A flexible and accurate method for electroencephalography rhythms extraction based on circulant singular spectrum analysis

Hai Hu [1] , **Zihang Pu** [1] , **Peng Wang** [Corresp. 1]

[1] Department of Precision Instrument, Tsinghua University, Beijing, China

Corresponding Author: Peng Wang
Email address: peng@mail.tsinghua.edu.cn

Rhythms extraction from electroencephalography (EEG) signals can be used to monitor the physiological and pathological states of the brain and has attracted much attention in recent studies. A flexible and accurate method for EEG rhythms extraction was proposed by incorporating a novel circulant singular spectrum analysis (CSSA). The EEG signals are decomposed into the sum of a set of orthogonal reconstructed components (RCs) at known frequencies. The frequency bandwidth of each RC is limited to a particular brain rhythm band, with no frequency mixing between different RCs. The RCs are then grouped flexibly to extract the desired EEG rhythms based on the known frequencies. The extracted brain rhythms are accurate and no mixed components of other rhythms or artifacts are included. Simulated EEG data based on the Markov Process Amplitude EEG model and experimental EEG data in the eyes-open and eyes-closed states were used to verify the CSSA-based method. Results showed that the CSSA-based method is flexible in alpha rhythms extraction and has a higher accuracy in distinguishing between the eyes-open and eyes-closed states, compared with the basic SSA method, the wavelet decomposition method, and the infinite impulse response filtering method.

A Flexible and Accurate Method for Electroencephalography Rhythms Extraction Based on Circulant Singular Spectrum Analysis

Hai Hu[1], Zihang Pu[1], Peng Wang[1]

[1] State Key Laboratory of Precision Measurement Technology and Instruments, Tsinghua University, Beijing, China

Corresponding Author:
Peng Wang
Department of Precision Instrument, Tsinghua University, Beijing, 100086, China
Email address: peng@mail.tsinghua.edu.cn

Abstract

Rhythms extraction from electroencephalography (EEG) signals can be used to monitor the physiological and pathological states of the brain and has attracted much attention in recent studies. A flexible and accurate method for EEG rhythms extraction was proposed by incorporating a novel circulant singular spectrum analysis (CSSA). The EEG signals are decomposed into the sum of a set of orthogonal reconstructed components (RCs) at known frequencies. The frequency bandwidth of each RC is limited to a particular brain rhythm band, with no frequency mixing between different RCs. The RCs are then grouped flexibly to extract the desired EEG rhythms based on the known frequencies. The extracted brain rhythms are accurate and no mixed components of other rhythms or artifacts are included. Simulated EEG data based on the Markov Process Amplitude EEG model and experimental EEG data in the eyes-open and eyes-closed states were used to verify the CSSA-based method. Results showed that the CSSA-based method is flexible in alpha rhythms extraction and has a higher accuracy in distinguishing between the eyes-open and eyes-closed states, compared with the basic SSA method, the wavelet decomposition method, and the infinite impulse response filtering method.

Introduction

Electroencephalograms (EEGs) are the electrical activity of the brain's neurons recorded at the scalp surface(Henry 2006). They consist of several rhythm bands: delta (1–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz) and gamma (>30 Hz). Because the rhythms reflect different physiological and pathological information, EEG rhythms extraction has been widely applied in many areas. Examples include portable and wearable EEG devices(Hwang et al. 2018; Maskeliunas et al. 2016), mental fatigue assessment(Taran & Bajaj 2017), disease diagnosis(Babiloni et al. 2016; Gupta & Pachori 2019), and brain computer interface systems(Jeunet et al. 2019; Liu et al. 2020).

The accuracy of EEG rhythms extraction determines the physiological and pathological information it provides. Various methods have been proposed to extract the desired EEG rhythms. Filtering components have the ability to restrict a signal to a specific frequency band, and such bandpass filters were first used to extract EEG rhythms(Pfurtscheller et al. 1997). This method performed well in EEGs of high signal-to-noise ratio (SNR). Then, the wavelet transform (WT) method was used for EEG rhythms extraction(Duque-Muñoz et al. 2015). By estimating the rhythms with a customized wavelet, the WT method can extract time-varying EEG rhythms with changes in brain state. To facilitate the EEG rhythms extraction, the independent component analysis(ICA) method was then introduced(Kavuri et al. 2018). By incorporating priori information about the desired rhythms as reference signals, the ICA method can extract EEG rhythms automatically. However, the extracted rhythms using the bandpass filter, WT, and ICA methods were contaminated by noise and artifacts overlapping in time–frequency space. In recent years, to improve the accuracy of EEG rhythms extraction, the singular spectrum analysis (SSA) method has been used(Akar et al. 2015; Mohammadi et al. 2016). This nonparametric method enables the separation of different sources even when they overlap in time–frequency space(Mohammadi et al. 2015).

In the basic SSA method, the grouping rule is important for SSA reconstruction. However, because of the lack of the information about the amplitude and frequency of the reconstructed components (RCs), there is no general grouping rule. Different grouping rules have been proposed depending on the research target, the types of signals, and noise. The conventional SSA grouping is performed according to the magnitudes of eigenvalues related to the power of each RC(Teixeira et al. 2005). Mohammadi et al. proposed a new grouping rule based on eigenvalue pairs to extract the main rhythms from sleep EEG signals(Mohammadi et al. 2016). Hu et al. proposed another efficient grouping rule based on the similarity between the eigenvalues and the peak frequency of RC, which makes SSA adaptive to EEG signals containing different levels of artifacts and rhythms(Hai et al. 2017). However, these grouping rules can only be applied to specific types of signals and must be incorporated with other methods (e.g., Fourier transform or wavelet decomposition) to pre-identify the frequencies of RCs, which is time-consuming and inflexible. Besides, the observed frequency mixing between different RCs leads to inaccurate EEG rhythms extraction(Xu et al. 2018).

In this paper, we introduce a novel circulant singular spectrum analysis (CSSA) method to improve the flexibility and accuracy of EEG rhythms extraction. Compared with the basic SSA method, the CSSA method has the advantage of avoiding the need for pre-identifying the frequencies of RCs. A set of orthogonal vectors are obtained by decomposing the circulant matrix, and the EEG signals can be decomposed into the sum of a set of orthogonal RCs of known frequencies. The RCs can be grouped automatically and flexibly to extract the specific EEG rhythms based on their frequencies. In addition, because the frequency bandwidth of each RC is limited to a particular band of the brain rhythm of interest, the extracted brain rhythms are accurate and no mixed components of other rhythms and artifacts are included.

Methods

The CSSA method is a nonparametric signal extraction method proposed by Juan Bógalo(Bógalo et al. 2020). CSSA consists of four steps: embedding, decomposition, diagonal averaging, and grouping. As in the basic SSA method, in the time-delay embedding step, the single-channel EEG time series $\mathbf{s} = (s_1, s_2, \ldots, s_N)^T$ (superscript T denotes the transpose of a vector) is mapped onto a multidimensional trajectory matrix $\mathbf{X}$ using a sliding window(Takens 1981):

$$\mathbf{X} = (\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_K) = \begin{pmatrix} s_1 & s_2 & \cdots & s_K \\ s_2 & s_3 & \cdots & s_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_L & s_{L+1} & \cdots & s_N \end{pmatrix} \tag{1}$$

where L denotes the window length (or embedding dimension), $K=N-L+1$, and $\mathbf{S}_i$ denotes the lagged vector.

In the decomposition step, the trajectory matrix is decomposed into elementary matrices of rank 1 that are associated with different frequencies. To do so, a related circulant matrix $\mathbf{C}_L$ is built based on the second order moments of the time series(Bógalo et al. 2020):

$$\mathbf{C}_L(f) = \begin{pmatrix} c_0 & c_1 & c_2 & \cdots & c_{L-1} \\ c_{L-1} & c_0 & c_1 & \cdots & c_{L-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & c_3 & \cdots & c_0 \end{pmatrix} \tag{2}$$

where

$$c_m = \frac{L-m}{L}\gamma_m + \frac{m}{L}\gamma_{L-m}, \gamma_m = \frac{1}{N-m}\sum_{t=1}^{T-m} s_t s_{t+m}, \quad m = 0,1,...,L-1 \tag{3}$$

the eigenvalues and eigenvectors of $\mathbf{C}_L$, respectively, are given(Gray & Robert 2006)

$$\begin{aligned} &\lambda_k = \sum_{m=0}^{L-1} c_m \exp(i2\pi m \frac{k-1}{L}) = f(\frac{k-1}{L}), k = 1,2,...,L \\ &\mathbf{u}_k = L^{-1/2}(u_{k,1}, u_{k,2}, ..., u_{k,L})^H, u_{k,j} = \exp(-i2\pi(j-1)\frac{k-1}{L}), k = 1,2,...,L \end{aligned} \tag{4}$$

where $f(\cdot)$ denotes the power spectral density of the signal, and H indicates the conjugate transpose of a matrix. The k-th eigenvalue and the corresponding eigenvector are associated with the given frequencies by

$$f_k = \frac{k-1}{L} f_s \tag{5}$$

where f_s is the sampling rate of the EEG signals. As a consequence, the diagonalization of $\mathbf{C}_L$ allows us to write $\mathbf{X}$ as the sum of the elementary matrices $\mathbf{X}_k$:

$$\mathbf{X} = \sum_{k=1}^{L} \mathbf{X}_k = \sum_{k=1}^{L} \mathbf{u}_k \mathbf{u}_k^H \mathbf{X} \tag{6}$$

The symmetry of the power spectral density leads to $\lambda_k = \lambda_{L+2-k}$. The corresponding eigenvectors given by (4) are complex; therefore, they are paired with complex conjugates, $\mathbf{u}_k = \mathbf{u}^*_{L+2-k}$ where $\mathbf{u}^*$ indicates the complex conjugate of a vector $\mathbf{u}$. Then, $\mathbf{X}_k$ and $\mathbf{X}_{L+2-k}$ correspond to the same harmonic period.

