

Research Article

Gaussian Elimination-Based Novel Canonical Correlation Analysis Method for EEG Motion Artifact Removal

Vandana Roy,¹ Shailja Shukla,² Piyush Kumar Shukla,³ and Paresh Rawat⁴

¹Department of Electronics and Communication, Jabalpur Engineering College, Jabalpur 482011, India

²Department of Computer Sciences and Engineering, Jabalpur Engineering College, Jabalpur 482011, India

³Department of Computer Sciences and Engineering, University Institute of Technology, Bhopal 462023, India

⁴Department of Electronics and Communication, Truba Group of Institute, Bhopal 462023, India

Correspondence should be addressed to Piyush Kumar Shukla; pphdwss@gmail.com

Received 7 December 2016; Accepted 7 June 2017; Published 8 October 2017

Academic Editor: João Manuel R. S. Tavares

Copyright © 2017 Vandana Roy et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The motion generated at the capturing time of electro-encephalography (EEG) signal leads to the artifacts, which may reduce the quality of obtained information. Existing artifact removal methods use canonical correlation analysis (CCA) for removing artifacts along with ensemble empirical mode decomposition (EEMD) and wavelet transform (WT). A new approach is proposed to further analyse and improve the filtering performance and reduce the filter computation time under highly noisy environment. This new approach of CCA is based on Gaussian elimination method which is used for calculating the correlation coefficients using backslash operation and is designed for EEG signal motion artifact removal. Gaussian elimination is used for solving linear equation to calculate Eigen values which reduces the computation cost of the CCA method. This novel proposed method is tested against currently available artifact removal techniques using EEMD-CCA and wavelet transform. The performance is tested on synthetic and real EEG signal data. The proposed artifact removal technique is evaluated using efficiency matrices such as del signal to noise ratio (DSNR), lambda (λ), root mean square error (RMSE), elapsed time, and ROC parameters. The results indicate suitability of the proposed algorithm for use as a supplement to algorithms currently in use.

1. Introduction

EEG signal is widely used for exploring the human brain activity and is preferred over other physiological signals because they can be used to detect directly brain electrical activity changes over spans of millisecond time, whereas functional magnetic resonance imaging (fMRI) has time resolutions in seconds or minutes. Usually, EEG signal suffers from various motion artifacts generated at the capturing time. There are two main sources of artifact in neural signals other than the machine and environment. These are the muscular and ocular activities of the individual which generate low-amplitude, low-frequency electrical pulses that fall in filter range of sensors and recording equipment.

The EEG signal is contaminated by various artifacts as electrocardiogram (ECG), electrooculogram (EOG), and electromyogram (EMG). The EMG artifact is of more interest as it is having higher amplitude, broad spectrum, and variable topographical distribution than EEG signal [1]. The EMG artifacts have lower autocorrelation than EEG signal due to its wide frequency spectrum. Moreover, these artifacts resemble as temporal white noise. Therefore, artifact rejection is a fundamental research topic and is well researched in [2]. The EEMD is a data-driven and noise-assisted approach which is applied to remove motion artifacts from single-channel EEG signal [3].

A component-based automated separator of artifacts is required to linearly decompose the signals into source components. The components give the individual nature of

information, where artifact information combines into separate sources and reconstruction of signals without these sources are claimed as artifact-free information. The performance of blind source separation (BSS) methods as independent component analysis (ICA) and CCA for EEG signal eye blink artifact removal is compared in [4] and concluded that CCA is more accurate and faster than ICA. The CCA algorithm performs better than ICA for muscle artifact removal because these artifacts are generated due to movement of body muscle group. Moreover, these artifacts do not represent stereotyped topography [5]. Anastasiadou et al. [6] applied CCA algorithm to remove muscle artifact from EEG signal. Moreover, artifact removal approach is improved by applying WT after CCA algorithm for automatic detection and removal of muscle artifact from EEG signal [7]. The cascaded combination of EEMD and CCA techniques is applied for single-channel EEG artifact removal in [8, 9]. The single-channel EEG signal is converted into the multidimensional signal by EEMD technique. In CCA, a second-order statistics is applied to segregate the artifact components from the input signal and its performance is compared and presented better in comparison of existing wavelet denoising and EEMD-ICA cascaded algorithms. Chen et al. [10] have improved the filtering approach by applying EEMD and MCCA (multiset CCA) to remove EMG with more computational time. To remove ocular artifacts automatically, the cascaded combination of CCA and WT algorithms is proposed in [11] and demonstrated that the proposed method removes artifacts significantly with preserving neural activity of the original signal. Safieddine et al. [12] suggested that EEMD algorithm outperforms over other stochastic and deterministic artifact removal approaches and WT algorithm performs well in the case of less noisy data.

