

Article

A Preliminary Study of Muscular Artifact Cancellation in Single-Channel EEG

Xun Chen ^{1,2}, Aiping Liu ², Hu Peng ^{1,*} and Rabab K. Ward ²

¹ Department of Biomedical Engineering, School of Medical Engineering, Hefei University of Technology, Hefei 230009, China; E-Mail: xunchen@ece.ubc.ca

² Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada; E-Mails: aipingl@ece.ubc.ca (A.L.); rababw@ece.ubc.ca (R.K.W.)

* Author to whom correspondence should be addressed; E-Mail: hpeng@hfut.edu.cn;
Tel.: +86-139-5513-1982.

External Editor: António Manuel de Jesus Pereira

Received: 31 May 2014; in revised form: 19 August 2014 / Accepted: 23 September 2014 /

Published: 1 October 2014

Abstract: Electroencephalogram (EEG) recordings are often contaminated with muscular artifacts that strongly obscure the EEG signals and complicates their analysis. For the conventional case, where the EEG recordings are obtained simultaneously over many EEG channels, there exists a considerable range of methods for removing muscular artifacts. In recent years, there has been an increasing trend to use EEG information in ambulatory healthcare and related physiological signal monitoring systems. For practical reasons, a single EEG channel system must be used in these situations. Unfortunately, there exist few studies for muscular artifact cancellation in single-channel EEG recordings. To address this issue, in this preliminary study, we propose a simple, yet effective, method to achieve the muscular artifact cancellation for the single-channel EEG case. This method is a combination of the ensemble empirical mode decomposition (EEMD) and the joint blind source separation (JBSS) techniques. We also conduct a study that compares and investigates all possible single-channel solutions and demonstrate the performance of these methods using numerical simulations and real-life applications. The proposed method is shown to significantly outperform all other methods. It can successfully remove muscular artifacts without altering the underlying EEG activity. It is thus a promising tool for use in ambulatory healthcare systems.

Keywords: EEG; single-channel; muscular artifacts; EEMD; JBSS

1. Introduction

The electroencephalogram (EEG) signals are often contaminated by various physiological activities of non-interest, such as the electrocardiogram (ECG), electrooculogram (EOG) and electromyogram (EMG). While ECG and EOG artifacts can be effectively removed by using adaptive filters and blind source separation (BSS) techniques [1], the artifacts induced by muscular activity (e.g., biting, chewing and frowning) are particularly difficult to correct [2]. The main reason lies in the fact that EMG artifacts have a higher amplitude than the EEG signals, a wide spectral distribution and a variable topographical distribution [2]. These muscular artifacts obscure the EEG signals and make EEG interpretation extremely complicated or almost impossible [3].

Low-pass filters are commonly employed to remove muscular artifacts. However, since the frequency spectrum of muscular artifacts significantly overlaps with that of brain signals, these filters unfortunately suppress the brain signals of interest during the suppression of the muscular artifacts [4]. Recently, using the popular BSS technique, independent component analysis (ICA) has been extensively explored for this purpose [5–7]. ICA utilizes higher-order statistics to separate the EEG recordings into statistically independent components (ICs). Clean EEG data can then be reconstructed by removing the artifact-related ICs from the raw EEG data. In some studies, however, muscular artifacts seriously contaminate most ICs. This results in a clearly observable crosstalk between brain and muscle activities [8,9]. One possible reason is that ICA only exploits the spatial structure of source signals. Thus, it is suitable when source signals are temporally statistically independent [10]. However, artifacts typically have certain temporal structures, which can be exploited for better source separation. To this end, a canonical correlation analysis (CCA) method has been proposed as a more suitable BSS approach for separating EMG artifacts from EEG signals [11]. Due to their broad frequency spectrum, EMG artifacts resemble temporal white noise and, thus, have lower autocorrelation compared to EEG signals. The CCA method exploits this characteristic for EMG cancellation and has been shown to outperform ICA on simulated data. Later, these results were also been documented by Gao *et al.* [12].

In recent years, biomedical signal measurement and processing techniques have been increasingly deployed in ambulatory situations, particularly in healthcare applications, where minimal instrumentation and low computational complexity are required [13–15]. To reduce the complexity, many ambulatory systems use only one single EEG channel [14,15]. However, almost all current methods for muscular artifact cancellation have been designed to handle multichannel EEG datasets and will fail to isolate the muscle activity in situations where only single-channel EEG recordings are available.

To address this issue, we propose a simple, yet effective, method to achieve muscular artifact cancellation in single-channel EEG cases. This method has a two-step strategy. The first step decomposes the single-channel EEG into multichannel datasets. To implement this step, empirical mode decomposition (EMD) is a suitable option. EMD is a single-channel technique that decomposes

nonstationary and nonlinear time series into a finite number of intrinsic mode functions (IMFs) [16]. Compared with other decomposition methods (e.g., wavelet transform), EMD is completely data-driven, *i.e.*, it decomposes a signal without requiring prior knowledge. It has been shown to be efficient in many biomedical applications, e.g., denoising electrohysterogram (EHG) signals [17] and removing eye blink artifacts from EEG recordings [18]. It should be noted that a noise-assisted version of EMD, called ensemble EMD (EEMD), was recently proposed and shown to have better performance than the original EMD [19]. EEMD extracts IMFs in a manner such that the mode mixing disadvantage of the EMD method is corrected. Sweeney *et al.* has utilized this new decomposition method with CCA to remove the motion artifacts from functional near-infrared spectroscopy (fNIRS) and EEG data [20].

In the second step, the emerging joint BSS (JBSS) techniques are formulated to separate the muscle artifacts from the multidimensional datasets obtained in the first step. JBSS algorithms attempt to achieve blind source separation on multiple datasets simultaneously by balancing two criteria: (1) maximizing the independence of the estimated sources within each dataset; and (2) maximizing the source dependence across datasets. To utilize JBSS for the blind source separation purpose, the original dataset and its time-delayed version are used as the input to the JBSS methods. The advantage of using the JBSS techniques instead of the BSS methods is that besides extracting statistically independent or uncorrelated sources, JBSS also exploits the temporal structure of the sources by examining their dependence with their time-delayed version. When explored by second order statistics (SOS), the stronger dependence indicates higher autocorrelation. Thus, the separation of muscle and brain activity components can be achieved due to the relative low autocorrelation of muscular artifacts in comparison with brain activity [11]. In this work, the two most popular JBSS methods, CCA and independent vector analysis (IVA) [21,22], will be explored with EEMD. While both CCA and IVA exploit SOS for the dependence, CCA and IVA separately employ SOS and higher order statistics (HOS) for source estimation. We denote the two EEMD-JBSS combinations as EEMD-CCA and EEMD-IVA, respectively.

In this paper, we also conduct a comparison study by examining other possible single-channel techniques, which have been devised for other purposes. Single-channel ICA (SCICA) is an adaptation of ICA to single-channel signals [23]. This method assumes that the signal is stationary and is composed of spectrally disjoint sources. The combination of EEMD and ICA, denoted as EEMD-ICA, is another popular method developed for source separation of single-channel recordings [24]. The main contribution of this work are the proposed practical solutions for the muscular artifact cancellation problem in single-channel EEG. This is of special importance at present, as ambulatory healthcare continues to draw increasing attention.

We examine the performance of the proposed EEMD-JBSS methods on both synthetic data and real data. We first validate the methods on simulated data. We then apply them to a real ictal EEG dataset and a real EEG dataset collected from subjects, while riding a stationary bicycle. The EEG signals are contaminated with muscular artifacts. We note that while the EEMD-JBSS method has been proposed to remove muscle activity from the single-channel EEG case, it is generally applicable to cases when one dataset contains relatively few channels (e.g., two or three). This is done by first applying EEMD to each channel and then utilizing JBSS on the integrated signals after decomposition.

2. Materials and Methods

2.1. Methods

In this section, we first briefly introduce the existing techniques. Then, we describe the proposed two EEMD-JBSS methods.

Notations: Scalars are denoted by lowercase italic letters ($a, b, etc.$), vectors by lowercase boldface letters ($\mathbf{a}, \mathbf{b}, etc.$), matrices by boldface capitals ($\mathbf{A}, \mathbf{B}, etc.$) and the number of rows and columns by italic capitals ($T, N, etc.$). A matrix or vector transposition is denoted by an uppercase superscript T (e.g., $\mathbf{X}^T, \mathbf{v}^T$). The vector \mathbf{x} (with size $1 \times T$) is used to represent the original single-channel signal with $\mathbf{x}(t)$ ($t = 1, 2, \dots, T$) denoting the signal value at the time point t .