To obtain the elementary matrices by frequency, we first form the groups of two elements $B_k = \{k, L+2-k\}, k = 2,3,...,M, M = \lfloor (L+1)/2 \rfloor$, with $B_1 = \{1\}$ and $B_{L/2+1} = \{L/2+1\}$ if L is even. Second, we compute the real elementary matrix for frequency $\mathbf{X}_{B_k}$ as the sum of the two elementary matrices $\mathbf{X}_k$ and $\mathbf{X}_{L+2-k}$, which are associated with eigenvalues λ_k, λ_{L+2-k} and frequency ω_k, given by (5)

$$\mathbf{X}_{B_k} = \mathbf{X}_k + \mathbf{X}_{L+2-k} = \mathbf{u}_k \mathbf{u}_k^H \mathbf{X} + \mathbf{u}_{L+2-k} \mathbf{u}_{L+2-k}^H \mathbf{X} = (\mathbf{u}_k \mathbf{u}_k^H + \mathbf{u}_k^* \mathbf{u}_k^{'})\mathbf{X} = 2(R_{\mathbf{u}_k} R_{\mathbf{u}_k}^{'} + I_{\mathbf{u}_k} I_{\mathbf{u}_k}^{'})\mathbf{X} \tag{7}$$

where $R_{\mathbf{u}_k}$ and $I_{\mathbf{u}_k}$ denote the real and imaginary parts of $\mathbf{u}_k$, respectively, and the matrices $\mathbf{X}_{B_k}$ are real.

Then, in the diagonal averaging step(Vautard et al. 1992), several time series are reconstructed from the corresponding real elementary matrices $\mathbf{X}_{B_k}$. The reconstructed time series are generally called RCs. Theoretically, the frequencies of the RCs are given by (5). Finally, the alpha rhythm (8–13 Hz) can be extracted automatically by

$$V_{alpha} = \sum_{i=\lceil 1+8L/f_s \rceil}^{\lfloor 1+13L/f_s \rfloor} RC_i \tag{8}$$

The frequency bandwidth of each RC can be roughly expressed by(A et al. 2010; Xu et al. 2018)

$$f_b = f_s / L \tag{9}$$

As a consequence, the frequency bandwidth of each RC is limited to f_s / L . Considering the frequency of each RC given by equation (5), there is no frequency mixing between different RCs, and the extracted alpha rhythms do not contain mixed components of other rhythms or artifacts.

The pseudo-code of the CSSA method is shown in Algorithm 1.

Simulation Results and Discussion

Markov Process Amplitude EEG model

Simulated spontaneous EEG signals were used to verify the validity of the CSSA method in alpha rhythm extraction. Spontaneous EEG signals were generated based on the Markov Process Amplitude (MPA) EEG model(Bai et al. 2001; Nishida et al. 1986). The MPA EEG model is a powerful and widely used method to simulate and interpret EEG signals. With a few parameters, the model can represent the two major characteristics of EEG signals: rhythmic oscillation and randomness. Rhythmic oscillation is represented by sinusoidal waves, and randomness was represented by the stochastic process amplitude of the first-order Markov process. In recent years, the MPA EEG model has been applied in several studies analyzing spontaneous EEG signals, which have employed such techniques as feature expression, quantitative analysis(Nakamura et al. 1997), and algorithm verification(Xu et al. 2018).

In the MPA EEG model, EEG signals consist of several rhythmic oscillations expressed by a sinusoidal wave

$$s(n\Delta t) = \sum_{i=1}^{K} a_i(n\Delta t)\sin(2\pi f_i n\Delta t + \theta_i) \tag{10}$$

where n is the number of samples, Δt is the time interval, K is the number of rhythms, f is the dominant frequency of rhythm, θ is the initial phase (zero), and a is the rhythmic amplitude obtained from the following first-order Gauss–Markov process:

$$a_i\left[(n+1)\Delta t\right] = \gamma_i a_i(n\Delta t) + \xi_i(n\Delta t) \tag{11}$$

where γ is the coefficient of the first-order Markov process, and ξ is a random increment of Gaussian distribution with mean zero and variance σ^{ξ} . Therefore, the rhythmic amplitude at the succeeding time $\left(n+1\right)\Delta t$ depends only on the amplitude at time Δt and is determined only by two parameters: γ and σ^{ξ}. The parameters of the MPA EEG model are determined in the frequency domain to achieve the maximum likelihood with respect to the power spectrum of real EEG. H_i is defined as the amplitude, and B_i is the frequency width at half of H_i of the EEG power spectrum. Based on the literature(Bai et al. 2001), H_i, B_i can be described as

$$\begin{cases} H_i = \dfrac{\Delta t\left(\sigma_i^{\xi}\right)^2}{4\left(1-\gamma_i\right)^2} \\ B_i = \dfrac{1}{\pi\Delta t}\cos^{-1}\dfrac{4\gamma_i-1-\left(\gamma_i\right)^2}{2\gamma_i} \end{cases} \tag{12}$$

The simulation procedures of spontaneous EEG signals based on the MPA EEG model are shown in Fig. 1. First, the power spectrum of a real EEG signal with sampling rate f_s= 200 Hz is calculated in Fig. 1A. Then, f_i, H_i, and B_i, which represent the peak frequencies, amplitude, and the frequency width at half of amplitude of the EEG rhythms (delta, theta, alpha, and beta), respectively, are obtained according to the power spectrum. Based on equation (12), the parameters of the first-order Gauss–Markov process (γ and σ^{ξ}) are obtained. All parameters of the MPA EEG model are shown in Table 1. Then, the delta, theta, alpha, and beta rhythms are simulated based on the determined parameters, as shown in Fig. 1B. Finally, the simulated EEG signal (shown in Fig. 1C) is generated as the sum of the four rhythms. The simulated spontaneous EEG lasts for 8 s with a 5-ms Δt interval.

Circulant singular spectrum analysis of simulated EEG signals

The simulated EEG signal is processed by the CSSA method with the embedding dimension set to L=40 and L=80. Figure 2 shows the power spectrum density (PSD) of the first six reconstructed components (RCs). When L=40, as shown in Fig. 2A, every RC falls on the theoretical frequency derived by equation (5). Furthermore, the bandwidth of each RC is limited to f_s / L=5 Hz. Similarly, when L=80, as shown in Fig. 2B, all RCs fall on the theoretical frequencies with the bandwidth limited to 2.5 Hz. However, the simulated EEG signal is processed by the basic SSA method with the embedding dimension set to L=40. The PSD of the first six RCs is shown in Fig. 3. The frequency of each RC is unknown. To group the RCs by frequency, other algorithms like Fourier transform are introduced to calculate the frequency of RCs. Besides, according to the PSD of RC5 and RC6, some components fall outside of the bandwidth limit of f_s / L=5 Hz. This phenomenon is called component mixing.

Because the frequencies of the RCs processed by the CSSA method are known, the alpha rhythm of the EEG signal can be extracted by equation (8). The error parameter for evaluating the performance of the alpha rhythm extraction is defined as(Xu et al. 2018)

$$\varepsilon_{ave} = \frac{1}{N}\sum_{i=1}^{N}\left| P_{\alpha}(i) - P_e(i) \right| \tag{13}$$

where ε_{ave} is the average error of the PSD between the simulated alpha rhythm and the extracted alpha rhythm, $P_{\alpha}(i)$ is the PSD of the simulated alpha rhythm, $P_e(i)$ is the PSD of the extracted alpha rhythm, and *N* is the length of the PSD.

Figure 4A shows the extracted alpha rhythm of the simulated EEG signal by the CSSA method with the embedding dimension set to *L*=40. Figure 2A shows that RC3 represents the alpha rhythm. The PSD of the simulated and extracted alpha rhythm by the CSSA method when *L*=40 is shown in Fig. 4B. There was component mixing (slash shadow), and the error of the extracted alpha rhythm was $\varepsilon = 0.43\ \mu V^2 / Hz$. Similarly, RC5 and RC6 represent the alpha rhythm from Fig. 2B with an embedding dimension of *L*=80. The alpha rhythm extracted from the combination of RC5 and RC6 is shown in Fig. 4C. The PSD of the simulated and extracted alpha rhythm by the CSSA method when *L*=80 is shown in Fig. 4D. There was less component mixing (slashed shadow) than that observed when *L*=40, and the error of the extracted alpha rhythm was $\varepsilon = 0.27\ \mu V^2 / Hz$. In the basic SSA method, the alpha rhythm is extracted according to the adaptive grouping rule(Hai et al. 2017). The extracted alpha rhythm by the basic SSA method is the sum of RC3, RC4, RC5, and RC6, as shown in Fig. 4E. The PSD of the simulated and extracted alpha rhythms by basic SSA, shown in Fig. 4F, illustrates the presence of more component mixing (slashed shadow) than that found with the CSSA method, and the error of the extracted alpha rhythm was $\varepsilon = 0.89\ \mu V^2 / Hz$.

To compare the performance of the alpha rhythm extraction with that of other methods, the alpha rhythms were extracted by the infinite impulse response (IIR) filtering methods and the wavelet decomposition (WDec) method. The PSDs of the extracted alpha rhythms by the IIR and WDec methods are shown in Fig. 5A and 5B, respectively. The PSD of the extracted alpha rhythm by the IIR method had a higher magnitude, and there was more component mixing than that by the CSSA method (Fig. 5A). The error of the extracted alpha rhythm by the IIR method was $\varepsilon = 0.57\ \mu V^2 / Hz$, which was higher than that obtained by the CSSA method. Figure 5B shows that the extracted alpha rhythm by the WDec method consisted almost entirely of mixed components. The error of the extracted alpha rhythm by the WDec method was $\varepsilon = 1.06\ \mu V^2 / Hz$, which was much higher than that obtained by the CSSA method.

We conclude that the RCs of simulated EEG signals processed by the CSSA method fall on the theoretical frequencies limited to the selected bandwidth ranges. The alpha rhythm can be extracted automatically based on the frequency feature. The alpha rhythm extraction by the CSSA method performed better than that of the basic SSA, IIR, and WDec methods. Therefore, the processing results of the simulated EEG verify the validity of the CSSA method's performance in alpha rhythm extraction. Furthermore, the error of the extracted alpha rhythm by the CSSA method varied with the embedding dimension. The calculated errors of the extracted alpha rhythm by the CSSA method with different embedding dimensions are shown in Table 2. The error attained a minimum value at *L*=80. Therefore, the embedding dimension of the CSSA method for alpha extraction was set to *L*=80.

