Biomedical Signal Processing and Control 30 (2016) 79–85

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Motion artifact removal from single channel electroencephalogram signals using singular spectrum analysis

Ajay Kumar Maddirala, Rafi Ahamed Shaik*

Department of Electronics and Electrical Engineering, Indian Institute of Technology Guwahati, Guwahati 781 039, India

ARTICLE INFO

Article history:
Received 8 January 2016
Received in revised form 16 May 2016
Accepted 27 June 2016
Available online 6 July 2016

Keywords:
Electroencephalogram (EEG)
Ensemble empirical mode decomposition (EEMD)
Canonical correlation analysis (CCA)
Singular value decomposition (SVD)
Singular spectrum analysis (SSA)

ABSTRACT

In ambulatory electroencephalogram (EEG) health care systems, recorded EEG signals often contaminated by motion artifacts. In this paper, we proposed a singular spectrum analysis (SSA) technique with new grouping criteria to remove the motion artifact from a single channel EEG signal. In order to remove the motion artifact from a single channel EEG signal, we considered the eigenvectors (basis vectors) corresponding to motion artifact are grouped or identified based on their local mobility, which is a signal complexity measure. However, as the local mobility of eigenvectors associated to the motion artifact are small, a threshold of 0.1 is set to identify them. The motion artifact signal is estimated using the identified eigenvectors and subtracted from the contaminated EEG signal to obtain the corrected EEG signal. The proposed technique is tested on 21 single channel real EEG signals contaminated by motion artifact and compared the results with the existing combined ensemble empirical mode decomposition and canonical correlation analysis (EEMD-CCA) technique. The simulation results show that the proposed modified SSA enjoys an improvement in the signal to noise ratio and the percentage reduction in artifact. Moreover, as the ambulatory EEG systems are battery operated, use of high computational signal processing techniques will reduce the battery lifetime. Hence, low computational signal processing techniques are greatly demanded in such applications. Thus, we have also evaluated the computational complexity of the proposed technique and compared with EEMD-CCA. We found that the proposed modified SSA technique significantly reduces the computational complexity and thereby lower power consumption compared to the EEMD-CCA.

1. Introduction

Ambulatory electroencephalogram (EEG) test is commonly performed to record the electrical activity of the brain for a long period and is often preferred where the diagnosis is unclear. Since the ambulatory EEG test allows the subject to move around, recorded EEG signals often contaminated by motion artifacts along with the common artifacts, such as ocular and muscle artifacts [1]. Independent component analysis (ICA) is a blind source separation technique often used to remove the artifacts from multichannel EEG signals [2,3]. In literature, few works have been reported to remove the motion artifacts from the multichannel EEG signals [1,4]. In [4], canonical correlation analysis (CCA) [5] has been employed to observe the extent to which the motion artifact is reflected in skin-electrode contact impedance. The use of CCA on multichannel EEG signals to remove the muscle artifacts has been presented in [6]. The difference between ICA and CCA is that the former uses the higher order statistics (HOS) to extract the source signals, whereas the later uses second order statistics (SOS). Since ICA uses HOS, the computational complexity of ICA is more than CCA.

In general, the ambulatory EEG system uses few EEG channels to reduce the cumbersome to the subject and maintain the minimum instrumentation complexity [7–9]. However, the techniques presented in [1–4], cannot be implemented for analysis of single channel EEG signals. The application of ICA on single channel signals has been proposed in [10]. However, this technique is not suitable to remove the artifacts from a single channel EEG signal due to the following two constraints: first, the signal of interest should be a stationary signal and secondly, the interested source signals should be disjoint in frequency domain. In [11], the ensemble empirical mode decomposition (EEMD) [12] and ICA techniques are jointly used to remove different artifacts from a single channel EEG and electromyogram (EMG) signals. In ambulatory EEG systems, as high computational signal processing algorithm consumes more power, use of such algorithms will reduce the battery

* Corresponding author.
E-mail address: rafiahamed@iitg.ernet.in (R.A. Shaik).

http://dx.doi.org/10.1016/j.bspc.2016.06.017
1746-8094/

lifetime [13]. So, low computational, as well as the single channel operated signal processing techniques, are great demand in such applications. Even though ICA has been employed for the removal of different artifacts from a single channel EEG signals, it suffers from high computational complexity. Since CCA involves less computations compared with ICA, use of such algorithm reduces the computational complexity of artifact removal system. However, the CCA technique is mostly suitable for analysis of multichannel EEG signals. In order to employ CCA on single channel EEG signals, first, the single channel signal has to be mapped into multivariate data. Recently, in [14], a combined EEMD and CCA technique, namely EEMD-CCA, has been proposed to remove the motion artifact from a single channel EEG signal. In this technique, first, the single channel signal is mapped into multivariate signal or data by decomposing it into a set of oscillating components, also called intrinsic mode functions (IMFs), using EEMD. Later, CCA extracts the source signals from the multivariate data and its delayed version. However, EEMD involves computationally intensive operations and hence increases the computational complexity of the overall EEMD-CCA technique. Moreover, EEMD-CCA technique exhibits poor performance to remove the low-frequency motion artifact.

In this paper, we propose a modified singular spectrum analysis (SSA) technique [15,16] with new grouping criteria to remove the motion artifact from a single channel EEG signal and compared its computational complexity with the existing EEMD-CCA technique. The proposed modified SSA technique is tested on the single channel EEG signals, contaminated by motion artifact, and is compared with the existing EEMD-CCA technique in terms of computational complexity, signal to noise ratio (SNR) and the percentage reduction in artifact. The simulation results show that an improvement in signal to noise ratio (SNR), as well as the percentage reduction in the artifact, is achieved with lower computational complexity than EEMD-CCA.