2.1.1. Independent Component Analysis

As ICA is a well-known BSS technique in the literature, we only briefly describe its basic concept. Suppose the mixed signals are stored in one matrix, \mathbf{X} , with size $P \times T$, where P indicates the number of channels and T indicates the number of observations per channel. The goal of ICA is to separate the mixed signals \mathbf{X} into their independent sources \mathbf{S} without any other prior knowledge using the linear model:

$$\mathbf{X} = \mathbf{A}\mathbf{S} \quad (1)$$

where \mathbf{A} is the mixing matrix and \mathbf{S} is of dimension $P \times T$. It is possible to estimate the underlying sources from the mixture signals provided they are statistically independent. Several algorithms have been developed to solve this problem. In this study, we employed the popular FastICA algorithm [25]. It is based on a fixed-point iteration scheme for maximizing the non-Gaussianity of the sources. Using this algorithm, the mixing matrix \mathbf{A} and the underlying sources $\hat{\mathbf{S}}$ can be estimated. Those sources that are deemed to be artifacts can be removed by setting the corresponding row of the matrix $\hat{\mathbf{S}}$ to zero. The artifact-free signals can then be reconstructed.

2.1.2. Single-Channel ICA

SCICA is an adaptation of ICA to single-channel signals [23]. The algorithm is described as follows. First, the observed signal \mathbf{x} is broken up into a sequence of contiguous blocks $\mathbf{b}(k)$ of length N :

$$\mathbf{b}(k) = [\mathbf{x}(k\tau), \dots, \mathbf{x}(k\tau + N - 1)]^T \quad (2)$$

where k is the block index, τ is a time delay and $(K\tau + N - 1)$ is the length of the original signal. Then, the matrix \mathbf{X} is formed as a set of observations $\mathbf{b}(k)$ ($k = 1, 2, \dots, K$) as below:

$$\mathbf{X} = [\mathbf{b}(1), \dots, \mathbf{b}(K)]^T \quad (3)$$

It should be noted that the performance of SCICA significantly depends on the parameters chosen. The authors in [23] suggest that the users select those parameters empirically. Finally, the FastICA algorithm can be applied to the matrix \mathbf{X} to obtain the mixing matrix \mathbf{A} and the underlying sources $\hat{\mathbf{S}}$. From the above procedure, it can be seen that SCICA assumes the signal \mathbf{x} to be stationary and composed of spectrally disjoint sources. These assumptions, however, do not always hold in practical applications.

2.1.3. Canonical Correlation Analysis

Suppose two zero-mean datasets are stored in two matrices, \mathbf{X}_1 with size $P_1 \times T$ and \mathbf{X}_2 with size $P_2 \times T$, where P_1 and P_2 indicate the numbers of channels in \mathbf{X}_1 and \mathbf{X}_2 , respectively, and T denotes the number of observations per channel. The aim of CCA is to find linear combinations of both \mathbf{X}_1 and \mathbf{X}_2 channels that have the maximum correlation coefficient with each other [26]. This leads to the following objective function with constraints:

$$\begin{aligned} \max_{\mathbf{v}_1, \mathbf{v}_2} \quad & (\mathbf{v}_1^T \mathbf{X}_1 \mathbf{X}_2^T \mathbf{v}_2)^2 \\ \text{s.t.} \quad & \mathbf{v}_1^T \mathbf{X}_1 \mathbf{X}_1^T \mathbf{v}_1 = 1, \quad \mathbf{v}_2^T \mathbf{X}_2 \mathbf{X}_2^T \mathbf{v}_2 = 1 \end{aligned} \quad (4)$$

where \mathbf{v}_i 's ($i = 1, 2$) are the weight vectors.

The solutions to this problem are the eigenvectors of the matrices $(\mathbf{X}_1 \mathbf{X}_1^T)^{-1} \mathbf{X}_1 \mathbf{X}_2^T (\mathbf{X}_2 \mathbf{X}_2^T)^{-1} \mathbf{X}_2 \mathbf{X}_1^T$ and $(\mathbf{X}_2 \mathbf{X}_2^T)^{-1} \mathbf{X}_2 \mathbf{X}_1^T (\mathbf{X}_1 \mathbf{X}_1^T)^{-1} \mathbf{X}_1 \mathbf{X}_2^T$, respectively. The canonical variates (CV) \mathbf{U}_i 's ($i = 1, 2$) can be calculated directly from the original matrices \mathbf{X}_i 's as $\mathbf{U}_i = \mathbf{V}_i^T \mathbf{X}_i$. The corresponding rows of \mathbf{U}_1 and \mathbf{U}_2 are highly correlated, while the rows within each individual \mathbf{U}_i are uncorrelated with each other. The detailed derivation can be referred to [27].

CCA has been further extended to solve the BSS problem in a functional magnetic resonance imaging (fMRI) study by assuming the source components to be maximally autocorrelated and mutually uncorrelated [28]. In this setting, let \mathbf{X}_1 be the observed data matrix \mathbf{X} with P mixtures and T samples, and let \mathbf{X}_2 be a temporally delayed version of the original data matrix $\mathbf{X}_2(t) = \mathbf{X}(t - 1)$. Thus, CCA can separate the recorded data into the self-correlated and mutually uncorrelated sources. As a potential alternative for the most widely used ICA method, CCA has been previously tested with a number of ICA algorithms. The CCA-based methods were shown to outperform the ICA-based techniques for EEG/fNIRS artifact removal [11,12,20]. Due to the usage of second-order statistics (SOS), they were more computationally efficient when having similar qualitative results for EEG/fMRI source separation [28,29].

2.1.4. Independent Vector Analysis

IVA is an extension of ICA from one to multiple datasets. In [21], IVA was formulated as a general JBSS framework to ensure that the extracted sources are independent within each dataset and well correlated across multiple datasets. In IVA, the concept of source component vector (SCV) is defined across multiple datasets [30]. The p -th SCV, $\mathbf{s}_p = [s_p^{[1]}, s_p^{[2]}, \dots, s_p^{[M]}]^T$ ($p = 1, 2, \dots, P$), is a random vector independent of all other SCVs and the components within each SCV are dependent. The symbol $s_p^{[m]}$ ($m = 1, 2, \dots, M$) represents the p -th underlying source component in the m -th dataset. The goal of IVA is to identify the independent SCVs from multiple multidimensional datasets. This can be achieved by minimizing the mutual information among the estimated SCVs $\hat{\mathbf{s}}_p$'s [21]:

$$\begin{aligned} I_{IVA} &\triangleq I[\hat{\mathbf{s}}_1; \hat{\mathbf{s}}_2; \dots; \hat{\mathbf{s}}_P] \\ &= \sum_{p=1}^P H[\hat{\mathbf{s}}_p] - H[\hat{\mathbf{s}}_1, \hat{\mathbf{s}}_2, \dots, \hat{\mathbf{s}}_P] \end{aligned} \quad (5)$$

where H denotes the entropy. The detailed derivation can be found in [21]. By solving the above optimization problem, each estimated SCV \hat{s}_p is independent of all other estimated SCVs, and meanwhile, the components within each SCV are dependent, e.g., $s_p^{[1]}$ and $s_p^{[2]}$ are highly correlated.

The implementation algorithms involve the selection of specific probability distributions for the SCVs. The most popular methods include IVA-L [30] and IVA-G [21]. IVA-L assumes that each SCV follows a multivariate Laplace distribution that is isotropic and possesses no second-order correlation, while IVA-G exploits second-order statistical information across datasets by assuming that each SCV is multivariate Gaussian distributed. In some applications, the second-order information across datasets may be minimal, such as in the frequency domain BSS for speech recognition. However, in some other applications, it is expected to have a much larger correlation, for instance in group fMRI studies. In this work, we will utilize IVA-G by taking into account the importance of the second-order information, which can exploit the temporal structure of muscular artifacts. Therefore, with a similar setting in CCA, IVA can separate the recorded data into the self-correlated and mutually-independent sources. The possible advantage of IVA over CCA is that IVA is able to extract independent sources rather than uncorrelated ones by using HOS. However, this requires more computational time. Moreover, IVA assumes the underlying sources to follow specific distributions, which may not be true in practice.