Experimental Results and Discussion

Results and discussion of database EEG signals

The database EEG signals reported in literature(Trujillo et al. 2017) were used. 22 subjects (11 female, 11 male, mean age=21.1±0.52 years, age range=18–26 years) underwent 8 min of resting state EEG recording while sitting quietly in a comfortable padded chair in a darkened room (4 min eyes open and 4 min eyes closed interleaved in 1-min intervals; the order of eyes open/closed was balanced across participants). The EEG signals of one subject (subject #6) were removed because of a technical recording error. 72 channels of continuous EEG signals were recorded using active Ag/AgCl electrodes mounted in a BioSemi electrode cap with international 10/5 system locations. All channels were amplified by a BioSemi Active II amplifier system in 24-bit DC mode at an initial sampling rate of 2,048 Hz (400-Hz bandwidth) downsampled online to 256 Hz.

Channel Fpz was selected for EEG analysis. The EEG data of the Fpz channel were divided into 8-s (2048-sample) epochs with 50% overlap, initially producing 91 epochs of eyes-open and eyes-closed conditions for each subject. This was done because artifacts, including those resulting from electrooculogram, electromyography, baseline drift, and stochastic noise, interfere with the rhythm extraction. The adaptive SSA method(Hai et al. 2017) was used to remove the artifacts, and the results are shown in Fig. 6. Figure 6A and 6B show EEG epochs of the eyes-open and eyes-closed conditions of subject #17, respectively. Figure 6C and 6D show the corrected EEG signals after artifact removal. The electrooculogram artifacts were removed from the EEG signals of the eyes-open condition, and the spontaneous EEG signals were preserved in both conditions.

The EEG signals after artifact removal were processed by the CSSA method with the embedding dimension set to L=80. The first six RCs of the EEG signals in the eyes-open and eyes-closed conditions are shown in Fig. 7A and 7B, respectively. Each RC falls on the theoretical frequency derived by equation (5), and the bandwidth of the RCs is limited to f_s / L =3.2 Hz, which is agreement with the simulation results. Figure 7A and 7B show that RC4 and RC5 represent the alpha rhythm in both the eyes-open and eyes-closed conditions, respectively. Thus, the alpha rhythms of the EEG signals can be extracted automatically as the sum of RC4 and RC5, which is agreement with equation (8). Figure 8A and 8B show the extracted alpha rhythms of the EEG signals in the eyes-open and eyes-closed conditions, respectively. The amplitude of the alpha rhythm in the eyes-open condition was lower than that in the eyes-closed condition. This was consistent with the results of previous studies, in which the alpha rhythm in the resting state in the eyes-open condition with visual stimulation was much weaker than that in the eyes-closed condition(Barry et al. 2007). Figure 8C illustrates the spectrogram of alpha rhythms in the eyes-open and eyes-closed conditions, which is the square of the rhythm's amplitude as a function of time and frequency. It illustrates a significant difference between the eyes-open and eyes-closed states.

The performance of alpha rhythm extraction by the CSSA method was compared with that of three other methods: the basic SSA method, the WDec method, and the IIR method. The alpha rhythms under the eyes-open and eyes-closed conditions were extracted using the CSSA, basic SSA, WDec, and IIR methods. The PSD of the extracted alpha rhythms by the four methods under the eyes-closed and eyes-open conditions are shown in Fig. 9. Figure 9A and 9D show that the extracted alpha rhythms using the CSSA method were within the alpha band (8–13 Hz) under both the eyes-open and eyes-closed conditions. In addition, the power of the extracted alpha rhythm under the eyes-open condition was lower than that under the eyes-closed condition. Therefore, the alpha rhythm extracted using the CSSA method could represent the real EEG alpha rhythm. However, the alpha rhythm extracted by the basic SSA method contained frequency components outside the alpha band because of component mixing under both the eyes-open and eyes-closed conditions. Especially in the eyes-open condition, most components of the extracted alpha rhythm fall outside the alpha band, inconsistently with reality. Similarly, the alpha rhythms extracted using the WDec method contained many components outside the alpha band (component mixing), as shown in Fig. 9B and 9E. Figure 9C and 9F show that the alpha rhythms extracted using the IIR method fell into the alpha band, and that the power of the extracted alpha rhythm was stronger than that extracted using the CSSA method under both the eyes-open and eyes-closed conditions. This is in agreement with the simulation results shown in Fig. 5A because the IIR method was unable to remove artifacts and noise from the alpha rhythm with an overlapping frequency spectrum. Therefore, the CSSA method performed better than the basic SSA, WDec, and IIR methods at alpha rhythm extraction.

To further verify the CSSA method's performance, the extracted alpha rhythms were used to distinguish between the eyes-open and eyes-closed states, and the classification results produced by the CSSA method were compared with those by the basic SSA, IIR, and WDec methods. In this study, the power ($P = \sum_{i=1}^{N} V_i^2 / N$) and the mean of the absolute value ($\overline{V} = \sum_{i=1}^{N} | V_i | / N$) were selected as the features of the alpha rhythm(Mohammadi et al. 2015), where V_i represents the amplitude of the extracted alpha rhythm, and N represents the number of samples. Figure 10A shows the values of the power and the mean of the absolute value of the extracted alpha rhythm by the CSSA method for subject #17. The power value and the mean of the absolute value under the eyes-open condition were lower than those under the eyes-closed condition. Then, the support vector machine method was used to classify the features under the eyes-open and eyes-closed conditions. The classification accuracy was 92.31%. Figure 10B–10D show the power values and mean absolute values of the extracted alpha rhythms by the basic SSA, IIR, and WDec methods, respectively. Similar to the results obtained by the CSSA method, the power and mean of the absolute value of the extracted alpha signal under the eyes-open condition were generally lower than those under the eyes-closed condition. The classification accuracy of feature extraction by the basic SSA, IIR, and WDec methods was 90.11%, 91.21%, and 91.21%, respectively, and these values were lower than the accuracy obtained by the CSSA method.

We calculated the power values and means of the absolute value of the extracted alpha rhythms of all 21 subjects, and the classification results are shown in Table 3. The classification accuracy varied greatly between different subjects because of individual differences in EEG signals. The mean and standard deviation of the classification accuracy was calculated for all subjects to compare classification performance between the CSSA, basic SSA, IIR, and WDec methods. The mean value of the classification accuracy for all subjects by the CSSA method was 92.36%, which was higher than those obtained by the basic SSA (88.38%), IIR (91.89%), and WDec (90.47%) methods. The standard deviation of the classification accuracy across all subjects by the CSSA method was 7.05%, which was lower than those obtained by the basic SSA (10.35%), IIR (7.50%), and WDec (10.88%) methods. Therefore, the CSSA method's classification performance was better and more robust than that by the basic SSA, IIR, and WDec methods.

Results and discussion of experimental EEG signals

Additional experimental EEG signals were recorded and used to further verify the validity of the CSSA method. The experiments were approved with a protocol (NO. 20170010) by the Institutional Review Board of Tsinghua University and the written informed consent was obtained from the subject. One male subject aged 29 years participated in the experiments and abstained from psychoactive substances for at least 4 h prior to the experiments. The experiments were carried out with the subject sitting on a comfortable chair in a room with normal lightness. The experimental EEG signals were recorded using the MP160 data acquisition and analysis system (BIOPAC Systems, Inc., Goleta, CA, USA). A three-electrode system was used to improve the common mode rejection ratio of the measurement setup. Ag/AgCl was the material of the recording electrode, which was flushed with conductive gel and then attached to the frontal region of the subject's scalp. The other two electrodes, which served as ground and reference, were attached to the earlobe and mastoid, respectively, as shown in Fig. 11B. The experimental procedures were as follows. First, the subject relaxed with eyes closed for 10 min. Next, the subject opened his eyes and focused on a cross symbol displayed on the computer screen. Finally, the subject kept his eyes open for 30 s, followed by a period with eyes closed for 30 s, and repeated this procedure 57 times. Throughout the experiment, the real EEG signal was recorded at a sampling rate of 200 Hz. To obtain the desired segments of the eyes-open and eyes-closed states, segments lasting 8 s were extracted from the middle of each period (Fig. 11B). Consequently, 57 segments each of the eyes-open and eyes-closed states were obtained. Artifacts were removed from each EEG segment using the adaptive SSA method(Hai et al. 2017).

The experimental EEG signals were processed by the CSSA method, and the alpha rhythms recorded under the eyes-open and eyes-closed conditions were extracted. The power and mean absolute values of the extracted alpha rhythms were calculated as features for classification using the support vector machine method. The classification accuracy by the CSSA method was 91.23%, which was higher than that obtained by the basic SSA (89.47%), IIR (89.47%), and WDec (88.60%) methods (Fig. 12). We therefore concluded that the CSSA method's classification performance was better than that by the basic SSA, IIR, and WDec methods.

Conclusions

In this paper, a flexible and accurate method based on CSSA was proposed for alpha rhythm extraction from EEG signals. By decomposing the EEG signals into a set of orthogonal reconstructed components (RCs) at specific bandwidths of frequencies, the alpha rhythm can be extracted flexibly and accurately from EEG signals. The proposed method performed well on both simulated EEG data generated from the MPA EEG model and experimental EEG data, as well as the EEG data obtained from a public database. Features of the alpha rhythms extracted from experimental EEG signals were calculated to distinguish between the eyes-open and eyes-closed states. The CSSA-based method showed higher classification accuracy and robustness than that of the basic SSA, IIR and WDec methods.

References

A EB, B RC, and A DF. 2010. Relationship between Singular Spectrum Analysis and Fourier analysis: Theory and application to the monitoring of volcanic activity. *Computers & Mathematics with Applications* 60:812-820.

Akar SA, Kara S, Latifolu F, and Bilgi V. 2015. Investigation of the noise effect on fractal dimension of EEG inschizophreni.

Bógalo J, Poncela P, and Senra E. 2020. Circulant Singular Spectrum Analysis: A new automated procedure for signal extraction. *arXiv*.

Babiloni C, Triggiani AI, Lizio R, Cordone S, Tattoli G, Bevilacqua V, Soricelli A, Ferri R, Nobili F, Gesualdo L, Millán-Calenti JC, Buján A, Tortelli R, Cardinali V, Barulli MR, Giannini A, Spagnolo P, Armenise S, Buenza G, Scianatico G, Logroscino G, Frisoni GB, and del Percio C. 2016. Classification of Single Normal and Alzheimer's Disease Individuals from Cortical Sources of Resting State EEG Rhythms. *Frontiers in Neuroscience* 10. 10.3389/fnins.2016.00047

Bai O, Nakamura M, Nishida S, Ikeda A, and Shibasaki H. 2001. Markov process amplitude EEG model for spontaneous background activity. *Journal of Clinical Neurophysiology Official Publication of the American Electroencephalographic Society* 18:283.

