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

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

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

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

where $|\rho_{\mathrm{u,v}}^{(i)}|$ 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.* $|\rho_{\mathrm{u,v}}^{(i)}| \approx |\rho_{\mathrm{u,v}}^{(k)}|$, $1\le i<k\le J$. 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