2.1.5. Ensemble Empirical Mode Decomposition

EMD is a single-channel decomposition method for nonstationary and nonlinear signals [16]. EMD decomposes a signal into a finite number of IMFs that represent fast to slow oscillations. An IMF is a function that satisfies two conditions [16]: (1) the number of extrema and the number of zero crossings must either be equal or differ by at most one; and (2) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero. To obtain an IMF from the original signal \mathbf{x} , a sifting process is performed [16]. First, all extrema of the original signal \mathbf{x} need to be identified. All local maximum points are connected by a cubic spline line to form the upper envelope \mathbf{e}_u . Additionally, all local minimum points are connected similarly to form the lower envelope \mathbf{e}_l . The mean of \mathbf{e}_u and \mathbf{e}_l , \mathbf{a}_1 , is calculated as:

$$\mathbf{a}_1 = \frac{\mathbf{e}_u + \mathbf{e}_l}{2} \quad (6)$$

The difference between the signal and the mean is defined as the first component \mathbf{h}_1 as:

$$\mathbf{h}_1 = \mathbf{x} - \mathbf{a}_1 \quad (7)$$

In the second sifting process, \mathbf{h}_1 is treated as the signal, and the mean \mathbf{a}_{11} of its local maxima and local minima is found. We then have:

$$\mathbf{h}_{11} = \mathbf{h}_1 - \mathbf{a}_{11} \quad (8)$$

Subsequently, we can repeat this sifting procedure k times until \mathbf{h}_{1k} is an IMF, with:

$$\mathbf{h}_{1k} = \mathbf{h}_{1(k-1)} - \mathbf{a}_{1k} \quad (9)$$

Therefore, the first IMF component derived from the original signal is designated as:

$$\mathbf{c}_1 = \mathbf{h}_{1k} \quad (10)$$

A criterion for stopping the sifting process when obtaining an IMF has been established by limiting the size of the standard deviation (SD), calculated from the two consecutive sifting sequences as below:

$$SD = \sum_{t=1}^T \left\{ \frac{[\mathbf{h}_{1(k-1)}(t) - \mathbf{h}_{1k}(t)]^2}{\mathbf{h}_{1(k-1)}^2(t)} \right\} \quad (11)$$

A typical value for SD can be set between 0.2 and 0.3 [16].

To extract the 2nd IMF component, we remove c_1 from the original signal x :

$$\mathbf{r}_1 = \mathbf{x} - \mathbf{c}_1 \quad (12)$$

The residual r_1 is treated as a new signal, and the same sifting process is applied to obtain the 2nd IMF component c_2 and the residual:

$$\mathbf{r}_2 = \mathbf{r}_1 - \mathbf{c}_2 \quad (13)$$

This procedure is repeated on the subsequent residuals \mathbf{r}_j 's, until the final residual \mathbf{r}_J no longer contains any oscillation information,

$$\mathbf{r}_j = \mathbf{r}_{j-1} - \mathbf{c}_j \quad (14)$$

By summing up Equations (12)–(14), we can obtain:

$$\mathbf{x} = \sum_{j=1}^J \mathbf{c}_j + \mathbf{r}_J \quad (15)$$

Thus, we decompose the original signal \mathbf{x} into J empirical modes \mathbf{c}_j 's and a residue \mathbf{r}_J .

However, the original EMD algorithm is highly sensitive to noise. Recently, Huang *et al.* introduced a new noise-assisted data analysis method, called EEMD [19]. The method defines the true IMF components as the mean of an ensemble of trials. Each trial consists of the signal plus an additive independent identically distributed white noise of the same standard deviation. In this case, although each individual trial may produce noisy results, the noise is canceled out in the ensemble mean of sufficient trials, since the noise in each trial is assumed independent. Regarding the ensemble number I , it is found that the performance of the technique becomes fairly consistent when using ten or more ensembles in our application. This is an acceptable number in practice considering the computational cost. The noise standard deviation has been suggested empirically to be 0.2-times the standard deviation of the original signal [19].

2.1.6. EEMD-ICA

The idea of combining EEMD with ICA for source separation from single-channel recordings was first proposed in [24] and was employed for the removal of ECG from EMG and also EMG/EOG artifacts from EEG. This is the only work we have found related to muscular artifact cancellation in single-channel EEG. However, the relevant results were limited. In this method, the EEMD technique can be used to create a multichannel signal matrix \mathbf{X} , comprised of IMFs from a single-channel recording \mathbf{x} . This matrix \mathbf{X} can then be employed as the input to the FastICA algorithm with the aim of estimating the underlying sources $\hat{\mathbf{S}}$. The sources deemed as artifacts can be removed by setting the corresponding row of the matrix $\hat{\mathbf{S}}$ to be zero. The source matrix is then passed through the mixing matrix \mathbf{A} to return the cleaned multichannel signals $\hat{\mathbf{X}}$, which are now, ideally, free of artifacts. The artifact-free single-channel recording $\hat{\mathbf{x}}$ can be determined by summing the recovered IMFs in the matrix $\hat{\mathbf{X}}$.

2.1.7. The Proposed EEMD-JBSS

To deal with the muscular artifact cancellation problem in single-channel EEG, we propose taking advantage of both EEMD and JBSS by exploring their combination. In fact, we propose a two-step strategy, operating in a similar manner to the EEMD-ICA technique. In the first step, EEMD is employed to decompose the single-channel EEG signal \mathbf{x} and to derive a set of averaged IMFs. All of the IMF components and the final residual are placed into a matrix \mathbf{X} . The size of \mathbf{X} is $P \times T$, where $P = J + 1$. In the second step, the matrix \mathbf{X} and its temporally delayed version matrix $\mathbf{X}(t - 1)$ are employed as the input to CCA or IVA. Then, the underlying sources $\hat{\mathbf{S}}$ in \mathbf{X} can be extracted and ordered in terms of their autocorrelations from high to low. The sources with low autocorrelation correspond to muscular artifacts and can be removed by setting the corresponding row of the matrix $\hat{\mathbf{S}}$ to be zero. The artifact-free multichannel signals $\hat{\mathbf{X}}$ can be reconstructed by using the updated source matrix and the mixing matrix \mathbf{A} . The recovered single-channel signal without muscular artifacts $\hat{\mathbf{x}}$ can be determined by simply summing the new IMFs components in the matrix $\hat{\mathbf{X}}$. After these two steps, the muscle activity is removed from the single-channel EEG.

2.2. Data Description

2.2.1. Synthetic Data

To demonstrate the performance of the proposed EEMD-JBSS methods, we generated synthetic single-channel EEG signals with two types of muscular artifacts. We employed some measures to test the performance, since the ground truth is known.

Traditionally, the “ground truth” EEG signals without muscular artifacts are selected by visual inspection of experienced neurophysiologists. However, not only is it difficult to obtain clean EEG signals, but there is also no guarantee that the signals are completely free of muscle activity when relying solely on visual inspection. Thus, in this study, we tend to use synthetic EEG data. A single-channel EEG data series can be generated according to the phase-resetting theory [31,32]. Similar to Makinen *et al.* [31], we generated our simulated data by summing 4 such sinusoids, whose frequencies were chosen randomly in the range 4–20 Hz. The sampling frequency was 250 Hz. Ten trial EEG data were generated, and each trial dataset was 1 s long. Thus, a 10-s series \mathbf{x}_{EEG} could be formed by concatenating the 10 trial datasets, containing mainly theta, alpha and beta activities. It should be noted that while each trial dataset included 4 distinct frequencies, the frequencies chosen for different trial data were also independent, which means that there was rich frequency information in the 10-s series.

To simulate real-life situations, obtaining pure muscle activity is necessary. It is insufficient to distinguish muscular artifacts directly from the EEG signal, as it contains both muscle and brain activity. To remove the brain activity and acquire the muscle activity, ICA was utilized to decompose a real EEG dataset with 21 channels. A neurophysiologist labeled the eye blink artifacts, eye movement artifacts and muscular artifacts from all of the decomposed ICs by inspecting some features, such as the power spectral density and topography. It is important to note that a large number of ICs contained both EMG and ongoing EEG activities. Nevertheless, there existed one component containing pure EMG

activity, denoted by \mathbf{x}_{EMG} . Since we focus on single-channel issues, it is not necessary to reconstruct the component with the corresponding field distribution.

Simulated muscle activity: To extensively investigate the performance of the methods, a mount of synthetic muscle artifacts have also been generated according to the work of Delorme *et al.* [33]. The muscle activity was modeled using random noise, band-pass filtered between 20 and 60 Hz. In this study, we generated 100 independent transient muscle artifact segments with a sampling rate of 250 Hz and a length of 10 s. Each individual segment is denoted by \mathbf{x}_{EMG} .

