Journal of Neuroscience Methods 353 (2021) 109104

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

A hybrid method for muscle artifact removal from EEG signals

Qiang Chen [a,c], Yingying Li [b,c,*], Xiaohui Yuan [c]

[a] Department of Biomedical Engineering, Hefei University of Technology, Hefei, Anhui 230009, China
[b] School of Electronics and Information Engineering, Anhui Jianzhu University, Hefei 230601, China
[c] Department of Computer Science and Engineering, University of North Texas, Denton, TX 76203, USA

ARTICLE INFO

Keywords:
Variational mode decomposition (VMD)
Canonical correlation analysis (CCA)
Electroencephalogram (EEG)
Signal artifact

ABSTRACT

Background: Electroencephalogram (EEG) signals may be contaminated with muscle artifacts that are usually difficult to be removed.
New method: In this article, a new hybrid method for suppressing muscle artifacts is proposed. Our method leverages variational mode decomposition (VMD) and canonical correlation analysis (CCA) algorithms. Each channel of EEG is decomposed into intrinsic mode functions (IMFs) with VMD to achieve an extended data set that contains more channels than the original data set. The potential artifact components are decomposed by CCA for further isolation.
Results: The proposed method is evaluated with semi-simulation and real contaminated EEG signals. The results show that the performance of removing artifacts for VMD-CCA exceeds the comparison methods.
Comparison with existing methods: Regardless of the number of EEG channels and the signal-to-noise ratio of the signal, the VMD-CCA approach is superior to the existing methods. As the number of EEG channels decreases, the average de-artifact effects of VMD-CCA and the comparison approaches are basically the same, but the randomness increases.
Conclusions: The VMD-CCA method can effectively isolate muscle artifacts in EEG in case of multiple channels or few channels.

1. Introduction

Electroencephalogram (EEG) signal is often contaminated by non-cerebral artifacts, e.g. muscle artifacts, that cause confusion in the analysis and interpretation process (Song and Sepulveda, 2018; Ren et al., 2019; Lopez-Larraz et al., 2018; Nathan and Contreras-Vidal, 2016). Muscle artifacts exhibit a high amplitude with variable topographical distributions, and span almost the entire spectrum of the EEG signals (Goncharova et al., 2003; McMenamin et al., 2011), which makes muscle artifact removal a challenging problem.

Blind source separation (BSS) techniques, such as independent component analysis (ICA) (Goncharova et al., 2003; Urigueen and Garcia-Zapirain, 2015; Minguillon et al., 2017; Mannan et al., 2018; Jiang et al., 2019) or canonical correlation analysis (CCA) (Nam et al., 2002; De Clercq et al., 2006), extracts components from EEG for artifact isolation. Yet, a clean separation of cerebral and non-cerebral components is still issue (McMenamin et al., 2011). In many real-world applications (Chaudhary et al., 2016; Minguillon et al., 2017), EEG equipments with a small number of electrodes or single electrode are often used due to cost and complexity concerns. In such settings, the multichannel decomposition methods face challenges because the number of underlying signal sources (e.g., muscle groups) are often more than the number of EEG channels.

Alternatively, single-channel decomposition methods separate each channel of EEG signals into several components. Using wavelet transform, EEG signals are decomposed into time-frequency representations and the artifact-like coefficients are suppressed using thresholding (Govindan et al., 2014; Yong et al., 2012). While the empirical mode decomposition (EMD) (Huang et al., 1999) and the ensemble EMD (EEMD) (Wu and Huang, 2009) require no pre-set basis function and are therefore suitable for non-stationary signal analysis.

Hybrid methods have been proposed to exploit the advantage of different methods (Mannan et al., 2018; Jiang et al., 2019). For single-channel EEG contaminated by muscle artifacts, the EEMD-ICA (Mijovic et al., 2010) approach combining EEMD and ICA demonstrated an improved performance in muscle artifacts suppression. Similar approaches, such as EEMD-CCA (Sweeney et al., 2013), further improve denoising effect. In multichannel situation (Zeng et al., 2016;

* Corresponding author at: School of Electronics and Information Engineering, Anhui Jianzhu University, Hefei 230601, China.
E-mail addresses: chenqiang@hfut.edu.cn (Q. Chen), yyhgd0520@gmail.com (Y. Li), xyuan@cse.unt.edu (X. Yuan).

https://doi.org/10.1016/j.jneumeth.2021.109104
Received 14 July 2020; Received in revised form 10 February 2021; Accepted 13 February 2021
Available online 20 February 2021

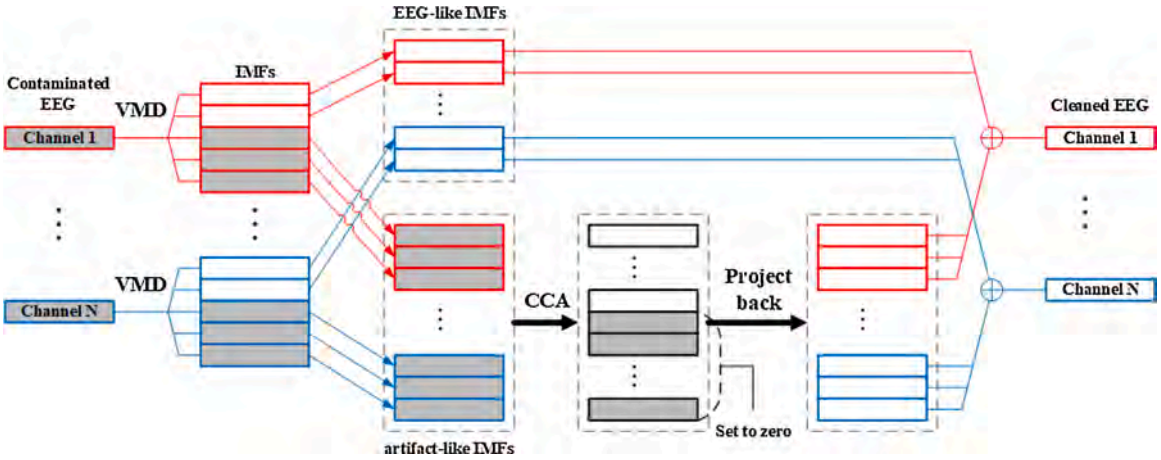

Fig. 1. Flowchart for the proposed VMD-CCA approach. The gray rectangles indicate the components suspected of containing artifacts, and the white rectangles indicate no artifacts.

