-dimensional brain

signal. Finally, multi-dimensional brain signals are superimposed and averaged as final frequency domain filtered signals. In this work, we compared the SSVEP feature recognition performance after time-domain and frequency-domain filtering of data from public datasets, and verified the effectiveness of using frequency-domain filtering in the signal pre-processing stage to improve SSVEP classification accuracy. To the best of our knowledge, it is the first time, in the literature, that the feasibility of SSVEP signal filtering in frequency domain is investigated.

The organization of this paper is as follows. In Section Methods, the data and methods used in this paper are introduced. Section Results analyzes the influence of parameters on FDF method and the feature recognition performances of FDF method on SSVEP signals, followed by discussions and suggestions for future work in Section Discussion. Finally, we conclude the work in Section Conclusions.

II. Methods

A. Dataset

The public dataset [35] used in this study includes SSVEP recordings of 35 healthy subjects (the data file of subject 5 in the Benchmark Dataset was damaged and could not be downloaded) focusing on 40 characters flickering at different frequencies (8-15.8 Hz with an interval of 0.2 Hz). For each subject, the experiment consisted of six blocks, where each block contained 40 trials corresponding to all 40 characters presented in a random order. The length of each trial is 5 s, and the sampling frequency of the data is 250 Hz. The SSVEP signal analysis channels selected in this study were O1, O2, Oz, PO3, PO4, POz, PO5, and PO6.

B. Canonical Correlation Analysis (CCA)

For EEG data X recorded from multiple channels and the reference signal Y, CCA analysis aims to find two projection vectors w_x and w_y so that the linear combination signals $w_x^{\mathrm{T}}X$ and $w_y^{\mathrm{T}}Y$ have the largest correlation. The correlation coefficient obtained using CCA is:

$$\rho = max \frac{E(w_x^T X Y^T w_y)}{\sqrt{E(w_x^T X X^T w_x) E(w_y^T Y Y^T w_y)}} \tag{1}$$

The reference signal Y were constructed at the stimulation frequency f:

$$Y_f = \left\{ \begin{array}{c} cos2\pi ft \\ sin2\pi ft \\ \cdots \\ cos2k\pi ft \\ sin2k\pi ft \end{array} \right\}, t = 1/f_s, \ldots, N_s/f_s \tag{2}$$

where k is the number of harmonics, f_s is the sampling rate, and Ns represents the number of sample points. By calculating the correlation coefficients between X and the reference signals at all stimulus frequencies, the corresponding target with the maximum correlation coefficient is identified as the focused target.

C. Filter Bank Canonical Correlation Analysis (FBCCA)

FBCCA divides EEG signal X into n sub-band signals $(X_1, X_2, \ldots, X_n)$ through filter banks and then calculates the correlation coefficient ρ_i between each sub-band signal X_i and the reference signal using CCA. The final discrimination coefficient is determined by integration of n correlation coefficients by:

$$\tilde{\rho} = \sum_{i=1}^{n} w(i) \cdot (\rho_i)^2 \tag{3}$$

where $w(i)$ refers to the weight corresponding to the correlation coefficient of the i-th sub-band signal. This can be calculated by:

$$w(i) = i^{-a} + b, i \in [1n] \tag{4}$$

Herein, a and b are 1.25 and 0.25, respectively. After obtaining ensemble coefficients at all stimulus frequencies, FBCCA considers the target corresponding to the maximum coefficient as the focused target.

D. Task-Related Component Analysis (TRCA)

The subject's training data is denoted as $Z \in \mathrm{R}^{Ns \times Nc \times Nf \times Nt}$, where N_s represents the number of sampling points, N_c represents the number of channels, N_f represents the number of stimulation frequencies, and N_t represents the number of blocks. TRCA extracts task-related components by spatially filtering the training data. The spatial filter $w_f \in \mathrm{R}^{Nc \times 1}$ at stimulation frequency f can be calculated by:

$$\underset{w_f}{argmax} \frac{w_f^T A^T A w_f}{w_f^T B^T B w_f} \tag{5}$$

where $A \in \mathrm{R}^{Ns \times Nc}$ represents the result of averaging N_t blocks data at the frequency f in Z:

$$A = \frac{1}{N_t} \sum_{i=1}^{N_t} Z_{i,f} \tag{6}$$

$B = [Z_{1,f}^T Z_{2,f}^T \; Z_{3,f}^T \ldots Z_{Nt,f}^T]^T \in R^{Nt \cdot Ns \times Nc}$, and $Z_{i,f} \in \mathrm{R}^{Ns \times Nc}$ represents the EEG signal with stimulation frequency f in the i-th block. After calculating the spatial filter w_f at the frequency f, TRCA uses the Pearson correlation coefficient between the spatial filtered training signal and the spatial filtered test signal as the final discriminant coefficient:

$$p' = corr(X w_f, A w_f) \tag{7}$$

After obtaining the discrimination coefficients at all stimulus frequencies, the target corresponds to the maximum coefficient as the focused target.

E. Frequency Domain Filtering Method

1) Empirical Mode Decomposition (EMD): EMD can decompose a signal into a superposition of several intrinsic mode functions (IMFs) based on the signal's own feature scale. The IMF needs to meet two conditions: 1) The number of extreme value points and the number of crossing zero points in the whole data set must be equal or differ by at most one; 2) The mean value of the upper and lower envelopes formed by the local extreme and local minimal points is zero at all times,