The organization of the paper is as follows: In Section 2, we briefly discuss the existing EEMD-CCA technique. Proposed SSA technique is discussed in Section 3. The computational complexity analysis of existing and the proposed SSA techniques are discussed in Section 4 showing the superiority of the proposed method over EEMD-CCA. Simulation studies of both techniques are presented in Section 5. Finally, Section 6 concludes the paper.

2. Existing method

2.1. EEMD-CCA

Blind source separation (BSS) is a method of separating the source signals from a group of mixed signals without knowing the prior information about the source signals. The CCA finds the solution to the BSS problem by imposing the constraint that the sources are maximally auto-correlated and mutually un-correlated [6]. Consider a data matrix $\mathbf{U}(t)$ mixed with J sources, having N number of samples and let the data matrix $\mathbf{V}(t)$, one sample delayed version of $\mathbf{U}(t)$, *i.e.* $\mathbf{V}(t)=\mathbf{U}(t-1)$. Then, CCA finds the basis vector $\mathbf{w_u}$ and $\mathbf{w_v}$ corresponding to $\mathbf{U}$ and $\mathbf{V}$ respectively, such that the correlation coefficient ρ between the variates $\mathbf{x}=\mathbf{w_u}^T\mathbf{U}$ and $\mathbf{y}=\mathbf{w_v}^T\mathbf{V}$ is maximized. The expression for correlation coefficient ρ is given by

$$\rho=\frac{\mathbf{w_u}^T\mathbf{C_{uv}}\mathbf{w_v}}{\sqrt{(\mathbf{w_u}^T\mathbf{C_{uu}}\mathbf{w_u})(\mathbf{w_v}^T\mathbf{C_{vv}}\mathbf{w_v})}} \tag{1}$$

where $\mathbf{C}_{uu}$ and $\mathbf{C}_{vv}$ are auto-covariance matrices of $\mathbf{U}$ and $\mathbf{V}$ respectively, and $\mathbf{C}_{uv}$ is the cross-covariance matrix of $\mathbf{U}$ and $\mathbf{V}$. The maximum value of ρ is obtained by taking the derivative of (1) with respect to $\mathbf{w}_u$ and $\mathbf{w}_v$, and equating to zero. The resulting equations following the two eigenvalue problems [17] are

$$\begin{aligned}\mathbf{C}_{uu}^{-1}\mathbf{C}_{uv}\mathbf{C}_{vv}^{-1}\mathbf{C}_{vu}^{T}\hat{\mathbf{w}}_u&=\rho^2\hat{\mathbf{w}}_u\\ \mathbf{C}_{vv}^{-1}\mathbf{C}_{vu}\mathbf{C}_{uu}^{-1}\mathbf{C}_{uv}^{T}\hat{\mathbf{w}}_v&=\rho^2\hat{\mathbf{w}}_v\end{aligned} \tag{2}$$

Since the data matrices $\mathbf{U}(t)$ and $\mathbf{V}(t)$ differ by one sample, finding the basis vector $\hat{\mathbf{w}}_\mathbf{u}$ is sufficient to extract the source signals.

Using CCA, the minimum condition to extract the source signals from the two data sets $\mathbf{U}$ and $\mathbf{V}$ as given in [18,19] can be represented by

$$\mid \rho_{\mathrm{u,v}}^{(i)} \mid \;\neq\; \mid \rho_{\mathrm{u,v}}^{(k)} \mid \quad 1\le i<k\le J \tag{3}$$

where $\mid \rho_{\mathrm{u,v}}^{(i)} \mid$ represents the correlation coefficient between the ith source from the data set $\mathbf{U}$ and $\mathbf{V}$. Since the CCA operates on multichannel EEG signals, this technique cannot be used for single channel EEG signals. In order to extract the source signals from a single channel EEG signal using CCA, first, the single channel EEG signal needs to be converted into a multivariate signal before applying to CCA. However, such conversion from a single channel signal into a multivariate signal can be efficiently performed using EEMD. Hence, the combined EEMD-CCA technique can be used to separate the sources from a single channel EEG signal too.

However, in the process of removing low-frequency motion artifact signal from a single channel EEG, the EEMD-CCA exhibits poor performance because of the fact that the correlation coefficients of the two hidden sources are approximately the same, *i.e.* $\mid \rho_{\mathrm{u,v}}^{(i)} \mid\approx\mid \rho_{\mathrm{u,v}}^{(k)} \mid$, $1\le i<k\le J$. Moreover, the EEMD-CCA algorithm also suffers from the high computational complexity resulting from computationally intensive operations such as the matrix inversion involved in (2) and hence makes the real time processing difficult.

3. Proposed singular spectrum analysis technique with new a grouping criteria

3.1. SSA

In general, removal of artifact from a measured signal is considered as an inverse problem, *i.e.*, reconstructing the desired signal from contaminated signal. The model assumed for the contaminated EEG signal is as follows: consider the measured EEG signal $d(n)=s(n)+a(n)$, $n=1, 2, \ldots, N$, where, $s(n)$ and $a(n)$ are samples of true EEG and motion artifact signals respectively and N is the number of samples. The removal of motion artifact from contaminated EEG signal using SSA involves four basic steps: embedding, decomposition, grouping and reconstruction. The embedding step involves the mapping of a single channel signal, $\mathbf{d}=[d(1), d(2), \ldots, d(N)]$ into a multivariate signal represented by a trajectory matrix

$$\mathbf{D}=\begin{bmatrix} d(1) & d(2) & \cdots & d(K)\\ d(2) & d(3) & \cdots & d(K+1)\\ \vdots & \vdots & \cdots & \vdots\\ d(M) & d(M+1) & \cdots & d(N)\end{bmatrix} \tag{4}$$

where M is the window length, $K=N-M+1$. The window length M is chosen based on the criteria $M>f_s/f$, where, f_s is the sampling frequency and f is the frequency of the signal of interest [20]. Let $\mathbf{S}$ and $\mathbf{A}$ are the trajectory matrices of the desired EEG and motion artifact signals $s(n)$ and $a(n)$ respectively. Then the trajectory matrix of a measured signal $d(n)=s(n)+a(n)$ is given by $\mathbf{D}=\mathbf{S}+\mathbf{A}$, where the trajectory matrix $\mathbf{A}$ has to be estimated from $\mathbf{D}$.