Barry RJ, Clarke AR, Johnstone SJ, Magee CA, and Rushby JA. 2007. EEG differences between eyes-closed and eyes-open resting conditions. *Clinical neurophysiology* 118:2765-2773.

Duque-Muñoz L, Pinzon-Morales RD, and Castellanos-Dominguez G. 2015. EEG Rhythm Extraction Based on Relevance Analysis and Customized Wavelet Transform. International Work-conference on the Interplay Between Natural & Artificial Computation.

Gray, and Robert M. 2006. Toeplitz and Circulant Matrices: A review. *Foundations & Trends in Communications & Information Theory* 2:155-239.

Gupta V, and Pachori RB. 2019. Epileptic seizure identification using entropy of FBSE based EEG rhythms. *Biomedical Signal Processing and Control* 53:101569.

Hai H, Guo S, Ran L, and Peng W. 2017. An adaptive singular spectrum analysis method for extracting brain rhythms of electroencephalography. *Peerj* 5:e3474.

Henry JC. 2006. Electroencephalography: basic principles, clinical applications, and related fields. *Neurology* 67:2092-2092-a.

Hwang S, Jebelli H, Choi B, Choi M, and Lee S. 2018. Measuring workers' emotional state during construction tasks using wearable EEG. *Journal of Construction Engineering and Management* 144:04018050.

Jeunet C, Glize B, McGonigal A, Batail J-M, and Micoulaud-Franchi J-A. 2019. Using EEG-based brain computer interface and neurofeedback targeting sensorimotor rhythms to improve motor skills: Theoretical background, applications and prospects. *Neurophysiologie Clinique* 49:125-136.

Kavuri SS, Veluvolu KC, and Chai QH. 2018. Evolutionary Based ICA With Reference for EEG mu Rhythm Extraction. *IEEE Access* 6:19702-19713.

Liu T, Huang G, Jiang N, Yao L, and Zhang Z. 2020. Reduce brain computer interface inefficiency by combining sensory motor rhythm and movement-related cortical potential features. *Journal of neural engineering* 17:035003.

Maskeliunas R, Da Masevicius R, Martisius I, and Vasiljevas M. 2016. Consumer-grade EEG devices: are they usable for control tasks? *Peerj* 4.

Mohammadi SM, Enshaeifar S, Ghavami M, and Sanei S. 2015. Classification of awake, REM, and NREM from EEG via singular spectrum analysis. 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC): IEEE. p 4769-4772.

Mohammadi SM, Kouchaki S, Ghavami M, and Sanei S. 2016. Improving Time-Frequency Domain Sleep EEG Classification via Singular Spectrum Analysis. *Journal of Neuroence Methods*:96-106.

Nakamura M, Takakura S, Sugi T, Nishida S, Ikeda A, and Shibasaki H. 1997. Quantitative analysis of slow wave characteristics of rhythmicity by use of EEG model. *Biomed Eng* 35:392-399.

Nishida S, Nakamura M, and Shibasaki H. 1986. [An EEG model expressed by sinusoidal waves with the Markov process amplitude]. *Iyō Denshi to Seitai Kōgaku Japanese Journal of Medical Electronics & Biological Engineering* 24:8.

Pfurtscheller G, Neuper C, Andrew C, and Edlinger G. 1997. Foot and hand area mu rhythms. *International Journal of Psychophysiology Official Journal of the International Organization of Psychophysiology* 26:121.

Takens F. 1981. Detecting strange attractors in turbulence. *Springer Berlin Heidelberg*.

Taran S, and Bajaj V. 2017. Rhythm-based identification of alcohol EEG signals. *IET Science, Measurement & Technology* 12:343-349.

Teixeira AR, Tome AM, Lang EW, Gruber P, and Silva A. 2005. On the use of clustering and local singular spectrum analysis to remove ocular artifacts from electroencephalograms. IEEE International Joint Conference on Neural Networks.

Trujillo LT, Stanfield CT, and Vela RD. 2017. The Effect of Electroencephalogram (EEG) Reference Choice on Information-Theoretic Measures of the Complexity and Integration of EEG Signals. *Frontiers in Neuroscience* 11. 10.3389/fnins.2017.00425

Vautard R, Yiou P, and Ghil M. 1992. Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. *Physica D* 158:95-126.

Xu S, Hu H, Ji L, and Peng W. 2018. Embedding Dimension Selection for Adaptive Singular Spectrum Analysis of EEG Signal. *Sensors* 18:697.

Table 1(on next page)

The pseudo-code of the CSSA method for alpha rhythm extraction

Algorithm 1 CSSA Method for Alpha Rhythm Extraction

input s, $\boldsymbol{L}$: single-channel EEG time series **s** and embedding dimension L

output α: extracted alpha rhythm

procedures

(1) $\mathbf{X}$: the trajectory matrix is constructed by Eq. 1

(2) $\mathbf{C}_L$: the circulant matrix is built by Eqs. 2 and 3

(3) λ_k, $\mathbf{u}_k$: the circulant matrix $\mathbf{C}_L$ is decomposed, and a set of eigenvalues and eigenvectors is derived by Eq. 4

(4) $\mathbf{X}_{B_k}$: the real elementary matrices are derived by Eq. 7

(5) **α**: the RCs are grouped by Eq. 8 to obtain the alpha rhythm

return α

Table 2(on next page)

The parameters of the MPA EEG model

1 Table 1. MPA EEG model parameters

Symbol	Value	Comments
f_1 (Hz)	3.71	Delta rhythm
σ_1^{ξ}	3.53	
γ_1	0.98	
f_2 (Hz)	7.62	Theta rhythm
σ_2^{ξ}	4.35	
γ_2	0.95	
f_3 (Hz)	10.45	Alpha rhythm
σ_3^{ξ}	1.65	
γ_3	0.99	
f_4 (Hz)	15.43	Beta rhythm
σ_4^{ξ}	0.24	
γ_4	0.99	

2

Table 3(on next page)

Error of alpha rhythm extraction by the CSSA method with different embedding dimensions

Table 2. Error of alpha rhythm extraction by the CSSA method with different embedding dimensions

L	$\boldsymbol{\varepsilon(\mu V^2/\mathrm{Hz})}$	**L**	$\boldsymbol{\varepsilon(\mu V^2/\mathrm{Hz})}$	**L**	$\boldsymbol{\varepsilon(\mu V^2/\mathrm{Hz})}$	**L**	$\boldsymbol{\varepsilon(\mu V^2/\mathrm{Hz})}$
20	0.78	70	0.50	120	0.47	170	0.53
30	0.62	80	0.27	130	0.40	180	0.43
40	0.43	90	0.46	140	0.47	190	0.52
50	0.59	100	0.51	150	0.53	200	0.54
60	0.32	110	0.43	160	0.46		

Table 4(on next page)

Classification accuracy for all subjects of the CSSA, basic SSA, IIR, and WDec methods

Table 3. Classification accuracy for all subjects of the CSSA, basic SSA, IIR, and WDec methods.

Subject #	**CSSA**	**basic SSA**	**IIR**	**WDec**
Subject 1	80.22%	76.92%	79.12%	81.32%
Subject 2	95.60%	93.41%	95.60%	97.80%
Subject 3	96.70%	57.14%	94.51%	95.60%
Subject 4	95.60%	94.51%	95.60%	96.70%
Subject 5	98.90%	94.51%	98.90%	97.80%
Subject 7	96.70%	98.90%	94.51%	96.70%
Subject 8	100%	100%	100%	100%
Subject 9	81.32%	82.42%	83.52%	80.22%
Subject 10	100%	99%	100%	97.80%
Subject 11	96.70%	95.60%	96.70%	95.60%
Subject 12	76.92%	70.33%	74.73%	78.02%
Subject 13	100%	92.31%	100%	98.90%
Subject 14	87.91%	89.01%	87.91%	85.71%
Subject 15	95.60%	86.81%	95.60%	92.31%
Subject 16	86.81%	85.71%	86.81%	86.81%
Subject 17	92.31%	90.11%	91.21%	91.21%
Subject 18	95.60%	91.21%	97.80%	96.70%
Subject 19	83.52%	79.12%	79.12%	51.65%
Subject 20	98.90%	97.80%	98.90%	97.80%
Subject 21	94.51%	95.60%	93.41%	95.60%
Subject 22	85.71%	85.71%	85.71%	85.71%
Average	**92.36%**	**88.38%**	**91.89%**	**90.47%**
STD	**7.05%**	**10.35%**	**7.50%**	**10.88%**

Figure 1

Procedures of the spontaneous EEG simulation based on the MPA EEG model

(A) The power spectrum of a real EEG. The peak frequencies ( ), amplitude (B_i) and the frequency width (B_i) at half of amplitude of EEG rhythms were determined based on the power spectrum. (B) The simulated four rhythms: delta, theta, alpha and beta, based on the determined parameters. (C) The simulated spontaneous EEG generated by a combination of the four rhythms.