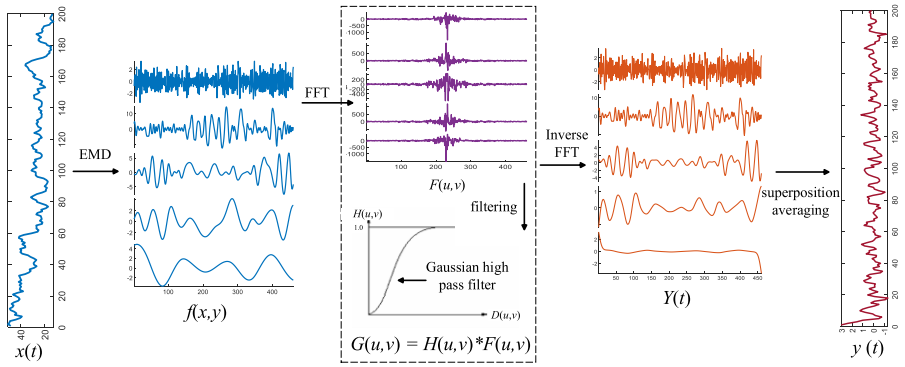

Fig. 1. The algorithm flow of the frequency filtering method.

i.e., the upper and lower envelopes are symmetric about the time axis. Assuming that the signal to-be-decomposed is $x(t)$, the detailed steps of the EMD procedure are as follows:

a) Determine all the local extreme points (including extreme and minimal points) of the original signals $x(t)$, and fit all the extreme points with the cubic spline algorithm to obtain the extreme value envelope $e_{max}(t)$ (upper envelope) and the minimal value envelope $e_{min}(t)$ (lower envelope), so that all the data points of signal $x(t)$ are between these two envelopes. The mean value of the upper and lower envelopes is used as the mean envelope of the original signals $m(t)$:

$$m(t) = \frac{e_{max}(t) + e_{min}(t)}{2} \tag{8}$$

b) Subtract the mean values of the upper and lower envelopes from the signal-to-be decomposed $x(t)$:

$$h_1(t) = x(t) - m(t) \tag{9}$$

Determine whether $h_1(t)$ satisfies the two qualifying conditions of the natural mode function, if not, then $h_1(t)$ is signal to-be-decomposed. Where, $h_1(t)$ is used instead of $x(t)$, and then the above steps are repeated until $h_1(t)$ satisfies the qualifying conditions of the natural mode function. This is written as:

$$c_1(t) = h_1(t) \tag{10}$$

c) $c_1(t)$ is the first intrinsic mode function (or first-order IMF) obtained from the EMD, and the residual sequence (or residual component) $r_1(t)$ is obtained by subtracting $c_1(t)$ from the original signals $x(t)$:

$$r_1(t) = x(t) - c_1(t) \tag{11}$$

d) Repeat steps a) - c) for $r_1(t)$ as the new signal to-be-decomposed, thus obtaining the second, third and up to n-th IMF, denoted as $c_1(t)$, $c_2(t)$, ..., $c_n(t)$. This process is called "screening", and the screening process is repeated until satisfying a predefined stopping criterion. Here, we used the energy ratio as the criterion for stopping EMD decomposition:

$$Energy\ Ratio = 10log10(\frac{\|x(t)\|_2}{\|r_i(t)\|_2}) \tag{12}$$

EMD decomposition is terminated when the energy ratio calculated in the i-th decomposition is greater than 20 or the number of decomposed natural mode functions is greater than 10. After decomposition steps a) - d), the original signals $x(t)$ are decomposed into the sum of several IMF components and one residual component:

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \tag{13}$$

2) Frequency-Domain Filtering Method Based on EMD and 2-D FFT: Fig. 1 illustrates the proposed frequency-domain filtering method algorithm procedure. First, the IMFs obtained by the EMD of the time-domain signal $x(t)$ are constructed as a 2-D function $f(x, y) = [c_1(t)c_2(t)\ldots c_n(t)]$:

$$f(x, y), 0 \le x \le M-1, 0 \le y \le N-1 \tag{14}$$

where M and N represent the dimensions of $f(x, y)$. The 2-D FFT is then applied to $f(x, y)$:

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(\frac{xu}{M} + \frac{yv}{N})} \tag{15}$$

The goal of the pre-processing of SSVEP signal is to remove the low-frequency components in the signal. In this step, the Gaussian high-pass filter is used to filter $F(u, v)$, and the corresponding Gaussian filter function can be expressed as:

$$H(u, v) = 1 - e^{-D^2(u,v)/(2 \cdot D_0^2)} \tag{16}$$

where $D(u, v)$ denotes the distance from point (u, v) to the center of the spectrum center $(M/2, N/2)$:

$$D(u, v) = \sqrt{(u - M/2)^2 + (v - N/2)^2} \tag{17}$$

where D_0 represents the passband radius, i.e., signal components larger than D_0 range are retained. The spectrum function $F(u, v)$ is multiplied by the filter function $H(u, v)$ for high-pass filtering:

$$G(u, v) = H(u, v) * F(u, v) \tag{18}$$

Next, a 2-D FFT inverse transform is performed on $G(u, v)$:

$$g(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} G(u, v) e^{j2\pi(\frac{xu}{M} + \frac{yu}{N})} \tag{19}$$

The real part of $g(x, y) \in \mathrm{R}^{M \times N}$ is as the frequency-domain high-pass filtered multi-dimensional brain signal, and is denoted as $Y(t) = real\ (g(x,y)) = [c_1'(t)\ c_2'(t)\ldots c_n'(t)]$, where $c_i'(t)$ represents the frequency-domain filtered intrinsic mode function. The filtered intrinsic mode functions can then

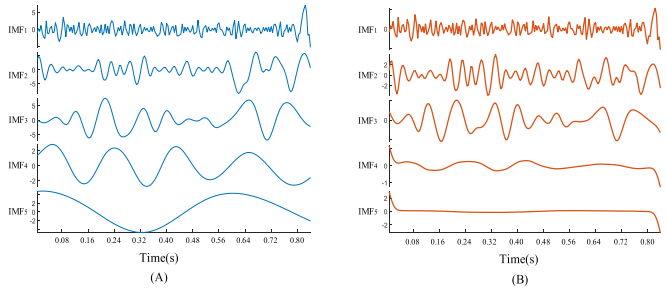

Fig. 2. Fig 2-A. The multi-dimensional signal decomposed by EMD. Fig 2-B. The multi-dimensional signal filtered by 2-D FFT.