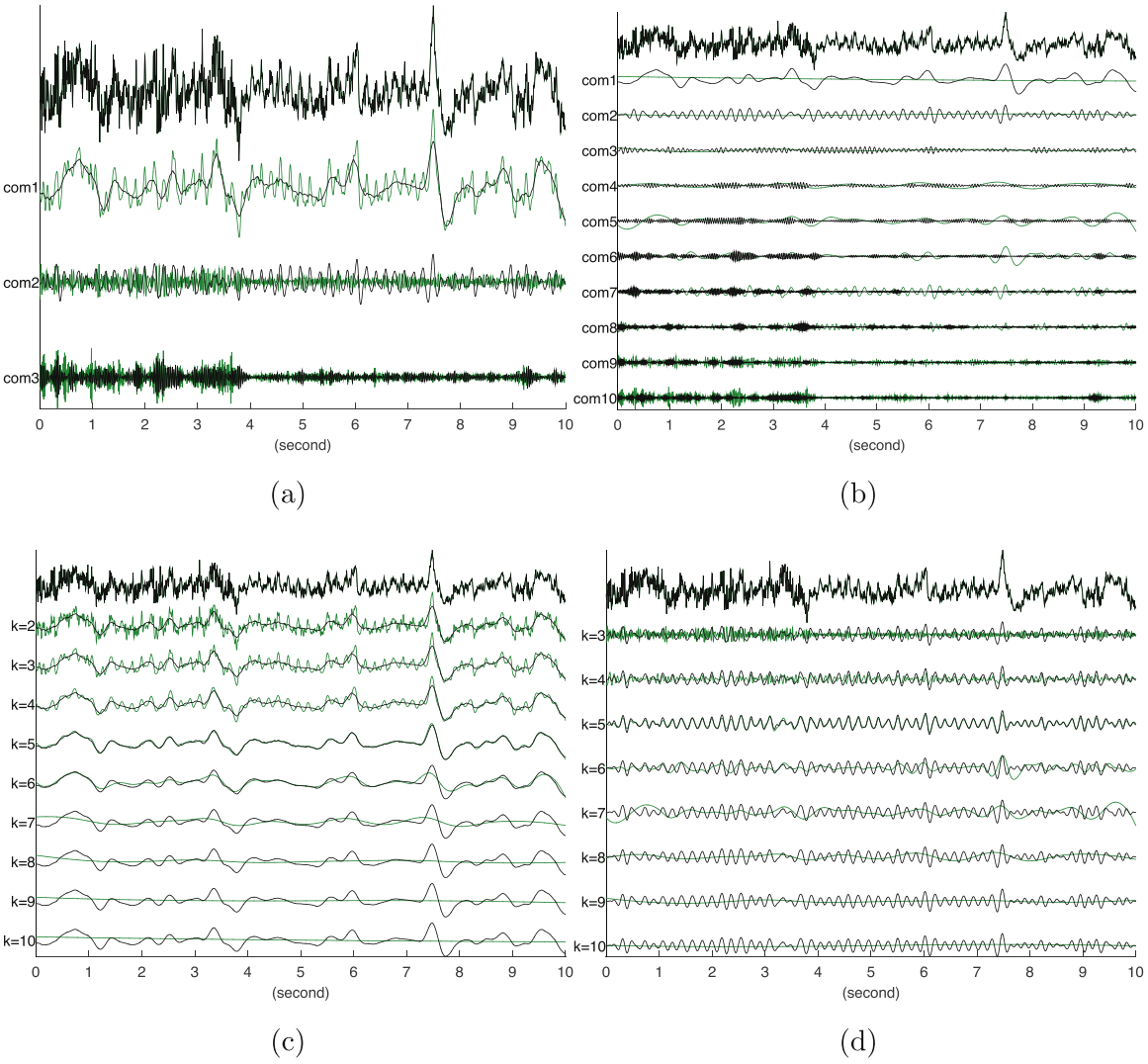

Fig. 2. A demo of comparison of VMD and EEMD decomposition effects. (a) A EEG signal is decomposed into 3 components by VMD and EEMD respectively (black: VMD components, green: EEMD components, the same below), (b) the same signal is decomposed into 10 components, (c) decompose the signal into 2, 3, …, 10 components respectively, and display the first component obtained by each decomposition, (d) decompose the signal into 3, 4, …, 10 components respectively, and display the second component obtained by each decomposition.

Safi et al., 2018; Chen et al., 2019), every channel is decomposed into several components, then all or artifact-like components from all channels are put together as the input for the following BSS procedure. EEG and muscular sources are further separated in BSS. Finally, those cerebral-related are used to reconstruct pure EEG.

However, EMD/EEMD lacks a solid mathematical foundation. Variational mode decomposition (VMD) (Dragomiretskiy and Zosso, 2014) is an alternative to EMD, which can adaptively decompose the signal into some limited bandwidth and approximately orthogonal IMFs. It has been proved that VMD is better than EMD in signal decomposition and noise robustness (Wang and Markert, 2016). In this paper, we propose a hybrid approach that leverage VMD and CCA. In the VMD-CCA approach, each channel of EEG is decomposed with intrinsic mode functions (IMFs) for enriched data set. The data consisting of artifacts is decomposed using CCA, and the artifact components are removed. In the artifact removal step, information between channels are used. A clean EEG is reconstructed from the non-artifact components.

The organization of the rest of this paper is as follows. Section 2 presents the proposed hybrid approach. Section 3 details the performance of VMD-CCA using a comparison study on both synthetic and real-life signals. Section 4 concludes the paper with a summary.

2. Methods

Fig. 1 illustrates the architecture of the proposed hybrid approach, which consists the following procedures: (1) decompose each EEG channel into IMFs, denoted as $\mathbf{F}_c$, using VMD and select IMFs with artifact, $\mathbf{F}_{c\text{Artifact}}$, using autocorrelation thresholding; (2) obtain a cleaned $\widetilde{\mathrm{Y}}$ with CCA and replace each $\mathbf{F}_{c\text{Artifact}}$ with the corresponding components in $\widetilde{\mathrm{Y}}$ to get cleaned IMFs, denoted as $\mathbf{F}_{c\text{Clean}}$, then reconstruct an artifact-free EEG by integrating $\mathbf{F}_{c\text{Clean}}$.

2.1. Signal decomposition and artifact concentration

2.1.1. VMD

VMD is a data-driven algorithm, which decomposes a signal $f(t)$ into a given number of IMFs, or modes u_k, and each mode has limited bandwidth with different central frequencies (Dragomiretskiy and Zosso, 2014).

In VMD, the problem is described as the sum of the estimated bandwidths of the modes being minimized under the constraint that the sum of the modes is equal to the original signal $f(t)$. The constrained variational formulation of VMD can be expressed as

$$\min_{\{u_k\},\{\omega_k\}} \left\{ \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \text{s.t.} \quad \sum_{k=1}^{K} u_k = f \tag{1}$$

where $\delta(t)$ denotes the Dirac distribution, * and ∂ denote the convolution and partial differential operators, and u_k and ω_k denote the kth ($k = 1, 2, \ldots, K$) mode of VMD and corresponding center frequency respectively.

The solution is obtained as the saddle point of the augmented Lagrangian $\mathcal{L}$ as follows:

$$\mathcal{L}(\{u_k\},\{\omega_k\},\lambda) = \alpha \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_{k=1}^{K} u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_{k=1}^{K} u_k(t) \right\rangle \tag{2}$$

where α is the penalty term. To find the saddle point of $\mathcal{L}$, alternate direction method of multipliers (ADMM) is used. The optimal $u_k(\omega)$ is directly updated by Wiener filtering in Fourier domain.

2.1.2. Over-segmentation problem

When using single-channel decomposition technology, a common problem is how many components should the signal be decomposed into. The number of decomposed components in the VMD, denoted as K in Eq. (1), needs to be predefined. Obviously, if the K is too small, it is not conductive to separation of signal and artifact. But if the K is too large, so-called over-segmentation, what will be impact in addition to increasing the calculation load? Fortunately, a good feature of the VMD is that when the signal is over segmented, the redundant modes are basically noise components (Dragomiretskiy and Zosso, 2014).

Fig. 2 shows the comparison of the decomposition effect of VMD and EEMD on a sample EEG signal, which contains obvious muscle artifact and may contain other noises, shown in the first line of Fig. 2. As shown in Fig. 2a, when the signal is decomposed by VMD, the first two components are low-frequency component and quasi-periodic signal respectively, and the third one is high-frequency component. In EEMD decomposition result, the second and third components are all high frequency signals. Fig. 2b shows the same signal decomposed into 10 components by VMD and EEMD respectively. In the decomposition result of VMD, the first component is still a low frequency term, the other components increase in frequency in turn, and the latter components are definitely high frequency artifacts or noise in terms of frequency. While the first six components of EEMD are all in low frequency, which means that they are the over decomposition of EEG baseline. Meanwhile the component 9 and 10 are obviously high frequency noise. This indicates that with the same K, the components belonging to the EEG band in the VMD result is more than the result of EEMD.

We use VMD and EEMD to decompose signals into 2, 3, …, 10 components respectively, then list all the first component obtained from each decomposition, as shown in Fig. 2c. It can be seen that although the K varies, all the components obtained by VMD are almost the same. In contrast, the first component obtained by EEMD is quite different when the K changes. The second components obtained by VMD and EEMD also have the same properties, as shown in Fig. 2d. This indicates that the shape of the first few components of VMD decomposition are relatively stable. So we don't need to worry the over decomposition of VMD.