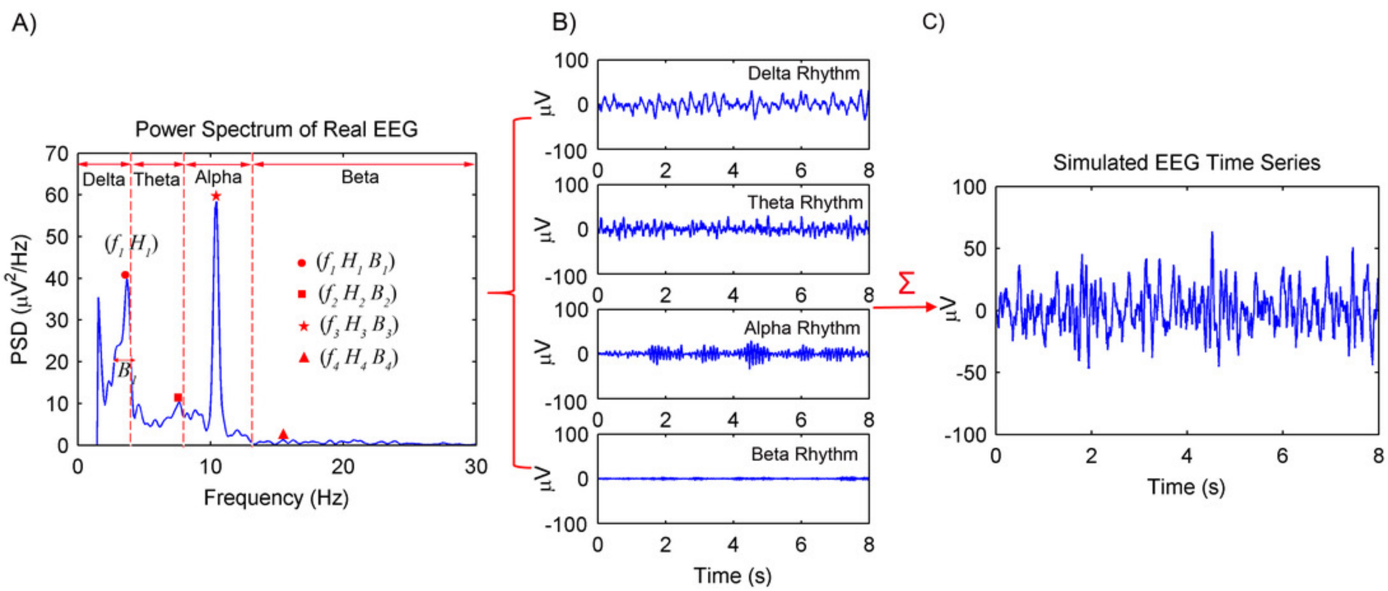
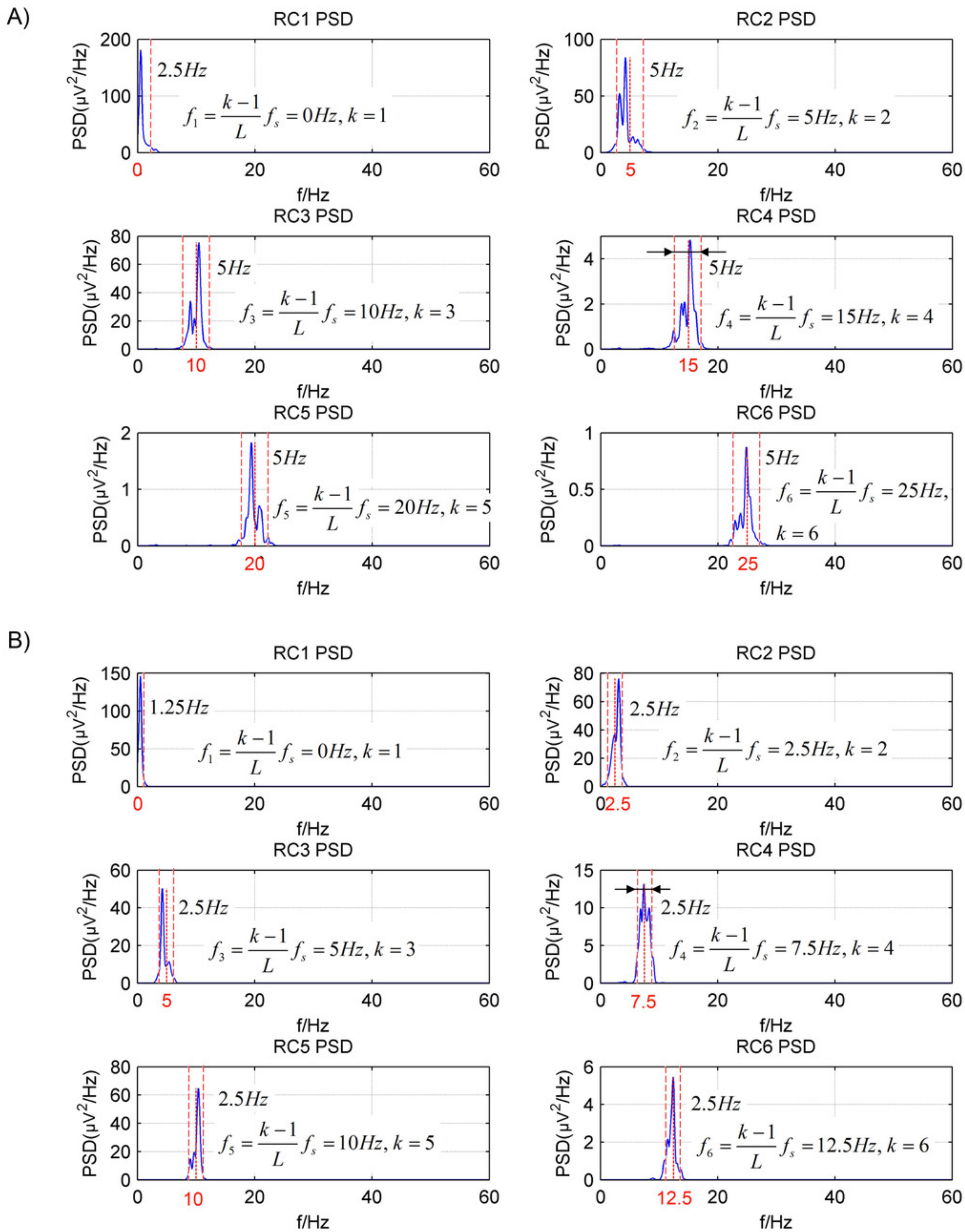


Figure 2

The power spectrum density of the first six reconstructed components of the simulated EEG signal processed by the CSSA method

(A) L=40; (B) L=80



A)
RC1 PSD
RC2 PSD
RC3 PSD
RC4 PSD
RC5 PSD
RC6 PSD
2.5Hz
5Hz
$f_1 = \frac{k-1}{L} f_s = 0Hz, k=1$
$f_2 = \frac{k-1}{L} f_s = 5Hz, k=2$
$f_3 = \frac{k-1}{L} f_s = 10Hz, k=3$
$f_4 = \frac{k-1}{L} f_s = 15Hz, k=4$
$f_5 = \frac{k-1}{L} f_s = 20Hz, k=5$
$f_6 = \frac{k-1}{L} f_s = 25Hz, k=6$
B)
1.25Hz
2.5Hz
$f_1 = \frac{k-1}{L} f_s = 0Hz, k=1$
$f_2 = \frac{k-1}{L} f_s = 2.5Hz, k=2$
$f_3 = \frac{k-1}{L} f_s = 5Hz, k=3$
$f_4 = \frac{k-1}{L} f_s = 7.5Hz, k=4$
$f_5 = \frac{k-1}{L} f_s = 10Hz, k=5$
$f_6 = \frac{k-1}{L} f_s = 12.5Hz, k=6$
PSD(μV²/Hz)
f/Hz

Figure 3

The power spectrum density of the first six reconstructed components of the simulated EEG signal processed by the basic SSA method with the embedding dimension L=40

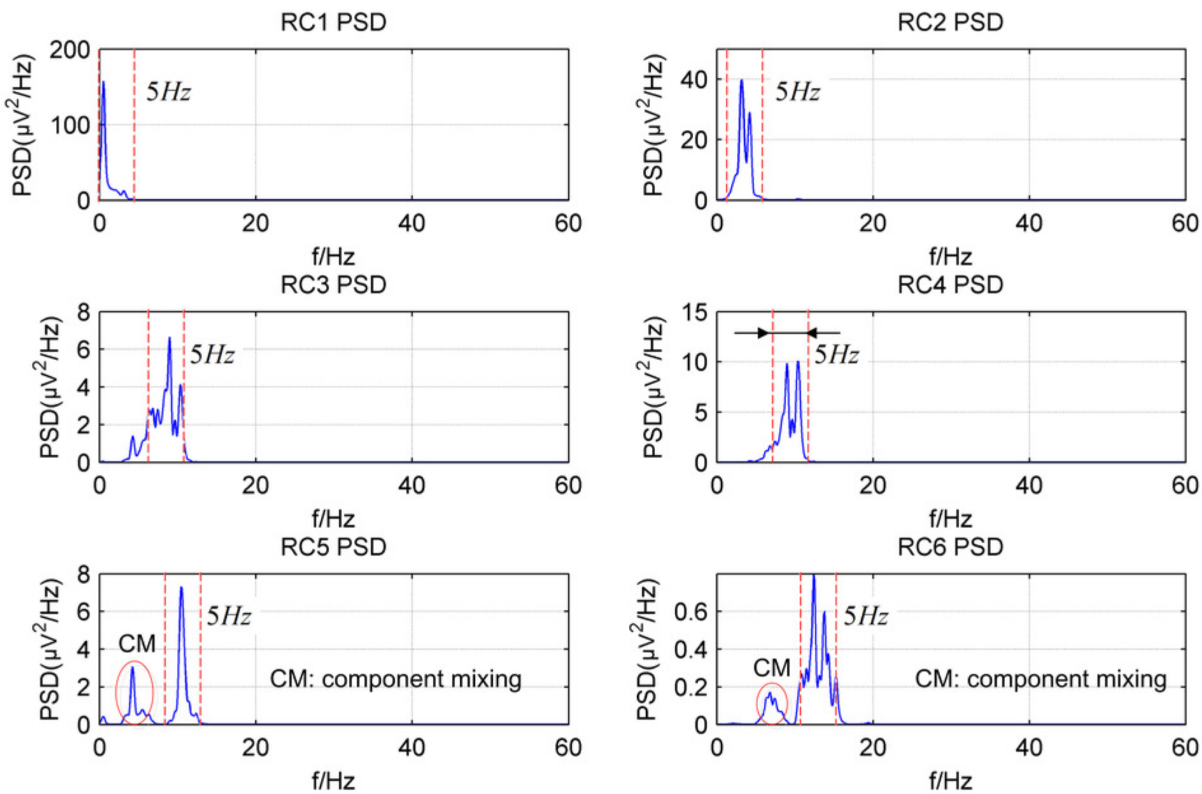


Figure 4

The extracted alpha rhythms of the simulated EEG signal and the PSD of the simulated and extracted alpha rhythms

(A) The extracted alpha rhythm of the simulated EEG signal by the CSSA method with the embedding dimension set to be L=40. RC3 represents the alpha rhythm. (B) The PSD of the simulated and extracted alpha rhythm by CSSA method when L=40. The slash shadow part is the component mixing and the error of extracted alpha rhythms is 0.43 uV^2/Hz . (C) The extracted alpha rhythm of the simulated EEG signal by the CSSA method with the embedding dimension set to be L=80. RC3 and RC4 represent the alpha rhythm. (D) The PSD of the simulated and extracted alpha rhythm by CSSA method when L=80. The error of extracted alpha rhythms is 0.27 uV^2/Hz . (E) The extracted alpha rhythm of the simulated EEG signal by the basic SSA method. RC3, RC4, RC5 and RC6 represent the alpha rhythm. (F) The PSD of the simulated and extracted alpha rhythm by the basic SSA method. The error of extracted alpha rhythms is 0.89 uV^2/Hz.