be superimposed and averaged to obtain the final frequency-domain filtered signal:

$$y(t) = \frac{1}{n}\sum_{i=1}^{n} c_i'(t) \tag{20}$$

III. Results

A. Analysis of Filtering Performance of Frequency-Domain Filter on SSVEP Signal

The proposed frequency-domain filtering method first expands the time-domain signal into a multi-dimensional signal using EMD, and then performs frequency-domain transform filtering on the multi-dimensional signal, as shown in Fig. 1. Fig. 2-A shows the multi-dimensional signal obtained from the PO_4 channel signal of subject 4 by EMD, where each dimension corresponds to a intrinsic mode function. Fig. 2-B shows the multi-dimensional signal filtered by 2-D FFT, using a passband radius of Gaussian high-pass filter set to 10. According to the multi-dimensional signal before and after frequency-domain filtering, the waveform of low-order intrinsic mode function becomes smoother after filtering (e.g., IMF_4 and IMF_5). The frequency-domain filter suppresses the signal components below the passband range for each intrinsic mode function. The low-order intrinsic mode function mainly contains low-frequency components in the signal, so the high-pass filtering of the frequency-domain filter is mainly required for the low-order intrinsic mode function. In the method developed here, the superposition average of all intrinsic mode functions in frequency-domain filtered multi-dimensional signal is calculated as the frequency-domain filtered time-domain signal.

Fig. 3 A-C shows the original signals, time-domain filtered signals and frequency-domain filtered signals of the eight channels of subject 4. As shown in Fig. 3-B, each channel was subjected to Chebyshev type I filtering with a bandpass range of 6 - 90 Hz, and Fig. 3-C shows the proposed frequency-domain filtering for each electrode channel. According to Fig. 3-A, the waveform of the original signals exhibits a large fluctuation that deviates from the baseline. This fluctuation is caused by baseline drift caused by interference such as muscle movement and the presence of baseline drift will seriously affect the accuracy of the signal. Therefore, it is necessary to carry out high-pass or bandpass filtering on the SSVEP signal in the preprocessing stage. Fig. 3-B shows that after time-domain filtering of the SSVEP signal, the waveform of the signal is uniformly distributed along the baseline, effectively eliminating trend term seen in the original signals. As shown in Fig. 3-B and 3-C, the frequency-domain filtered signal waveform exhibits smaller fluctuations than the time-domain filtered signal waveform, such as the channels PO_4, PO_6, and Oz. The variance was calculated for the original, time-domain filtered and frequency-domain filtered signal waveforms for each of the eight channels, and the results are shown in Fig. 4. As observed, the frequency-domain filtered signal exhibited the smallest variance, indicating that the frequency-domain filtered signal waveform has a more uniform distribution along the baseline, indicating that the applied frequency-domain filtering method effectively eliminated the baseline drift of SSVEP signals. EMD can decompose the time domain signal into the sum of several intrinsic mode functions and one residual component, which can be considered the trend term of the signal (e.g., IMF_5 in Fig. 2-A). As shown in Fig. 2-B, the frequency-domain filtering method can smooth the trend terms in the signal. Because the proposed frequency-domain filtering method first determines the trend term of the signal through EMD, the frequency-domain filter directly eliminates the signal trend term for an improved baseline drift removal effect relative to that achieved by the time-domain filter.

Fig. 5 A-C shows the results of power spectrum analysis of original signals, time-domain filtered signals, and frequency-domain filtered signals, with red circles indicating the stimulus frequencies and their harmonics. As shown, there is more low-frequency interference in the original signals, and the amplitude of the low-frequency components is much higher than the amplitude at the stimulation frequency and its harmonics, which will seriously impede accurate analysis of the SSVEP signal. As shown in Fig. 5-B, after time-domain filtering, the low-frequency components in the time domain filtered signal spectrum are suppressed, allowing the accurate detection of the main frequency components of the SSVEP signal. Fig. 5-C shows the frequency-domain filtered signal power spectra, with the effective removal of the low-frequency components of the signal. The spectral data after time-domain and frequency-domain filtering show that the frequency-domain filtering method exhibits a more significant suppression effect on the signal components at non-stimulated frequencies (e.g., 5 - 10 Hz), making the stimulated frequency components of SSVEP more prominent.

Overall, frequency-domain filtering can more effectively remove baseline drift from the signal than time-domain filtering, with a more significant suppression effect on non-stimulated frequency components.

B. Effect of Frequency-Domain Filter Passband Radius on SSVEP Feature Recognition Performance

After transformation to the frequency domain, the signal must be filtered using a Gaussian high-pass filter, and the passband radius of the filter affects filtering performances. The time-domain filter can set a clear filtering range, but the proposed frequency-domain filtering method requires setting the passband radius to determine the filtering range, i.e., the passband radius does not directly correspond to the filtering frequency. This section analyzes the filtering performance of

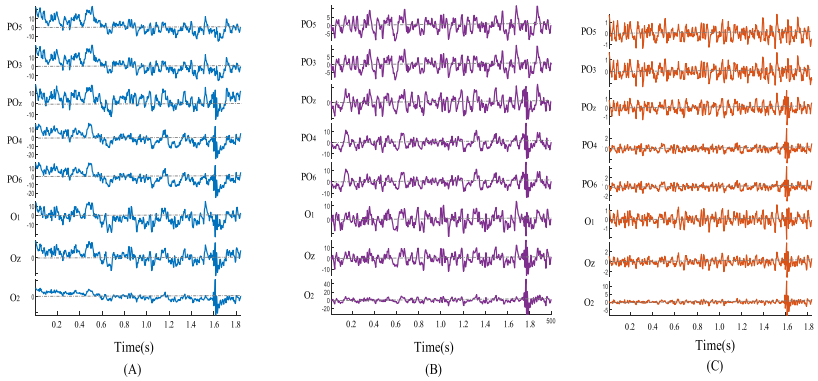

Fig. 3. Fig 3-A. The original signals of the eight channels. Fig 3-B. The time-domain filtered signals of the eight channels. Fig 3-C. The frequency-domain filtered signals of the eight channels.