The second step in SSA is performing the singular value decomposition (SVD) on the trajectory matrix $\mathbf{D}=\mathbf{Q\Sigma R}$, where, $\mathbf{Q}$ and $\mathbf{R}$ are left and right orthogonal matrices, whose columns are the eigenvectors of the covariance matrix of $\mathbf{D}$ and $\mathbf{\Sigma}$ is the rectangular diagonal matrix, whose elements are square roots of the eigenvalues. In general, SVD of a rectangular matrix is commonly performed using Golub and Reinsh algorithm [21], which involves $\mathcal{O}(M^2K + MK^2 + K^3)$ computations for an $M \times K$ matrix. The computation of SVD of a trajectory matrix can be performed using the eigenvalue decomposition (EVD) of the covariance matrix $\mathbf{C} = \mathbf{D}\mathbf{D}^T$. Let the eigenvalues and eigenvectors of the covariance matrix $\mathbf{C}$ are $\lambda_1, \lambda_2, \ldots, \lambda_M$ and $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_M$, respectively with $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M \geq 0$. Then the kth eigenvector of the right orthogonal matrix $\mathbf{R}$ can be expressed as

$$\mathbf{r}_k = \mathbf{D}^T\mathbf{q}_k / \sqrt{\lambda_k} \tag{5}$$

where $k = 1, 2, \ldots, M$. The trajectory matrix for the kth component present in the measured signal $\mathbf{d}$, can be defined as

$$\mathbf{D}_k = \sqrt{\lambda_k}\mathbf{q}_k\mathbf{r}_k^T \tag{6}$$

After replacing $\mathbf{r}_k$ in (6) by (5), kth trajectory matrix $\mathbf{D}_k$ can be expressed as

$$\mathbf{D}_k = \mathbf{q}_k\mathbf{q}_k^T\mathbf{D} \tag{7}$$

The term $\mathbf{q}_k\mathbf{q}_k^T$ in (7) forms, a subspace for the kth component present in signal vector $\mathbf{d}$. Finally, the trajectory matrix in (4) decomposed into M number of $\mathbf{D}_k$'s such that $\mathbf{D} = \mathbf{D}_1 + \mathbf{D}_2 + \cdots + \mathbf{D}_M$.

In the third step of the conventional SSA, the trajectory matrices $\mathbf{D}_k$, $k = \{1, 2, \ldots, M\}$, are clustered or split into L groups. In ambulatory EEG, compared to other artifacts, the motion artifact is mostly predominant, thus measured EEG signal treated as a mixer of two signals viz the desired EEG and the motion artifact and hence it is customary to assume that $L = 2$. In practice, the grouping step in conventional SSA is performed based on the magnitudes of eigenvalues. For example, to group the trajectory matrices corresponding to the high energy signal, first, the arguments of the eigenvalues, whose magnitudes were large, should be identified and all the trajectory matrices corresponding to the identified arguments are added, which results a trajectory matrix for the high energy signal. However, this has to be performed manually. In [20], minimum description length (MDL) criteria [22] has been used for automatic grouping of trajectory matrices corresponding to the desired signal. However, such criteria can be applicable in the cases where the signal of interest is well defined [23]. For example, the ocular artifact present in the EEG signal causes an abrupt change in the baseline and hence can be treated as a well defined signal, that means the boundaries of the ocular artifact signal is well defined. But, in the case of motion artifact reduction, use of such criteria will not work well due to the fact that the motion artifact signal is a slowly time varying signal.

In this paper, we propose a new efficient grouping criteria based on local mobility of the eigenvectors. As the eigenvectors represent the frequency components present in the given signal, we exploit the local mobility of each eigenvector in grouping step. In order to define the local mobility s_m of kth eigenvector vector $\mathbf{q}_k = [q_k(1), q_k(2), \ldots, q_k(M)]$, we define the difference signal $y(n)$ as $y(n) = q_k(n) - q_k(n-1)$, $n = 2, 3, M$. Then the local mobility s_m of eigenvector $\mathbf{q}_k$ as in [24] is given by

$$s_m = \frac{f_1}{f_0} \tag{8}$$

where $f_1 = \sqrt{\frac{\sum_{n=1}^{M-1} y(n)}{M-1}}$ and $f_0 = \sqrt{\frac{\sum_{n=1}^{M} q_k(n)}{M}}$. It is clear that f_1, which represents the average of the difference signal, increases with frequency of $\mathbf{q}_k$. Hence, the local mobility is small for lower frequency signals and large for high frequency signals. We exploited this fact to separate the low-frequency motion artifact signal from a contaminated single channel EEG signal by setting an appropriate threshold. In order to determine a threshold, we investigated the local mobility of a sinusoidal signal whose frequency is set to the maximum frequency component of the motion artifact, *i.e.* 4 Hz, and found that $s_m = 0.1$. After the application of embedding step, for window length $M = 32$ and SVD step of SSA on the contaminated EEG signal, the local mobility of each eigenvector is computed and plotted in Fig. 1, where the dotted line indicates the threshold. From Fig. 1, it is clear that the eigenvector correspond to the index 1, *i.e.* $\mathbf{q}_1$ represents the basis vector to extract the motion artifact signal. The estimated trajectory matrix of the motion artifact signal $\hat{\mathbf{A}} = \mathbf{q}_1\mathbf{q}_1^T\mathbf{D}$ can be obtained by projecting the data matrix $\mathbf{D}$ onto the subspace span by the eigenvector $\mathbf{q}_1$.