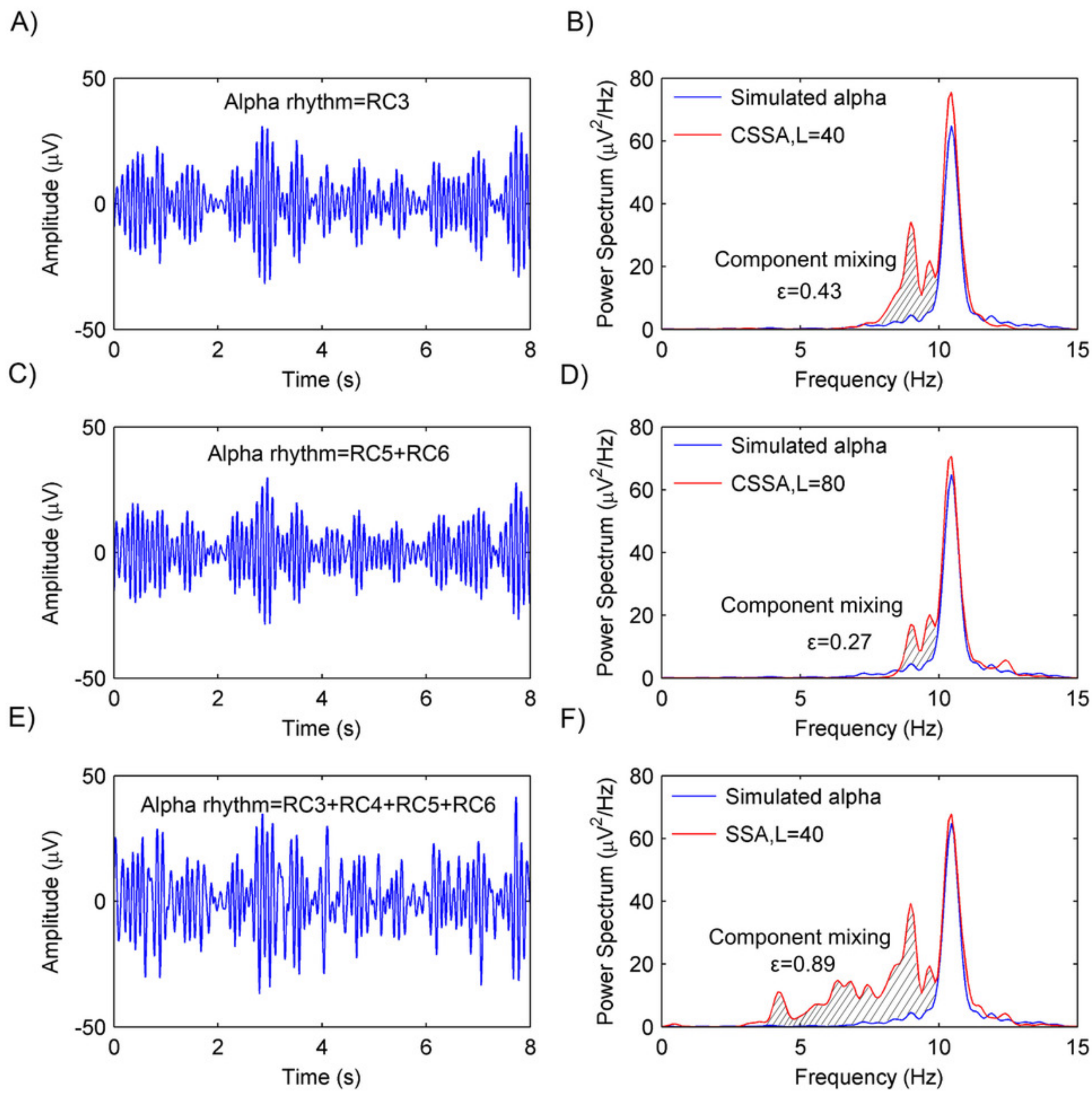

A)
B)
C)
D)
E)
F)
Alpha rhythm=RC3
Alpha rhythm=RC5+RC6
Alpha rhythm=RC3+RC4+RC5+RC6
Amplitude (μV)
Time (s)
Simulated alpha
CSSA,L=40
CSSA,L=80
SSA,L=40
Component mixing
ε=0.43
ε=0.27
ε=0.89
Power Spectrum (μV²/Hz)
Frequency (Hz)

Figure 5

The PSD of the extracted alpha rhythms by IIR and WDec method

(A) The PSD of the simulated alpha rhythm and the extracted alpha rhythms by CSSA and IIR. (B) The PSD of the simulated alpha rhythm and the extracted alpha rhythms by CSSA and WDec

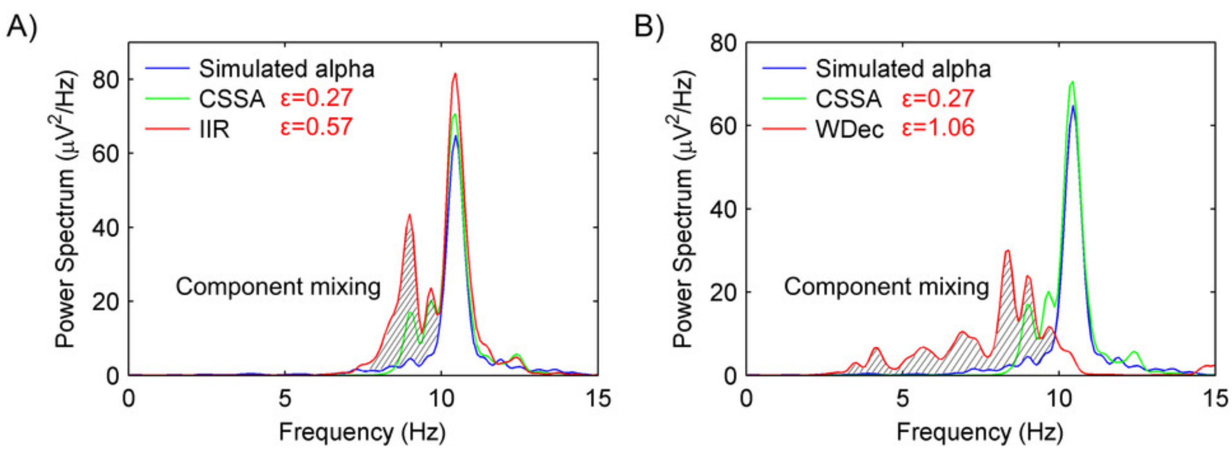


Figure 6

Artifacts removal of EEG signals

(A) A raw EEG epoch of subject 17# in eyes-open condition. (B) A raw EEG epoch of subject 17# in eyes-closed condition. (C) The corrected eyes-open EEG signal after artifact removal. (D) The corrected eyes-closed EEG signal after artifact removal