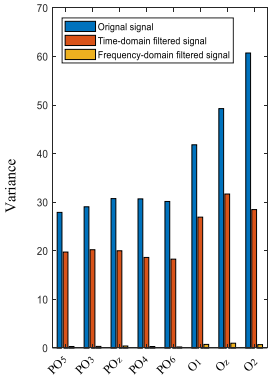

Fig. 4. The variance for the original, time-domain filtered and frequency-domain filtered signals for each of the eight channels.

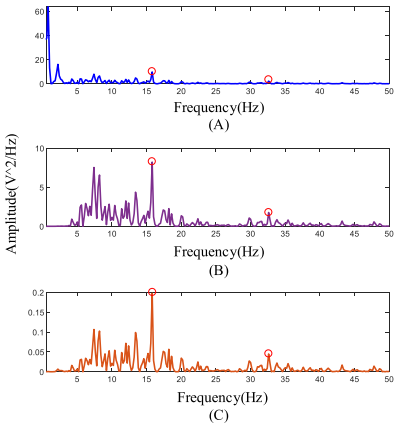

Fig. 5. Fig 5-A. The power spectrum of original signals. Fig 5-B. The power spectrum of time-domain filtered signals. Fig 5-C. The power spectrum of frequency-domain filtered signals.

the frequency-domain filter with different passband radius values for improved selection of the appropriate passband radius parameter.

Fig. 6 shows the power spectra obtained from the PO5 channel of subject 4 for passband radius of 5 - 50, in increments of 5, and the stimulus frequency and its harmonic frequency are indicated by red circles. With increasing passband radius, the low-frequency component of the signal is gradually eliminated, but the stimulus frequency of the signal is also suppressed. As the passband radius increases, the amplitude of the fundamental frequency also decreases. The stimulus frequency range of the focused target in the dataset used in this study was 8 - 15.8 Hz, so the passband radius should be appropriately set to avoid filtering out signal components with frequency above 8 Hz. As shown in Fig. 6, the use of a passband radius of 5 did not remove low-frequency interference below 2 Hz. However, the use of a passband radius of 20 suppressed signal components with frequency above 8 Hz, indicating that the appropriate passband range is in the range of 5 - 20. However, it is difficult to establish the relationship between passband radius and filtering frequency by spectrum analysis, as it is impossible to determine the corresponding filtering frequency for a passband radius of 10 or other values. For this reason, the SSVEP signal feature recognition accuracy was calculated for different passband radius values. Fig. 7 shows the recognition accuracy calculated for each subject using CCA, for 1 s of data analysis. The vertical coordinate of Fig. 7 represents the subject number, the horizontal coordinate represents the passband radius, and the color value represents the accuracy. As observed, with increasing passband radius, the accuracy tends to increase and then decrease. This is because the low-frequency interference in signal cannot be removed if the passband radius is too small, but the effective component of the signal can be removed if the passband radius is too large. Optimized recognition accuracy for the data of most subjects occurs using a passband radius of 10, so the optimized passband radius was set to 10.

C. Improvement Effect of Frequency Domain Filtering Method on SSVEP Recognition Accuracy

The SSVEP recognition accuracies obtained using time-domain filtering (TDF), wavelet threshold method and frequency-domain filtering (FDF) in the pre-processing stage were compared using data from the public datasets. The average recognition accuracy was calculated for each subject by superimposing and averaging the accuracy of six blocks of data, and then the average recognition accuracies of the 34 subjects were superimposed and averaged as the final analysis results. Paired-t test was used to determine significant differences (define as $p<0.05$) in accuracy for different methods.

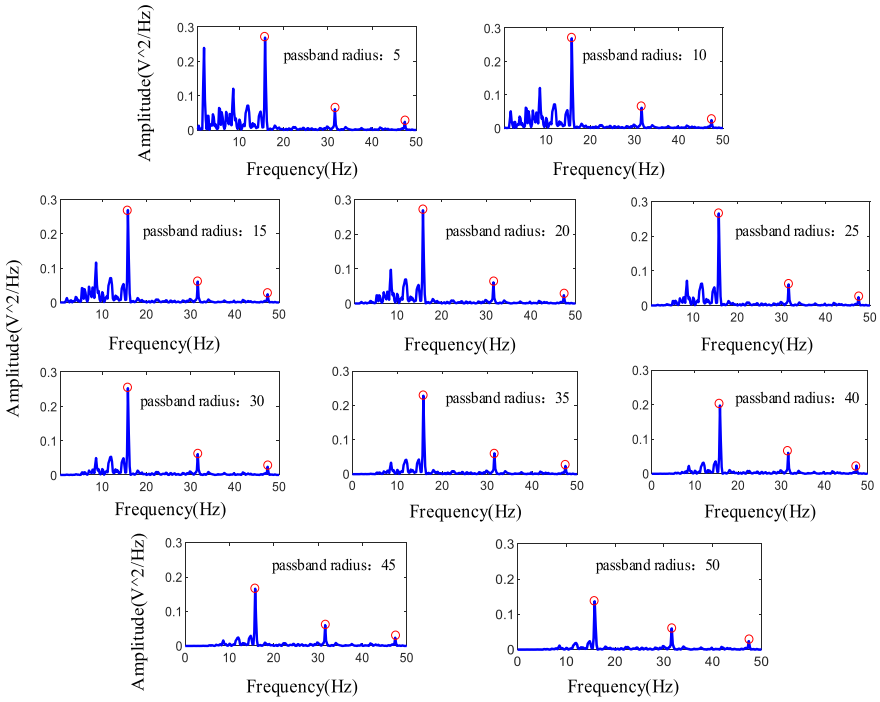

Fig. 6. The power spectra obtained from the PO_5 channel of subject 4 for passband radius of 5 - 50.