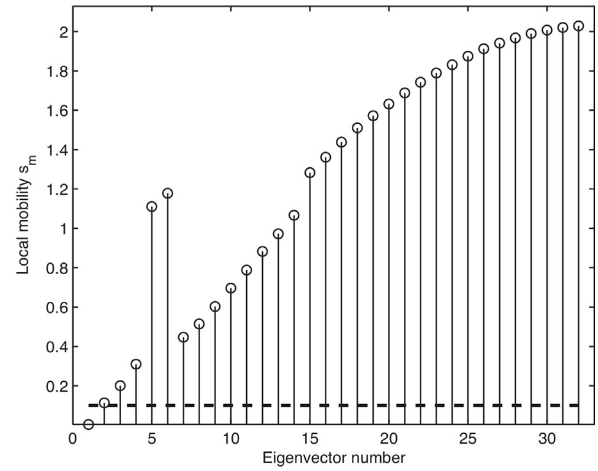

Fig. 1. Local mobilities of M eigenvectors; dashed line indicates the threshold (0.1).

In the reconstruction step of SSA, the estimated trajectory matrix of the signal of interest, *i.e.* $\hat{\mathbf{A}}$, is mapped into a single channel signal. For example, let $\hat{a}_{kj}$ is an element of kth row and jth column of the trajectory matrix $\hat{\mathbf{A}}$, then the mathematical expressions for the reconstruction of a single channel motion artifact signal $\hat{a}(n)$ from the estimated trajectory matrix $\hat{\mathbf{A}}$ can be expressed as follows.

$$\hat{a}(n) = \begin{cases} \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} \hat{a}_{k,n-k+1} & \text{for } 1 \geq n < M \\ \dfrac{1}{M}\displaystyle\sum_{k=1}^{M} \hat{a}_{k,n-k+1} & \text{for } M \geq n \leq K \\ \dfrac{1}{N-n+1}\displaystyle\sum_{k=n-K+1}^{N-K+1} \hat{a}_{k,n-k+1} & \text{for } K < n \leq N \end{cases} \tag{9}$$

Finally, the estimate of corrected EEG signal $\hat{s}(n)$ is extracted by subtracting the estimated motion artifact signal $\hat{a}(n)$ from the measured EEG signal $d(n)$, is given by

$$\hat{s}(n) = d(n) - \hat{a}(n) = s(n) + a(n) - \hat{a}(n) \tag{10}$$

4. Computational complexity issues

In order to compute the computational complexities of EEMD, CCA, EEMD-CCA and the proposed modified SSA techniques, we considered N sampled signal vector. As the first step in the proposed modified SSA technique involves mapping of univariate signal into multivariate data, no computations are required. In the second step of the proposed modified SSA, the SVD of $\mathbf{D}$ is computed in two steps: first, finding the covariance matrix $\mathbf{C}$ and secondly computing the EVD of $\mathbf{C}$. The computation of the covariance matrix of $\mathbf{D}$ of size $M \times K$, involves $\mathcal{O}(M^2K)$ computations. In addition, $\mathcal{O}(M^3)$ computations are needed to find EVD of $\mathbf{C}$, as the covariance matrix $\mathbf{C}$ is a square matrix. The third step in the proposed modified

Table 1
Computational complexities of EEMD, CCA, EEMD-CCA and proposed modified SSA techniques.

Methods	Computational complexity
EEMD	$N_eN_s(41N_{imf}N)$
CCA	$5NJ^2 + 5NJ + 19J^3/3$
EEMD-CCA	$N_eN_s(41N_{imf}N) + 5NJ^2 + 5NJ + 19J^3/3$
Proposed modified SSA	$M^2K + M^3 + M^2(1 + K) + NM$

SSA is estimating the trajectory matrix of motion artifact signal using (7), which involves $\mathcal{O}(pM^2(1+K))$ computations, where p represents the number of eigenvectors whose local mobilities are less than the specified threshold. For evaluation of the final step of proposed modified SSA, $\mathcal{O}(NM)$ computations are needed to map the estimated trajectory matrix into single channel signal. To compare the computational complexity of the proposed modified SSA technique with EEMD-CCA, we define the free parameters N_e, N_s, N_{imf} and J as the number of ensembles, the number of sifting operations, the number of IMFs and the number of observations, respectively. Then, the computational complexity of the EEMD as reported in [25] is $\mathcal{O}(41N_eN_sJN)$ computations. In practice, to make use of CCA on a single channel signal, first it should be mapped into multivariate data. In EEMD-CCA, this will be performed by EEMD. As the output of EEMD is given as input to the CCA, the number of observations $J = N_{imf}$. So, the computational complexity of CCA as it is claimed in [26] is $\mathcal{O}(5NJ^2 + 5NJ + 19J^3/3)$ computations. Finally, the overall computational complexity of EEMD-CCA is equal to the sum of individual complexities of both EEMD and CCA. Table 1 provides the overall computational complexities of EEMD, CCA, EEMD-CCA and proposed modified SSA techniques. For $N = 2560$, $M = 32$, $p = 1$, $N_e = 5$, $N_s = 5$ and $J = N_{imf} = 11$, proposed modified SSA technique reduces the computational complexity by factor of *five* times compared with the EEMD. Although CCA reduces the computational complexity *three* times compared with the proposed modified SSA technique, CCA alone cannot be used for single channel signal analysis. For single channel EEG signal analysis, CCA has to be used in combination with EEMD which can increase the overall complexity. From Table 1, it is clear that for the parameters mentioned above, the proposed modified SSA technique reduces the computational complexity by a factor of approximately *six* times compared with EEMD-CCA.