An efficient cascaded approach EEMD-CCA-SWT is found successful for EEG motion artifact removal [13]. The combination of EEMD, CCA, and SWT approaches has been applied for effective suppression of the motion artifact from EEG signal. This three-stage cascaded approach removes the artifacts effectively with increased computational cost. This computational complexity is reduced in this research paper by developing an existing correlation-based algorithm with Gaussian elimination (GE) and inserted at the cascade of EEMD-SWT, leading to EEMD-GECCA-SWT which is the combination of EEMD and an improved approach GECCA (Gaussian elimination canonical correlation analysis) with SWT. This increased computational cost is effectively reduced due to applying GECCA approach in place of CCA. The left matrix division applied in GECCA allows better estimates for the matrix inversion; therefore, it improves the SWT filtering efficiency and thus improves the overall efficiency of the EEG motion artifact removal.

The paper organization is as the following. The artifact removal methods are recalled in Section 2, and then the proposed algorithm is discussed in Section 3. Details of applied EEG dataset is given in Section 4; results obtained by methods are presented, compared in tabular form, and detailed discussed in Section 5. Finally, conclusion is in Section 6.

2. Artifact Removal Methods

2.1. Ensemble Empirical Mode Decomposition. The EEMD algorithm decomposes a signal into a number of intrinsic mode functions (IMFs) through an iterative method termed as sifting [12]. At first level, the IMF1 is the mean of upper and lower envelop of original EEG signal $X(t)$. Then residual signal is obtained by subtracting IMF1 from $X(t)$. This process is iterated till stopping criterion is fulfilled (residual signal energy content is close to zero). The remaining residual signal is

$$P_n(t) = P_{n-1}(t) - \text{IMF}_n(t), \quad (1)$$

where $P_n(t) = X(t)$.

Finally, the signal is reconstructed by adding all IMFs and residual signal as

$$X(t) = P_n(t) + \sum_{i=1}^N \text{IMF}_i(t). \quad (2)$$

2.2. Gaussian Elimination Canonical Correlation Analysis (GECCA). The standard existing CCA algorithm starts with an assumption that $X[n]$ and $Y[n]$ are two sets of random variables [4]. $X[n]$ is the input vector of matrix and $Y[n]$ is defined as temporally correlated by 2-D valid convolution operator from $X[n]$ vectors using the linear convolution mask $[1 \ 0 \ 1]$ as

$$Y[n] = \text{conv2}(X, [1 \ 0 \ 1]). \quad (3)$$

Merging both input vectors as

$$Z = [X; Y]. \quad (4)$$

If ρ is the maximum canonical correlation, C_{xx} , C_{yy} are auto covariances of vectors X and Y , respectively, and C_{xy} and C_{yx} are the cross covariance between vectors X and Y . The various possible correlation matrixes are

$$C_{xx} = C(1 : sx, 1 : sx) + \beta * \text{eye}(sx), \quad (5)$$

where sx is the size of the X , sy is the size of the Y , β is the predefined small residual constant which is set to 10^{-8} in this paper which is used by standard CCA method [2], and eye is the identity matrix with diagonal terms as 1 and all other terms are zero. The term $\beta * \text{eye}(sx)$ is used in order to set and maintain the initial nonzero value of the covariance matrix C_{xx} . The small value of β does not affect the covariance matrix values much.

$$C_{xy} = C(1 : sx, sx + 1 : sx + sy),$$

$$C_{yx} = C_{xy}',$$

$$C_{yy} = C(sx + 1 : sx + sy, sx + 1 : sx + sy) + \zeta * \text{eye}(sy),$$

$$C_{yy}^{-1} = \text{inv}(C_{yy}).$$

(6)

Two different canonical solutions are obtained from Z by calculating the covariance as

$$C = \text{cov}(Z.^'). \quad (7)$$

This is equivalent to two linear equations in X and Y vector directions, respectively, as

$$\begin{aligned} p[n] &= a * X[n], \\ q[n] &= b * Y[n], \end{aligned} \quad (8)$$

where a and b are weight vectors and $p[n]$ and $q[n]$ are canonical variates correspondent to $X[n]$ and $Y[n]$, respectively.

