2015-07-15

A Wavelet-Based Artifact Reduction from Scalp EEG for Epileptic Seizure Detection

Islam, Md Kafiul

IEEE

A Wavelet-Based Artifact Reduction from Scalp EEG for Epileptic Seizure Detection

Md Kafiul Islam, Amir Rastegarnia, and Zhi Yang

***Abstract*—This paper presents a method to reduce artifacts from scalp EEG recordings to facilitate seizure diagnosis/detection for epilepsy patients. The proposed method is primarily based on stationary wavelet transform and takes the spectral band of seizure activities (i.e. 0.5 - 29 Hz) into account to separate artifacts from seizures. Different artifact templates have been simulated to mimic the most commonly appeared artifacts in real EEG recordings. The algorithm is applied on three sets of synthesized data including fully simulated, semi-simulated and real data to evaluate both the artifact removal performance and seizure detection performance. The EEG features responsible for the detection of seizures from non-seizure epochs have been found to be easily distinguishable after artifacts are removed and consequently the false alarms in seizure detection are reduced. Results from an extensive experiment with these datasets prove the efficacy of the proposed algorithm, which makes it possible to use it for artifact removal in epilepsy diagnosis as well as other applications regarding neuroscience studies.**

***Index Terms*—Artifact, Scalp EEG, Epilepsy, Seizure detection, Stationary Wavelet Transform.**

I. Introduction

Approximately 2% of the world population suffer from epilepsy seizures. The occurrence of seizure is almost uncertain which is the main cause of disability associated with epilepsy [1]. To reduce this uncertainty, a recording system that provides early and accurate seizure detection with immediate warning is highly desired. One way to achieve that is to use long-term EEG recording to detect the characteristic EEG waveforms during seizures. The prolonged EEG recordings not only can increase the chance of detecting an ictal event (seizure) or an interictal epileptic discharge, but also useful in the diagnosis of non-epileptic paroxysmal disorders compared to a routine EEG. Unfortunately, EEG recordings are often contaminated by different forms of artifacts such as artifacts due to electrode displacement and pop-up, motion artifacts, ocular artifacts and EMG artifacts from muscle activity, which reduce the accuracy of recorded EEG signal. Besides, some artifacts may increase the false positive rate during seizure detection while some certain types of seizures can be misdiagnosed as non-epileptic events when they are submerged/masked under artifacts. Thus, in order to correctly diagnose epilepsy, it is extremely important to remove such offending artifacts automatically, prior to seizure detection. However, automatic detection and removal of artifacts in such applications is a great challenge problem since, the artifacts overlap with background EEG rhythms and seizure events in both temporal and spectral domain. On the other hand, the artifacts are of various types related to their origins, waveform shapes, frequency characteristics which make it difficult to differentiate them from the signal of interest.

Many traditional approaches have been proposed to remove or attenuate artifacts from recorded EEG signals [2]–[14]. The most widely used methods for attenuating artifacts in EEG signals are based on blind source separation such as independent component analysis (ICA) and canonical correlation analysis (CCA) [9], [11], [15]–[18]. The BSS-based algorithms assume that the observations are linear mixing of the sources and the number of sources is equal or less than the number of observations. Another assumption is that the sources have to be either independent for ICA based methods or maximally uncorrelated for CCA based methods. Beside these assumptions, there are some issues that affect the usefulness of BSS-based methods such as

- Some ictal events can only be found in few channels if it is a focal seizure.
- Some of the artifacts are localized in a single channel, resulting in failure to identify the artifact source in the cross-channel analysis.
- Detection of artifactual IC is not automatic or semi-automatic given that the reference channel that records the artifactual source separately is available [3].
- The artifactual independent component is often found to be mixed with neural signals and therefore complete rejection of such IC results in serious signal distortion.

The methods in [8], [9], [12], [19] rely on adaptive filtering to remove artifacts from EEG signal. These methods are applicable only when there is any reference artifact channel available. However, due to diversity of artifacts for different movements and in different surrounding environments, such reference channel is not feasible. Other filtering methods like Kalman, Wiener and Particle filters, however, do not require an extra reference channel, but they need a-priori user input to function which may not be feasible always [6]. Some other limitations of the existing artifact removal methods are:

- Most of the methods remove single type of artifact and unable to handle other types e.g. ocular artifact [10], [17], [20]–[22], motion artifact [18] and muscle artifact [23].
- Many studies have proposed methods to remove artifacts

M. K. Islam and Z. Yang are with,Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117583, Singapore. E-mail: kafiul_islam@nus.edu.sg, eleyangz@nus.edu.sg.

A. Rastegarnia is with Department of Electrical Engineering, Malayer University, Malayer 65719-95863, Iran. Email: a_rastegar@ieee.org.

Manuscript received 2014.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2015.2457093, IEEE Journal of Biomedical and Health Informatics

for general purpose [24]–[34] and they do not consider specific target application. As a result, it brings unnecessary complexity in their algorithms and also results in over-correction of data.

In this paper, we develop an automated algorithm to remove artifacts as much as possible without distorting the signal of interests. The proposed algorithm is based on the stationary wavelet transform (SWT) and takes the spectral band of seizure activities into account to separate artifacts from seizures. The reason of choosing wavelet transform over other methods (e.g. BSS, EMD, Adaptive Filtering, etc.) is its ability to decompose single-channel EEG data into different frequency bands with high temporal resolution followed by easier denoising technique [35]. This is done with reasonable computational complexity compared with BSS or EMD and without requiring any reference channel unlike adaptive filtering. In addition, the choice of SWT (also known as Undecimated Wavelet Transform) over discrete wavelet transform (DWT) is the factor that SWT is translational-invariant since it involves upsampling of the filter coefficients instead of downsampling unlike in DWT [36]. Therefore small shifts in a signal can't cause large changes in the wavelet coefficients and large variations in the distribution of energy in the different wavelet scales in SWT unlike in DWT and consequently denoising with DWT often results in introducing of artifacts in the signal near discontinuities during signal reconstruction [23], [37], [38].

The proposed method is evaluated for both real and simulated EEG data where both data consist of epileptic seizures and artifacts. By extensive testing, it has been shown that the proposed algorithm can reduce artifacts to an extent that can significantly increase the performance of a seizure detector/classifier, which proves its suitability to use in such EEG applications for epilepsy diagnosis.

The rest of this paper is organized as follows. Section III provides the methods of data collection and synthesis. Section II discribes the proposed method. In Section IV, formulation and analysis for performance evaluation are presented. Section V provides the simulation results and discusses about the performance of the proposed algorithm. Section VI gives concluding remarks.

II. Proposed Algorithm

The first priority of the proposed artifact removal algorithm is not to distort any seizure waveforms at any cost and then to remove artifacts as much as possible. The proposed algorithm has total four stages out of which stage-0, i.e. *reference generation* can be obtained offline prior to the incoming of EEG data. A block diagram for the proposed method is shown in Fig. 1. The rest of the stages can only function during the incoming stream of data. The description of the stages are given below.

A. Reference Generation

This stage generates a reference seizure epoch of length N (i.e. duration of N/F_s second) either from an available seizure-type specific labeled seizure database or from simulating a particular seizure-type epoch by simple mathematical model. For example, the neonatal seizure events can be simulated from a free online database available at [39]. This EEG simulator has mainly two parts: a background simulator and a seizure simulator [40], [41]. On the other hand, if a seizure-type specific database (epilepsy patient database) is available where the seizure events are labeled by the clinicians, then we can also use such database to generate the reference seizure to be used for subsequent stages. However, some preprocessing steps are necessary before starting to use such database. One of them is to band-pass filter the raw database from 0.5 Hz to 30 Hz to eliminate other signal components and to amplify the desired seizure activities (since frequency band of seizure is 0.5$\sim$29 Hz [42], [43]. Let $\mathbf{x}x_{bp}$ be the bandpass-filtered epoch which will be utilized in stage-3 for similarity check.