5. Results

In order to validate the performance of the proposed modified SSA technique, we considered data on the contaminated single channel EEG signals recorded using a novel technique proposed in [27]. This technique involves recording of two highly correlated EEG signals measured using two active electrodes chosen from an array of 256 electrode cap secured to the scalp of the subject. The two recordings are taken from electrode positions Fpz and Fp_1 over the scalp. To make proper contact with the scalp, an electrode gel is applied under each electrode. Using these two electrodes, labeled as channel 1 and channel 2, two EEG signals were recorded. As the spacing between them is 20 mm, recorded two EEG signals were highly correlated. During the EEG signals recording, subjects were advised not to perform any activity. Moreover, in order to minimize the other artifacts such as ocular artifact and movement artifacts by the head movement, subjects were also instructed to keep their eyes closed and stationary head position throughout the experiment. In order to induce an electrode motion artifact in one of the measured EEG signals, for every 2-min interval, channel 2's electrode is mechanically disturbed by pulling on the connective lead. As the electrode cap is manufactured by a fabric material, the movement of one electrode (channel 2) can not alter the position of the adjacent electrode (channel 1) which makes the EEG signal recorded from first electrode more immune to the induced motion artifact.

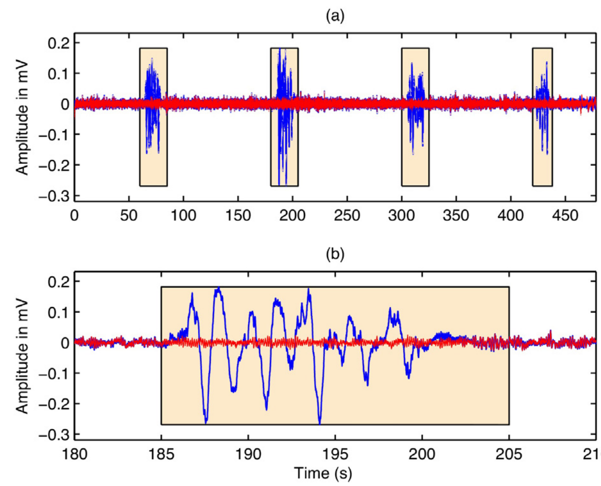

Fig. 2. (a) Superimposition of contaminated EEG signal (blue) and the ground truth EEG signal (red), (b) zoomed portion of plot (a) in the time interval 180–210 s. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

The EEG signals so obtained are maintained in PhysioBank Database [28]. This data comprises two highly correlated EEG signals: one is an artifact free EEG signal, *i.e.* "ground truth" and the other is the motion artifact contaminated EEG signal. As these EEG signals recorded at high sampling rate ($f_s = 2048$ Hz), the computational load at the signal processing stage is increased, thus results in increased processing time [29]. To reduce the processing time, first the raw EEG signals were down-sampled to 256 Hz. However, as the EEG signals were recorded from the frontal lobe of each subject with eyes closed, most of the EEG components will fall in the frequency band 0.5–30 Hz. Hence, the raw EEG signals were filtered by an appropriate 2nd order Butterworth band-pass filter with lower and upper cut-off frequencies 0.5 Hz and 30 Hz respectively. The contaminated and ground truth EEG signals, and it's zoomed version between the time interval 180–210 s are shown in Fig. 2(a) and (b) respectively. Before employing the EEMD-CCA on a single channel EEG signal, the free parameters of EEMD-CCA such as the

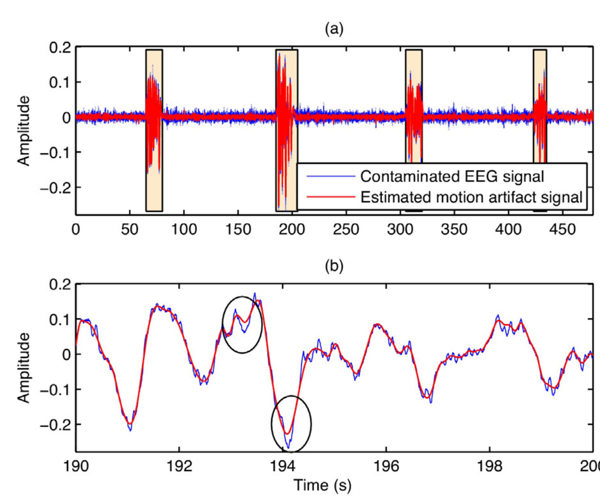

Fig. 3. (a) Superimposition of contaminated EEG signal and the estimated motion artifact signal by EEMD-CCA, (b) the zoomed portion of the above plot in the time interval 190–200 s. (For interpretation of the references to color in the text, the reader is referred to the web version of the article.)

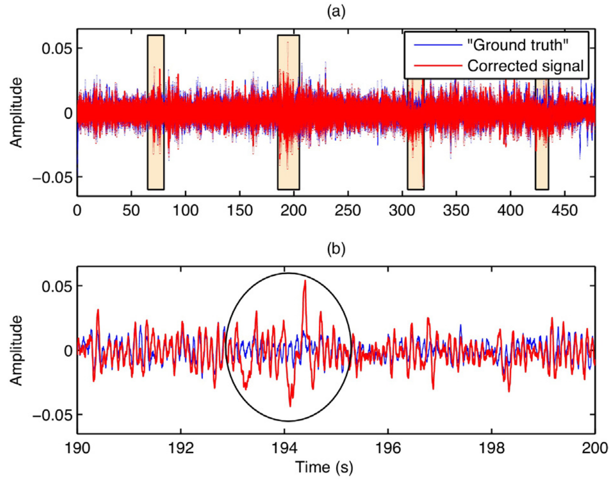

Fig. 4. (a) Superposition of ground truth EEG signal, and the corrected EEG signal by EEMD-CCA technique, (b) the zoomed version of the above plot in the time interval 190–200 s.