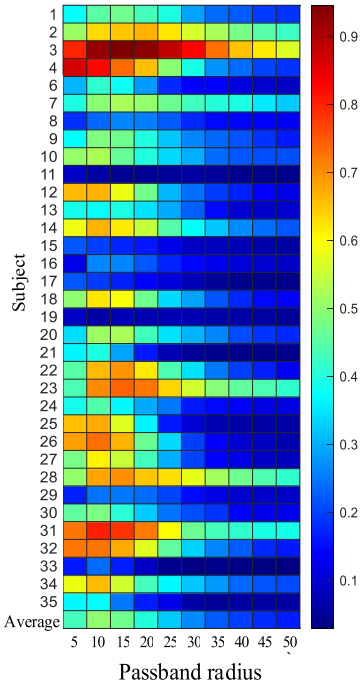

Fig. 7. The accuracy calculated for each subject with different passband radius.

Fig. 8 shows the accuracies of time-domain filtered signals, wavelet filtered signals and frequency-domain filtered signals obtained by the CCA method, with data analysis lengths of 0.8 s, 1 s, and 1.2 s. The number of harmonic of the reference signal in the CCA method was set to 5. For the time-domain filtering method, the SSVEP signals were subjected to Chebyshev type I filtering with a band-pass range of 6 - 90 Hz. For wavelet threshold method, the signal denoising is implemented by Matlab library functions *ddencmp* and *wdencmp*. For the frequency-domain filtering method, the Gaussian high-pass filter passband radius was set to 10. Fig. 8 shows that compared with the time-domain filtering method, the recognition accuracy of the frequency-domain filtering method was improved by 9.76%, 11.7%, and 11.35% for stimulus durations of 0.8 s, 1 s, and 1.2 s, respectively, and the recognition accuracy of the frequency-domain filtering method was improved by 8.52%, 11.25%, and 9.97%, compared with the wavelet threshold method, respectively. Paired t-tests showed significant differences between the accuracies obtained using frequency-domain filtered signals, wavelet filtered signals and time-domain filtered signals ($^{*}p$ <0.05, $^{**}p$ <0.01, $^{***}p$ <0.001), suggesting that the use of frequency filtering in the preprocessing stage improves SSVEP feature recognition effect relative to use of time-domain filtering and wavelet filtering.

To analyze the reasons for the excellent performance of frequency domain filtering in more detail, the first block data of subject 3 was selected and the classification results obtained by time domain filtering and frequency domain filtering methods are shown in Fig. 9 A-B. Each subplot in Fig. 9 A-B represents the classification result of a single trial, where the vertical coordinate represents the canonical correlation coefficient (obtained by CCA using Eq. 1) and the horizontal coordinate represents the stimulus target number. The canonical correlation coefficient corresponding to the subject true focused target is indicated with a red dot. As shown in Fig. 9-A, the correlation coefficients obtained using time-domain filtered signals did not differ much between focused and non-focused targets. For example, the canonical correlation coefficients corresponding to the focused target (marked with red dot in the figure) and the non-focused targets

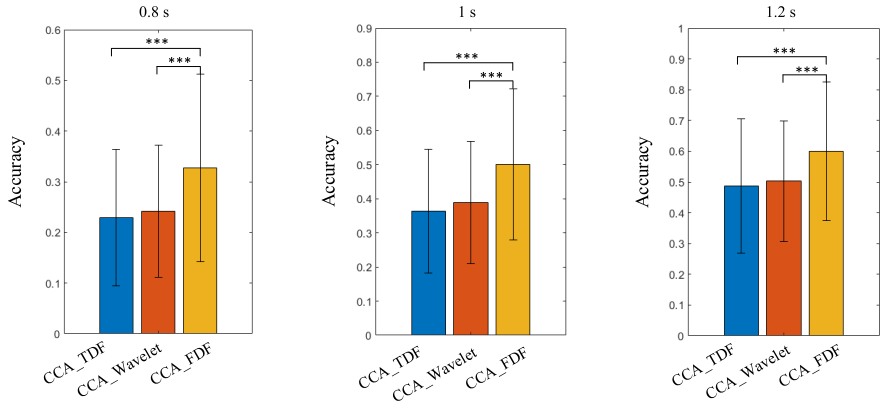

Fig. 8. The accuracy of time-domain filtered signals and frequency-domain filtered signals obtained by the CCA.