Through experiments, we found that pretty results can be achieved if $K \geq 5$.

2.1.3. Muscle artifact judgment

Compared to EEG, muscle artifacts are similar to random noise and their spectrums are wider (Goncharova et al., 2003; McMenamin et al., 2011), so their delay-1 autocorrelation coefficients are lower than those of EEG. In order to detect muscle artifacts, the threshold is set relative higher here (0.95 for 10-s segment empirically) (Chen et al., 2019). In the VMD and the following CCA step, any IMF with a delay-1 autocorrelation coefficient below the threshold is considered to contain muscle artifacts.

2.2. Artifact isolation and signal reconstruction

CCA, based on second order statistics (SOS), is a well-known technique to search underlying correlation between two multidimensional signals (Hardoon et al., 2004; De Clercq et al., 2006). The goal of CCA is to find the basis vectors of two data sets so that the correlation between the projections of the variables on the basis vectors are maximized with each other. Let $\mathrm{Y}_1(t)$ be the EEG signal and $\mathrm{Y}_2(t) = \mathrm{Y}_1(t-1)$ be a temporally delayed version of $\mathrm{Y}_1(t)$. CCA solves the BSS problem by obtaining two projection vectors, $\mathbf{w}_1$ and $\mathbf{w}_2$, which make the correlation between the projections of $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are mutually maximized (De Clercq et al., 2006). The objective function is

$$max_{\mathrm{w}_1,\mathrm{w}_2} \frac{\mathrm{w}_1^T \boldsymbol{\Sigma}_{12} \mathrm{w}_2}{\sqrt{\mathrm{w}_1^T \boldsymbol{\Sigma}_{11} \mathrm{w}_1} \sqrt{\mathrm{w}_2^T \boldsymbol{\Sigma}_{22} \mathrm{w}_2}} \tag{3}$$

where $\boldsymbol{\Sigma}_{12}$ is the crosscovariance matrix of $\mathbf{Y}_1$ and $\mathbf{Y}_2$, and $\boldsymbol{\Sigma}_{11}$ and $\boldsymbol{\Sigma}_{22}$ the autocovariance matrices of $\mathbf{Y}_1$ and $\mathbf{Y}_2$ respectively.

The weight vectors and their corresponding canonical variates can be derived from the optimization function (3) (Hardoon et al., 2004).

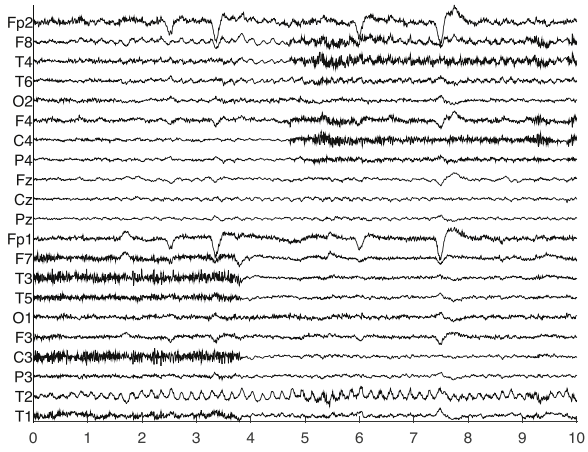

Fig. 3. A real-life ictal EEG recording contaminated with muscle artifacts. The horizontal axis represents time with unit second.

Projections onto $\mathbf{w}_1$ and $\mathbf{w}_2$ are called canonical variants, which are maximally correlated and the canonical variates within each data set are mutually uncorrelated. CCA decomposes an EEG signal into several mutually uncorrelated components, which are arranged in descending order according to the autocorrelation values.

Due to the low autocorrelation of muscle artifacts, components with autocorrelation value below a threshold can be considered to contain muscle artifacts, then these components are set to zero. Then, according to the inverse process of the previous signal decomposition process, the non-artifact EEG signal can be reconstructed, as shown in Fig. 1.

3. Experiment results

3.1. Datasets and experiment settings

3.1.1. Semi-simulation datasets

To evaluate the performance of the proposed VMD-CCA, we created semi-simulation data sets with real clean EEG and real electromyogram (EMG) signals.

Muscle artifact-free EEG data were recorded from 20 healthy subjects. The 19-channel EEG data sets were recorded from an EEG QuickCap and a NuAmps amplifier (Compumedics Neuroscan, El Paso, TX), with the sampling rate 500 Hz and a bandpass filter in the range of 1 and 70 Hz. Twenty 10-s muscle artifact-free EEG epochs were selected through visual inspection by an independent neurophyiologist. These clean EEG data sets, denoted as $\mathrm{X}_{\mathrm{EEG}}^{(i)}$ ($i = 1, 2, \ldots, 20$), are used for subsequent data generating.

It is difficult to extract pure muscle activity from contaminated EEG, thus real EMG signals are used in this work. Ten-second EMG data segments were collected from healthy volunteers by placing electrodes on forearms with a Trigno wireless surface EMG system (Delsys Inc., Natick, MA). To simulate realistic situations, we collected not only continuous muscle artifacts, but also transient ones by controlling the contraction intervals. To simulate massive muscle artifacts, a 19-channel muscle artifacts data matrix, denoted as $\mathrm{X}_{\mathrm{EMG}}^{(i)}$ is generated with a random mixing matrix and a 19-channel EMG source matrix (Chen et al., 2016), which is randomly chosen from the EMG data set. As mentioned above, the multi-channel BSS step of the algorithm framework separates EEG and muscle artifacts from suspected artifact-like components to retrieve useful information as much as possible. ICA, CCA, or other BSS algorithms need to use some correlation between signal channels.

A contaminated EEG matrix can be obtained according to the following:

$$\mathrm{X}_i = \mathrm{X}_{\mathrm{EEG}i} + \lambda \mathrm{X}_{\mathrm{EMG}i} \tag{4}$$

where the parameter λ is used to adjust the signal-to-noise (SNR) of $\mathbf{X}_i$. The meaning of SNR is

$$\mathrm{SNR} = \frac{\mathrm{RMS}(\mathrm{X}_{\mathrm{EEG}})}{\mathrm{RMS}(\lambda \mathrm{X}_{\mathrm{EMG}})} \tag{5}$$

where the root mean squared (RMS) is defined in equation. (7).

A demo of semi-simulation data set is shown in Fig. 4a and b.

3.1.2. Real dataset

A real-life EEG data set is also employed to evaluate the performance of these approaches. The ictal EEG data set is a 21-channel, 10-s scalp EEG recording from a long-term epilepsy monitoring unit, with sampling rate of 250 Hz and bandpass filtering at the range of 0.3–35 Hz. Obvious epileptic waveform can be observed in the channels near the right temporal lobe, including T2, T4, T6 and F8. But the ictal EEG is severely contaminated with muscle activity, as shown in Fig. 3. The data set can be found at http://www.esat.kuleuven.be/sista/members/biomedng/biosource.htm.

3.2. Evaluation metrics

In synthesized data processing, the ground truth $\mathbf{X}_{\mathrm{EEG}}$ is known. For quantitative evaluation of algorithms performance, the relative root-mean-squared error (RRMSE) is used as an evaluation index of muscle artifact rejection, defined as

$$\mathrm{RRMSE} = \frac{\mathrm{RMS}(\mathrm{X}_{\mathrm{EEG}} - \widetilde{\mathrm{X}}_{\mathrm{EEG}})}{\mathrm{RMS}(\mathrm{X}_{\mathrm{EEG}})} \tag{6}$$

where the $\widetilde{\mathrm{X}}_{\mathrm{EEG}}$ is the processed EEG data, the **RMS** is defined as

$$\mathrm{RMS}(\mathrm{X}) = \sqrt{\frac{1}{C \cdot T} \sum_{c=1}^{C} \sum_{t=1}^{T} \mathrm{X}^2(c,t)} \tag{7}$$

where C denoting the number of EEG channels and T the number of time samples.