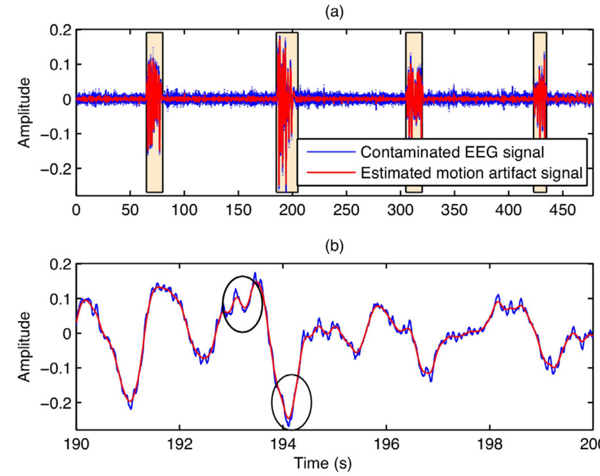

Fig. 6. (a) Superimposition of contaminated EEG signal and the estimated motion artifact signal by the proposed modified SSA technique, (b) the zoomed portion of the above plot in the time interval 190–200 s.

number of ensembles and the noise parameter (np) are set to 5 and 0.1 times the standard deviation (SD) of the signal respectively. Fig. 3(a) shows the super imposition of the contaminated EEG signal (blue) and the estimated motion artifact signal (red). In order to have a clear view of the estimated motion artifact signal obtained by EEMD-CCA technique, the magnified version of Fig. 3(a) between the time interval 190-200 s is shown in Fig. 3(b). It is clear from Fig. 3(b) (encircled portion) that the EEMD-CCA technique failed to extract the motion artifact faithfully. Fig. 4(a) shows superimposition of the corrected EEG signal (resulted by subtraction of the estimated motion artifact from the contaminated EEG signal) and the ground truth signal. The corresponding magnified version of Fig. 4(a) between time interval 190–200 s is shown in Fig. 4(b). It is observed from Fig. 4(b) that remnants of the motion artifact signal were present in the corrected EEG signal, marked in the circle. This is because of the fact that the correlation coefficient ρ between the sources hidden in the lower set of IMFs, for example, IMF14 to IMF16 as shown in Fig. 5, its delayed version is almost same and hence the condition in (3) may not be hold.

This problem can be overcome by using the proposed modified SSA technique. The estimated motion artifact and corrected EEG signals using proposed modified SSA technique are shown in Figs. 6 and 7 respectively. From Fig. 6(b), it is observed that the proposed modified SSA technique efficiently extracts the motion

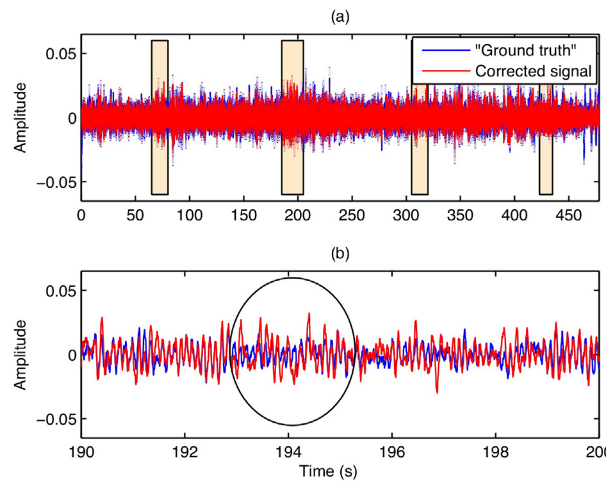

Fig. 7. (a) Superposition of ground truth EEG signal and the corrected EEG signal by the proposed modified SSA technique, (b) the zoomed version of the above plot in the time interval 190–200 s.

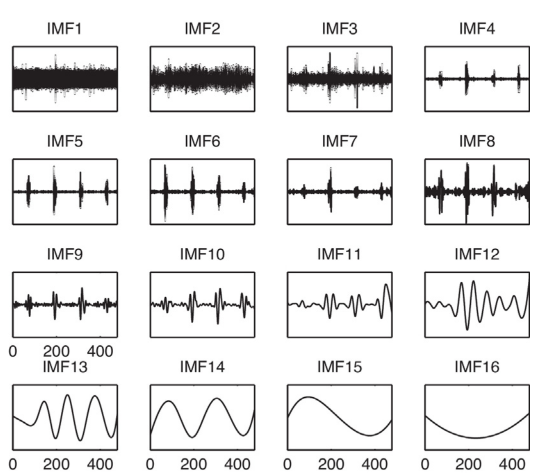

Fig. 5. EEMD decomposition of motion artifact contaminated EEG signal.

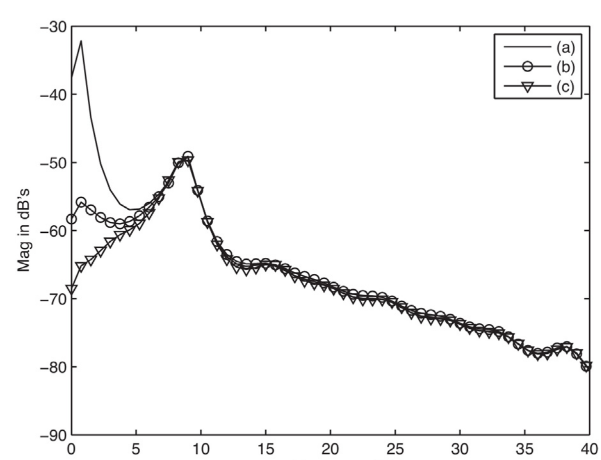

Fig. 8. PSD of (a) contaminate EEG signal, (b) corrected EEG signal by EEMD-CCA and (c) corrected EEG signal by the proposed modified SSA respectively.