B. Preprocessing

To begin with, let $x_{\text{raw}}(n)$ denote the sampled raw EEG signal which is sampled at F_s Hz where n is the discrete-time index. We assume that the power-line interference of 50/60 Hz and the baseline of raw EEG have already been removed prior to this preprocessing stage. In the preprocessing, the incoming signal is firstly divided into *non-overlapping* epochs with size of N. Then, the jth epoch is given by

$$\mathbf{x}_j = \begin{bmatrix} x_{\text{raw}}(jN-1) \\ x_{\text{raw}}(jN-2) \\ \vdots \\ x_{\text{raw}}(jN-N) \end{bmatrix} \tag{1}$$

Note that the choice of epoch duration plays an important role in both amount of artifact removal and amount of distortion made to the signal of interest (i.e. seizure events). In addition, in case of automated seizure detection method to work after our proposed algorithm is applied, the value of N will determine the minimum time delay for seizure detection after its onset. This epoch by epoch processing will allow almost no distortion to signal of interest with the penalty of less amount of artifacts to be removed. If N is too low (e.g. $\frac{N}{F_s} < 1$ sec.), then such short duration epoch may not represent seizure waveform properly (the typical duration of seizure event may be several seconds in general) and likely to be confused with artifact waveform (artifacts tend to be more transient than seizure). When N is too high (e.g. $\frac{N}{F_s} > 5$ sec.), then there is high chance that lot of artifacts will be missed to be detected and hence will lower the amount of artifact removal. In our algorithm, we have found $\frac{N}{F_s}$ as 3-sec to be optimum after trying different values empirically. Although EEG signal is nonstationary but it can be considered as stationary for shorter duration epochs (e.g. 1 sec). Therefore for each of such epoch, the statistical properties of time-frequency representation achieved by SWT can be considered as stationary too. Now, sometimes there are some slow artifacts (e.g. ocular or movement artifacts) that last for > 1 sec, e.g. 2-3 sec. So in order to capture the full duration of artifacts the epoch size may be required to more than 1 sec, i.e. 2 or 3 sec. Again, if the epoch size is too large (e.g. > 3

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2015.2457093, IEEE Journal of Biomedical and Health Informatics

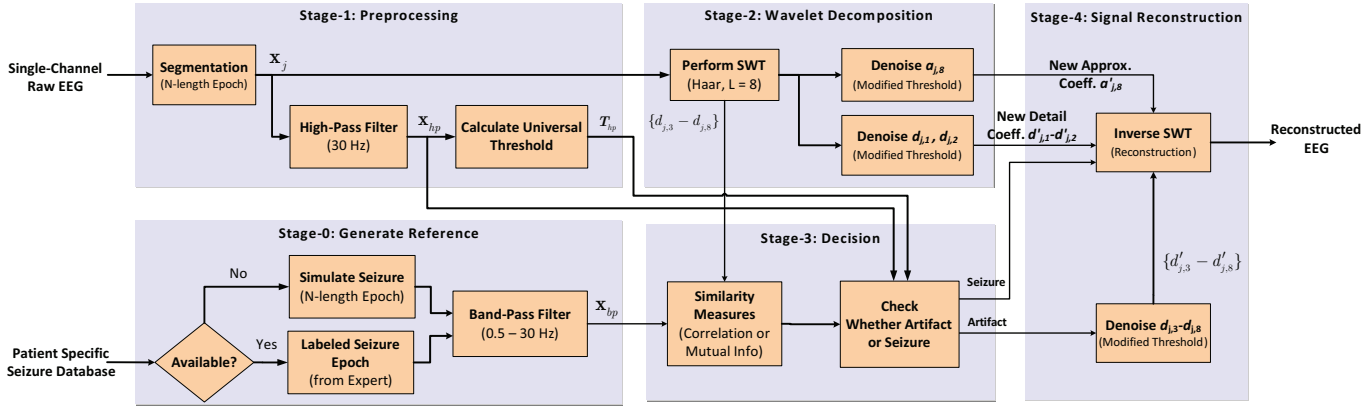

Fig. 1: The overall process flow of the proposed method.

sec), then the stationary assumption of EEG will no longer valid, and consequently brings unavoidable errors in detection followed by removal of artifacts. Another point to note that, in order for automated seizure detection to work in real-time processing of epilepsy data, the epoch size cant be too long; otherwise it would create non-acceptable amount of delay in seizure detection applications.

After segmentation, each epoch is passed through a high-pass filter of 30 Hz to obtain signals which likely to have least seizure information (as seizure activities lie between $0.5 - 29$ Hz band [42], [43]) but contain high frequency artifacts and gamma waves. Then, for every filtered epoch $\mathbf{x}_{hp}$ its corresponding universal threshold [44], [45] is calculated. The reason, computation and use of such threshold will be discussed in stage-3. Finally, both the high-pass filtered epoch and threshold value are passed to stage-3 for double check to decide whether an epoch is artifactual or seizure.

C. Wavelet Decomposition and Denoising

1) Wavelet decomposition: Wavelet decomposition and subsequently removing unwanted artifacts by applying threshold is a familiar denoising process in biomedical signals [38]. Usually, the denoising process refers to removing high frequency noise by thresholding the detail coefficients after wavelet decomposition. However, in this paper, by using the term denoising, we refer to removing artifactual components from neural signals in the wavelet domain, irrespective of whether it is high-frequency or low-frequency artifacts. The objective of this stage is to decompose and analyse the raw epoch with a reasonable time-scale resolution in wavelet domain for possible identification of artifactual components in the later stage. To this end, stationary wavelet transform is performed on the epochs $\{\mathbf{x}_j\}_{j\geq 1}$ with level-8 decomposition by $Haar$ as basis wavelet which results in final approximate $a_{j,8}$ and detail coefficients $d_{j,1}, d_{j,2}, \cdots, d_{j,8}$. Although there are many types of wavelet transform (e.g. DWT, CWT, SWT, etc.), we chose SWT for its advantage of being translational invariant [38]. The choice for level of decomposition is mainly inspired from the bandwidth of EEG signal (i.e. 0.05 - 128 Hz) and dominant frequency band of seizure activities (i.e. 0.5 - 29 Hz) in order to have enough number of frequency

Seizure Activities

SWT Coeff.	$d_{j,1}$	$d_{j,2}$	$d_{j,3}$	$d_{j,4}$	$d_{j,5}$	$d_{j,6}$	$d_{j,7}$	$d_{j,8}$	$a_{j,8}$
Freq. Band (Hz)	64-128	32-64	16-32	8-16	4-8	2-4	1-2	0.5-1	0-0.5

Fig. 2: The frequency bands of wavelet coefficients after performing level-8 SWT on the raw EEG data which has a typical sampling frequency of 256 Hz. Coefficients that correspond to the seizure activities are from $d_{j,3}$ to $d_{j,8}$.

sub-bands to make decisions on where to denoise carefully and where not. Fig. 2 shows the frequency sub-bands of the decomposed detail coefficients and the final-stage approximate coefficient after 8-level SWT is performed on the raw EEG data. It is clear that $\{d_{j,3}, \cdots, d_{j,8}\}$ correspond to the seizure frequency band and hence during denoising process, we need to be very careful to handle these coefficients. The other three coefficients, i.e. $d_{j,1}$, $d_{j,2}$ and $a_{j,8}$ can be denoised by applying the modified universal threshold directly without requiring the decision stage.

2) Denoising: We use non-negative *garrote shrinkage function* during denoising since it has some appealing properties of being less sensitive to input change, having lower bias and being continuous [46]. This is a nice trade-off between soft and hard threshold function in terms of amount of artifact removal and signal distortion [46] and is given by

$$g(j,\ell) = \begin{cases} d_{i,j} & |d_{j,\ell}| \leq t_{j,\ell} \\ \frac{t^2_{j,\ell}}{d_{j,\ell}} & |d_{j,\ell}| > t_{j,\ell}. \end{cases} \quad (2)$$

where $g(j,\ell)$ is the garrote threshold function at each decomposition level of ℓ for epoch j, and $t_{j,\ell}$ denotes the threshold value. It is from the fact that hard threshold function is discontinuous that produces large variance (i.e. very sensitive to small changes in the input data) and hence it induces artifact itself when there is a spike-like transient artifacts. On the other hand, soft threshold has large bias in the denoised signal which results-in under-correction of artifacts. Therefore we decided to choose Garrote threshold function which is a balanced approach between hard and soft threshold and doesnt have the mentioned disadvantages [46].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2015.2457093, IEEE Journal of Biomedical and Health Informatics

To denoise the critical coefficients $\{d_{j,3}, \cdots, d_{j,8}\}$, we have used modified universal threshold reported by [45]

$$t'_{j,\ell} = K\alpha_{j,\ell}\sqrt{2\ln N}, \tag{3}$$

where in (3) N is the length of epoch and $\alpha_{j,\ell}$ is the estimated noise variance for $w_{j,\ell}$ which is usually calculated by following formula [23]

$$\alpha_{j,\ell} = \frac{\text{median}(|w_{j,\ell}|)}{0.6745}. \tag{4}$$

where $w_{j,\ell}$ is the wavelet coefficient at the ℓth decomposition level (i.e. $w_{j,\ell} = a_{j,\ell}$ for approximation coefficient and $w_{j,\ell} = D_{j,\ell}$ for detail coefficient.). The new parameter K in (4) comes from the empirical observations [45]. It is given as

$$K = \begin{cases} K_A \ (0 < K_A < 1) & \text{for thresholding } a_{j,8} \\ K_D \ (1 < K_D < 3) & \text{for thresholding } d_{j,\ell} \end{cases} \tag{5}$$

where, $K = K_A$ is selected for thresholding approximate coefficient $a_{j,8}$ and select $K = K_D$ to threshold all the detail coefficients ($d_{j,\ell},\ \ell = 1, 2, \cdots, 8$). The tuning of parameter K is discussed in Appendix A.