The maximum correlation between variables a and b are calculated as [8]

$$\max(\rho) = \frac{a^T C_{xy} b}{\sqrt{a^T C_{xx} a} \sqrt{b^T C_{yy} b}}. \quad (9)$$

Now, the demixing matrix W is calculated by simplifying (9) as

$$\rho^2 a = \text{inv}(C_{xx}) * C_{xy} * \text{inv}(C_{yy}) * C_{yx} * a, \quad (10)$$

$$\rho^2 b = \text{inv}(C_{xx}) * (C_{xy}) * \text{inv}(C_{yy}) * C_{yx} * b, \quad (11)$$

where ρ^2 is calculated by obtaining Eigen values (k) of (10) and (11) and further, a and b variables are equal to Eigen vectors for the greatest value of k (Eigen value).

These weight vectors a and b are replaced in (8) to calculate canonical variates $p[n]$ and $q[n]$. Moreover, the demixing matrix W can be calculated by inverse of estimated weight vector ($W = a^{-1}$). Finally, the source $S[n]$ is estimated by applying W and $p[n]$ or $q[n]$ as

$$S[n] = W^T p[n]. \quad (12)$$

Thus, it is observed that standard CCA algorithm uses the matrix inverse operation to solve the linear equations by obtaining the maximum value of ρ^2 for determining the demixing matrix W .

In this research paper, a fast and efficient modified CCA algorithm is proposed which uses the backslash or left matrix division operator for solving the linear (10) and (11) to calculate Eigen values. This backslash operator is fast and efficient than the inverse operator. The remaining algorithm is similar to the CCA approach. Let us define the linear (10) as

$$G = \rho^2 = \text{inv}(C_{xx}) * C_{xy} * \text{inv}(C_{yy}) * C_{yx}, \quad (13)$$

$$G = \rho^2 = A * B, \quad (14)$$

where

$$A = \text{inv}(C_{xx}) * C_{xy} = C_{xx}^{-1} * C_{xy}, \quad (15)$$

$$B = \text{inv}(C_{yy}) * C_{yx} = C_{yy}^{-1} * C_{yx}. \quad (16)$$

The parameters A and B from linear (15) and (16) are defined as

$$\begin{aligned} A * C_{xx} &= C_{xy}, \\ B * C_{yy} &= C_{yx}. \end{aligned} \quad (17)$$

Standard CCA method uses matrix inverse to obtain the value of A and B . In this paper, the left matrix division is proposed to solve these linear (15) and (16).

The left division operator as $C_{xx} \setminus C_{xy}$ can be defined as a matrix division of C_{xx} into C_{xy} , which is equal to the solution of $\text{inv}(C_{xx}) * C_{xy}$. If C_{xx} is an N -by- N matrix and C_{xy} is a column vector with N components or a matrix containing several such columns, then $A * C_{xx} = C_{xy}$ equation can be defined in terms of left division operator as $A = C_{xx} \setminus C_{xy}$.

The similar solution can be defined for the matrix B . The solution of (15) and (16) can be written as

$$A = C_{xx} \setminus C_{xy}, \quad (18)$$

$$B = C_{yy} \setminus C_{yx}. \quad (19)$$

Equations (18) and (19) suggest that the role of inverse operation is avoided to obtain the value of parameters A and B . This will simplify the equation and thus computation time too.

2.2.1. The Justification of Left Division Operator Employed in GECCA. In order to prove left division operator, the numerator and denominator are multiplied by numerator transposed as

$$\begin{aligned} A &= C_{xx}^T * C_{xx} \setminus C_{xx}^T * C_{xy}, \\ B &= C_{yy}^T * C_{yy} \setminus C_{yy}^T * C_{yx}. \end{aligned} \quad (20)$$

By shuffling the order of operator by using Gaussian elimination concept,

$$A = C_{xx}^T * (C_{xx} * C_{xx}^T) \setminus C_{xy}, \quad (21)$$

$$B = C_{yy}^T * (C_{yy} * C_{yy}^T) \setminus C_{yx}, \quad (22)$$

since C_{xx} and C_{yy} are N -by- N symmetric matrix and C_{xy} and C_{yx} are column vectors with N components or a matrix with several such columns.

Since $C_{xx} * C_{xx}^T = C_{yy} * C_{yy}^T = 1$, therefore, (21) and (22) can be written as

$$A = C_{xx}^T * (1) \setminus C_{xy}, \quad (23)$$

$$B = C_{yy}^T * (1) \setminus C_{yx}. \quad (24)$$

For symmetric and square matrix, $C_{xx}^T = C_{xx}$ and $C_{yy}^T = C_{yy}$.