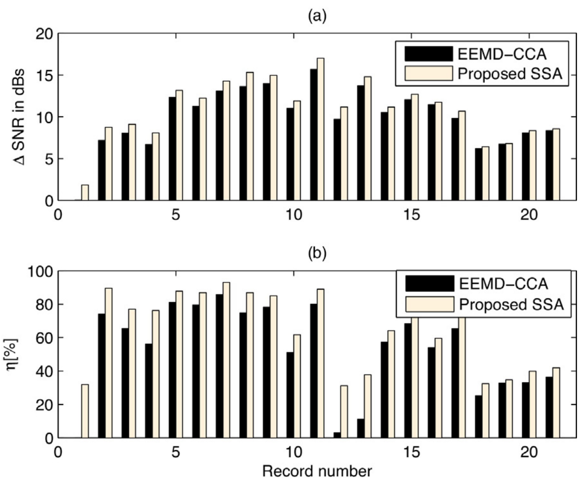

Fig. 9. Performance comparison of EEMD-CCA and proposed modified SSA techniques in terms of (a) ΔSNR and (b) η.

artifact and hence the corrected EEG is not having any remnants of the motion artifact, as shown in Fig. 7(b). We have also plotted the power spectral density (PSD) of contaminated and corrected EEG signals obtained by EEMD-CCA and the proposed modified SSA techniques shown in Fig. 8. From Fig. 8, it is clear that the remnants of the motion artifact (0.5–4 Hz) still present in the corrected EEG signal, obtained by EEMD-CCA and are not present in the corrected EEG signal obtained using the proposed modified SSA techniques.

To quantify the ability of the proposed modified SSA technique in removal of the motion artifact from a single channel EEG signal, we consider two performance measures, namely, the difference in the signal to noise ratio ΔSNR and the percentage reduction in artifact η. The difference in SNR is given by

$$\Delta SNR = 10\log_{10}\left(\frac{\sigma_a^2}{\sigma_{r_{after}}^2}\right) - 10\log_{10}\left(\frac{\sigma_a^2}{\sigma_{r_{before}}^2}\right) \tag{11}$$

where σ_a^2 is the variance of the "ground truth signal" and $\sigma_{r_{before}}^2$ and $\sigma_{r_{after}}^2$ are the variance of the error signal before and after performing the artifact removal operation. Here, the error signal is the difference between the contaminated and ground truth EEG signals. The percentage reduction in artifact η is given by

$$\eta = \left(1 - \frac{C_{clean} - C_{after}}{C_{clean} - C_{before}}\right) \times 100 \tag{12}$$

where C_{before} is the correlation between the ground truth signal and the motion artifact contaminated EEG signal, C_{after} is the correlation between the ground truth signal and corrected EEG signal and C_{clean} is the correlation between the contaminated and ground truth EEG signals over an artifact free time interval and on an average for all records it is found to be 0.9. It is worth noting that efficient artifact reduction is achieved if $C_{after} \approx C_{clean}$. However, this depends on how well the artifact removal technique estimates the motion artifact signal.

To quantify the results, 23 single channel EEG records are considered. After visual inspection of all records, it is found that no brain activity is registered in 12 and 15 records. Moreover, the poor correlation coefficient was found over the clean epochs, *i.e.* non shaded region as shown in Fig. 2. Hence, we tested the proposed modified SSA technique on the remaining 21 motion artifact contaminated single channel EEG records in which α component of the EEG signal is registered. Fig. 9 shows the comparison of proposed modified SSA and the EEMD-CCA technique in terms of ΔSNR, and η. It is clear from Fig. 9 that for all the records there is an improvement in ΔSNR, as well as in η, using the proposed modified SSA technique. On average, there will be 0.91507 dB improvement in ΔSNR and 11.388% improvement in η.

6. Conclusion

In this paper, we proposed modified SSA technique to remove the motion artifact from single channel EEG signals. The proposed modified SSA technique gains 0.91507 dB improvement in the signal to noise ratio and 11.388% improvement in percentage reduction in artifact. Moreover, the proposed modified SSA technique reduces the computational complexity by a factor of approximately *six* times compared with EEMD-CCA, which enable the low power implementation of ambulatory EEG health care systems.