D. Decision

The most important part of our artifact removal algorithm is stage-3, i.e. *Decision*. Depending upon this stage, the decision of whether an epoch is to be detected as artifactual or seizure is made. In addition, if there is possibility for an epoch to be both artifactual and seizure, how carefully that particular epoch to be denoised to remove artifacts, is also decided in this stage. The first step of this stage is to measure the similarity between epochs of decomposed coefficients $\{d_{j,3}, \cdots, d_{j,8}\}$ coming from stage-2 and reference epochs of x_{bp} coming from stage-0. The similarity is measured in terms of either correlation value or mutual information. Depending on the similarity values, we choose two levels of threshold: one is upper limit T_{high} and the other one is lower limit T_{low}. Hence three conditions arise which results in three decisions: if it is high likelyhood to be a seizure, then denoising is not performed on that epoch; if it is in between seizure and artifacts, then we carefully denoise the epoch and finally if it is least likely to be seizure then we fully denoise that epoch. A pseudo code for this decision stage is provided in Table I. To double check apart from the similarity based decision, we also take input from the output of stage-1 where we have a highpass filtered epoch and its threshold value. Since we assume that the epoch x_{hp} is less likely to be seizure and most likely to be artifacts if the value exceeds the calculated threshold value T_{hp}, so if any of the three decisions made from similarity based condition contradicts with this hypothesis, then to be in the safe side, the epoch is not denoised in order to preserve the seizure events all the time. However, in such case, where the epoch is actually artifactual and not seizure, but due to the decision made not to denoise the epoch, we pay the penalty of less artifact reduction.

TABLE I: Pseudo code for the separation of seizures from artifacts. The decision is made by the similarity based thresholding.

Decision Making	Remarks/Comment
If $\|C_i\| \geq T_{C_{high}}$ $k_i = 3$	Don't denoise the epoch
Else if $T_{C_{low}} \leq \|C_i\| \leq T_{C_{high}}$ *if* $\|x_{hp}\| > T_{hp}$ $k_i = 1$ *Else* $k_i = 1.5$	Denoise carefully
Else $k_i = 1$ *End*	Fully denoise

E. Reconstruction

In the final stage of reconstruction, based on the decision stage, we either apply thresholding (fully or carefully) or let the coefficients $\{d_{j,3}, \cdots, d_{j,8}\}$ remain same. Finally with all the new set of coefficients obtained from stage-2 (i.e. $d'_{j,1}$ - $d'_{j,2}$and $a'_{j,8}$) and the ones obtained from the first step of this stage (i.e. $\{d'_{j,3}, \cdots, d'_{j,8}\}$), we apply inverse SWT to reconstruct the EEG epochs. Thus a new sequence of data so called *reconstructed data* is obtained.

III. Methods and Experiments

A. Data Collection

Real EEG recordings are downloaded from CHB-MIT Scalp EEG Database [47] which was collected from the Children's Hospital Boston. The database consists of EEG recordings from pediatric subjects with intractable seizures and the patients were monitored for up to several days. The signals are sampled at 256 Hz with 16-bit resolution. Apart from that, we have also performed some simple experiments to record 32-channel EEG data with a healthy subject by using the commercial Mitsar-EEG-202 recorder as shown in Fig. 3. The subject is asked to perform specific task in order to record and characterize some common artifacts, e.g. chewing, swallowing, head movement, body movement, eye blinking, eye movement, etc. The timing of those tasks are noted down and later confirmed with the corresponding recorded signals.

B. Data Synthesis

1) Semi-Simulated: We have synthesized an artifact-free EEG sequence of 5 min long from real EEG collected from CHB-MIT database as ground truth and different types of simulated artifact waveforms to test our artifact removal algorithm. The process of data synthesis is shown in Fig. 4.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2015.2457093, IEEE Journal of Biomedical and Health Informatics

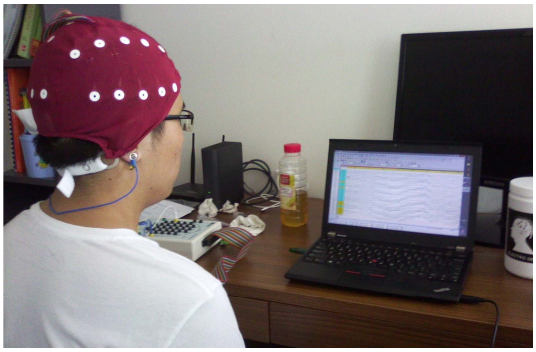

Fig. 3: EEG experiment performed.

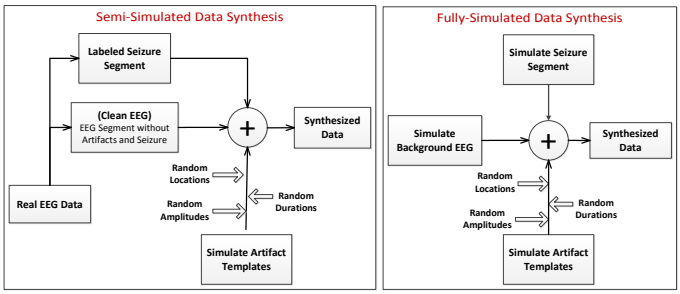

Fig. 4: Illustration of the synthesis process to generate artifactual EEG data with seizure segment.

2) Fully-Simulated: In this dataset, we have simulated all three data components: artifacts, seizure event and EEG background activity (i.e. EEG rhythm), and then combined them together to make artifactual EEG dataset with seizure events. The simulated EEG data have been generated according to the classical theory of Event Related Potentials (ERP) as described in [48]. The MATLAB code we used to generate such simulated EEG is available to download for free from [39].

IV. Perfromance Evaluation

A fair performance evaluation of any artifact removal algorithm has often been an issue because of few reasons; e.g. lack of ground truth data, insuffcicient amount of data used, casual choice of performance metrics and so on. Therefore, it is often seen that only qualitative evaluation is available in time domain plot and/or in spectral domain in terms of PSD plot [20], [22], [37], [49]. In order to have a fair and complete evaluation, enough quantitative results are required along with the traditional qualtitative approach. In addition, further analysis of later stage signal processing is required to observe the afteraffect of artifact removal. Hence this section deals on the way of performance evaluation for both artifact removal and seizure detection accuracy by quantitatively as well as in qualitative manner.

To define the quantitative metrics used for artifact removal and seizure detection, we define $x_{\rm ref}(n)$, $x_{\rm art}(n)$ and $x_{\rm rec}(n)$ as the discrete time signals of length L representing clean reference signal (artifact-free), artifactual and reconstructed signal respectively. Then, the error signal before and after artifact removal can be defined respectively as

$$e_{\rm br}(n) = x_{\rm art}(n) - x_{\rm ref}(n), \tag{6}$$
$$e_{\rm ar}(n) = x_{\rm rec}(n) - x_{\rm ref}(n). \tag{7}$$

In the sequel, we introduce the metrics used for artifact removal and seizure detection in more detail.

A. Metrics for Artifact Removal

The performance of the proposed algorithm on the artifact removal has been evaluated both in terms of the amount of artifact reduction and the amount of distrotion it brings into the signal of interest, specially to the seizure events. Several efficiency metrics have been calculated in both time and spectral domain to quantify such evaluation. In order to have fair evaluation and clear idea we also have considered the amount and duration of artifacts present in the signals. The metrics are described as follows:

1) $\boldsymbol{\lambda}$: The reduction in artifact is calculated using the following formula [18]

$$\boldsymbol{\lambda} = 100\left(1 - \frac{c_{\rm ref} - c_{\rm rec}}{c_{\rm ref} - c_{\rm art}}\right), \tag{8}$$

Here, $c_{\rm ref}$ denotes the auto-correlation coefficient of the reference signal at time lag 1, $c_{\rm art}$ and $c_{\rm rec}$ are the cross-correlation coefficient between reference signal with artifactual and reconstructed signal respectively.

2) $\mathbf{\Delta SNR}$: Assuming the signals have zero mean, then $\mathbf{\Delta SNR}$ is the difference in SNR before and after artifact removal is given by the following formula [18]

$$\mathbf{\Delta SNR} = 10\log_{10}\left(\frac{\sigma^2_{x_{\rm ref}}}{\sigma^2_{e_{\rm br}}}\right) - 10\log_{10}\left(\frac{\sigma^2_{x_{\rm ref}}}{\sigma^2_{e_{\rm ar}}}\right), \tag{9}$$

where $\sigma^2_{x_{\rm ref}}$, $\sigma^2_{e_{\rm br}}$ and $\sigma^2_{e_{\rm ar}}$ be the variance of reference signal, error signal before and after artifact removal respectively.

3) $\mathbf{RMSE}$: The root mean square error is calculated as follows

$$\mathbf{RMSE} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}[e_{\rm ar}(n)]^2}. \tag{10}$$

4) $\mathbf{P_{dis}}$: Denote $\mathcal{P}_{\rm ref}(f)$, $\mathcal{P}_{\rm art}(f)$ and $\mathcal{P}_{\rm rec}(f)$ the power spectral densities of reference signal, artifactual signal and reconstructed signal respectively, the spectral distortion $\mathbf{P_{dis}}$ is calculated as follows

$$\mathbf{P_{dis}} = \frac{\sum_{f=1}^{F_s/2}\left(\mathcal{P}_{\rm rec}(f)\right)^2}{\sum_{f=1}^{F_s/2}\left(\mathcal{P}_{\rm ref}(f)\right)^2}. \tag{11}$$

5) $\mathbf{\Delta Cor}$: Correlation is the measure of similarity between two time series in time domain. In order to calculate the improvement in correlation $\mathbf{\Delta Cor}$ due to artifact removal, the following equation is used

$$\mathbf{\Delta Cor}(\%) = \frac{c_{\rm rec} - c_{\rm art}}{c_{\rm art}} \times 100 \tag{12}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2015.2457093, IEEE Journal of Biomedical and Health Informatics

where $c_{\rm art}$ and $c_{\rm rec}$ are the cross-correlation coefficients between reference signal with artifactual and reconstructed signal respectively.