Then (23) and (24) are simplified as

$$A = C_{xx} \setminus C_{xy}, \quad (25)$$

$$B = C_{yy} \setminus C_{yx}. \quad (26)$$

Replacing the values of A and B from (25) and (26) to (14), the solution of G is obtained as

$$G = (C_{xx} \setminus C_{xy}) * (C_{yy} \setminus C_{yx}). \quad (27)$$

The value of G is obtained without inverse operation by just applying left division operator. This solution not only saves the computation time but also increases the efficiency of the CCA method with minimizing the error. The obtained

$G = \rho^2$ is the Eigen value with k in descending order, that is, $k_1 > k_2 > k_n$; a and b are the Eigen vector for the greatest value of k .

Then the CCA component $p[n] = [p_1, p_2, \dots, p_n]$ are obtained by placing the vector $X[n]$ in (8), and sources are expected by using the weighted demixing matrix from X by using (12).

If A is a $M \times N$ matrix with $M \leq N$ and B is a column vector with M components, or a matrix with several such columns, then $X = A \setminus B$ is the solution in the least squares sense to the under or over determined system of equations $A * X = B$. The effective rank K of A is determined from the QR decomposition with pivoting. The solution X is computed which has at most K nonzero components per column.

2.2.2. Benefits of GECCA over CCA Algorithm. The advantage of GECCA Algorithm over CCA are as follows:

- (1) The operation of CCA algorithm depends on the inverse operation to find the demixing matrix. If the input matrix is not a square matrix, then this CCA algorithm fails. However, left division operator-based GECCA algorithm provides a solution if input matrix is not square.
- (2) The CCA algorithm involves the inverse of the matrix for their operation. The inverse operation is computationally complex and time-consuming. However, proposed GECCA algorithm is faster than CCA for obtaining the correlation-based source separation by employing left division operator in place of inverse operation.
- (3) The proposed GECCA algorithm employs left division operator. This operator allows better estimates of the matrix inversion. These efficiently estimated coefficients once applied to SWT filter improve the SWT filtering efficiency thus improving the overall efficiency of the motion artifact removal algorithm.
- (4) The CCA algorithm operation is not consistent in each situation. This issue can be overcome by the proposed GECCA algorithm.

The above discussed efficient and fast correlation-based algorithm applied with WT algorithm is discussed in the subsequent section.

2.3. Wavelet Transform (WT). The artifacts available in the EEG signal are suppressed with cascaded approach of EEMD and GECCA. Although, some brain actions may be disturbed due to high-frequency sensor noise and low-pass and high-pass filter components. As these noise signal frequencies will get overlap with the brain signals, then conventional filtering technique cannot be utilized and thus this WT is applied to take away unwanted noises from EEG signal [12].

The most frequently applied wavelet transform is discrete wavelet transform (DWT). In the case of EEG artifact removal, preserving neural information of the signal is determinant. Thus, some latest research shows that stationary

wavelet transform (SWT) is a powerful tool to remove artifacts of the signal with preserving the neural information of original EEG signal [14].

SWT algorithm is translation invariant, so no down-sampling of the data is involved. Translation-invariance is achieved by removing the down- and upsamplers as in the DWT and upsampling the filter coefficients by a factor of $2^{(j-1)}$ at the j th level of the algorithm. SWT is preferred as it removes unpredictable behavior and noise randomness in the EEG signal due to the motion artifacts remaining after two-stage filtering of EEG signals. Therefore, EEG signal gets smoothed over the length with containing all their fundamental properties. Finally, proposed algorithm based on discussed algorithm is discussed in the next section.

3. Proposed Algorithm

The proposed artifact removal algorithm is as follows:

- (1) Define the reference EEG data from the multichannel data set for correlation match.
- (2) The input artifact EEG channel is passed through EEMD to decompose with 3 ensembles to convert single-channel EEG into multichannel EEG data.
- (3) Generated IMFs are passed to GECCA algorithm for source separation.
- (4) The GECCA output contains traces of artifacts in the form of randomness and noise. Thus, the output of GECCA is applied further to SWT algorithm for effective artifact removal. The decomposition through SWT is performed on the frequency domain of the signal with Rigrsure thresholding.
- (5) Pearson's correlation coefficient is used for artifact recognition and their suppression and finally reconstructed to find the artifact-free signal.
- (6) Parametric evolution is carried out based on SNR and correlation improvement. The correlation is carried out with respect to reference original EEG data initially defined.
- (7) The parametric evaluation of the proposed method is done with existing motion artifact removal methods.