References

[1] J.T. Gwin, K. Gramann, S. Makeig, D.P. Ferris, Removal of movement artifact from high-density EEG recorded during walking and running, J. Neurophysiol. 103 (6) (2010) 3526–3534, http://dx.doi.org/10.1152/jn.00105.2010.
[2] T.-P. Jung, S. Makeig, C. Humphries, T.-W. Lee, M.J. McKeown, V. Iragui, T.J. Sejnowski, Removing electroencephalographic artifacts by blind source separation, Psychophysiology 37 (2) (2000) 163–178, http://dx.doi.org/10.1111/1469-8986.3720163.
[3] R.R. Vázquez, H. Velez-Perez, R. Ranta, V.L. Dorr, D. Maquin, L. Maillard, Blind source separation wavelet denoising and discriminant analysis for EEG artefacts and noise cancelling, Biomed. Signal Process. Control 7 (4) (2012) 389–400, http://dx.doi.org/10.1016/j.bspc.2011.06.005.
[4] A. Bertrand, V. Mihajlovic, B. Grundlehner, C. Van Hoof, M. Moonen, Motion artifact reduction in EEG recordings using multi-channel contact impedance measurements, in: Biomedical Circuits and Systems Conference (BioCAS), 2013 IEEE, 2013, pp. 258–261, http://dx.doi.org/10.1109/BioCAS.2013.6679688.
[5] H. Hotelling, Relations between two sets of variates, Biometrika (1936) 321–377, http://dx.doi.org/10.2307/2333955.
[6] W. De Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen, S. Van Huffel, Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram, IEEE Trans. Biomed. Eng. 53 (12) (2006) 2583–2587, http://dx.doi.org/10.1109/TBME.2006.879459.
[7] X. Chen, J. Wang, Design and implementation of a wearable, wireless EEG recording system, in: 5th IEEE International Conference on Bioinformatics and Biomedical Engineering (ICBBE) 2011, 2011, pp. 1–4, http://dx.doi.org/10.1109/icbbe.2011.5781501.
[8] S. Ridwan, R. Thompson, B. Jap, S. Lal, P. Fischer, Single channel wireless EEG: proposed application in train drivers, in: Third International Conference on Broadband Communications, Information Technology Biomedical Applications, 2008, 2008, pp. 58–63, http://dx.doi.org/10.1109/BROADCOM.2008.69.
[9] B. Koley, D. Dey, An ensemble system for automatic sleep stage classification using single channel EEG signal, Comput. Biol. Med. 42 (12) (2012) 1186–1195, http://dx.doi.org/10.1016/j.compbiomed.2012.09.012.
[10] M. Davies, C. James, Source separation using single channel ICA, Signal Process. 87 (8) (2007) 1819–1832, http://dx.doi.org/10.1016/j.sigpro.2007.01.011.
[11] B. Mijovic, M. De Vos, I. Gligorijevic, J. Taelman, S. Van Huffel, Source separation from single-channel recordings by combining empirical-mode decomposition and independent component analysis, IEEE Trans. Biomed. Eng. 57 (9) (2010) 2188–2196, http://dx.doi.org/10.1109/TBME.2010.2051440.
[12] Z. Wu, N.E. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method, Adv. Adapt. Data Anal. 01 (01) (2009) 1–41, http://dx.doi.org/10.1142/S1793536909000047.
[13] K. Patel, C.-P. Chua, S. Fau, C. Bleakley, Low power real-time seizure detection for ambulatory EEG, in: 3rd International Conference on Pervasive Computing Technologies for Healthcare, 2009. PervasiveHealth 2009, 2009, pp. 1–7, http://dx.doi.org/10.4108/ICST.PERVASIVEHEALTH2009.6019.
[14] K. Sweeney, S. McLoone, T. Ward, The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique, IEEE Trans. Biomed. Eng. 60 (1) (2013) 97–105, http://dx.doi.org/10.1109/TBME.2012.2225427.
[15] M. Ghil, M. Allen, M. Dettinger, K. Ide, D. Kondrashov, M. Mann, A.W. Robertson, A. Saunders, Y. Tian, F. Varadi, et al., Advanced spectral methods for climatic time series, Rev. Geophys. 40 (1) (2002) 1–3.
[16] N. Golyandina, V. Nekrutkin, A. Zhigljavsky, Analysis of Time Series Structure: SSA and Related Techniques, Chapman & Hall/CRC Monographs on Statistics & Applied Probability, CRC Press, 2001.
[17] O. Friman, Adaptive analysis of functional MRI data (Ph.D. thesis), Linköping University, Sweden, 2003, dissertation No. 836, ISBN 91-7373-699-6.

[18] Y.-O. Li, T. Adali, W. Wang, V. Calhoun, Joint blind source separation by multiset canonical correlation analysis, IEEE Trans. Signal Process. 57 (10) (2009) 3918–3929, http://dx.doi.org/10.1109/TSP.2009.2021636.
[19] X. Chen, C. He, H. Peng, Removal of muscle artifacts from single-channel EEG based on ensemble empirical mode decomposition and multiset canonical correlation analysis, J. Appl. Math. (2014), http://dx.doi.org/10.1155/2014/261347.
[20] A. Teixeira, A. Tom, E. Lang, P. Gruber, A.M. da Silva, Automatic removal of high-amplitude artefacts from single-channel electroencephalograms, Comput. Methods Progr. Biomed. 83 (2) (2006) 125–138, http://dx.doi.org/10.1016/j.cmpb.2006.06.003.
[21] G. Golub, C.E. Reinsch, Singular value decomposition and least squares solutions, Numer. Math. 14 (5) (1970) 403–420, http://dx.doi.org/10.1007/BF02163027.
[22] M. Wax, T. Kailath, Detection of signals by information theoretic criteria, IEEE Trans. Acoust. Speech Signal Process. 33 (2) (1985) 387–392, http://dx.doi.org/10.1109/TASSP.1985.1164557.
[23] S. Sanei, T. Lee, V. Abolghasemi, A new adaptive line enhancer based on singular spectrum analysis, IEEE Trans. Biomed. Eng. 59 (2) (2012) 428–434, http://dx.doi.org/10.1109/TBME.2011.2173936.
[24] K. Najarian, R. Splinter, Biomedical Signal and Image Processing, CRC Press, 2005.
[25] Y.-H. Wang, C.-H. Yeh, H.-W.V. Young, K. Hu, M.-T. Lo, On the computational complexity of the empirical mode decomposition algorithm, Physica A: Stat. Mech. Appl. 400 (2014) 159–167, http://dx.doi.org/10.1016/j.physa.2014.01.020.
[26] D. Safieddine, A. Kachenoura, L. Albera, G. Birot, A. Karfoul, A. Pasnicu, A. Biraben, F. Wendling, L. Senhadji, I. Merlet, Removal of muscle artifact from EEG data: comparison between stochastic (ICA and CCA) and deterministic (EMD and wavelet-based) approaches, EURASIP J. Adv. Signal Process. (1) (2012), http://dx.doi.org/10.1186/1687-6180-2012-127.
[27] K. Sweeney, H. Ayaz, T. Ward, M. Izzetoglu, S. McLoone, B. Onaral, A methodology for validating artifact removal techniques for physiological signals, IEEE Trans. Inf. Technol. Biomed. 16 (5) (2012) 918–926, http://dx.doi.org/10.1109/TITB.2012.2207400.
[28] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.-K. Peng, H.E. Stanley, PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, Circulation 101 (June (23)) (2000) e215–e220.
[29] J.-H. Lee, S. Oh, F.A. Jolesz, H. Park, S.-S. Yoo, Application of independent component analysis for the data mining of simultaneous EEG-fMRI: preliminary experience on sleep onset, Int. J. Neurosci. 119 (8) (2009) 1118–1136, http://dx.doi.org/10.1080/00207450902854627.