6) $\mathbf{\Delta Coh}$: Coherence is the measure of similarity between two time series in frequency domain and it is defined between two signals $x(t)$ and $y(t)$ as:

$$\mathbf{\Delta Coh} = \frac{|\mathcal{G}_{xy}|^2}{\mathcal{G}_{xx}\mathcal{G}_{yy}}, \tag{13}$$

where $|\mathcal{G}_{xy}|$ is the cross-spectral density between $x(t)$ and $y(t)$; $\mathcal{G}_{xx}$ and $\mathcal{G}_{yy}$ are the auto-spectral density of $x(t)$ and $y(t)$ respectively. Now, we assume $\mathbf{Coh_{bef}}$ be the coherence between reference and artifactual signal while $\mathbf{Coh_{aft}}$ be the coherence between reference and recosntructed signal, then the average improvement in coherence due to artifact removal denoted by $\mathbf{\Delta Coh}$ is calculated by following equation:

$$\mathbf{\Delta Coh}(\%) = \frac{\mathbf{Coh_{aft}} - \mathbf{Coh_{bef}}}{\mathbf{Coh_{bef}}} \times 100 \tag{14}$$

7) $\mathbf{SNDR}$: Signal to noise and distortion ratio in frequency domain is calculated as follows:

$$\mathbf{SNDR} = 10\log_{10}\left(\frac{\mathbf{P_{ref}}}{\mathbf{P_{e_r}}}\right). \tag{15}$$

where $\mathbf{P_{ref}}$ and $\mathbf{P_{e_r}}$ are the power spectral densities of $x_{ref}(n)$ and $e_r(n)$ respectively. $e_r(n)$ is $e_{br}(n)$ for before and $e_{ar}(n)$ for after artifact removal respectively.

8) $\mathbf{SNR_{art}}$: Artifact SNR is calculated considering artifact as signal and reference neural signal as noise using the following formula

$$\mathbf{SNR_{art}} = 10\log_{10}\left(\frac{\sigma^2_{e_{\rm br}}}{\sigma^2_{x_{\rm ref}}}\right). \tag{16}$$

9) $\mathbf{\Delta T_{art}}$: It denotes the artifact duration out of total data length in percentage and calculated as follows

$$\mathbf{\Delta T_{art}}(\%) = \frac{T_{\rm art}}{T_{\rm total}} \times 100, \tag{17}$$

where $T_{\rm art}$ and $T_{\rm total}$ are the time duration of artifact and whole data sequence respectively.

B. Metrics for Seizure Detection

Seizure detection/classification still is an active research problem in the epilepsy research community. There are several seizure detection methods available in the literature and none of them can claim to be robust for every patient and in every recording/surrounding enovironment as most of them are evaluated based on small quantity of dataset and do not consider the effects of all types of artifacts. However, since the purpose of this study is not to develope a seizure detection algorithm but to verify the performance of seizure analysis after the proposed artifact removal algorithm is applied, therefore in this section we will show some examples of simple seizure analysis or measurement available in the literature to prove the efficacy of the artifact removal algorithm.

Fig. 5: Process flow for validation of seizure detection.

1) Feature Extraction: Feature extraction is an improtant stage for classification in machine learning on which both classification performance and classifier complexity greatly depends. There are many ways of extracting EEG features for seizure classification that are mentioned in the literature [50]–[53]. Most of them use the statistical features (e.g. entropy [54], [55], kurtosis, skewness, line length [50], variance, min, max, maxima count, etc.) either from directly time domain, or from both time and frequency domain or even some combined spatial (channel-wise) domain based features along with time and frequency. Some recent literatures also use wavelet domain based features to extract the desired frequency sub-bands [52], [53]. In this paper, we use a single-dimension feature so called Sample Entropy which is described below.

- Sample Entropy: Sample entropy or SampEn which was introduced by [54], quantifies the complexity of a time series data and recently it has become an attractive measure in analysing non-linear physiological signals [55]. Unlike other entropy or complexity measures (e.g. approximate entropy or ApEn), the advantage of SampEn is that it is resistant to the short-duration transient interferences like spikes. It is the negative natural logarithm of an estimate of the conditional probability that if two sets of simultaneous data points of length m match point wise within a tolerance r then two sets of simultaneous data points of length $m+1$ also match pointwise within the tolerance r and represented as $SampEn(m, r, N)$ where m, r and N are the embedding dimension, tolerance and number of data points respectively [54].
  We assume a time-series data epoch of length $N = x_1, x_2, x_3, ..., x_N$; a template vector of length m, such that $X_m(i) = x_i, x_{i+1}, x_{i+2}, ..., x_{i+m-1}$ and the distance function $d[x_m(i), x_m(j)]$ for $i \neq j$. Now the number of vector pairs in template vectors of length m and $m+1$ having $d[x_m(i), x_m(j)] < r$ are denoted by B and A respectively. Thus the sample entropy is defined as

$$SampEn = -\log_e(\frac{A}{B}) \tag{18}$$

  where, A = no of template vector pairs having $d[x_m(i), x_m(j)] < r$ of length $m+1$
  and B = no of template vector pairs having $d[x_m(i), x_m(j)] < r$ of length m

2) SVM Classification: Support vector machine is a supervised learning based classifier which is widely used in simple binary linear classification [56]. In our problem of classifying seizure epoch from non-seizure epoch, we have used a simple SVM classifier whose input is the extracted SampEn feature. Initially the classifier is trained with the SampEn values calculated from each epoch as training samples. The benchmark epochs are obtained from reference signal where

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2015.2457093, IEEE Journal of Biomedical and Health Informatics

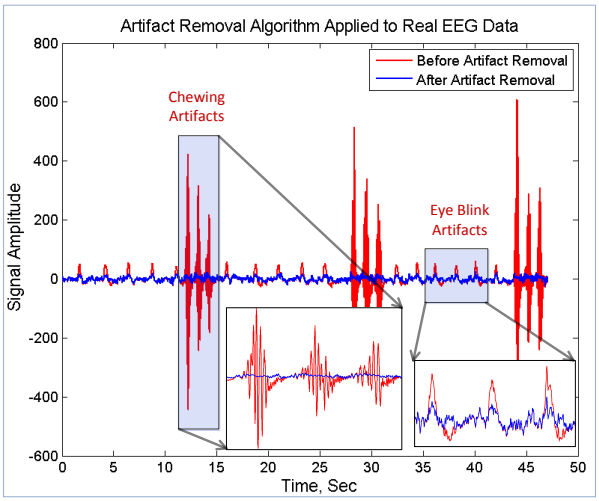

Fig. 6: The removal result after the proposed algorithm is applied on our recorded EEG.

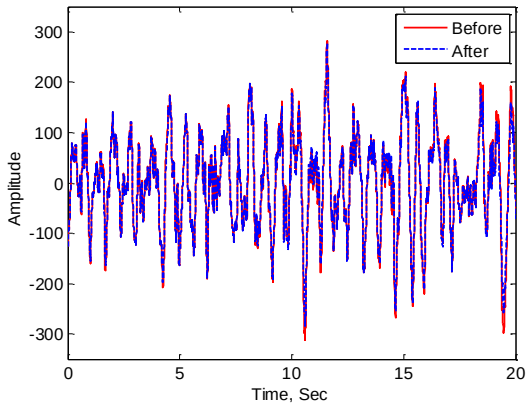

Fig. 7: The removal result after the proposed algorithm is applied on an artifact-free seizure segment labeled and collected from MIT-CHB dataset. The reconstruction is almost perfect when there is no visible artifact.

both artifact free seizure and artifact-free non-seizure epochs have known label. Then the classifier is tested with SampEn values calculated from artifactual signal epochs. Finally the same designed classifier is again tested with the SampEn values calculated from reconstructed signal epochs. In both cases, the no. of true positives and false positives are recorded. The process flow of seizure detection after artifact removal is shown in Fig. 5.

- $\mathbf{\Delta F}(\%)$: If $F_{\rm bef}$ and $F_{\rm aft}$ denote the number of false positives for seizure classification before and after artifact removal respectively, then the improvement in number of false positives, i.e. $\Delta \mathrm{F}(\%)$ is given by

$$\mathbf{\Delta F}(\%) = \frac{F_{\rm bef} - F_{\rm aft}}{F_{\rm bef}} \times 100 \qquad (19)$$

V. Results and Discussion

A. Qualitative Evaluation

1) Real Data: The proposed algorithm is applied to our recorded EEG data from a healthy subject with labeled *chewing* and *eye-blink* artifacts. The artifact removal result in terms of time-domain plot is shown in Fig. 6 for qualitative evaluation which suggests a satisfactory removal of both types of artifacts without distorting the background EEG signals in the non-artifactual region.

Another example of artifact removal result is illustrated in Fig. 7 where there is an artifact-free seizure segment is present. It is obvious from the time-course data that almost perfect reconstruction of seizure activities occurs.