4. EEG Signal Data Set

The EEG data is a multichannel data considered from standard MIT scalp data set with 24 channels recorded with different electrodes. This dataset is available on PhysioNet. The data is taken for 10 seconds and with 2560 samples/signals. Sampling frequency is taken as 256 Hz, and sampling interval is of 0.00390625 sec. These data samples are reduced to 16 channels to satisfy the international standard of 10–20 channels. Figure 1 shows the first 16 channels of EEG data set available online with chb01_01_edfm.mat available at PhysioNet provided by Shoeb [15]. It is observed from Figure 1

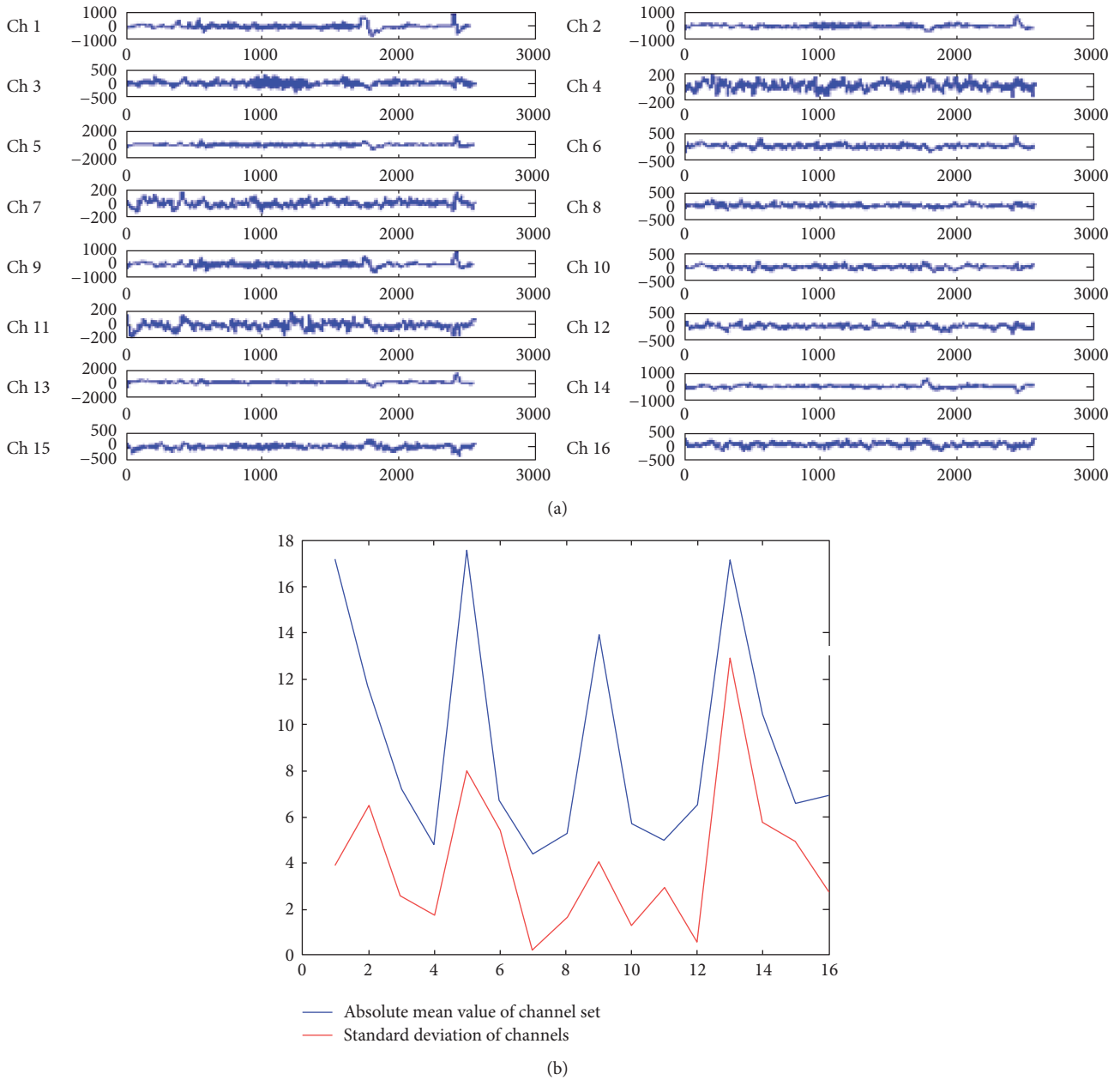


FIGURE 1: (a) Multichannel EEG data set containing the muscular and eye blink motions. (b) Statistical properties of EEG channels.

that each channel is recorded with different electrode having different motion characteristics.