The EMG activity was superimposed on the EEG signal as follows:

$$\mathbf{x} = \mathbf{x}_{EEG} + \varepsilon \mathbf{x}_{EMG} \quad (16)$$

where ε represents the contribution of muscle activity. Figure 1 shows the original EEG signal \mathbf{x}_{EEG} and the EEG containing muscular artifacts \mathbf{x} ($\varepsilon = 1.5$). The signal-to-noise ratio (SNR) can then be adjusted by changing the parameter ε :

$$\text{SNR} = \frac{\text{RMS}(\mathbf{x}_{EEG})}{\text{RMS}(\varepsilon \mathbf{x}_{EMG})} \quad (17)$$

where the root mean squared (RMS) value is defined as:

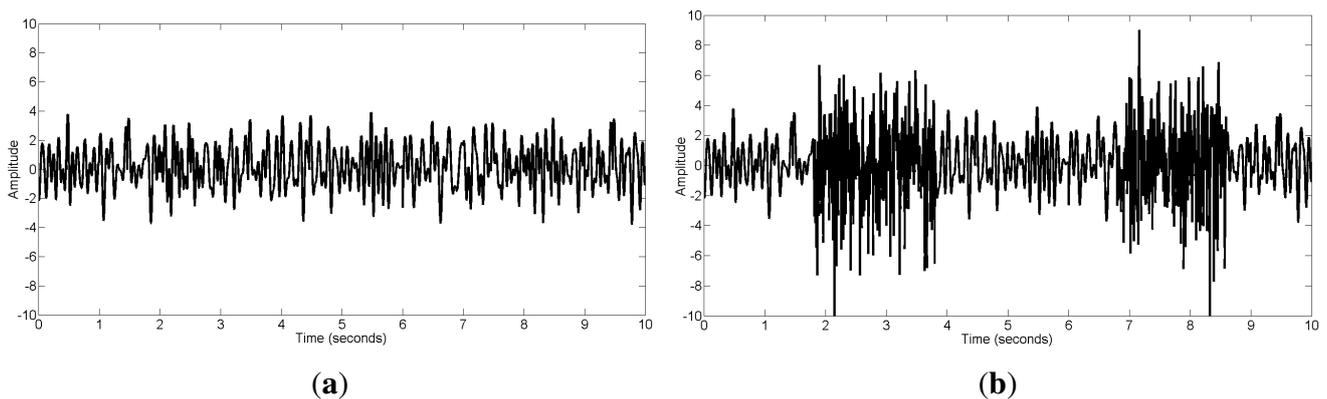
$$\text{RMS}(\mathbf{x}) = \sqrt{\frac{1}{T} \mathbf{x} \mathbf{x}^T} \quad (18)$$

To be consistent with previous EEG denoising studies [11,12], the SNR values spread from 0.25 to 3, and each SNR value corresponds to one ε value. The relative root-mean-squared error (RRMSE) is used as an evaluation measure of the effects of muscular artifact cancellation, which is defined as follows:

$$\text{RRMSE} = \frac{\text{RMS}(\mathbf{x}_{EEG} - \hat{\mathbf{x}})}{\text{RMS}(\mathbf{x}_{EEG})} \quad (19)$$

where $\hat{\mathbf{x}}$ is the estimated EEG signal after muscular artifact cancellation. To further measure the capability of the proposed method for preserving the original EEG signal, the correlation coefficient between the two waveforms \mathbf{x}_{EEG} and $\hat{\mathbf{x}}$ is also calculated. Hence, in this work, RRMSE and correlation coefficient (CC) serve as the main criteria for measuring the performance of muscular artifact cancellation.

Figure 1. (a) The original EEG data; (b) The contaminated EEG data by muscle activity.



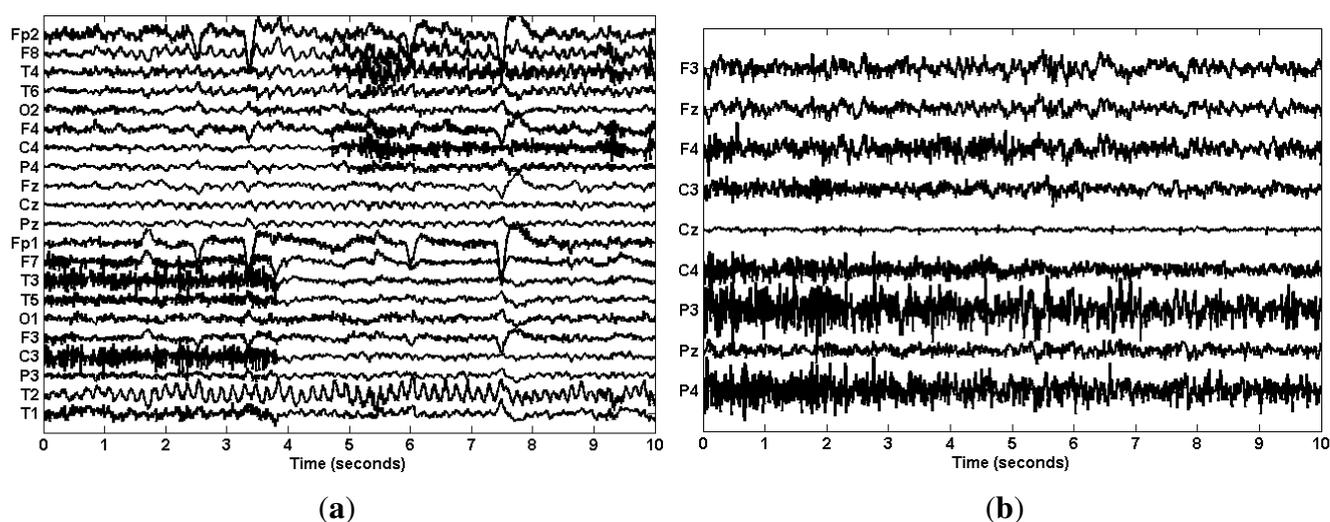
2.2.2. Real Data

For the real data case study, we used two EEG datasets. The first one was the public ictal (epilepsy) EEG data from the BioSource database established by Sabine Van Huffel (<http://www.esat.kuleuven.be/stadius/members/biomed/biosource.htm>). Ictal EEG signals are often severely contaminated with muscular artifacts, which make the determination and localization of the ictal onset complicated. Figure 2a shows the 10-s scalp EEG recordings with 21 channels obtained using a long-term epilepsy monitoring unit. This recording contains the ictal activity from a patient with mesial temporal lobe epilepsy. The sampling frequency was 250 Hz. The seizure activity was contaminated with muscular artifacts and eye blinks. Muscular artifacts can be observed between 0–3.9 s on channels F7, T3, T5, C3, T1 and between 5–10 s on channels F8, T4, F4, C4, P4.

The second dataset was collected from eight health subjects while stably cycling on an exercise bicycle. The EEG data were collected using an EEG cap (Quick-Cap, Compumedics, El Paso, TX, USA) with nine electrodes F3, Fz, F4, C3, Cz, C4, P3, Pz, P4 based on the International 10–20 system and using SynAmps2 amplifiers (NeuroScan, Compumedics, El Paso, TX, USA). The sampling rate was 1000 Hz. Data were later digitally band-pass filtered between 1~70 Hz. The University of British Columbia Ethics Board approved the study. EEG recordings during cycling were easily contaminated with muscle activity, and subsequent EEG signal processing, such as brain network study, may be complicated by the resulting EMG signals. As shown in Figure 2b, all channels of the 10-s scalp EEG were contaminated with muscle activity.

Although in both cases the single-channel technique is unnecessary, we can still apply the proposed EEMD-JBSS method to each channel individually and demonstrate its effectiveness for removing muscular artifacts from different places in the brain.

Figure 2. The original 10-s scalp EEG recordings for (a) ictal; (b) cycling.



3. Results and Discussion

3.1. The Synthetic Data Study

3.1.1. The Real Muscle Activity Case

We applied SCICA, EEMD-ICA and the two proposed methods to the synthetic single-channel data x . For a complete comparison, we should compare the performance of these methods at different SNR values in terms of RRMSE and CC. However, we found that SCICA and EEMD-ICA were unable to effectively separate muscular artifacts from brain activity, e.g., at SNR = 0.76 and SNR = 0.30, as shown in Figure 3a–d. These figures present the decomposition results of SCICA and EEMD-ICA at the two different SNR values. It can be clearly seen that most components contain both muscle activity and brain activity, e.g., IC7 in Figure 3a, IC10 in Figure 3b, IC4 in Figure 3c and IC3 in Figure 3d. One possible reason is due to the fact that muscular artifacts involve the movement of a group of muscles, which do not have a stereotyped topography [12]. Thus, the two ICA-based methods do not function correctly here, as too much brain activity has to be sacrificed to be able to sufficiently remove the muscular artifacts. Moreover, it is well-known that ICA has the permutation problem and cannot return a unique result, which will increase the difficulty in the selection of artifact components during data reconstruction. Therefore, SCICA and EEMD-ICA are unsuitable for muscular artifact cancellation in the single-channel EEG case. This will be further demonstrated by real data in Section 3.2.