2) Semi Simulated: An example of artifact removal algorithm applied on semi-simulated data is presented in Fig. 8. The artifactual data sequence is made up with real seizure and real background EEG data where simulated artifacts are superimposed. Its obvious that the algorithm can't remove all of the artifacts all the time, but can reduce them significantly most of the time and more importantly can still preserve the desired seizure activities pretty well. This qualitative illustration in time domain data shows a better visualization

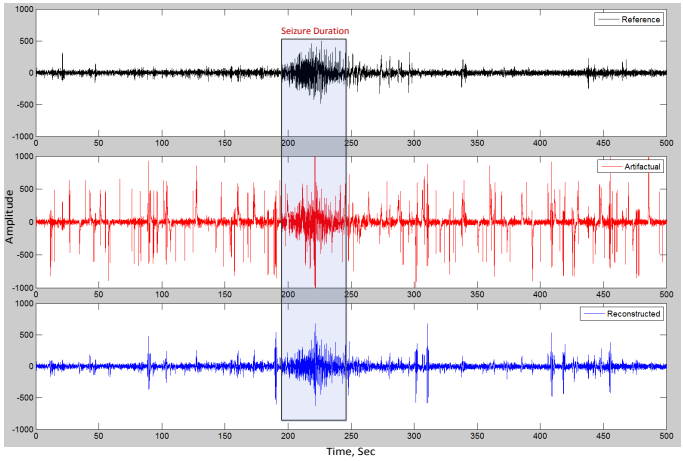

Fig. 8: The removal result after the proposed algorithm is applied on the semi-simulated EEG data.

to detect seizure offline after significant reduction of most of the artifacts as usually done by the clinicians.

3) Fully Simulated: The artifact removal result from a fully simulated data is illustrated in Fig. 9 where all three signal components are simulated. The synthesized artifactual data is severely contaminated with different types of artifacts and thus makes it difficult to detect the segment of seizure activities properly. Once most of the artifacts are reduced, its now easy to detect seizure segment which also increases the true positive detection.

Another example of all six types of simulated artifacts and their reduction is shown in Fig. 10. Here each plot shows each type of artifact contaminated segment before and after artifacts are removed along with the reference artifact-free segment. This qualitative illustration proves that most of the time when there is no seizure, the algorithm can significantly reduce each type of artifacts.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2015.2457093, IEEE Journal of Biomedical and Health Informatics

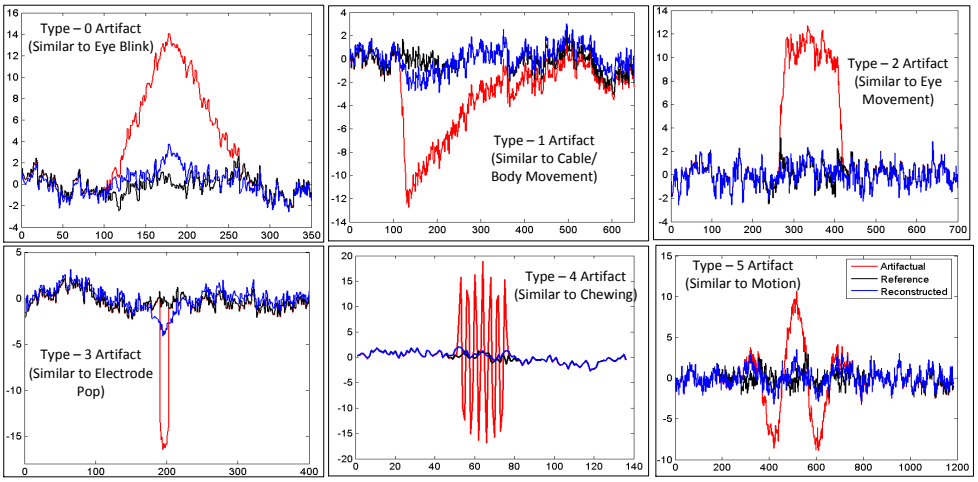

Fig. 10: Six types of different simulated artifacts that mimicking real artifacts found in the typical scalp EEG recording environments. The application of proposed artifact removal algorithm can almost successfully remove such artifacts most of the time without distorting the background EEG signals. The black, red and blue traces denoting reference, artifactual and reconstructed simulated EEG data respectively. The y-axis is normalized signal amplitude and x-axis is time index with sampling frequency of 256 Hz.

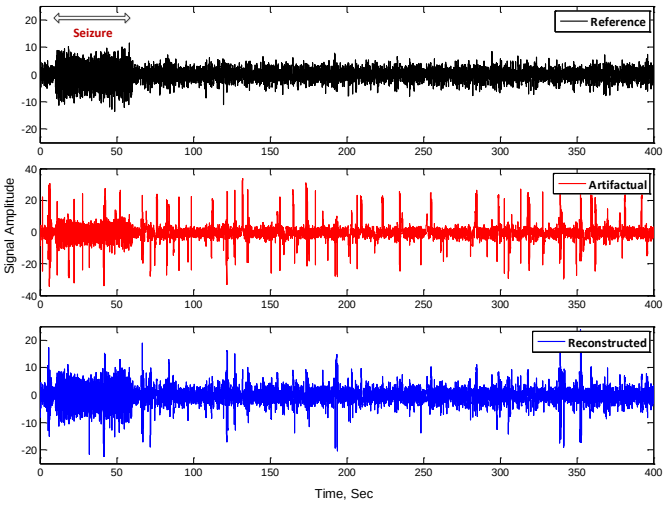

Fig. 9: Artifact removal result applied to semi-simulated dataset-1. The plot is a time course data where all six types of artifacts are present. Note that, not all of the artifacts are removed or attenuated. The reason is that in order to preserve the seizure events, the amount of artifact reduction has been compromised. The y-axis unit is normalized signal amplitude.

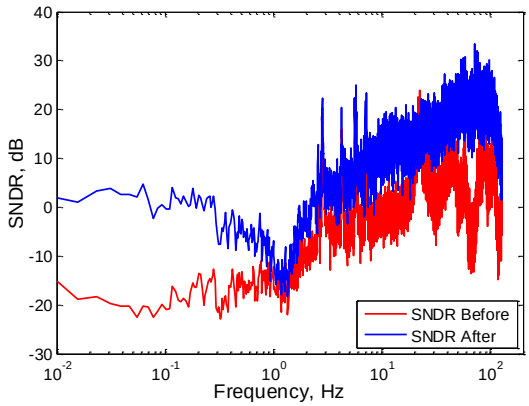

Fig. 11: SNDR for signals before and after artifact removal clearly shows the improvement in signal quality over the entire frequency band.

B. Quantitative Evaluation

This sub section quantifies the results obtained both in terms of artifact removal and the consequence of artifact removal, i.e. improvement in seizure detection.

1) Artifact Removal Results: As discussed in Section-V, we have calculated several time and frequency domain metrics to quantify both amount of artifacts removed as well as amount of distortion made with respect to both amount and intensity of artifacts present in the data. Fig. 11 shows the calculated **SNDR** over the entire frequency bandwidth of EEG data for fully simulated data sequence as in Fig. 9 before and after artifacts are removed. It is clearly seen that a significant improvement of **SNDR** on an average of 5-10 dB over the entire frequency is made due to artifact removal which proves the efficacy of the proposed algorithm. Table II presents the quantitative metrics of artifact removal with respect to the strength of artifacts, i.e. different artifact SNR $\mathbf{SNR_{art}}$. Table III presents the quantitative metrics of artifact removal with respect to different artifact duration $\mathbf{\Delta T_{art}}$.

2) Comparison with Other Methods: We have compared the performance our proposed method with few state-of-the-art artifact removal methods (i.e. wavelet-BSS and EMD-BSS based methods) in terms of both quantitative removal metrics (Figure 13) and computational time (Figure 14) to roughly illustrate a comparison the efficacy of our method compared with others. The process flow of wavelet-BSS and EMD-BSS

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2015.2457093, IEEE Journal of Biomedical and Health Informatics

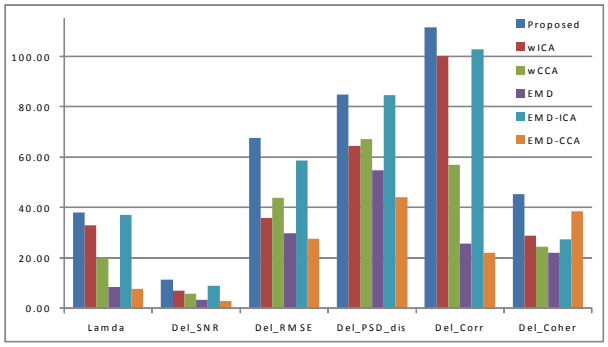

Fig. 13: Comparison of proposed method with respect to few available artifact removal methods in terms of the quantitative metrics.