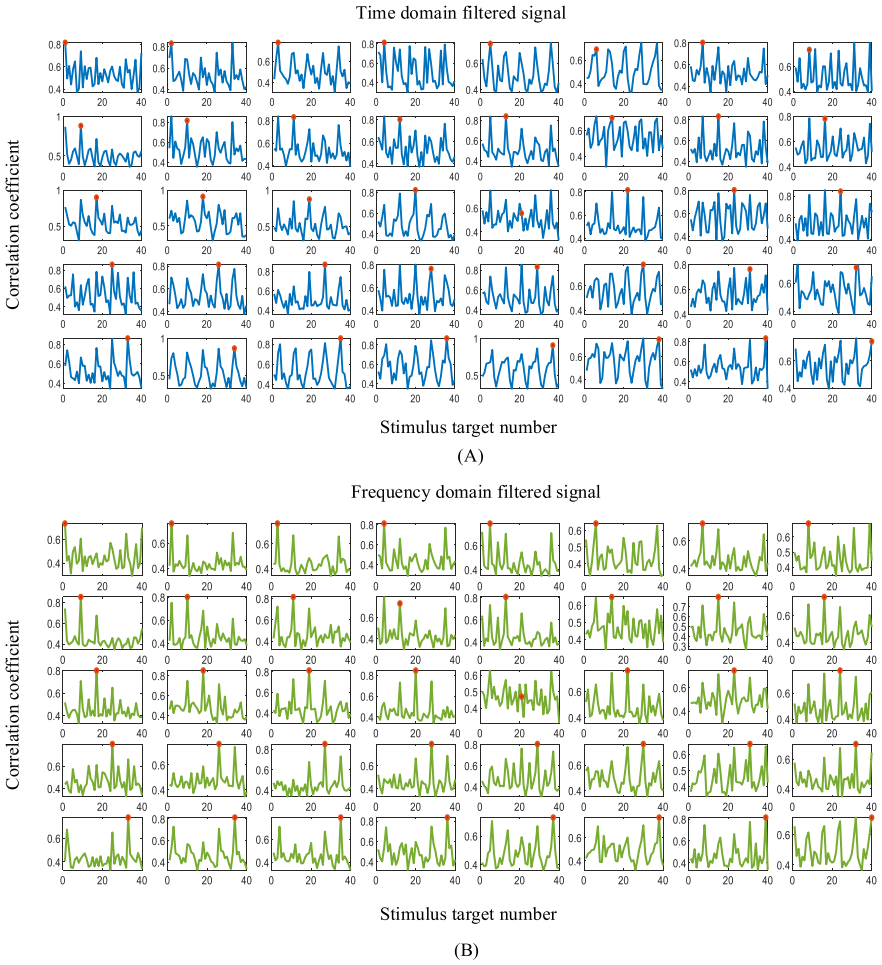

Fig. 9. Fig 9-A. The classification results of time-domain filtered signal using CCA. Fig 9-B. The classification results of frequency-domain filtered signal using CCA.

in subplots 2, 4, 5, 11, 13, 15, 16, 25 did not have a significant discrimination, indicating that time domain filtering method can easily misclassify a non-focused target as a focused target. In Fig. 9-B, for the same subplots, the frequency-domain filtering method allows improved distinction of canonical correlation coefficients between focused and non-focused targets, which will facilitate the accurate identification of the user focused target. These results show that frequency-domain filtering gives improved performance of SSVEP classification compared to the time-domain filtering method.

To test the proposed method further, the frequency-domain filtering method was also applied to FBCCA and TRCA methods. To combine frequency-domain filtering with FBCCA and TRCA methods, a feature ensemble scheme (denoted as the FBCCA_FDF and TRCA_FDF methods) was used in which the feature coefficients obtained from time-domain

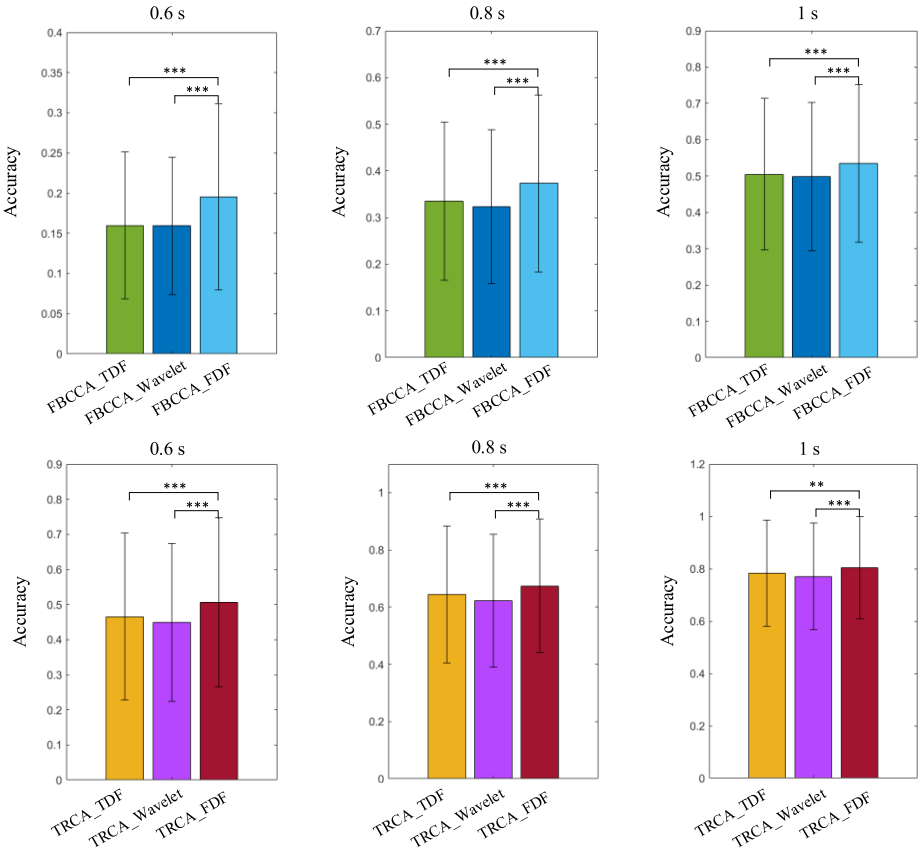

Fig. 10. The accuracies of the FBCCA, FBCCA_FDF, TRCA, and TRCA_FDF methods.

filtering and those obtained from frequency-domain filtering were integrated to obtain the final feature discriminant coefficients. For the FBCCA method, the number of filter banks was set to 3, and the harmonic number in the reference signal was set to 5. Let the feature coefficient obtained by CCA after the signal is processed by frequency-domain filter be p_1, and let the feature coefficients obtained by CCA after processing of the signal by three filter banks be p_2, p_3, and p_4, then the ensemble coefficient is:

$$p = \sum_{i=1}^{4} w_i \cdot p_i \tag{21}$$

For TRCA methods, let the feature coefficient of signal obtained by frequency-domain filtering using TRCA be p_1', and let the feature coefficient of signal obtained by time-domain filtering using TRCA be p_2', then the ensemble coefficient is:

$$p' = \sum_{i=1}^{2} w_i \cdot p_i' \tag{22}$$

where the value of $w(i)$ is the same as Eq. (4).