In Figure 1(a), the channels 1, 5, 9, and 13 correspond to the eye blinks as shown by negative high peaks (scale of 1000 and 2000). This was displayed and referenced by Shoeb [15]. These channels (1, 5, 9, and 13) have maximum motion or artifacts since they have higher standard deviation 172.0907, 175.417, 136.129, and 170.467, respectively. In this paper, channel numbers 4, 7, and 11 have the least artifact effects (only with scale of 200) out of the 16 channel EEG signal, this can be quantitatively justified by their lowest mean value and standard deviation (on division by 10 scale) as shown in Figure 1(b). The channel 7 in Figure 1(a) has the smallest standard deviation ($=4,364,903$) and the smallest

absolute mean value ($=0.228125$) as shown in Figure 1(b) which justifies the minimum motion artifacts. Thus, the channel number 7 is used as a reference signal for correlation-based artifact removal. This reference signal is termed as original signal in the paper.

Figure 2 compares the original or reference EEG data (channel number 7) with blue color and artifactual signals (channel number 5) with red color. The channel 5 is shown here for the comparison because it has maximum standard deviation of 175.417 with maximum peak variations with eye blink. It is observed from Figure 2 that raw EEG data may suffer from either muscular motion artifact (shown under green rectangle) or eye blink artifact (shown under yellow rectangle) or both of them. The eye blink artifact

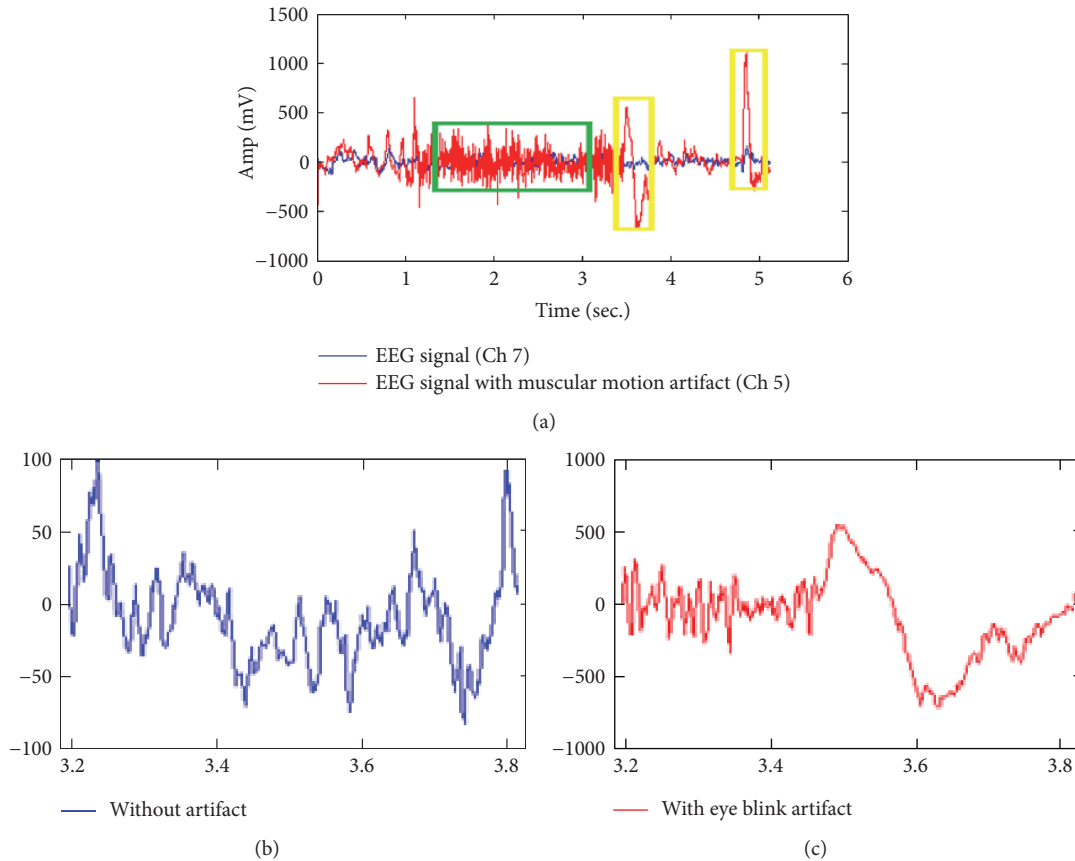


FIGURE 2: (a) Comparison of referenced channel 7 and artifact channel 5 for first 10 seconds. (b) Enhanced reference data channel 7. (c) Artifact channel 5 for 3-4 seconds respective to yellow rectangle.

may cause amplitude increase about 10 times greater than the original EEG data. The EOG artifact overlaps with EEG signal in impulsive form, while muscular motions cause random broad spectrum amplitude variations. Nature of the reference EEG data with and without artifacts is clearly shown in Figures 2(b) and 2(c).