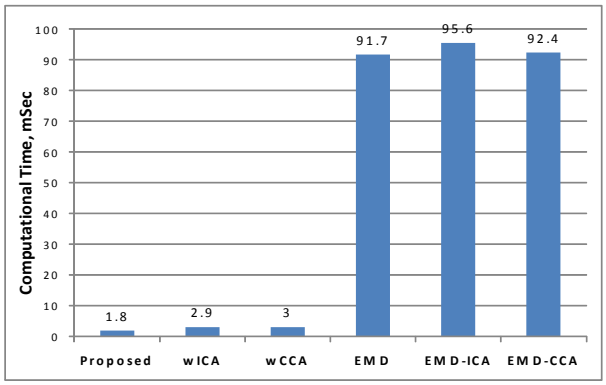

Fig. 14: Comparison of proposed method with respect to few available artifact removal methods in terms of the computational time required to process each 1 second of data in MATLAB simulation.

based methods is shown in Figure 12. Note that the epoch-by-epoch processing is only applied for our proposed method and for others we just process the whole sequence at once as it's difficult for BSS-based methods to function properly to separate components with small duration of data (i.e. epoch).

3) Seizure Detection Results: In order to show that artifact removal by this proposed algorithm not only makes offline analysis during seizure detection easier and more accurate, but also helps the availabe automated seizure detector (ASD) to improve their performance significantly. An example of false alarms due to artifacts is presented in Fig. 15. Here we present three sequences of fully simulated EEG data: reference, artifactual, reconstructed and their corresponding SamEn values calculated with 2-sec time window. For an ideal case, i.e. without artifact, seizure and non-seizure segments can be easily separated by comparing their average sample entropy. However, in practice, EEG sequence is always contaminated with different artifacts and hence it may introduce some false alarms due to artifacts. Once most of the artifacts are reduced, the no. of false alarms is also reduced significantly as illustrated clearly in the Fig. 15.

A quantitative representation of amount of improvement in seizure detection is illustrated in Fig. 16 where the no. of false positives are plotted for both before and after artifact removal.

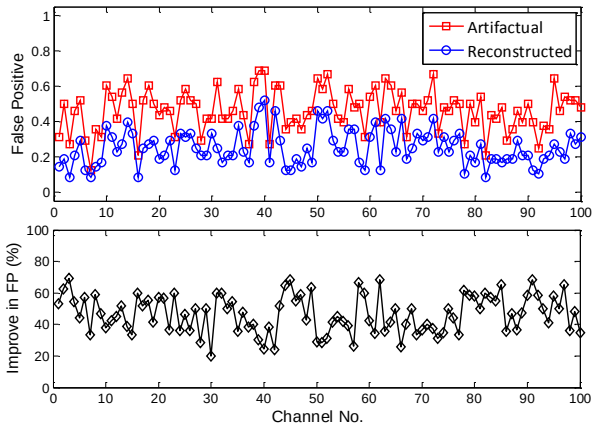

Fig. 16: False positive (FP) before and after artifact removal (shown on top) and improvement of false positive (ΔFP) in percentage (shown on bottom) are plotted with respect to different data sequence/channel no. For seizure detection purpose, SampEn is used as feature and SVM as classifier.

In addition the corresponding improvement in $\mathbf{\Delta F}(\%)$ which suggests that on average 20-80% improvement is possible after artifacts are reduced by proposed algoirthm. Please note that, for this quantification we have simulated 100 different data sequences each of 200-sec duration where 100-sec is seizure and rest 100-sec is non-seizure segment. Then for each epoch of 2-sec, SampEn is calculated and used as a feature for SVM classification to quantify the no. of FP. The no. of TP in this simulation remains always 100% due to the single feature selection. However, these values are not absolute and may differ depending on the type and size of features, type of classifier used, length of the seizure segment, epoch duration and so on. The results are given only for simple understanding of the fact that artifacts removal with proposed algorithm can significantly improve the performance of seizure detection.

Fig. 17 shows some common statistical features used in literature for differentiating seizure epochs from non-seizure ones for both before and after artifact removal. It is obvious from the plots that after artifact removal, the features are easier to distinguish than before artifact removal. Hence the proposed algorithm can also be useful in improving seizure detector performance in other seizure detection algorithms where combination of different statistical features are used for classification.

VI. Conclusions

The purpose of this research was to develop an artifact removal method in order to make the seizure analysis process easier for the clinicians and also to improve the performance of the available automated seizure detection algorithm. In addition, such artifact removal which preserves the seizure events, can greatly reduce the labor and complexity of seizure detection by making it easy to analyze underlying signal of interest. To ensure a fair performance evaluation of the proposed method, we performed extensive simulations on both real and synthesized data with several metrics to quantify

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2015.2457093, IEEE Journal of Biomedical and Health Informatics

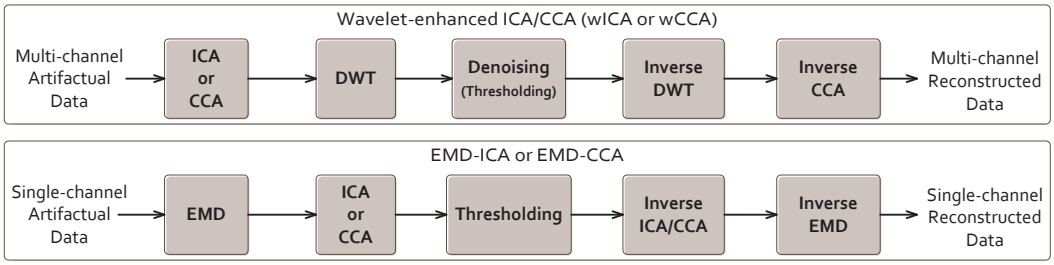

Fig. 12: General Process Flow of EMD-BSS and Wavelet-BSS Methods

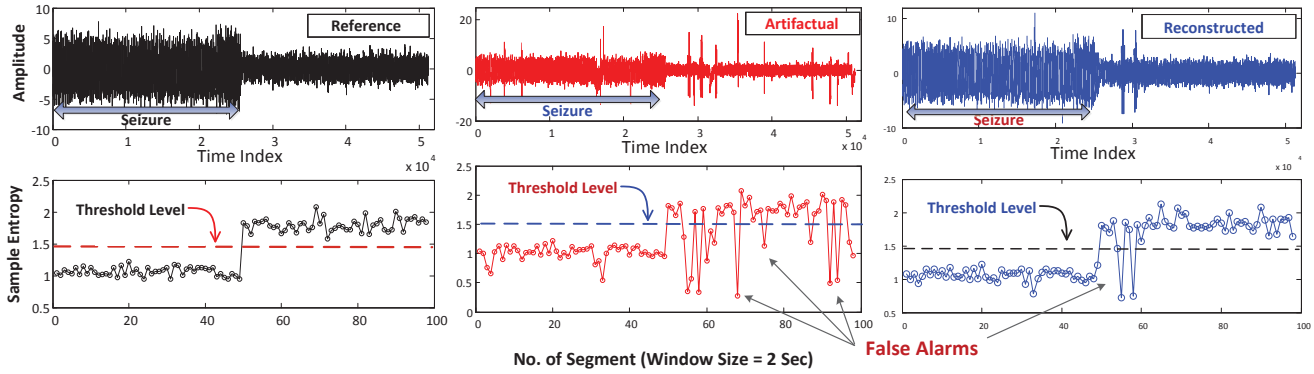

Fig. 15: An example of false alarms due to artifacts is illustrated where sample entropy is chosen as a feature to separate seizure from non-seizure (normal) events. Artifact removal can significantly reduce the false alarms by reducing the amount of artifacts. The y-axis unit is normalized signal amplitude in the top plots.

results. Also an analsyis of a simple seizure detection proves the efficacy of the method that seizure detection accuracy can be significantly improved. The results are impressive and further improvement of the current algorithm to be able to remove artifacts in real-time will surely be a breakthrough in epilepsy patients monitoring. It is, therefore, expected to have more analysis on this particular research to enhance the quality of epilepsy patients by ensuring proper seizure diagnosis and treatment.

APPENDIX A
TUNING OF PARAMETER K

The tuning of parameter K_A depends on the data distribution of $a_{j,8}$ epoch which less likely contains any seizure events (since its frequency band is 0-0.5 Hz), but contains both the *delta* wave and some low-frequency artifacts (mostly large amplitude slow movement artifacts). So when the histogram of the data has large deviation from its standard deviation (large tail on the histogram on either one side or both), it is more likely due to presence of such artifacts. Therefore a value less than 1 is chosen for K_A and if there is no such unusual tail present, then $K_A = 1$ is chosen that makes the threshold exactly same as the original universal threshold, i.e. $t'_{j,\ell} = t_{j,1}$. The criterion for the choice of K_A is given below

$$K_A = \begin{cases} 1 & \text{if } \max(|A_{j,8}|) > m \times \sigma_j, \\ 0 \leq k_A < 1 & \text{otherwise}, \end{cases} \tag{20}$$

where σ_j denotes the standard deviation of $A_{j,8}$. The value of m is based on the parameter tuning and can be obtained from some initial several seconds of incoming raw *EEG* data samples to update the threshold value. From the empirical studies, the value of m is found as minimum of 3, i.e. $3 < m < \infty$ (See Appendix A of [45]).