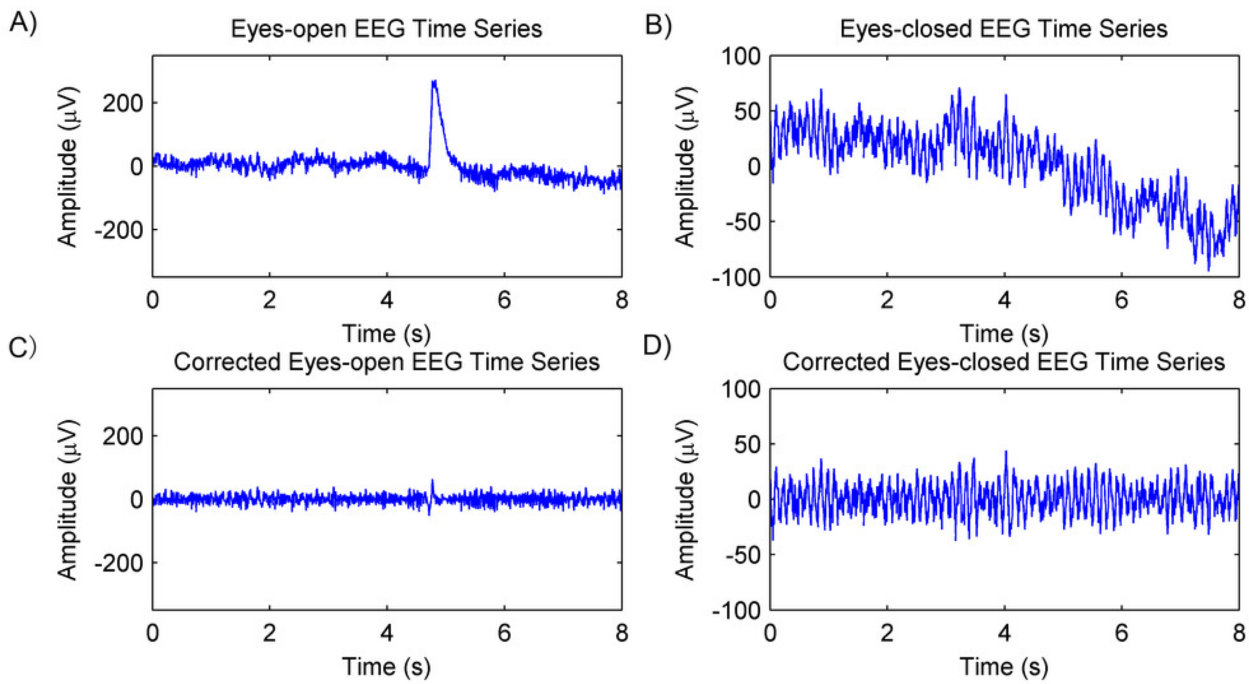


Figure 7

The PSD of first six RCs of real EEG signals processed by the CSSA method

(A) eyes-open condition and (B) eyes-closed condition

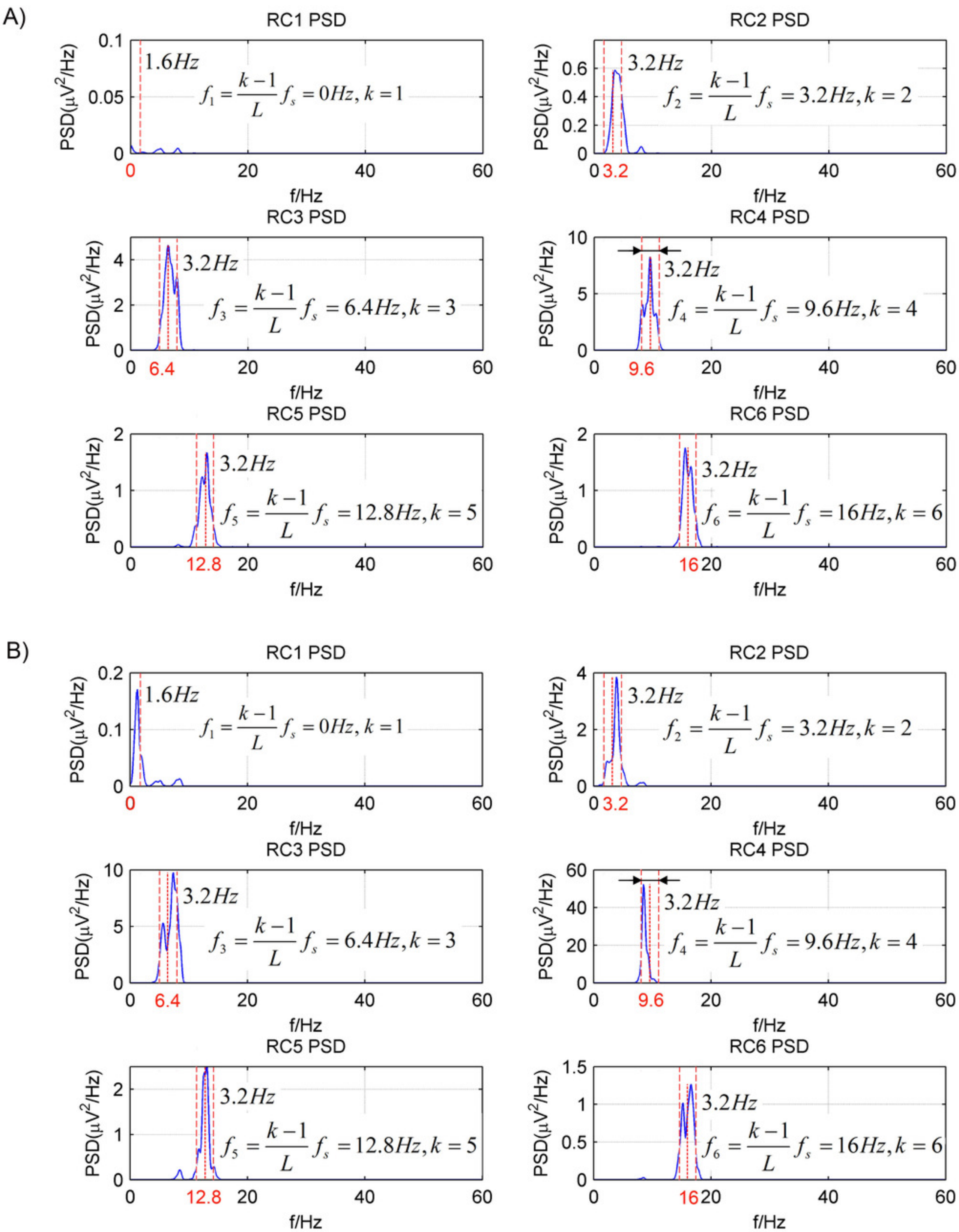

A)
RC1 PSD
1.6Hz
$f_1 = \frac{k-1}{L} f_s = 0Hz, k=1$
RC2 PSD
3.2Hz
$f_2 = \frac{k-1}{L} f_s = 3.2Hz, k=2$
RC3 PSD
3.2Hz
$f_3 = \frac{k-1}{L} f_s = 6.4Hz, k=3$
RC4 PSD
3.2Hz
$f_4 = \frac{k-1}{L} f_s = 9.6Hz, k=4$
RC5 PSD
3.2Hz
$f_5 = \frac{k-1}{L} f_s = 12.8Hz, k=5$
RC6 PSD
3.2Hz
$f_6 = \frac{k-1}{L} f_s = 16Hz, k=6$
PSD(μV²/Hz)
f/Hz
B)
RC1 PSD
1.6Hz
$f_1 = \frac{k-1}{L} f_s = 0Hz, k=1$
RC2 PSD
3.2Hz
$f_2 = \frac{k-1}{L} f_s = 3.2Hz, k=2$
RC3 PSD
3.2Hz
$f_3 = \frac{k-1}{L} f_s = 6.4Hz, k=3$
RC4 PSD
3.2Hz
$f_4 = \frac{k-1}{L} f_s = 9.6Hz, k=4$
RC5 PSD
3.2Hz
$f_5 = \frac{k-1}{L} f_s = 12.8Hz, k=5$
RC6 PSD
3.2Hz
$f_6 = \frac{k-1}{L} f_s = 16Hz, k=6$
PSD(μV²/Hz)
f/Hz

Figure 8

The extracted alpha rhythms of real EEG signals

(A) eyes-open and (B) eyes-closed condition and the (C) the spectrogram of alpha rhythms

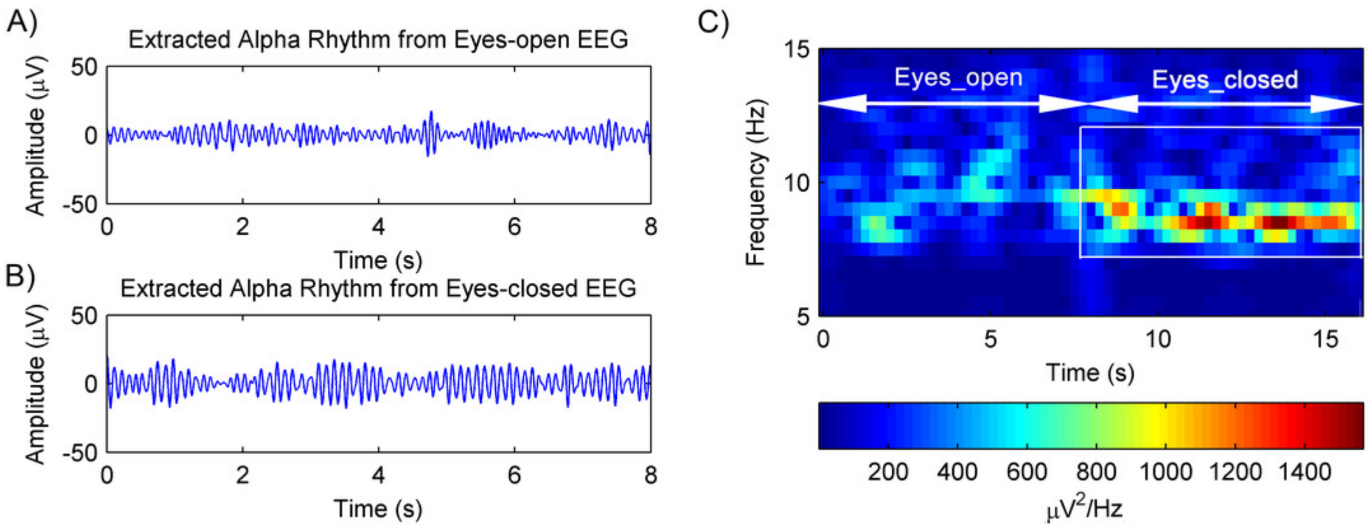


Figure 9

The PSD of the extracted alpha rhythms using the CSSA, basic SSA, WDec and IIR methods. The PSD of the extracted alpha rhythms using the CSSA and basic SSA method

(A) eyes-open and (D) eyes-closed conditions; The PSD of the extracted alpha rhythms using the CSSA and WDec method under (B) eyes-open and (E) eyes-closed conditions; The PSD of the extracted alpha rhythms using the CSSA and IIR method under (C) eyes-open and (F) eyes-closed conditions.