In contrast, EEMD-CCA and EEMD-IVA were able to effectively isolate muscle activity into the bottom components due to their low autocorrelations, as shown in Figure 3e–h. This shows the advantage of JBSS over ICA for solving the permutation problem and facilitates automatic artifact cancellation, such as setting a threshold for autocorrelation. To compare the two EEMD-JBSS methods, we examined their performance at different SNR values in terms of RRMSE and CC, as shown in Figure 4. It can be seen that EEMD-CCA and EEMD-IVA had similar performance irrespective of the used measures. Yet, EEMD-CCA slightly outperformed EEMD-IVA when the SNR values were very low. The possible reason is that IVA assumes the underlying sources follow a Gaussian distribution, which may not be satisfied in practice. Moreover, EEMD-CCA only employs SOS and, thus, has higher computational efficiency than EEMD-IVA, which utilizes HOS. Therefore, for practical reasons, we conclude that EEMD-CCA outperformed all other methods.

To see more details about the EEMD-CCA method, we also present the step-wise results in Figure 5. The IMF components extracted by EEMD are shown in Figure 5a, where those with small indexes correspond to components of high frequencies and *vice versa*. After applying CCA, the uncorrelated sources were ordered in terms of their autocorrelations, as displayed in Figure 5b. Muscle activity was present in the bottom two components with lowest autocorrelations in the CCA decomposition. Excluding the muscular artifact components in the reconstruction led to the cleaned EEG shown in Figure 5c. To further illustrate the performance, an amplified version, including both recovered and original EEG signals, is presented in Figure 5d. From this figure, we can see that the proposed method highly preserved the original brain activity.

Figure 3. The decomposition components of the synthetic data \mathbf{x} obtained using single-channel ICA (SCICA), EEMD-ICA, EEMD-canonical correlation analysis (CCA) and EEMD-independent vector analysis (IVA) at two different SNR values: **(a)** SCICA at $\varepsilon = 2$ (SNR = 0.7611); **(b)** SCICA at $\varepsilon = 5$ (SNR = 0.3044); **(c)** EEMD-ICA at $\varepsilon = 2$; **(d)** EEMD-ICA at $\varepsilon = 5$; **(e)** EEMD-CCA at $\varepsilon = 2$; **(f)** EEMD-CCA at $\varepsilon = 5$; **(g)** EEMD-IVA at $\varepsilon = 2$; **(h)** EEMD-IVA at $\varepsilon = 5$.

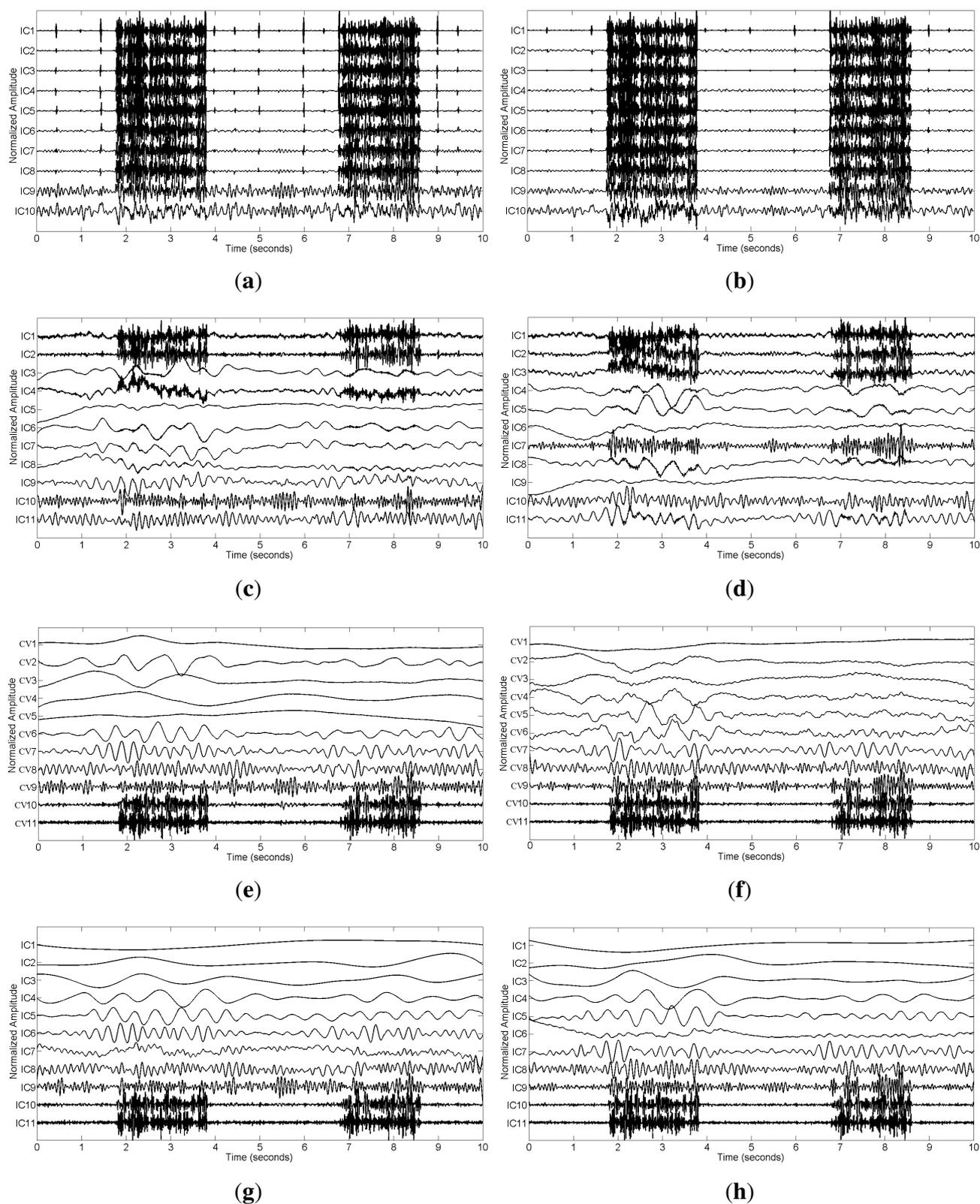


Figure 4. The performance measures for the two EEMD-JBSS methods at different SNR values for the real muscle activity case: (a) RRMSE; (b) CC.