In order to calculate the value of k_D, *Decision* stage helps. Since $D_{j,1}$ and $D_{j,2}$ contains higher frequency activities than the seizure frequency band (see Fig. 2), therefore value of k_D for first two detail coefficients is selected as 1, i.e. same as universal threshold. For rest of the detail coefficients, (i.e. $\{D_{j,3}, \cdots, D_{j,8}\}$) the value is chosen based on the decision stage's output as shown in I. The equation is given below:

$$K_D = \begin{cases} 3 & \text{case-1} \\ 1 < k_{j,\ell} \leq 1.5 & \text{case-2} \quad \text{where } \ell = 3, 4, \cdots, 8 \\ 1 & \text{case-3} \end{cases} \tag{21}$$

ACKNOWLEDGMENT

We thank our group members Peng Sun and Wing-kin Tam for participating in our EEG experiment to record artifacts. The experimental seizure data recorded from epilepsy patients are provided by MIT-CHB online database downloaded from *physionet.org*. We also express our gratitude to Dr. Chan Wei Shih Derrick from Dept of Paediatrics, Neurology Service, KK Women's and Children's Hospital, Singapore, for his expert opinion and suggestion regarding epilpesy seizure

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2015.2457093, IEEE Journal of Biomedical and Health Informatics

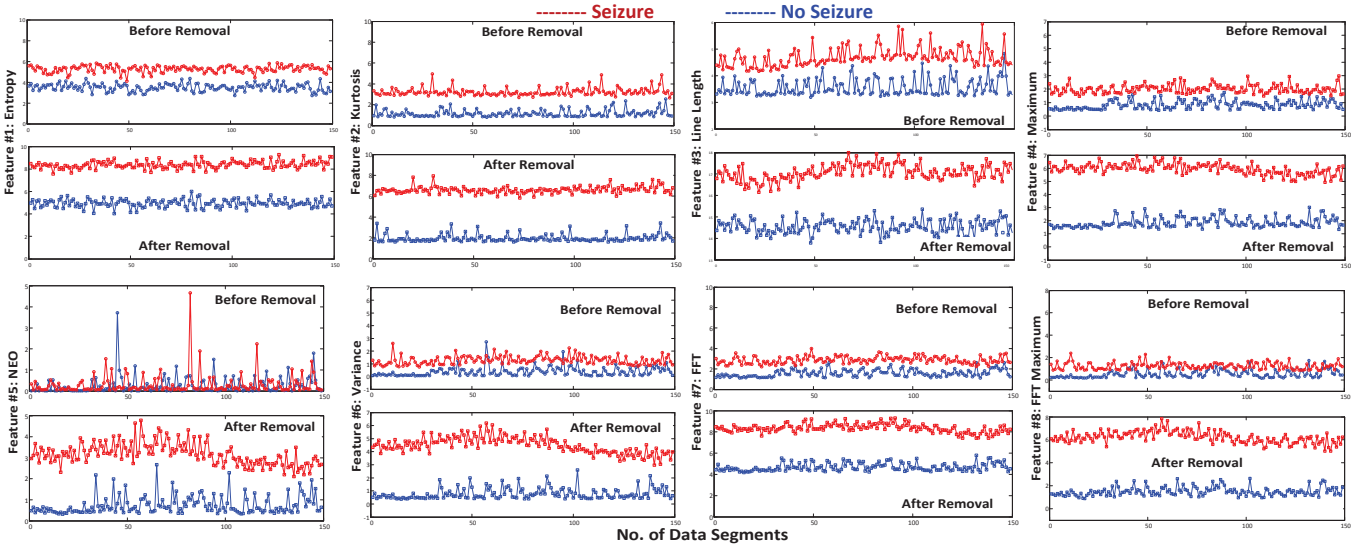

Fig. 17: Different features from EEG data for seizure and non-seizure events calculated for each data segment of time window 1 sec.

TABLE II: Quantitative Metrics of Artifact Removal Results for Different Artifact SNR (SNR_{Art}).

SNR_{Art} (dB)	Performance Metrics					
	λ	ΔSNR	ΔRMSE (%)	$\Delta\text{PSD}_{\text{dis}}$ (%)	ΔCorr (%)	ΔCoher (%)
5	62.3	8.5	62.2	90.5	63.8	25.5
10	48.5	9.6	67.1	98.8	110.5	53.8

TABLE III: Quantitative Metrics of Artifact Removal Results for Different Artifact Durations ($\Delta\text{T}_{\text{Art}}$).

$\Delta\text{T}_{\text{Art}}$ (%)	Performance Metrics					
	λ	ΔSNR	ΔRMSE (%)	$\Delta\text{PSD}_{\text{dis}}$ (%)	ΔCorr (%)	ΔCoher (%)
20	44.3	7.8	75.9	98.8	84.5	46.7
25	57.0	8.6	66.5	97.8	75.5	29.8
30	66.2	9.1	51.7	86.3	67.9	30.2
35	57.8	10.8	49.9	99.9	139.7	45.9
40	68.8	11.1	61.1	99.6	113.5	30.7
Mean 30.87%	Mean 54.71%	Mean 8.27%	Mean 63.92%	Mean 97.55%	Mean 98.91%	Mean 43.05%

detection and affect of artifacts on seizure detection from an Electroencephalogist's point of view.

The authors would like to acknowledge the funding support by A*STAR PSF Grant R-263-000-699-305, NUS YIA Grant R-263-000-A29-133, and MOE R-263-000-A47-112.

REFERENCES

[1] R. S. Fisher, B. G. Vickrey, P. Gibson, B. Hermann, P. Penovich, A. Scherer, and S. Walker, "The impact of epilepsy from the patients perspective i. descriptions and subjective perceptions," *Epilepsy Research*, vol. 41, no. 1, pp. 39 – 51, 2000.

[2] S. F. S. O. Regan and W. Marnane, "Automatic detection of eeg artefacts arising from head movements using eeg and gyroscope signals," *Medical Engineering and Physics*, vol. 35, pp. 867–874, 2013.

[3] A. Bertrand, V. Mihajlovic, B. Grundlehner, C. Van Hoof, and M. Moonen, "Motion artifact reduction in eeg recordings using multi-channel contact impedance measurements," in *Biomedical Circuits and Systems Conference (BioCAS), 2013 IEEE*, Oct 2013, pp. 258–261.

[4] J. T. Gwin, K. Gramann, S. Makeig, and D. P. Ferris, "Removal of movement artifact from high-density eeg recorded during walking and running," *Journal of Neurophysiology*, vol. 103, no. 6, pp. 3526–3534, 2010.

[5] X. Yong, M. Fatourechi, R. Ward, and G. Birch, "Automatic artefact removal in a self-paced hybrid brain- computer interface system," *Journal of NeuroEngineering and Rehabilitation*, vol. 9, no. 1, p. 50, 2012.

[6] K. Sweeney, T. Ward, and S. McLoone, "Artifact removal in physiological signals x2014;practices and possibilities," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 16, no. 3, pp. 488–500, May 2012.

[7] B. Noureddin, P. Lawrence, and G. Birch, "Online removal of eye movement and blink eeg artifacts using a high-speed eye tracker," *Biomedical Engineering, IEEE Transactions on*, vol. 59, no. 8, pp. 2103–2110, Aug 2012.

[8] P. He, G. Wilson, and C. Russell, "Removal of ocular artifacts from electro-encephalogram by adaptive filtering," *Medical and Biological Engineering and Computing*, vol. 42, no. 3, pp. 407–412, 2004. [Online]. Available: http://dx.doi.org/10.1007/BF02344717

[9] C. Guerrero-Mosquera and A. Navia-Vazquez, "Automatic removal of ocular artefacts using adaptive filtering and independent component analysis for electroencephalogram data," *Signal Processing, IET*, vol. 6, no. 2, pp. 99–106, April 2012.

[10] W.-Y. Hsu, C.-H. Lin, H.-J. Hsu, P.-H. Chen, and I.-R. Chen, "Wavelet-based envelope features with automatic {EOG} artifact removal: Application to single-trial EEG data," *Expert Systems with Applications*, vol. 39, no. 3, pp. 2743 – 2749, 2012.

[11] B. S. Raghavendra and D. N. Dutt, "Wavelet enhanced cca for minimization of ocular and muscle artifacts in eeg," vol. 5, no. 9, pp. 846 – 851, 2011.

[12] S. Puthusserypady and T. Ratnarajah, "H infin; adaptive filters for eye blink artifact minimization from electroencephalogram," *Signal Processing Letters, IEEE*, vol. 12, no. 12, pp. 816–819, Dec 2005.

[13] A. Schlgl, C. Keinrath, D. Zimmermann, R. Scherer, R. Leeb, and G. Pfurtscheller, "A fully automated correction method of EOG artifacts

in EEG recordings," *Clinical Neurophysiology*, vol. 118, no. 1, pp. 98 – 104, 2007.

[14] S. Puthusserypady and T. Ratnarajah, "Robust adaptive techniques for minimization of EOG artefacts from EEG signals," *Signal Processing*, vol. 86, no. 9, pp. 2351 – 2363, 2006, special Section: Signal Processing in UWB Communications.

[15] A. Flexer, H. Bauer, J. Pripfl, and G. Dorffner, "Using ICA for removal of ocular artifacts in EEG recorded from blind subjects," *Neural Networks*, vol. 18, no. 7, pp. 998 – 1005, 2005.