3.3. Comparison methods

3.3.1. Decomposition with EMD & EEMD

The EMD is a technique which reduce signal $f(t)$ into a sum of several subsignals, called IMFs (Huang et al., 1999). Firstly, EMD extracts all the local maxima and minima of the signal. Secondly, the upper and lower envelopes are produced by cubic spline interpolation of the maxima and minima respectively. Thirdly, the mean of two envelopes, i.e. the local trend, is substracted from the initial signal to obtain a IMF. To obtain the next IMF, new maxima and minima shall be extracted from the local trend and all above three steps will be repeated until the last local trend satisfies stopping criterion. $f(t)$ is expressed as follows:

$$f(t) = \sum_{j=1}^{N} \mathrm{IMF}_j(t) + r_N(t) \tag{8}$$

where N is the number of IMFs and $r_N(t)$ is the final residue. All extracted IMFs obey two properties: (a) the number of local minima and maxima differs at most by one, and (b) the mean value of IMF is zero.

EEMD (Wu and Huang, 2009), a known enhanced version of the standard EMD, was proposed to overcome the shortcoming of sensitivity of the original EMD respect to noise. In EEMD, the IMFs are obtained from the average of corresponding IMFs estimated by original EMD from an ensemble of the signal plus independent white noise. Unlike Fourier or wavelet transform, the EMD/EEMD decomposition is completely data-driven without any pre-set basis functions, and it is widely applied in nonlinear and non-stationary signal analysis.

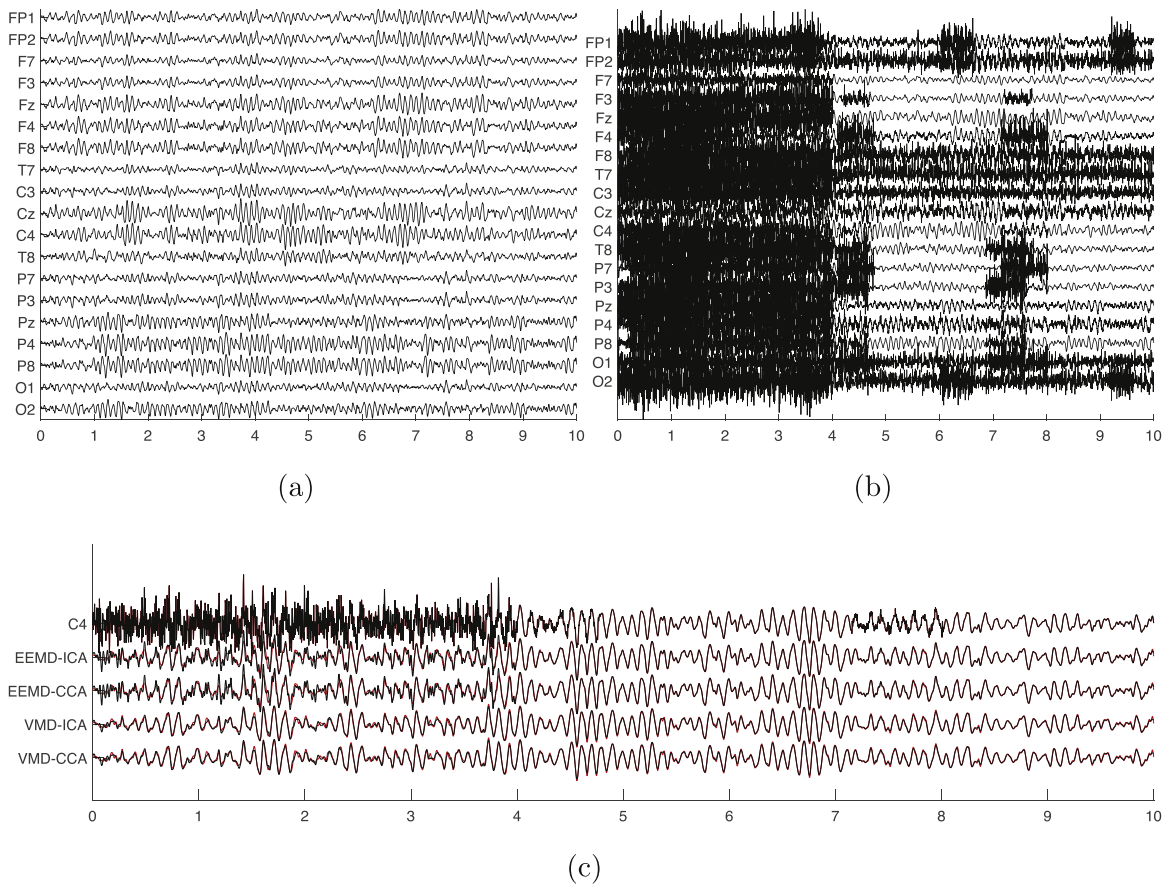

Fig. 4. Example of 19-channel semi-simulation EEG. (a) Real clean EEG, (b) semi-simulation EEG (SNR = 0.5), (c) channel C4 and comparison of the results of four methods. The black lines are the reconstructed signals, and the red ones are clean EEG of channel C4.

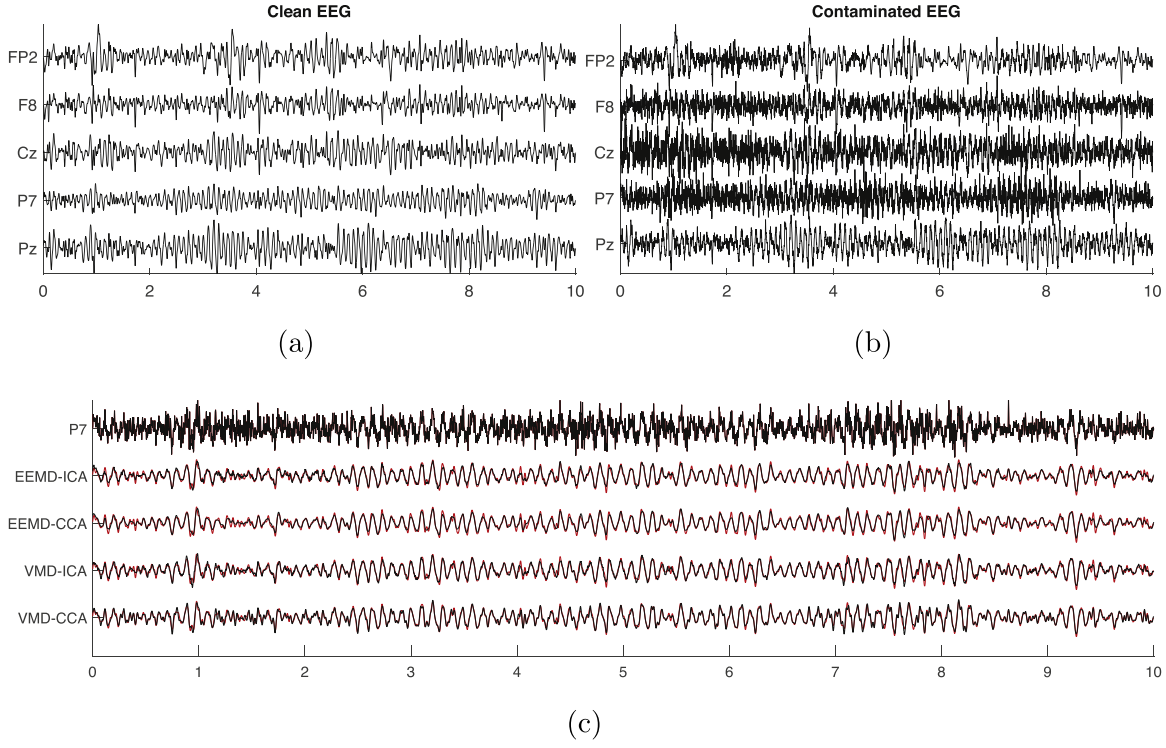

Fig. 5. Example of 5-channel semi-simulation EEG. the signal is composed of 5 channels randomly selected from the signal shown in Fig. 4. (a) Real clean EEG, (b) semi-simulation EEG, (c) channel P7 and comparison of the results of four methods. The black lines are results and red ones are clean EEG.