The eye blink artifact may cause amplitude increase about 10 times greater than the original EEG data. The EOG artifact overlaps with EEG signal in impulsive form, while muscular motions cause random broad spectrum amplitude variations. In this paper, the muscular motion artifact signals (as in channel 5) are naturally generated at the time of EEG data capturing with electrodes. Therefore, for the sake of clarity, the enhanced version of reference and artifact signal is shown in Figures 2(b) and 2(c).

5. Result and Discussion

The novel GECCA artifact removal approach is evaluated on artifactual EEG channel number 5 and compared with CCA algorithm. Comparison of the EEG artifact removal methods for channel number 5 using CCA and GECCA along with DWT is shown in Figure 3. The EEG signal channel number 7 is considered as pure EEG signal. The comparison of the green and yellow boxes in Figures 3(a)

and 3(b) suggests that blind source separation- (BSS-) based proposed GECCA method removes the motion artifacts significantly better than the existing CCA-based method. Green rectangles are correspondent to CCA, and yellow rectangles are correspondent to GECCA. It is observed that the eye blink peaks with GECCA-DWT are minimized to half that of the eye blink peaks with CCA-DWT (of the order of ± 200 is reduced to around ± 100) as mentioned in fourth row of Figure 3. Moreover, the pattern of the second blink is also much similar to the reference EEG signal after GECCA-DWT.

The SWT algorithm is more proficient for artifact removal of neural EEG signal in comparison to DWT algorithm; consequently, the efficacy of proposed BSS approach (GECCA) is also evaluated with SWT and compared with CCA-based approaches. Figure 4 provides the comparison of the EEG artifact removal methods using CCA and GECCA along with SWT filtering. The artifact removal approaches are applied on channel number 5, and performance is compared with reference channel 7 EEG signal. The comparison of the respective green and yellow boxes in Figures 4(a) and 4(b) suggests that the proposed GECCA method removes the motion artifacts significantly better than the existing CCA-based method. In addition, the vigilant observation and comparison of the respective green rectangles in Figure 4(a) and yellow rectangles in Figure 4(b) suggests that

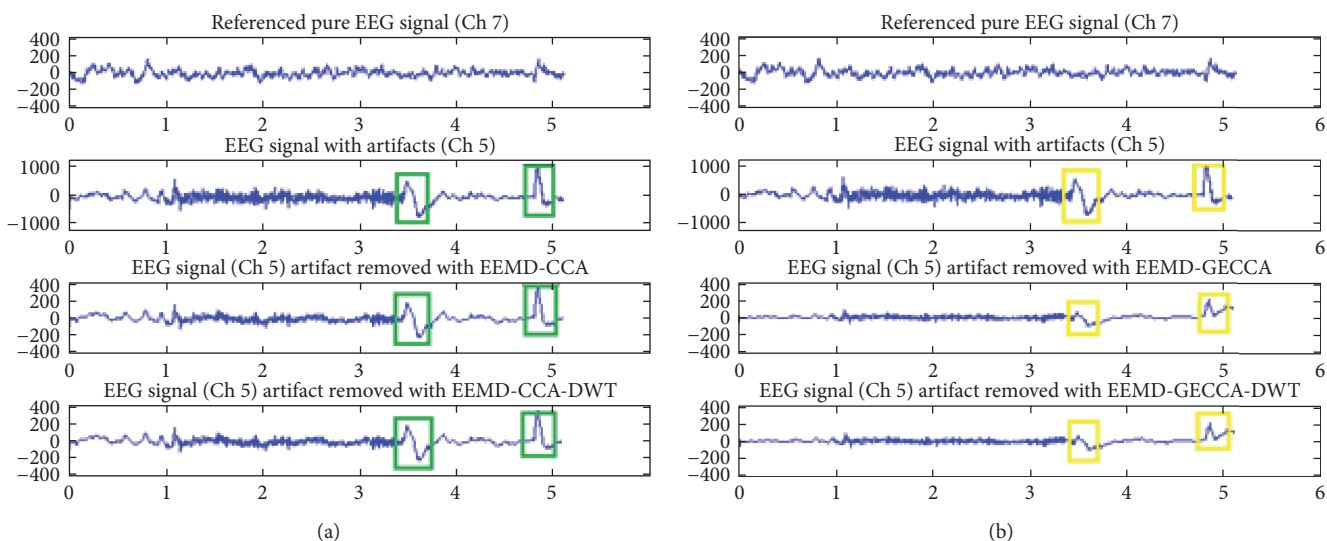


FIGURE 3: Comparisons of EEG artifact removal methods (a) with CCA and DWT and (b) with GECCA and DWT.