[16] N. P. Castellanos and V. A. Makarov, "Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis," *Journal of Neuroscience Methods*, vol. 158, no. 2, pp. 300 – 312, 2006.

[17] C. Zhao and T. Qiu, "An automatic ocular artifacts removal method based on wavelet-enhanced canonical correlation analysis," in *Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE*, 2011, pp. 4191–4194.

[18] K. Sweeney. PhD Thesis, National University of Ireland Maynooth, 2013.

[19] M. Zima, P. Tichavsk, K. Paul, and V. Krajca, "Robust removal of short-duration artifacts in long neonatal EEG recordings using wavelet-enhanced ICA and adaptive combining of tentative reconstructions," *Physiological Measurement*, vol. 33, no. 8, p. N39, 2012.

[20] A. Flexer, H. Bauer, J. Pripfl, and G. Dorffner, "Using ica for removal of ocular artifacts in eeg recorded from blind subjects," *Neural Networks*, vol. 18, no. 7, pp. 998–1005, 2005.

[21] C. Guerrero-Mosquera and A. Navia-Va zquez, "Automatic removal of ocular artefacts using adaptive filtering and independent component analysis for electroencephalogram data," *Signal Processing, IET*, vol. 6, no. 2, pp. 99–106, 2012.

[22] C. A. Joyce, I. F. Gorodnitsky, and M. Kutas, "Automatic removal of eye movement and blink artifacts from eeg data using blind component separation," *Psychophysiology*, vol. 41, no. 2, pp. 313–325, 2004.

[23] D. Safieddine, A. Kachenoura, L. Albera, G. n. l. Birot, A. Karfoul, A. Pasnicu, A. Biraben, F. Wendling, L. Senhadji, and 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. Sig. Proc.*, vol. 2012, p. 127, 2012.

[24] S.-Y. Shao, K.-Q. Shen, C. J. Ong, E. Wilder-Smith, and X.-P. Li, "Automatic eeg artifact removal: a weighted support vector machine approach with error correction," *Biomedical Engineering, IEEE Transactions on*, vol. 56, no. 2, pp. 336–344, 2009.

[25] J. Mateo, A. M. Torres, and M. A. García, "Eye interference reduction in electroencephalogram recordings using a radial basic function," *Signal Processing, IET*, vol. 7, no. 7, pp. 565–576, 2013.

[26] K. Nazarpour, Y. Wongsawat, S. Sanei, J. A. Chambers, and S. Oraintara, "Removal of the eye-blink artifacts from eegs via stf-ts modeling and robust minimum variance beamforming," *Biomedical Engineering, IEEE Transactions on*, vol. 55, no. 9, pp. 2221–2231, 2008.

[27] R. Mahajan and B. Morshed, "Unsupervised eye blink artifact denoising of eeg data with modified multiscale sample entropy, kurtosis and wavelet-ica," 2015.

[28] J. J. Kierkels, J. Riani, J. W. Bergmans, and G. J. van Boxtel, "Using an eye tracker for accurate eye movement artifact correction," *Biomedical Engineering, IEEE Transactions on*, vol. 54, no. 7, pp. 1256–1267, 2007.

[29] Z. Wang, P. Xu, T. Liu, Y. Tian, X. Lei, and D. Yao, "Robust removal of ocular artifacts by combining independent component analysis and system identification," *Biomedical Signal Processing and Control*, vol. 10, pp. 250–259, 2014.

[30] C. Burger and D. J. van den Heever, "Removal of eog artefacts by combining wavelet neural network and independent component analysis," *Biomedical Signal Processing and Control*, vol. 15, pp. 67–79, 2015.

[31] M. A. Klados, C. Papadelis, C. Braun, and P. D. Bamidis, "Reg-ica: A hybrid methodology combining blind source separation and regression techniques for the rejection of ocular artifacts," *Biomedical Signal Processing and Control*, vol. 6, no. 3, pp. 291–300, 2011.

[32] J. Ma, P. Tao, S. Bayram, and V. Svetnik, "Muscle artifacts in multichannel eeg: characteristics and reduction," *Clinical neurophysiology*, vol. 123, no. 8, pp. 1676–1686, 2012.

[33] N. Mammone and F. C. Morabito, "Enhanced automatic artifact detection based on independent component analysis and renyis entropy," *Neural networks*, vol. 21, no. 7, pp. 1029–1040, 2008.

[34] J. Hu, C.-s. Wang, M. Wu, Y.-x. Du, Y. He, and J. She, "Removal of eog and emg artifacts from eeg using combination of functional link neural network and adaptive neural fuzzy inference system," *Neurocomputing*, vol. 151, pp. 278–287, 2015.

[35] S. Mallat, *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*, 3rd ed. Academic Press, 2008.

[36] H. Guo and C. Burrus, "Convolution using the undecimated discrete wavelet transform," in *Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings., 1996 IEEE International Conference on*, vol. 3, 1996, pp. 1291–1294 vol. 3.

[37] M. K. I. Molla, M. R. Islam, T. Tanaka, and T. M. Rutkowski, "Artifact suppression from EEG signals using data adaptive time domain filtering," *Neurocomputing*, vol. 97, no. 0, pp. 297 – 308, 2012.

[38] R. R. Coifman and D. L. Donoho, *Translation-invariant de-noising*. Springer, 1995.

[39] http://www.cs.bris.ac.uk/rafal/phasereset/phase.zip.

[40] L. Rankine, N. Stevenson, M. Mesbah, and B. Boashash, "A nonstationary model of newborn eeg," *Biomedical Engineering, IEEE Transactions on*, vol. 54, no. 1, pp. 19–28, Jan 2007.

[41] B. Boashash, M. Mesbah, L. Rankine, and N. Stevenson, "Newborn eeg seizure simulation using time-frequency signal synthesis," in *WDIC 2005: APRS Workshop on Digital Image Computing: Workshop Proceedings*, B. Lovell and A. Maeder, Eds. Australia, Queensland, Brisbane, Southbank: University of Queensland, 2005, pp. 145–150. [Online]. Available: http://eprints.qut.edu.au/24536/

[42] J. Gotman, "Automatic recognition of epileptic seizures in the EEG," *Electroencephalography and Clinical Neurophysiology*, vol. 54, no. 5, pp. 530 – 540, 1982.

[43] M. Saab and J. Gotman, "A system to detect the onset of epileptic seizures in scalp EEG," *Clinical Neurophysiology*, vol. 116, no. 2, pp. 427 – 442, 2005.

[44] D. L. Donoho, "De-noising by soft-thresholding," *Information Theory, IEEE Transactions on*, vol. 41, no. 3, pp. 613–627, 1995.

[45] M. K. Islam, A. Rastegarnia, A. T. Nguyen, and Z. Yang, "Artifact characterization and removal for in vivo neural recording," *Journal of Neuroscience Methods*, vol. 226, no. 0, pp. 110 – 123, 2014.

[46] H. ye Gao, "Wavelet shrinkage denoising using the non-negative garrote," 1997.

[47] http://physionet.nlm.nih.gov/pn6/chbmit/.

[48] N. Yeung, R. Bogacz, C. B. Holroyd, and J. D. Cohen, "Detection of synchronized oscillations in the electroencephalogram: An evaluation of methods," *Psychophysiology*, vol. 41, no. 6, pp. 822–832, 2004.

[49] H. Hallez, M. De Vos, B. Vanrumste, P. Van Hese, S. Assecondi, K. Van Laere, P. Dupont, W. Van Paesschen, S. Van Huffel, and I. Lemahieu, "Removing muscle and eye artifacts using blind source separation techniques in ictal eeg source imaging," *Clinical Neurophysiology*, vol. 120, no. 7, pp. 1262–1272, 2009.

[50] R. Esteller, J. Echauz, T. Tcheng, B. Litt, and B. Pless, "Line length: an efficient feature for seizure onset detection," in *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE*, vol. 2, 2001, pp. 1707–1710 vol.2.

[51] N. Stevenson, J. OToole, L. Rankine, G. Boylan, and B. Boashash, "A nonparametric feature for neonatal EEG seizure detection based on a representation of pseudo-periodicity," *Medical Engineering & Physics*, vol. 34, no. 4, pp. 437 – 446, 2012.

[52] D. Wang, D. Miao, and C. Xie, "Best basis-based wavelet packet entropy feature extraction and hierarchical {EEG} classification for epileptic detection," *Expert Systems with Applications*, vol. 38, no. 11, pp. 14 314 – 14 320, 2011.

[53] N. Rafiuddin, Y. Uzzaman Khan, and O. Farooq, "Feature extraction and classification of eeg for automatic seizure detection," in *Multimedia, Signal Processing and Communication Technologies (IMPACT), 2011 International Conference on*, Dec 2011, pp. 184–187.

[54] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology - Heart and Circulatory Physiology*, vol. 278, no. 6, pp. H2039–H2049, 2000.

[55] Y. Song, J. Crowcroft, and J. Zhang, "Automatic epileptic seizure detection in EEGs based on optimized sample entropy and extreme learning machine," *Journal of Neuroscience Methods*, vol. 210, no. 2, pp. 132 – 146, 2012.

[56] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines: And Other Kernel-based Learning Methods*. New York, NY, USA: Cambridge University Press, 2000.