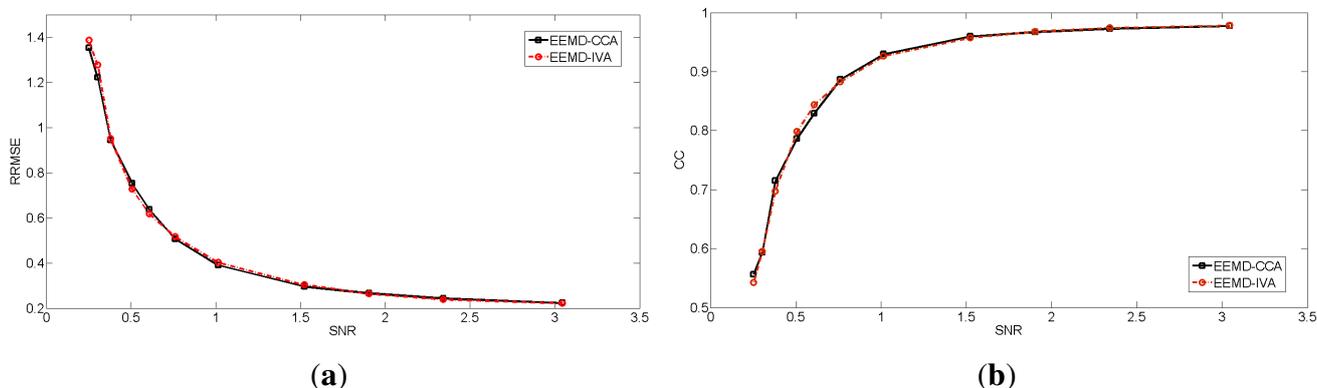
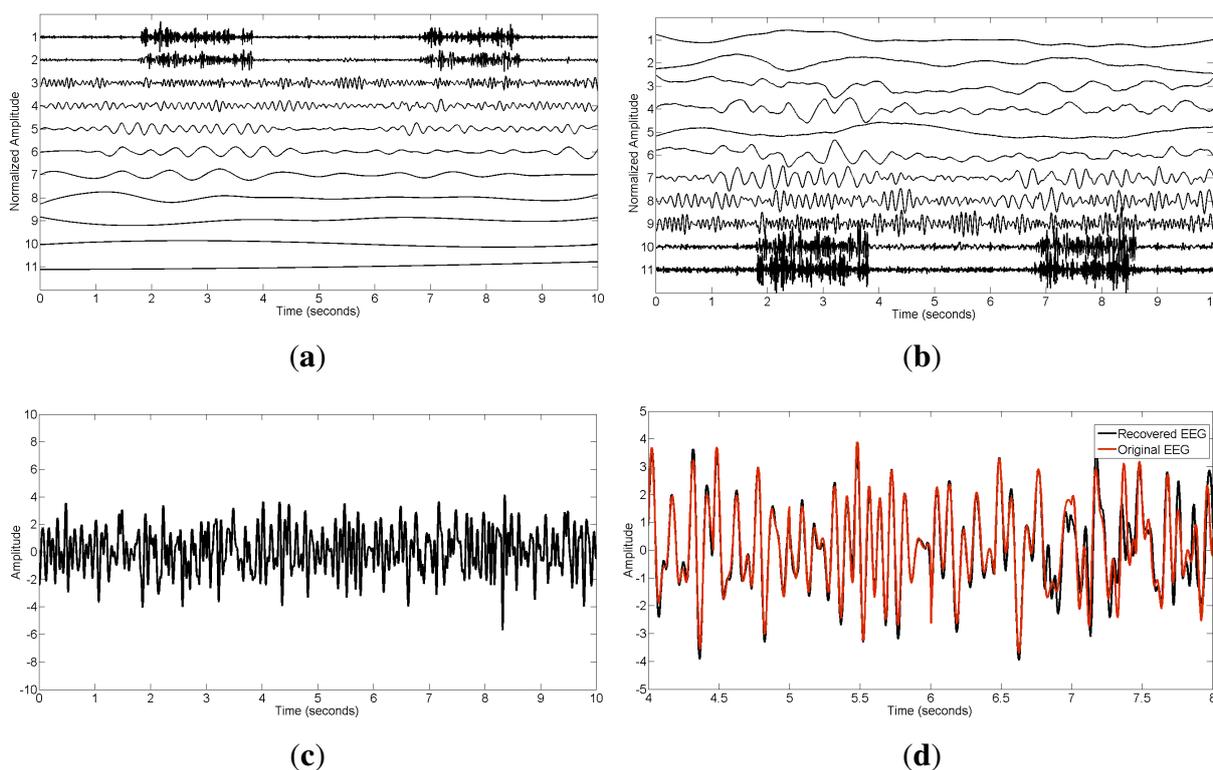


Figure 5. The step-wise results of the EEMD-CCA method: (a) the intrinsic mode function (IMF) components after applying EEMD to the single-channel EEG \mathbf{x} ($\varepsilon = 1.5$); (b) the canonical variates after using CCA; (c) the reconstructed EEG signal $\hat{\mathbf{x}}$ after muscular artifact cancellation; (d) the amplified version of $\hat{\mathbf{x}}$ compared with the original EEG \mathbf{x}_{EEG} .

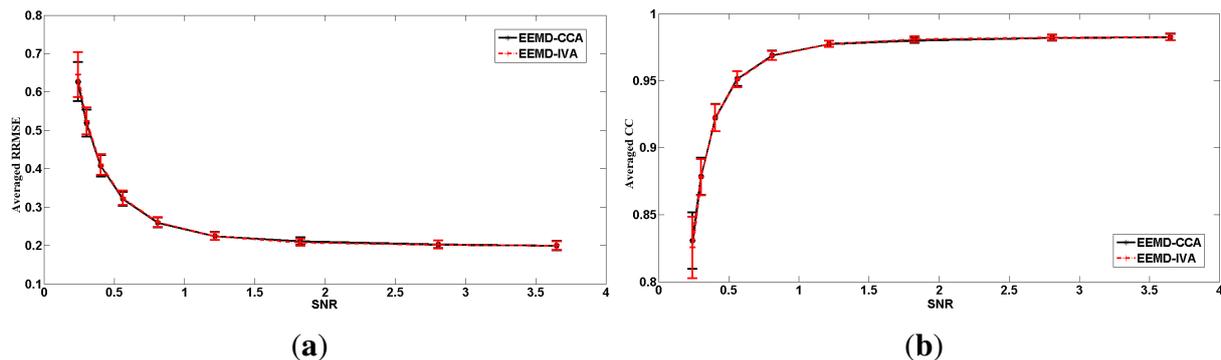


3.1.2. The Simulated Muscle Activity Case

To avoid the possibly subjective comparison between EEMD-CCA and EEMD-IVA, we simulated 100 independent segments that only contained transient muscle artifacts. Each segment was 10 s long and had a 2-s transient muscle artifact. Each individual segment \mathbf{x}_{EMG} was superimposed on the simulated EEG \mathbf{x}_{EEG} at different SNR values. Then, we evaluated the performance of the two proposed methods on the 100 segments. We obtained 100 RRMSE values and 100 CC values for each method at each SNR

value. The means and standard deviations are shown in Figure 6. Through this extensive numerical simulation, we can see that EEMD-CCA and EEMD-IVA still had almost the same performance at most SNR values. However, EEMD-CCA slightly outperformed EEMD-IVA at the lowest SNR value. The reasons are discussed in Section 3.1.1.

Figure 6. The performance measures for the two EEMD-JBSS methods at different SNR values for the simulated muscle activity case: (a) relative root-mean-squared error (RRMSE); (b) CC.