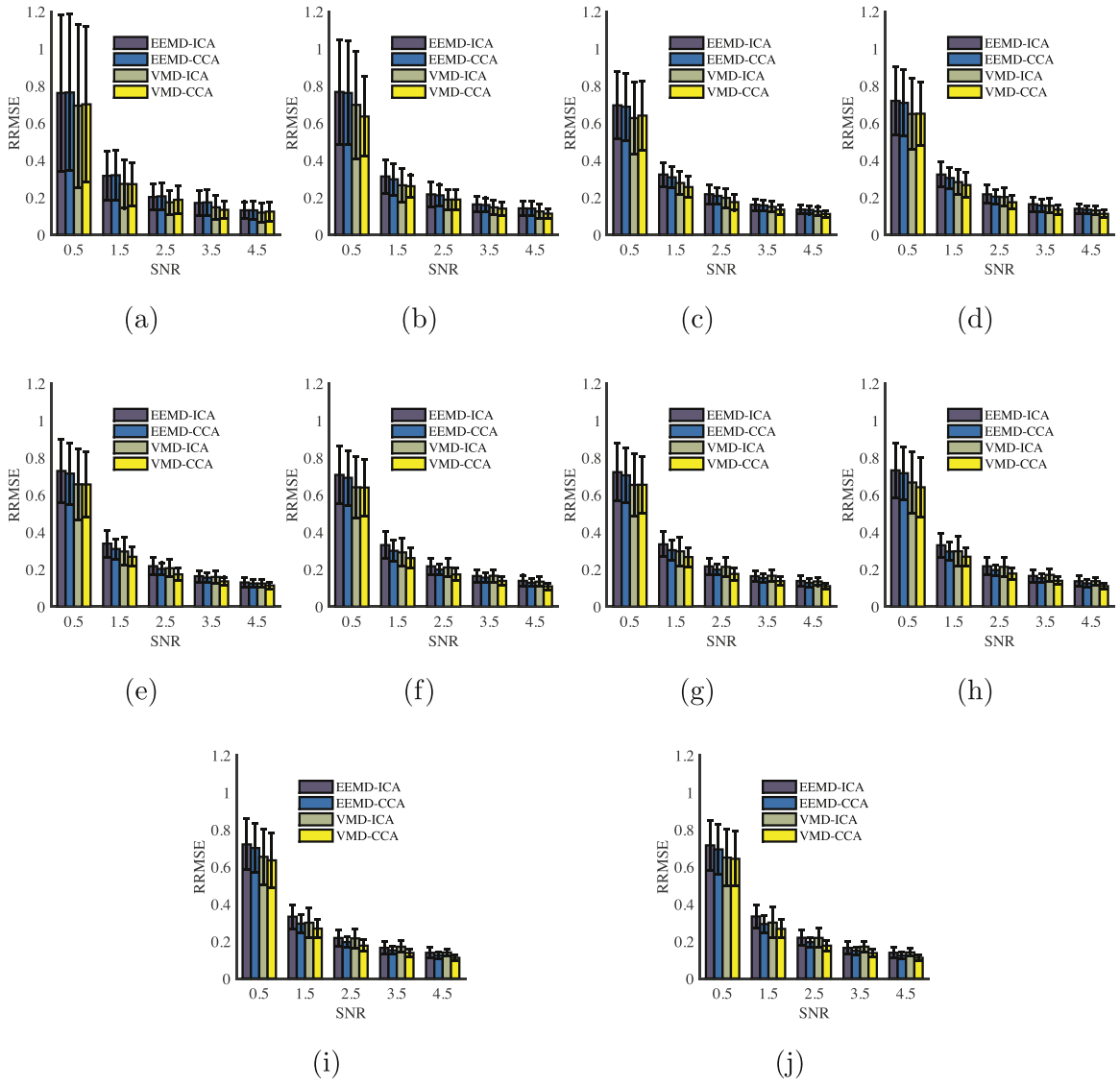

Fig. 6. The semi-simulated study: performance comparison of EEMD-ICA, EEMD-CCA, VMD-ICA and VMD-CCA at various SNR values in terms of RRMSE. C channels randomly chosen, (a) $C = 1$, (b) $C = 3$, (c) $C = 5$, (d) $C = 7$, (e) $C = 9$, (f) $C = 11$, (g) $C = 13$, (h) $C = 15$, (i) $C = 17$, (j) $C = 19$.

3.3.2. *ICA for multichannel EEG*

The ICA approaches assume that the observed multi-channel signals are a linear combination of multiple mutual-independent sources, in the form of $\mathbf{X} = \mathbf{AS}$, and attempt to make the estimated sources as independent as possible with higher order statistics (HOS) (Albera et al., 2012). In EEG measurement, it can be considered that the EEG components and the EMG artifacts are independent, so that the signal and artifact components can be separated by ICA, and the pure EEG can be obtained by removing the artifact components.

3.3.3. *EEMD-CCA approach*

The recently proposed EEMD-CCA approach also belongs to the framework of combining single-channel decomposition and multi-channel BSS as described above. The difference between EEMD-CCA and VMD-CCA is that EEMD is utilized to decompose each EEG channel. It is shown that the EEMD-CCA outperforms ICA, CCA and EEMD-ICA in muscle artifact cancelation (Chen et al., 2019).

3.3.4. *EEMD-ICA approach*

The EEMD-ICA approach proposed earlier uses EEMD to decompose each channel of EEG, but then uses ICA to separate EEG and muscle artifacts (Zeng et al., 2016).

3.3.5. *VMD-ICA approach*

Although (De Clercq et al., 2006; Chen et al., 2019) show that CCA is superior to ICA in terms of muscle artifact removing, this is not conclusive. So we also use ICA instead CCA in the framework, i.e. VMD-ICA, as a comparison.

3.4. *Results*

3.4.1. *Semi-simulation data*

As mentioned in Section 3.1.1, synthesized EEG signals with complex muscle artifact sources were generated, where 19 EMG sources were mapped to 19-channel clean EEG with a given SNR. The signal shown in Fig. 4b is reconstructed by four methods, and the results of channel C4 are shown in Fig. 4c. The closer the reconstructed result is to the real clean EEG, the better the reconstruction effect. It can be seen that the results of the four methods are basically the same in the signal segment with small artifacts (4–10 s). But in the signal segment with strong artifacts (the first 4 s), the result of VMD-CCA is closer to the original signal.