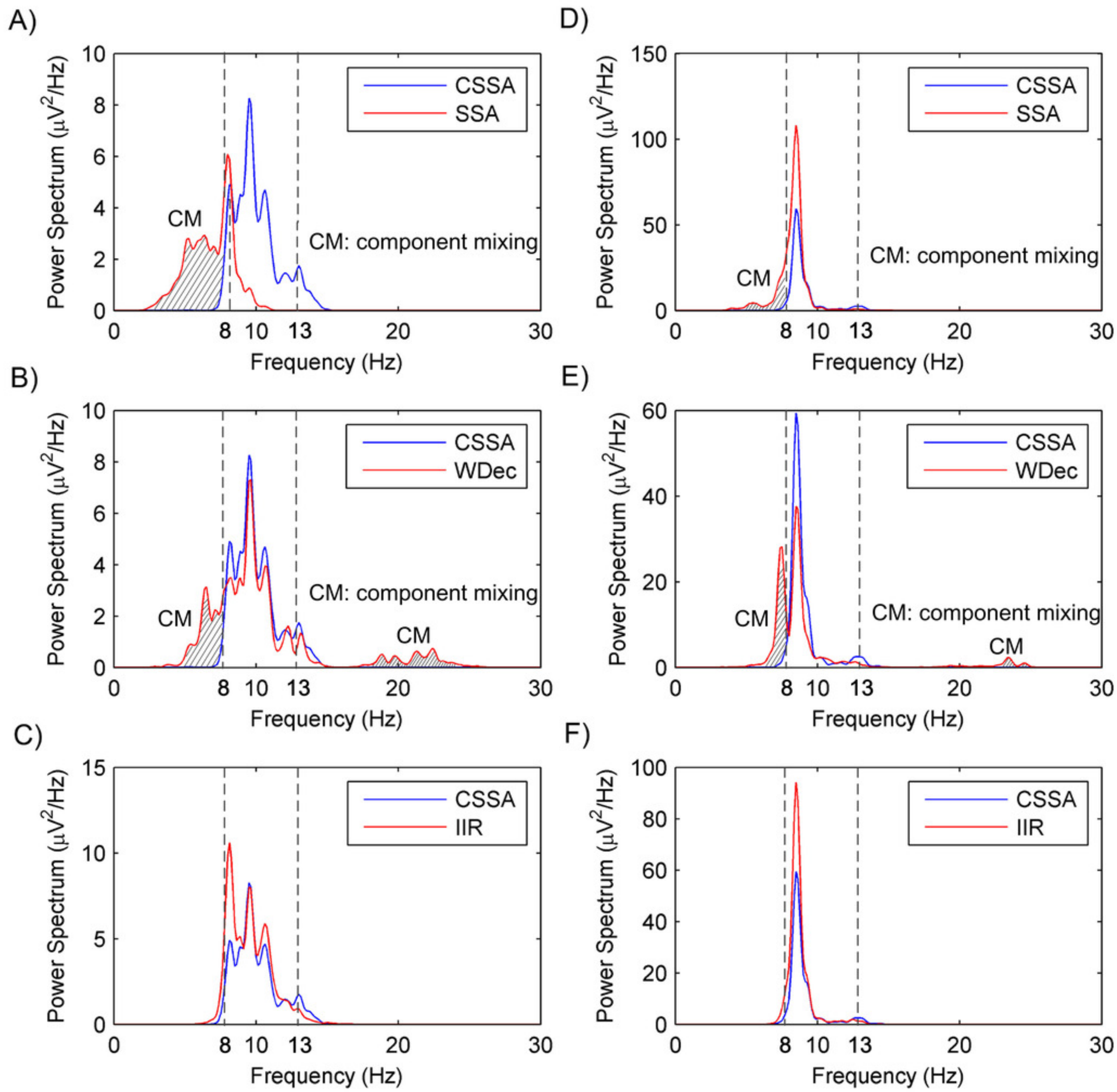

A)
Power Spectrum (μV²/Hz)
Frequency (Hz)
CSSA
SSA
CM
CM: component mixing
B)
CSSA
WDec
CM
CM: component mixing
CM
C)
CSSA
IIR
D)
CSSA
SSA
CM
CM: component mixing
E)
CSSA
WDec
CM
CM: component mixing
CM
F)
CSSA
IIR

Figure 10

Classification results for subject 17# between eyes-open and eyes-closed states

(A) the CSSA method, (B) the basic SSA method, (C) the IIR method and (D) the WDec methods

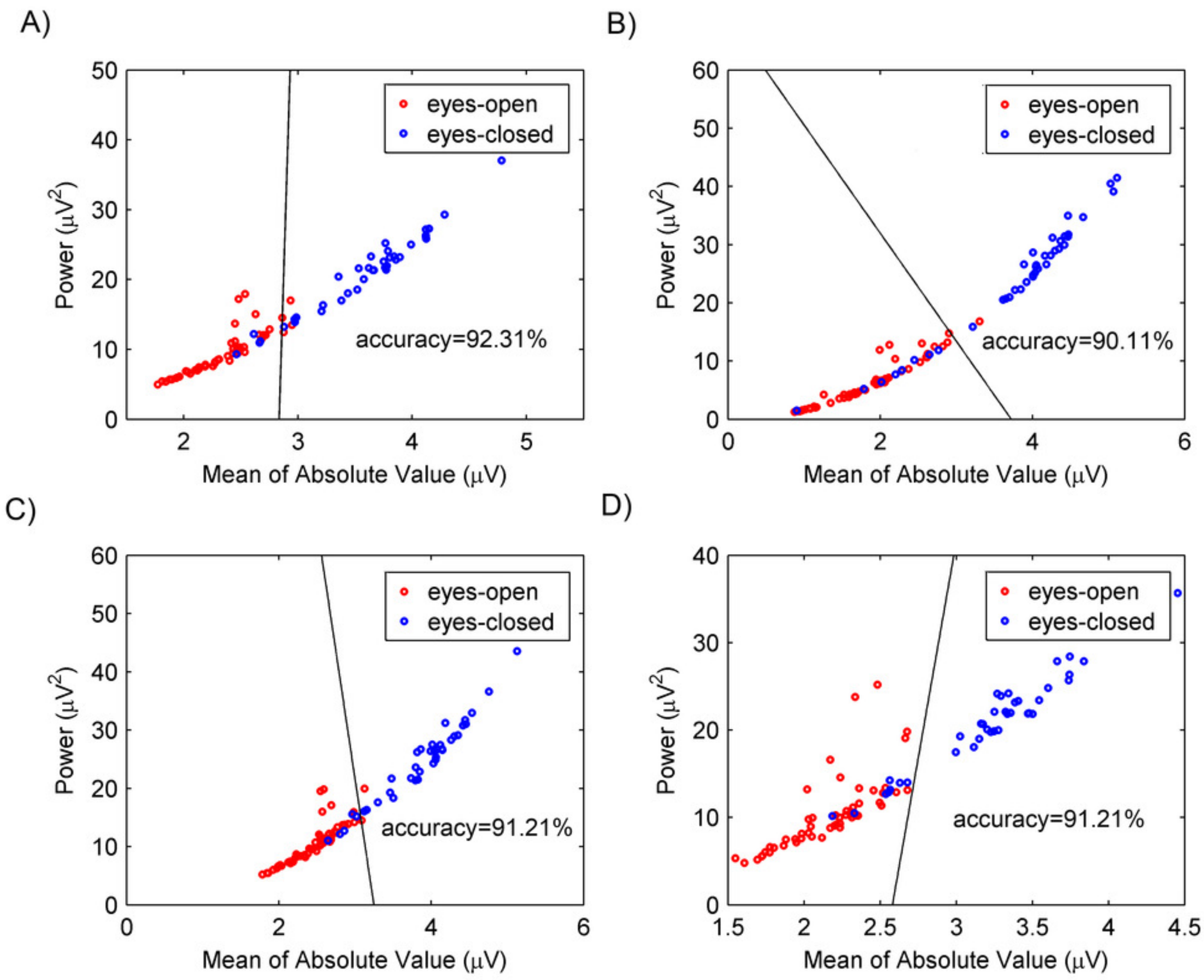


Figure 11

Set up of the experiment

(A) Schematic of the recorded EEG data and (B) The photograph of the experiment. 114 times of alternating periods of 30 s eyes open followed by 30 s eyes closed. The desired EEG segments were cut off in the middle of every period of the eyes-open and eyes-closed states. Each segment last for 8 s

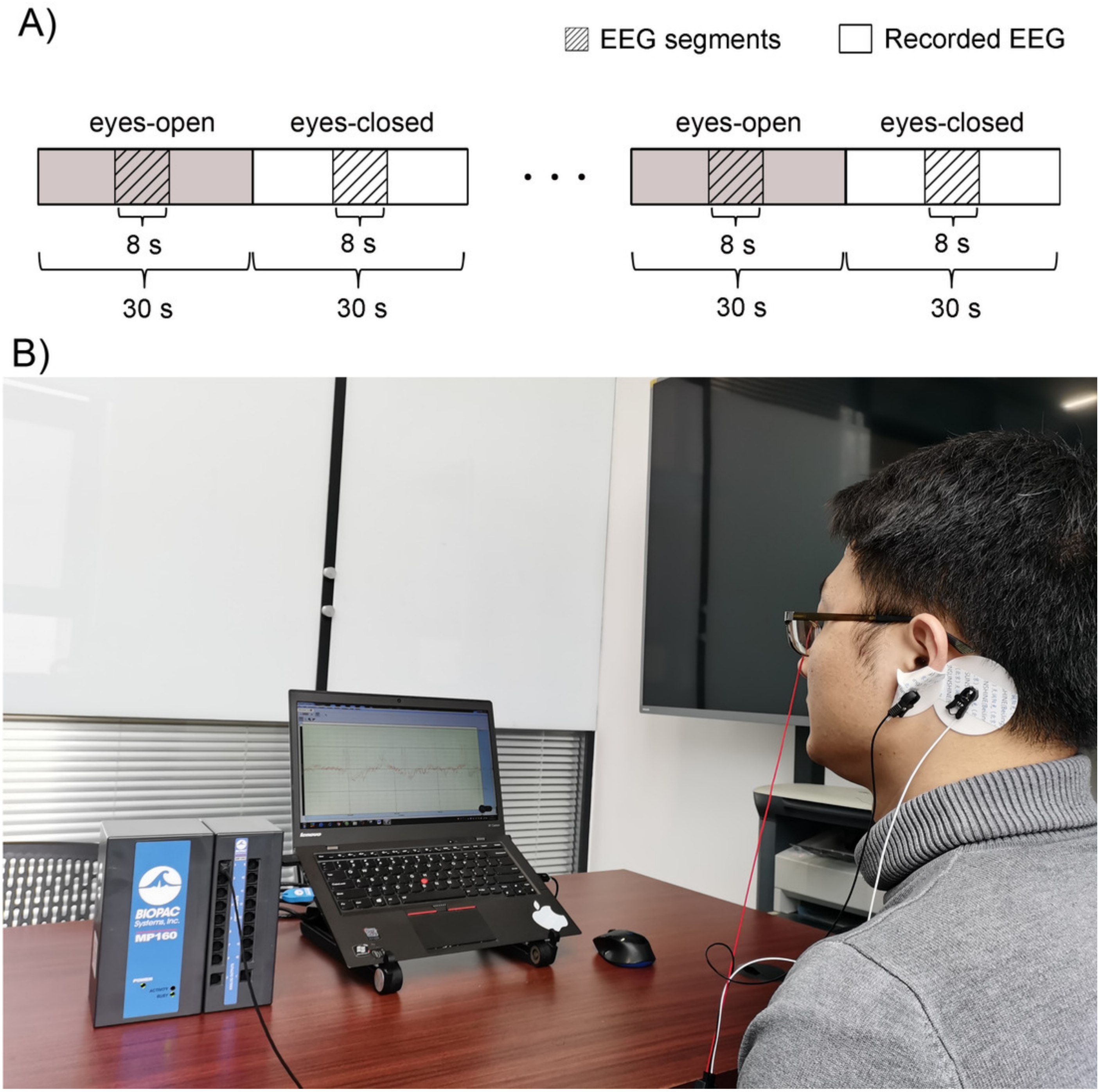

A)
EEG segments
Recorded EEG
eyes-open
eyes-closed
eyes-open
eyes-closed
8 s
8 s
8 s
8 s
30 s
30 s
30 s
30 s
B)
BIOPAC
Systems, Inc.
MP160

Figure 12

Classification results for experimental EEG signals between eyes-open and eyes-closed states

(A) the CSSA method, (B) the basic SSA method, (C) the IIR method and (D) the WDec methods