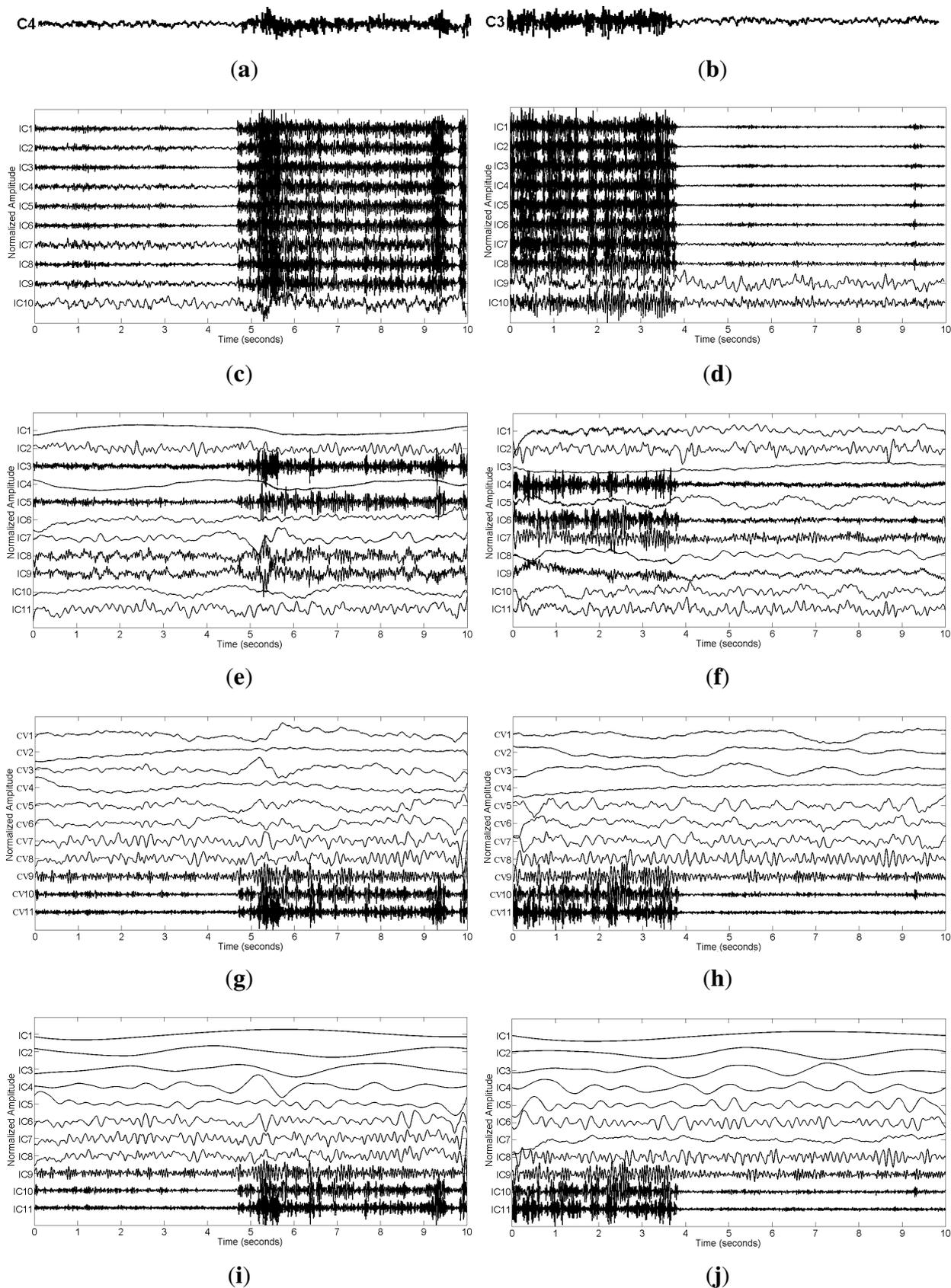
3.2. The Real Data Study

For the real data case, we first examined the ictal EEG data. The effect of muscular artifact cancellation highly depends on whether the methods can isolate muscle activity from brain activity. Hence, in Figure 7, we first present the decomposition results by applying these methods to the two channels, C4 and C3 (Figure 2a), which were severely contaminated with muscle activity.

From Figure 7c–f, we can see that SCICA and EEMD-ICA were unable to sufficiently extract the components for muscle activity; thus, it was impossible for them to remove muscular artifacts successfully. Moreover, if we attempt to remove them manually, we have to identify the components corresponding to muscle activity, which is quite time consuming and unsuitable for practical usage. However, by using EEMD-JBSS, we found that it was fairly easy to distinguish the muscular artifact components from the ones related to brain activity. As shown in Figure 7g–j, muscle activity is easily observed in the bottom two components with indexes 10–11 in the JBSS decomposition. By setting a proper threshold value for the autocorrelation, we can remove muscle activity automatically. Analogous to the simulation case, EEMD-CCA and EEMD-IVA obtained similar decomposition results.

We also tested the computational time cost of EEMD-CCA and EEMD-IVA over each of the 21 channels separately. The mean time for EEMD-CCA over each channel was 2.606 s with a standard deviation 0.0918, while the mean for EEMD-IVA was 2.756 s with a standard deviation 0.3253. The implementation was done in MATLAB (MathWorks Inc., Novi, MI, USA) and run under Microsoft Windows 8 × 64 OS on a computer with Dual Intel(R) Core(TM) i-3427U 1.80 GHz CPU and 8.00 GB RAM. The time cost is well acceptable for removing artifacts from 10-s EEG data, especially for ambulatory systems for which obtaining clean information and direct feedback in a fast fashion are essential. Considering such practical issues, EEMD-CCA becomes the best choice for this single-channel problem.

Figure 7. The decomposition components of EEG signals in (a) C4 and (b) C3 obtained using the four methods: SCICA for (c) C4 and (d) C3; EEMD-ICA for (e) C4 and (f) C3; EEMD-CCA for (g) C4 and (h) C3; EEMD-IVA for (i) C4 and (j) C3.



Finally, we applied the proposed EEMD-CCA method to each individual channel of the EEG recordings, as shown in Figure 2. When processing the EEG recordings of each single channel, muscle activity was present in the last two components in the ictal EEG and in the last four in the cycling EEG data. Excluding those components in the reconstruction of the EEG resulted in the cleaned EEG (red) shown in Figure 8. It can be seen that muscular artifacts were sufficiently removed, in contrast to the original EEG (black). In particular, for the ictal EEG, the ictal activity in each of the T2, F8, T4 and T6 electrodes was perfectly preserved. The ictal activity in F8 and T4, which originally was blurred by muscular artifacts, became visible by using the proposed EEMD-CCA method. It should also be noted that there existed some obvious EOG artifacts (marked) in ictal EEG, while their cancellation was beyond the scope of this paper. However, these EOG artifacts help demonstrate the superior performance of our proposed method due to the fact that they were preserved with little distortion.

Figure 8. The reconstructed EEG signals after muscular artifact cancellation (red) compared with the original EEG recordings (black): (a) ictal; (b) cycling. The purple arrows indicate EOG events.

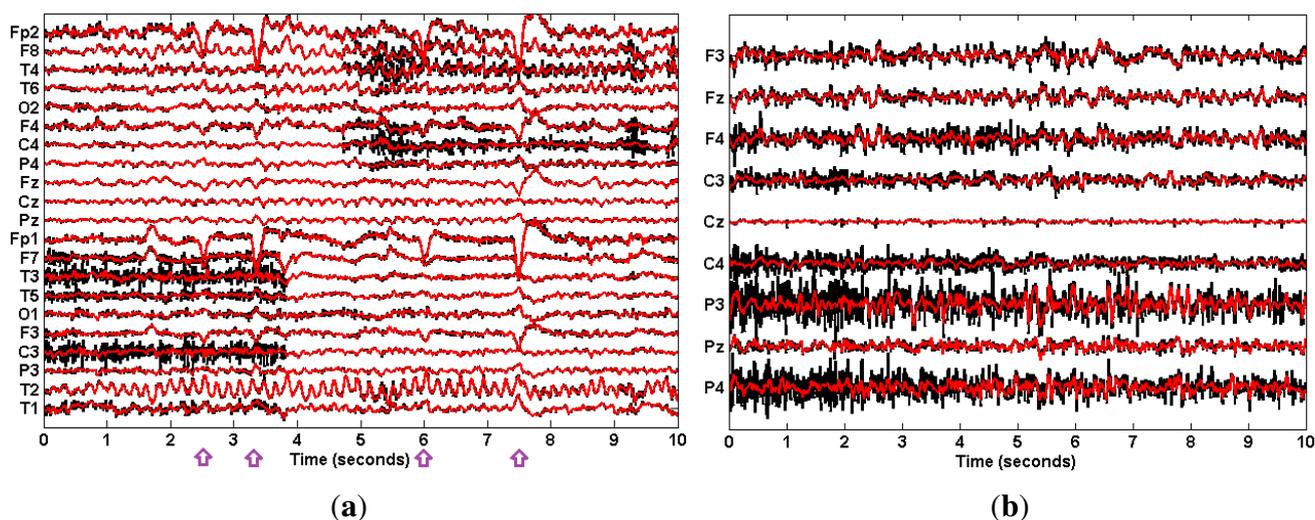
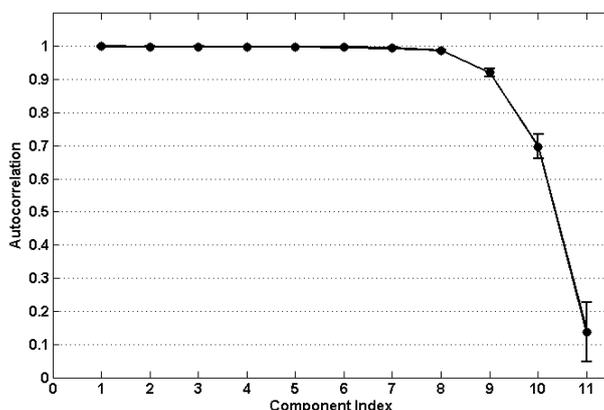


Figure 9. The averaged autocorrelation of the eleven EEMD-CCA components over all 21 ictal EEG channels.



To provide some practical guidance for the selection of the autocorrelation threshold value, we calculated the autocorrelation values for the eleven decomposition components of each individual ictal

EEG channel. Figure 9 presents the mean and standard deviation values averaged across the 21 ictal EEG channels. We suggest that the threshold should be set to no less than 0.9. The components with the value below 0.9 are deemed to be muscle artifacts.