In order to evaluate the performance of the proposed VMD-CCA algorithm in discarding muscle artifacts under different number of EEG channels, random combination of each 19-channel semi-simulation EEG

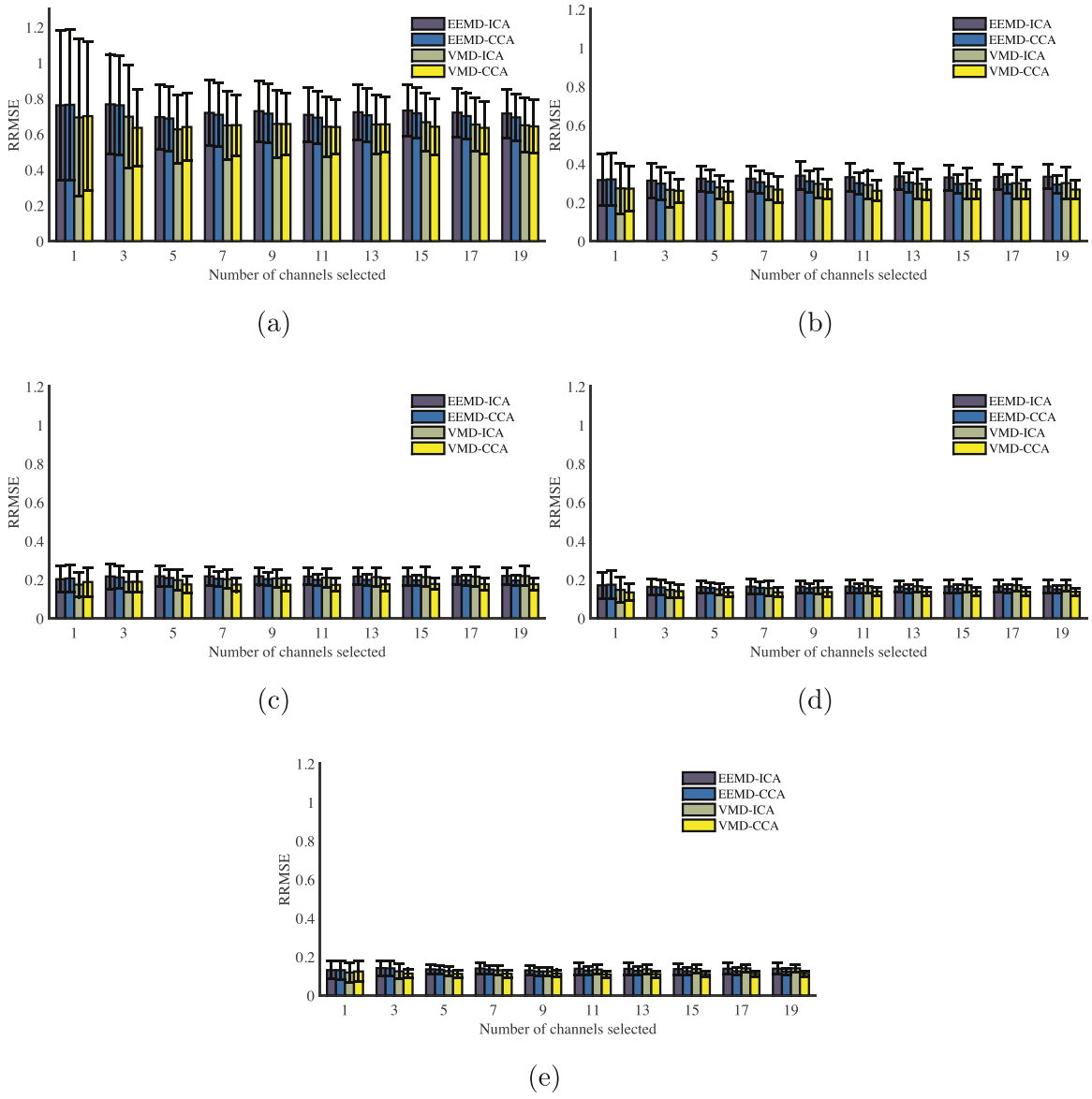

Fig. 7. The semi-simulated study: performance comparison of EEMD-ICA, EEMD-CCA, VMD-ICA and VMD-CCA at various number of selected channel in terms of RRMSE. (a) SNR = 0.5, (b) SNR = 1.5, (c) SNR = 2.5, (d) SNR = 3.5, (e) SNR = 4.5.

data set is performed. That is, C channel signal is randomly selected 5 times from each 19-channel semi-simulation data $\mathbf{X}_i$ to form a new data set $X_{i(C)}$ ($C = 1, 3, 5, \ldots, 15, 17, 19$). An example of $X_{i(5)}$ is shown in Fig. 5.

Fig. 6 shows the average and standard deviation of RRMSE as functions of SNR with VMD-CCA and comparison approaches, EEMD-ICA, EEMD-CCA and VMD-ICA, in removing muscle artifacts. Fig. 6a, b, …, i respectively show to the analysis results of $X_{i(1)}$, $X_{i(3)}$, …, $X_{i(17)}$, and Fig. 6j shows the results of $X_{i(19)}$, that is $\mathbf{X}$.

As the SNR increases, the mean and standard variation (STD) of RRMSE of all algorithms decrease, indicating that the de-artifact ability and stability of all the algorithms increase with the increase of SNR, regardless of the EEG channel number of data set.

For the data sets with the same number of EEG channels, the averaged RRMSE of VMD-CCA is less than that of EEMD-CCA and EEMD-ICA. This shows that since VMD can effectively separate EEG and muscle artifacts, the proposed VMD-CCA approach is good at dealing with minor or severe muscle artifacts in EEG regardless of the channel number. In addition, VMD-CCA performs slightly better than VMD-ICA, and EEMD-CCA outperforms EEMD-ICA (Chen et al., 2019), indicating that CCA may be more suitable for dealing with muscle artifact.

We grouped the results by SNR, as shown in Fig. 7. It can be seen that at different SNRs (from 0.5 to 4.5), the mean value of RRMSE for all the approaches have no significant change with respect to the number of EEG channel C, while the STD of RRMSE decreases as the C increases, especially between 1 and 5. This interesting result can be explained from the following aspects. Firstly, the VMD-CCA and its comparison approaches showed no significant correlation between the mean RRMSE and the C, while VMD-CCA/VMD-ICA consistently outperformed EEMD-CCA/EEMD-ICA, which indicates that the difference of mean RRMSE value is caused by the difference in the first step in the framework, namely the single-channel decomposition algorithm. As previously analyzed, VMD's ability to separate EEG and artifacts is stronger than EEMD. Secondly, the STD of RRMSE varies with C, but the difference in STD of RRMSE between different approaches is not obvious under the same C and the same SNR. Due to the volume conduction effect of the head, theoretically all EEG sources and muscle artifact sources project on all electrodes, but their effects decay with the distance from the source to the electrodes. Therefore, signals collected on electrodes that are closer together may be more relevant, and vice versa. In the semi-

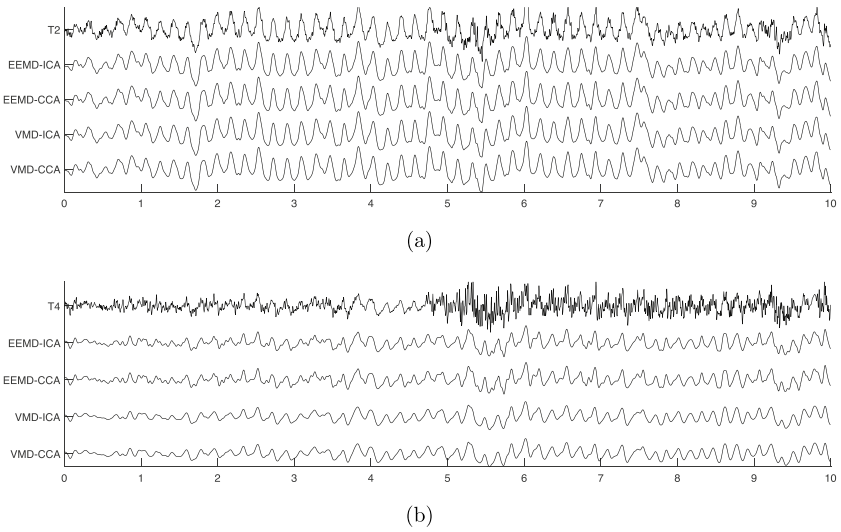

Fig. 8. A real data demo: result comparison of the same channel. (a) T2, (b) T4. The horizontal axis represents time with unit second.

simulation experiment, there is certainly no inter-channel information in the single-channel data set ($C = 1$), which limits the effect of algorithms. When the C is small, for example 3, there is a high probability of randomly selected channels weakly related to each other. In this case, the subsequent multichannel BSS process has almost no more available inter-channel information, and which is essentially equivalent to processing each single-channel signal separately. Conversely, if the randomly selected channels are more relevant, the advantages of subsequent CCA or ICA can be exploited, and the effect of artifact removing may be better. This can explain the large STD value of RRMSE under small C.