Fig. 10 shows the recognition accuracies of time-domain filtered signals, wavelet filtered signals and frequency-domain filtered signals obtained by the FBCCA and TRCA methods for data analysis lengths of 0.6, 0.8 and 1s. As observed, the recognition accuracy of frequency-domain filtering based FBCCA method (FBCCA_FDF) was improved by 3.56%, 3.82%, and 2.9%, compared with the time-domain filtering based FBCCA method (FBCCA_TDF) for stimulus lengths of 0.6 s, 0.8 s, and 1 s, respectively, and the recognition accuracy of the frequency-domain filtering based TRCA method (TRCA_FDF) was improved by 3.97%, 3.04%, and 2.1%, compared with the time-domain filtering based TRCA method (TRCA_TDF), respectively. The recognition accuracy of FBCCA_FDF method was improved by 3.62%, 4.96%, and 3.57%, compared with the wavelet filtering based FBCCA method (FBCCA_Wavelet) for stimulus lengths of 0.6 s, 0.8 s, and 1 s, respectively, and the recognition accuracy of the TRCA_FDF method was improved by 5.66%, 5.11%, and 3.29%, compared with the wavelet filtering based TRCA method (TRCA_Wavelet), respectively. In summary, the feature recognition performance of FBCCA and TRCA methods can be improved by using the frequency-domain filter.

IV. Discussion

SSVEP signal is widely used in the field of BCI, but also has good prospects for applications of medical testing and cognitive neuroscience research. Thus, SSVEP signal are important signals for neural information analysis. SSVEP signal collected from the scalp can contains various artifacts such as electrical activities related to the eye and muscles, so pre-processing of SSVEP signal is required before analysis. Band-pass or high-pass filtering in the time-domain is commonly used for SSVEP signal pre-processing. Because the frequency is the most important feature of SSVEP signal, in this work we developed a frequency-domain filter method for SSVEP.

In this method, the time-domain signal is expanded into multi-dimensional signal by EMD, which allows the subsequent frequency domain filtering of signal by image filtering. EMD decomposes the signal into several modal functions and one residual component, which facilitates the removal of artifacts in the signal at different scales. Comparison of the waveforms of multi-dimensional signal before and after 2-D FFT filtering shows that the frequency-domain filter can remove the low-frequency interference component from each order modal function. Since the low-order intrinsic mode function mainly includes signal low frequency components, the waveform of the low-order intrinsic mode function becomes smoother after frequency-domain filtering. Comparison of the waveforms of the original, time-domain, and frequency-domain filtered signals allows the calculation of the variance of the three signals. As observed, there are large fluctuations in the original signals that deviate from the baseline and have a large signal variance. This unstable distribution of signals is called trend term or baseline drift, and the presence of this trend term will seriously limit the accuracy of signal analysis. Time-domain filtering effectively removes the trend term, but the removal effect is weaker than that achieved by the frequency-domain filtering method. The experimental results show that the frequency filtered signal exhibits the smallest variance, so that after processing, the waveform is more uniformly distributed along the baseline. EMD can obtain the low-frequency component representing the signal trend term and frequency-domain filter can directly filter this trend term, thus the proposed frequency domain filtering method exhibits improved trend term removal effect. Comparison of the spectral results of the original signals, time-domain filtered signals, and frequency-domain filtered signals revealed a higher low-frequency interference component amplitude in the original signals than the amplitude at the stimulus frequency, and application of the frequency-domain filtering method provides an improved low-frequency interference component suppression effect.

Unlike the digital filter used in time-domain filtering, this proposed frequency domain filtering method requires the passband radius of the filter to determine the filtering range, as the passband radius does not correspond to the filtering frequency. Therefore, it is necessary to analyze the filtering performances for different passband radius values to select the appropriate passband radius. To do this, the power spectra of signals under different passband radius must be analyzed. The passband radius is set to remove low-frequency interference from the signal and to avoid filtering out the effective component of the signal. Here, the appropriate passband radius range was determined to be 10 - 20. Next, the SSVEP feature recognition accuracy was calculated for different passband radius, and the results showed that most subjects obtained optimized recognition performance at a passband radius of 10. We also analyzed the effect of frequency-domain filtering method on the performance of SSVEP feature recognition with different stimulus durations. The experimental results showed that the frequency-domain filtering method significantly improved the recognition accuracy of CCA, FBCCA and TRCA methods. Together, the experimental results indicate that the use of the frequency-domain filtering method in the pre-processing stage can facilitate the practical application of the SSVEP-BCI system. It is worth pointing out that the method proposed in this study has limitations, the frequency domain filter we designed cannot directly set the filter frequency, and the filter frequency needs to be determined by setting the passband radius.

V. Conclusion

Considering that the main feature of SSVEP signal is frequency, we proposed a frequency-domain filtering method using EMD and 2-D Fourier transform filtering. The performance of time-domain filtering and frequency filtering methods were compared, and the results showed that frequency-domain filtering can effectively remove the trend term in the signal and suppress low-frequency interference components. By analyzing the recognition accuracy of the frequency-domain filter for different passband radius values, the optimal passband radius for frequency-domain filter can be determined. Tests using data from public datasets show that the frequency-domain filtering method can significantly improve the recognition accuracy of CCA, FBCCA, and TRCA methods, demonstrating that frequency-domain filtering may be an appropriate pre-processing method for SSVEP signals.

References

[1] G. K. Anumanchipalli, J. Chartier, and E. F. Chang, "Speech synthesis from neural decoding of spoken sentences," *Nature*, vol. 568, no. 7753, pp. 493–498, Apr. 2019.

[2] C. Heelan et al., "Decoding speech from spike-based neural population recordings in secondary auditory cortex of non-human primates," *Commun. Biol.*, vol. 2, no. 1, p. 466, Dec. 2019.