4. Conclusions

As the popularity of using ambulatory devices in healthcare systems increases, more applications that rely on EEG signals are being developed. For practical reasons, these applications use only one EEG channel. For such ambulatory applications, muscle artifact removal from the EEG recordings becomes important. Although there exist two IC-based methods for performing source separation of a single-channel signal, they are found unsuitable for removing the artifacts arising from muscle activity. In this paper, we propose two effective methods for canceling muscular artifacts in single-channel EEG recordings. Each method utilizes the known EEMD and JBSS techniques, and they are denoted as EEMD-CCA and EEMD-IVA. Their implementation has two steps. In the first step, EEMD is used to decompose the single-channel EEG into multichannel datasets. In the second step, CCA or IVA is applied to separate the muscle artifacts from the multidimensional datasets obtained in the first step. The main difference between CCA and IVA is that CCA utilizes second order statistics, while IVA employs higher-order statistics and assumes a specific probability distribution. We examined the performance of the two proposed methods using synthetic data, as well as real-life data. We observed that both proposed methods were able to remove muscle activity while also preserving the brain activity very well. The performance of EEMD-CCA was, however, slightly better than that of EEMD-IVA, and its computational efficiency was significantly better. Therefore, EEMD-CCA is recommended in this work. It is worth noting that besides being effective in removing muscle activity in the single-channel EEG case, EEMD-CCA is also applicable in the multichannel case when few channels (e.g., two or three) are used.

Acknowledgments

The authors would like to thank Martin J. McKeown and his research group for their assistance in the preparation and examination of the clinical data. The work is supported by National Natural Science Foundation of China (no. 61172037), NSERC Canada and Qatar National Research Fund (QNRN No. NPRP 09-310-1-058).

Author Contributions

Xun Chen proposed this preliminary study and wrote this paper under the supervision of Hu Peng and Rabab K. Ward. Aiping Liu implemented the data analysis.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Shao, S.Y.; Shen, K.Q.; Ong, C.J.; Wilder-Smith, E.P.V.; Li, X.P. Automatic EEG Artifact Removal: A Weighted Support-Vector-Machine Approach With Error Correction. *IEEE Trans. Biomed. Eng.* **2009**, *56*, 336–344.
2. McMenamin, B.W.; Shackman, A.J.; Greischar, L.L.; Davidson, R.J. Electromyogenic artifacts and electroencephalographic inferences revisited. *NeuroImage* **2011**, *54*, 4–9.
3. Spencer, S.; Williamson, P.; Bridgers, S.; Mattson, R.; Cicchetti, D.; Spencer, D. Reliability and accuracy of localization by scalp ictal EEG. *Neurology* **1985**, *35*, 1567–1575.
4. Goncharova, I.; McFarland, D.; Vaughan, T.; Wolpaw, J. EMG contamination of EEG: Spectral and topographical characteristics. *Clin. Neurophysiol.* **2003**, *114*, 1580–1593.
5. Vigario, R.; Sarela, J.; Jousmaki, V.; Hamalainen, M.; Oja, E. Independent component approach to the analysis of EEG and MEG recordings. *IEEE Trans. Biomed. Eng.* **2000**, *47*, 589–593.
6. Castellanos, N.P.; Makarov, V.A. Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis. *J. Neurosci. Methods* **2006**, *158*, 300–312.
7. Crespo-Garcia, M.; Atienza, M.; Cantero, J.L. Muscle artifact removal from human sleep EEG by using independent component analysis. *Ann. Biomed. Eng.* **2008**, *36*, 467–475.
8. Nam, H.; Yim, T.G.; Han, S.; Oh, J.B.; Lee, S. Independent component analysis of ictal EEG in medial temporal lobe epilepsy. *Epilepsia* **2002**, *43*, 160–164.
9. Urrestarazu, E.; Iriarte, J.; Alegre, M.; Valencia, M.; Viteri, C.; Artieda, J. Independent component analysis removing artifacts in ictal recordings. *Epilepsia* **2004**, *45*, 1071–1078.
10. Belouchrani, A.; Abed-Meraim, K.; Cardoso, J.F.; Moulines, E. A blind source separation technique using second-order statistics. *IEEE Trans. Signal Process.* **1997**, *45*, 434–444.
11. Clercq, W.D.; Vergult, A.; Vanrumste, B.; van Paesschen, W.; van Huffel, S. Canonical Correlation Analysis Applied to Remove muscle artifacts From the Electroencephalogram. *IEEE Trans. Biomed. Eng.* **2006**, *53*, 2583–2587.
12. Gao, J.; Zheng, C.; Wang, P. Online removal of muscle artifact from electroencephalogram signals based on canonical correlation analysis. *Clin. EEG Neurosci.* **2010**, *41*, 53–59.
13. Chen, X.; Wang, Z.J. Design and implementation of a wearable, wireless EEG recording system. In Proceedings of the IEEE 5th International Conference on Bioinformatics and Biomedical Engineering, Wuhan, China, 10–12 May 2011; pp. 1–4.
14. Ridwan, S.D.; Thompson, R.; Jap, B.T.; Lal, S.; Fischer, P. Single channel wireless EEG: Proposed application in train drivers. In Proceedings of the 3rd International Conference on Broadband Communications, Information Technology and Biomedical Applications, Gauteng, South Africa, 23–26 November 2008; pp. 58–63.
15. Koley, B.; Dey, D. An ensemble system for automatic sleep stage classification using single channel EEG signal. *Comput. Biol. Med.* **2012**, *42*, 1186–1195.
16. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shin, H.H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. *Proc. R. Soc.* **1998**, *454*, 903–995.

17. Hassan, M.; Boudaoud, S.; Terrien, J.; Karlsson, B.; Marque, C. Combination of canonical correlation analysis and empirical mode decomposition applied to denoising the labor electrohysterogram. *IEEE Trans. Biomed. Eng.* **2011**, *58*, 2441–2447.
18. Zeng, H.; Song, A.; Yan, R.; Qin, H. EOG Artifact Correction from EEG Recording Using Stationary Subspace Analysis and Empirical Mode Decomposition. *Sensors* **2013**, *13*, 14839–14859.
19. Wu, Z.; Huang, N.E. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.* **2009**, *1*, 1–41.
20. Sweeney, K.T.; McLoone, S.F.; Ward, T.E. The Use of Ensemble Empirical Mode Decomposition With Canonical Correlation Analysis as a Novel Artifact Removal Technique. *IEEE Trans. Biomed. Eng.* **2013**, *60*, 97–105.
21. Anderson, M.; Adali, T.; Li, X.L. Joint blind source separation with multivariate Gaussian model: Algorithms and performance analysis. *IEEE Trans. Signal Process.* **2012**, *60*, 1672–1683.
22. Chen, X.; Liu, A.; McKeown, M.J.; Poizner, H.; Wang, Z.J. An EEMD-IVA Framework for Concurrent Multidimensional EEG and Unidimensional Kinematic Data Analysis. *IEEE Trans. Biomed. Eng.* **2014**, *61*, 2187–2198.
23. Davies, M.E.; James, C.J. Source separation using single channel ICA. *Signal Process.* **2007**, *87*, 1819–1832.
24. Mijovic, B.; Vos, M.D.; Gligorijevic, I.; Taelman J.; Huffel, S.V. Source separation from single-channel recordings by combining empirical-mode decomposition and independent component analysis. *IEEE Trans. Biomed. Eng.* **2010**, *57*, 2188–2196.
25. Hyvarinen, A.; Oja, E. Fast fixed-point algorithm for independent component analysis. *Neural Comput.* **1997**, *9*, 1483–1492.
26. Hotelling, H. Relations between two sets of variates. *Biometrika* **1936**, *28*, 321–377.
27. Chen, X.; Liu, A.; Wang, Z.J.; Peng, H. Corticomuscular Activity Modeling by Combining Partial Least Squares and Canonical Correlation Analysis. *J. Appl. Math.* **2013**, doi:10.1155/2013/401976.
28. Friman, O.; Borga, M.; Lundberg, P.; Knutsson, H. Exploratory fMRI analysis by autocorrelation maximization. *NeuroImage* **2002**, *16*, 454–464.
29. Borga, M.; Knutsson, H. *A Canonical Correlation Approach to Blind Source Separation*; Technical Report LiU-IMT-EX-0062; Linköping University: Linköping, Sweden, 2001.
30. Kim, T.; Eltoft, T.; Lee, T.W. Independent vector analysis: An extension of ICA to multivariate components. In *Independent Component Analysis and Blind Signal Separation*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 165–172.
31. Makinen, V.; Tiitinen, H.; May, P. Auditory event-related responses are generated independently of ongoing brain activity. *NeuroImage* **2005**, *24*, 961–968.
32. Yeung, N.; Bogacz, R.; Holroyd, C.B.; Nieuwenhuis, S.; Cohen, J.D. Theta phase resetting and the error-related negativity. *Psychophysiology* **2007**, *44*, 39–49.

33. Delorme, A.; Sejnowski, T.; Makeig, S. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *NeuroImage* **2007**, *34*, 1443–1449.

© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).