3.4.2. Real data

The real EEG example and the electrodes position are shown in Figs. 3 and 9b. The results obtained by the four algorithms are not significantly different overall. For example, as shown in Fig. 8a, the channel T2 processing results obtained by the four algorithms are almost the same. But the results of some channels are slightly different. As shown in Fig. 8b, under the same threshold setting, the processing results of VMD-ICA/VMD-CCA have the similar rhythm in 0–3 s as in 3–10 s, while the processing results of EEMD-ICA/EEMD-CCA have more high frequency components in 0–3 s than in 3–10 s.

Note that there is no ground truth in the real data set. But it can be seen from Fig. 3 that the rhythm of the less-contaminated channels, such as Fz, Cz, Pz and T2, are not changed significantly during the 10 s, so we can reasonably conclude that the rhythm of the contaminated channels should also be roughly constant. Therefore, the results of VMD-ICA/VMD-CCA may be slightly better than those of EEMD-ICA/EEMD-CCA.

As in the previous semi-simulation data study, we also compared the results of different algorithms with only a few channels selected. For example, we select two 3-electrodes sets, the first one including T5 and its adjacent electrodes T3 and C3, the second including T5 and two far apart electrodes F4 and O2, as shown in Fig. 9b. The signal of the T5 channel has obvious muscle artifacts at 0–4 s. The signal obtained from set 1 is denoted as $\mathbf{X}_{\text{near}}$, the signal from set 2 as $\mathbf{X}_{\text{far}}$. The channel T5 recovered from the $\mathbf{X}_{\text{near}}$ and $\mathbf{X}_{\text{far}}$ using the four algorithms are shown in Fig. 9d and e respectively. Fig. 9f–i show the spectral contrast of the results.

It can be seen from Fig. 9d and e that for both the $\mathbf{X}_{\text{near}}$ and $\mathbf{X}_{\text{far}}$, the signal obtained by EEMD-ICA/EEMD-CCA contain obvious muscle artifacts, while the muscle artifacts are not obvious in the results by VMD-CCA. This result is similar to the results of the previous simulation study.

It is worth noting that in the $\mathbf{X}_{\text{near}}$, the artifact intensity in T3 and C3 significantly exceeds the artifact intensity of F4 and O2 in $\mathbf{X}_{\text{far}}$. However, comparing the T5 recovered from $\mathbf{X}_{\text{near}}$ and $\mathbf{X}_{\text{far}}$ by EEMD-ICA, VMD-ICA and VMD-CCA, the muscle artifacts contained in T5 recovered from $\mathbf{X}_{\text{far}}$ are more obvious. It can also be seen from the spectrum comparison that the signal recovered from $\mathbf{X}_{\text{far}}$ are significantly more high-frequency in the spectrum range of EEG (10–30 Hz).

It can be clearly seen from the waveforms of $\mathbf{X}_{\text{near}}$ and $\mathbf{X}_{\text{far}}$ that the muscle artifacts in T5, T3 and C3 are synchronic, suggesting that they come from same artifact source, while the source of the muscle artifacts in T5, F4 and O2 are different. When dealing with the $\mathbf{X}_{\text{near}}$ dataset, it may be that the information provided by the association between the channels improves the effect of artifact canceling.

Although the proposed VMD-CCA has advantages over EEMD-CCA in muscle artifacts discarding, the limitation of this method is the processing speed. The mean time costs for EMD-CCA, VMD-ICA and VMD-CCA over 10-s 19-channel EEG were 5.52 s, 13.63 s and 13.19 s respectively. The test was done in MATLAB (MathWorks Inc. Novi, MI, USA) under Microsoft Windows 10 x64 OS on a notebook computer with Intel(R) Core(TM) i7-6500U 2.50GHz CPU and 8.00 GB RAM. The algorithm needs further optimization to achieve the effect of real-time processing.

4. Conclusion

Here we proposed a hybrid approach based on VMD and CCA to suppress muscle artifacts in EEG. We utilize VMD to decompose each channel of EEG signal into multiple IMFs, then use the autocorrelation value as a criterion to select the IMFs suspected of muscle artifacts, and then combine these IMFs as a new data set, using CCA decompose it into several uncorrelated components and the autocorrelation values decrease in turn, so that some EEG components can be retrieved. By setting the component determined to be artifacts to zero, the clean EEG can be reconstructed.

Semi-simulation data sets from real EEG and real EMG are used to evaluate the VMD-CCA algorithm. The results show that the proposed VMD-CCA algorithm outperforms the EEMD-ICA and EEMD-CCA in different SNR and a different number of EEG channels. From the simulation study, we also found that under the "single-channel decomposition + multichannel BSS" framework, such as EEMD-ICA, EEMD-CCA and our proposed VMD-CCA, the mean value of de-artification effect measured by RRMSE is almost independent of the number of channels, but the variance of the RRMSE increases as the number of channels decrease. The reason may be that the correlation of signals from adjacent electrodes is beneficial to recover some EEG components in BSS step. The real contaminated EEG signal analysis also made this suggestion.