[3] J. Kubanek, J. Brown, P. Ye, K. B. Pauly, T. Moore, and W. Newsome, "Remote, brain region–specific control of choice behavior with ultrasonic waves," *Sci. Adv.*, vol. 6, no. 21, May 2020, Art. no. eaaz4193.

[4] H. A. Lamti, M. M. Ben Khelifa, and V. Hugel, "Mental fatigue level detection based on event related and visual evoked potentials features fusion in virtual indoor environment," *Cogn. Neurodyn.*, vol. 13, no. 3, pp. 271–285, Jun. 2019.

[5] A. Spiegel, J. Mentch, A. J. Haskins, and C. E. Robertson, "Slower binocular rivalry in the autistic brain," *Current Biol.*, vol. 29, no. 17, pp. 2948–2953, Sep. 2019.

[6] W. Yan, G. Xu, J. Xie, M. Li, and Z. Dan, "Four novel motion paradigms based on steady-state motion visual evoked potential," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 8, pp. 1696–1704, Aug. 2018.

[7] M. J. Wieser, V. Miskovic, and A. Keil, "Steady-state visual evoked potentials as a research tool in social affective neuroscience," *Psychophysiology*, vol. 53, no. 12, pp. 1763–1775, Dec. 2016.

[8] W. Yan et al., "Enhancing detection of steady-state visual evoked potentials using channel ensemble method," *J. Neural Eng.*, vol. 18, no. 4, Aug. 2021, Art. no. 046008.

[9] W. Yan, C. Du, Y. Wu, X. Zheng, and G. Xu, "SSVEP-EEG denoising via image filtering methods," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1634–1643, 2021.

[10] W. Yan, G. Xu, Y. Du, and X. Chen, "SSVEP-EEG feature enhancement method using an image sharpening filter," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 115–123, 2022.

[11] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIS," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 6, pp. 1172–1176, Jun. 2007.

[12] X. Chen, Y. Wang, S. Gao, T.-P. Jung, and X. Gao, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface," *J. Neural Eng.*, vol. 12, no. 4, Aug. 2015, Art. no. 046008.

[13] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T.-P. Jung, and S. Gao, "High-speed spelling with a noninvasive brain–computer interface," *Proc. Nat. Acad. Sci. USA*, vol. 112, no. 44, pp. E6058–E6067, Nov. 2015.

[14] M. Nakanishi, Y. Wang, X. Chen, Y. Wang, X. Gao, and T.-P. Jung, "Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 1, pp. 104–112, Jan. 2018.

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

[16] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Multi-class brain–computer interface classification by Riemannian geometry," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 4, pp. 920–928, Apr. 2012.

[17] L. Mayaud et al., "Brain-computer interface for the communication of acute patients: A feasibility study and a randomized controlled trial comparing performance with healthy participants and a traditional assistive device," *Brain-Comput. Interfaces*, vol. 3, no. 4, pp. 197–215, Dec. 2016.

[18] B. A. S. Hasan and J. Q. Gan, "Hangman BCI: An unsupervised adaptive self-paced brain–computer interface for playing games," *Comput. Biol. Med.*, vol. 42, no. 5, pp. 598–606, May 2012.

[19] T. Zeyl, E. Yin, M. Keightley, and T. Chau, "Partially supervised P300 speller adaptation for eventual stimulus timing optimization: Target confidence is superior to error-related potential score as an uncertain label," *J. Neural Eng.*, vol. 13, no. 2, Apr. 2016, Art. no. 026008.

[20] C. Vidaurre, A. Schlögl, R. Cabeza, R. Scherer, and G. Pfurtscheller, "Study of on-line adaptive discriminant analysis for EEG-based brain computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 3, pp. 550–556, Mar. 2007.

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

[22] K.-J. Chiang, C.-S. Wei, M. Nakanishi, and T.-P. Jung, "Boosting template-based SSVEP decoding by cross-domain transfer learning," *J. Neural Eng.*, vol. 18, no. 1, Feb. 2021, Art. no. 016002.

[23] C. M. Wong et al., "Inter- and intra-subject transfer reduces calibration effort for high-speed SSVEP-based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 10, pp. 2123–2135, Oct. 2020.

[24] C. M. Wong et al., "Transferring subject-specific knowledge across stimulus frequencies in SSVEP-based BCIs," *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 2, pp. 552–563, Apr. 2021.

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

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

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

[28] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *J. Neural Eng.*, vol. 14, no. 1, Feb. 2017, Art. no. 016003.

[29] H. Cecotti and A. Graser, "Convolutional neural networks for P300 detection with application to brain-computer interfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 3, pp. 433–445, Mar. 2011.

[30] N. Lu, T. Li, X. Ren, and H. Miao, "A deep learning scheme for motor imagery classification based on restricted Boltzmann machines," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 6, pp. 566–576, Jun. 2017.

[31] S. Pouryazdian and A. Erfanian, "Detection of steady-state visual evoked potentials for brain-computer interfaces using PCA and high-order statistics," in *Proc. World Congr. Med. Phys. Biomed. Eng.*, Sep. 2009, pp. 480–483.

[32] Y. Wang, Z. Zhang, X. Gao, and S. Gao, "Lead selection for SSVEP-based brain-computer interface," in *Proc. 26th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 26, Sep. 2004, pp. 4507–4510.

[33] N. Mammone, F. La Foresta, and F. C. Morabito, "Automatic artifact rejection from multichannel scalp EEG by wavelet ICA," *IEEE Sensors J.*, vol. 12, no. 3, pp. 533–542, Mar. 2012.

[34] B. Mijović, M. De Vos, I. Gligorijević, J. Taelman, and S. Van Huffel, "Source separation from single-channel recordings by combining empirical-mode decomposition and independent component analysis," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 9, pp. 2188–2196, Sep. 2010.

[35] Y. Wang, X. Chen, X. Gao, and S. Gao, "A benchmark dataset for SSVEP-based brain-computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.* vol. 25, no. 10, pp. 1746–1752, Oct. 2017.