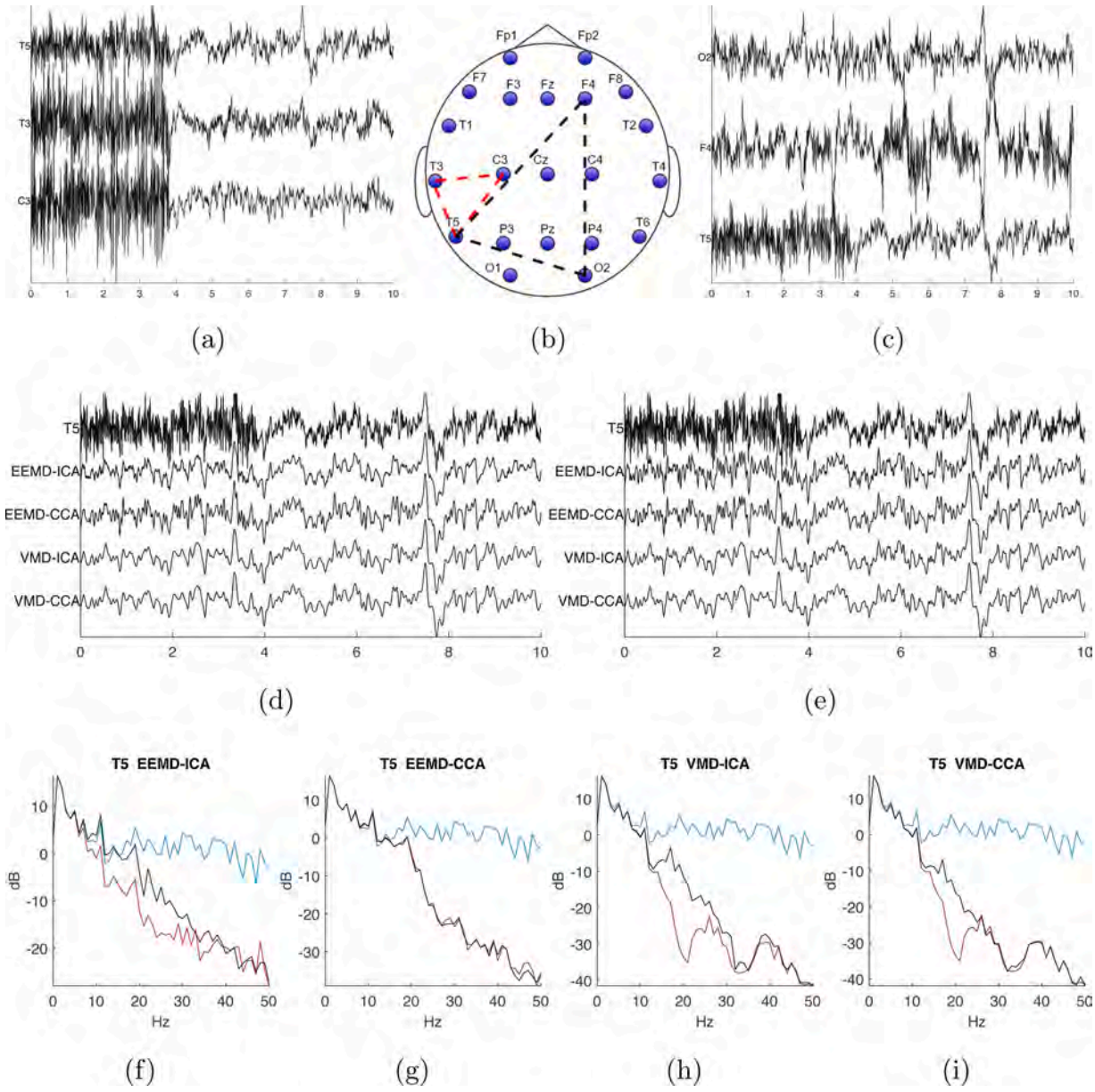

Fig. 9. A real data demo: result comparison of the same channel in two 3-channel sets. Set 1 includes 3 adjacent electrodes: T5, T3 and C3. Set 2 includes 3 non-adjacent electrodes: O2, F4 and T5. The common electrode is T5. (a) Signals from set 1, denoted as $\mathbf{X}_{\text{near}}$, (b) diagram of electrode position and the two sets, (c) signals from set 2, denoted as $\mathbf{X}_{\text{far}}$, (d) the results of T5 after using the 4 algorithms for the $\mathbf{X}_{\text{near}}$, (e) the results of T5 after using the 4 algorithms for the $\mathbf{X}_{\text{far}}$, (f–i) spectrum comparison of the original signal (blue), $\mathbf{X}_{\text{near}}$ (red) and $\mathbf{X}_{\text{far}}$ (black).

This implies that if the electrode position is set well, less number of electrodes may obtain good de-artification effect.

Author contributions

Qiang Chen: Conceptualization, methodology, software, visualization. Yingying Li: Supervision, resources, writing – original draft. Xiaohui Yuan: Methodology, writing – review & editing, visualization.

Conflict of interest

The authors declare that there is no conflict of interest.

Acknowledgments

This study was supported by grants from the State Scholarship Fund, China (201806695030), Natural Science Research Project of Educational Commission of Anhui Province of China (KJ2018A0521) and Young Talents Foreign Visiting and Training Program of Educational Commission of Anhui Province of China (GXGWFX2018047).

References

Albera, L., Kachenoura, A., Comon, P., Karfoul, A., Wendling, F., Senhadji, L., Merlet, I., 2012. ICA-based EEG denoising: a comparative analysis of fifteen methods. Bull. Polish Acad. Sci. Tech. Sci. 60, 407–418.

Chaudhary, U., Birbaumer, N., Ramos-Murguialday, A., 2016. Brain-computer interfaces for communication and rehabilitation. Nat. Rev. Neurol. 12, 513–525.

Chen, X., Chen, Q., Zhang, Y., Wang, Z.J., 2019. A novel EEMD-CCA approach to removing muscle artifacts for pervasive EEG. IEEE Sens. J. 8420–8431.

Chen, X., Liu, A., Chiang, J., Wang, Z.J., McKeown, M.J., Ward, R.K., 2016. Removing muscle artifacts from EEG data: multichannel or single-channel techniques? IEEE Sens. J. 16, 1986–1997.

De Clercq, W., Vergult, A., Vanrumste, B., Van Paesschen, W., Van Huffel, S., 2006. Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram. IEEE Trans. Biomed. Eng. 53, 2583–2587.
Dragomiretskiy, K., Zosso, D., 2014. Variational mode decomposition. IEEE Trans. Signal Process. 62, 531–544.
Goncharova, I., McFarland, D., Vaughan, T., Wolpaw, J., 2003. EMG contamination of EEG: spectral and topographical characteristics. Clin. Neurophysiol. 114, 1580–1593.
Govindan, S.M., Duraisamy, P., Yuan, X., 2014. Adaptive wavelet shrinkage for noise robust speaker recognition. Digit. Signal Process. 33, 180–190.
Hardoon, D., Szedmak, S., Shawe-Taylor, J., 2004. Canonical correlation analysis: an overview with application to learning methods. Neural Comput. 16, 2639–2664.
Huang, N.E., Shen, Z., Long, S.R., 1999. A new view of nonlinear water waves: the Hilbert spectrum. Annu. Rev. Fluid Mech. 31, 417–457.
Jiang, X., Bian, G.B., Tian, Z., 2019. Removal of artifacts from EEG signals: a review. Sensors 19, 987.
Lopez-Larraz, E., Figueiredo, T.C., Insausti-Delgado, A., Ziemann, U., Birbaumer, N., Ramos-Murguialday, A., 2018. Event-related desynchronization during movement attempt and execution in severely paralyzed stroke patients: an artifact removal relevance analysis. NeuroImage: Clinical 20, 972–986.
Mannan, M.M.N., Kamran, M.A., Jeong, M.Y., 2018. Identification and removal of physiological artifacts from electroencephalogram signals: a review. IEEE Access 6, 30630–30652.
McMenamin, B.W., Shackman, A.J., Greischar, L.L., Davidson, R.J., 2011. Electromyogenic artifacts and electroencephalographic inferences revisited. Neuroimage 54, 4–9.
Mijovic, B., De Vos, M., Gligorijevic, I., Taelman, J., Van Huffel, S., 2010. Source separation from single-channel recordings by combining empirical-mode decomposition and independent component analysis. IEEE Trans. Biomed. Eng. 57, 2188–2196.
Minguillon, J., Lopez-Gordo, M.A., Pelayo, F., 2017. Trends in EEG-BCI for daily-life: requirements for artifact removal. Biomed. Signal Process. Control 31, 407–418.
Nam, H., Yim, T.G., Han, S.K., Oh, J.B., Lee, S.K., 2002. Independent component analysis of ictal EEG in medial temporal lobe epilepsy. Epilepsia 43, 160–164.
Nathan, K., Contreras-Vidal, J.L., 2016. Negligible motion artifacts in scalp electroencephalography (EEG) during treadmill walking. Front. Hum. Neurosci. 9, 1–12.
Ren, S., Gliske, V.S., Brang, D., Stacey, W.C., 2019. Redaction of false high frequency oscillations due to muscle artifact improves specificity to epileptic tissue. Clin. Neurophysiol. 130, 976–985.
Safi, S.M.M., Pooyan, M., Nasrabadi, A.M., 2018. Improving the performance of the SSVEP-based BCI system using optimized singular spectrum analysis (OSSA). Biomed. Signal Process. Control 46, 46–58.
Song, Y., Sepulveda, F., 2018. A novel technique for selecting EMG-contaminated EEG channels in self-paced brain-computer interface task onset. IEEE Trans. Neural Syst. Rehabil. Eng. 26, 1353–1362.
Sweeney, K.T., McLoone, S.F., Ward, T.E., 2013. The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique. IEEE Trans. Biomed. Eng. 60, 97–105.
Urigueen, J.A., Garcia-Zapirain, B., 2015. EEG artifact removal-state-of-the-art and guidelines. J. Neural Eng. 12, 031001.
Wang, Y., Markert, R., 2016. Filter bank property of variational mode decomposition and its applications. Signal Process. 120, 509–521.
Wu, Z., Huang, N.E., 2009. Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv. Adapt. Data Anal. 1, 1–41.
Yong, X., Fatourechi, M., Ward, R.K., Birch, G.E., 2012. Automatic artefact removal in a self-paced hybrid brain-computer interface system. J. Neuroeng. Rehabil. 9.
Zeng, K., Chen, D., Ouyang, G., Wang, L., Liu, X., Li, X., 2016. An EEMD-ICA approach to enhancing artifact rejection for noisy multivariate neural data. IEEE Trans. Neural Syst. Rehabil. Eng. 24, 630–638.