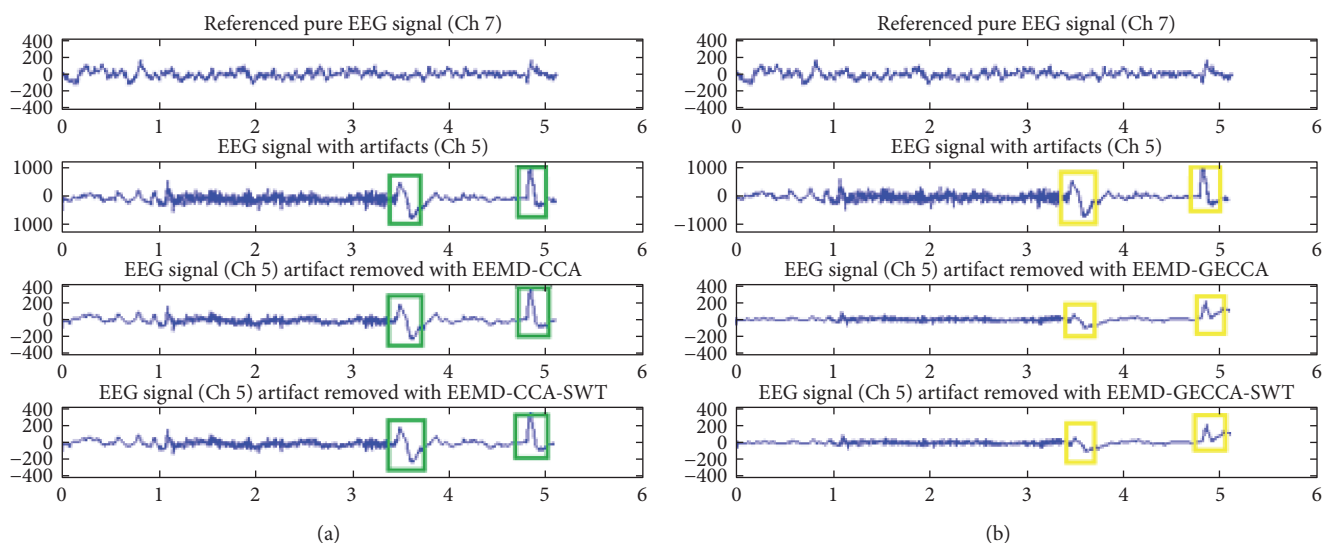


FIGURE 4: Comparisons of EEG artifact removal methods (a) with CCA and SWT and (b) with GECCA and SWT.

GECCA with SWT removes the eye blink effect better than CCA-SWT. It is observed that the eye blink peaks with GECCA-SWT are minimized to half that of the eye blink peaks with CCA-SWT (of the order of ± 200 is reduced to around ± 100) as mentioned in the fourth row of Figure 4. This could be quantitatively even better clear in the next sections with evaluation parametric comparison encapsulated in Table 1.

The 8 reconstructed signals after SWT-smoothed approach are compared for CCA and GECCA as shown in Figure 5. It is observed that GECCA method not only preserves the information but also provides better and fast estimates. This is concluded by comparing the 2nd and 4th frequency bands as shown in Figures 5(c) and 5(d).

Moreover, the proposed artifact removal method (EEMD-GECCA-SWT) outputs are compared for three

TABLE 1: Parametric comparison of the artifact removal methods with CCA and GECCA.

Algorithm Parameters	With DWT filtering		With SWT filtering	
	EEMD-CCA	EEMD-GECCA	EEMD-CCA	EEMD-GECCA
MSE	69.3283	51.2272	70.2484	50.1366
Correlation improvement	0.0019	0.0666	0.0103	0.0654
Lambda	66.8838	86.0016	66.2544	87.2759
DSNR	17.7248	29.0387	17.2621	30.2080

distinct EEG channel data, namely, Ch 9, Ch 4, and Ch 14, as presented in Figure 6. In Figure 6, blue color presents the original artifactual signal, whereas red color signal represents the filtered output by the proposed approach. It is demonstrated

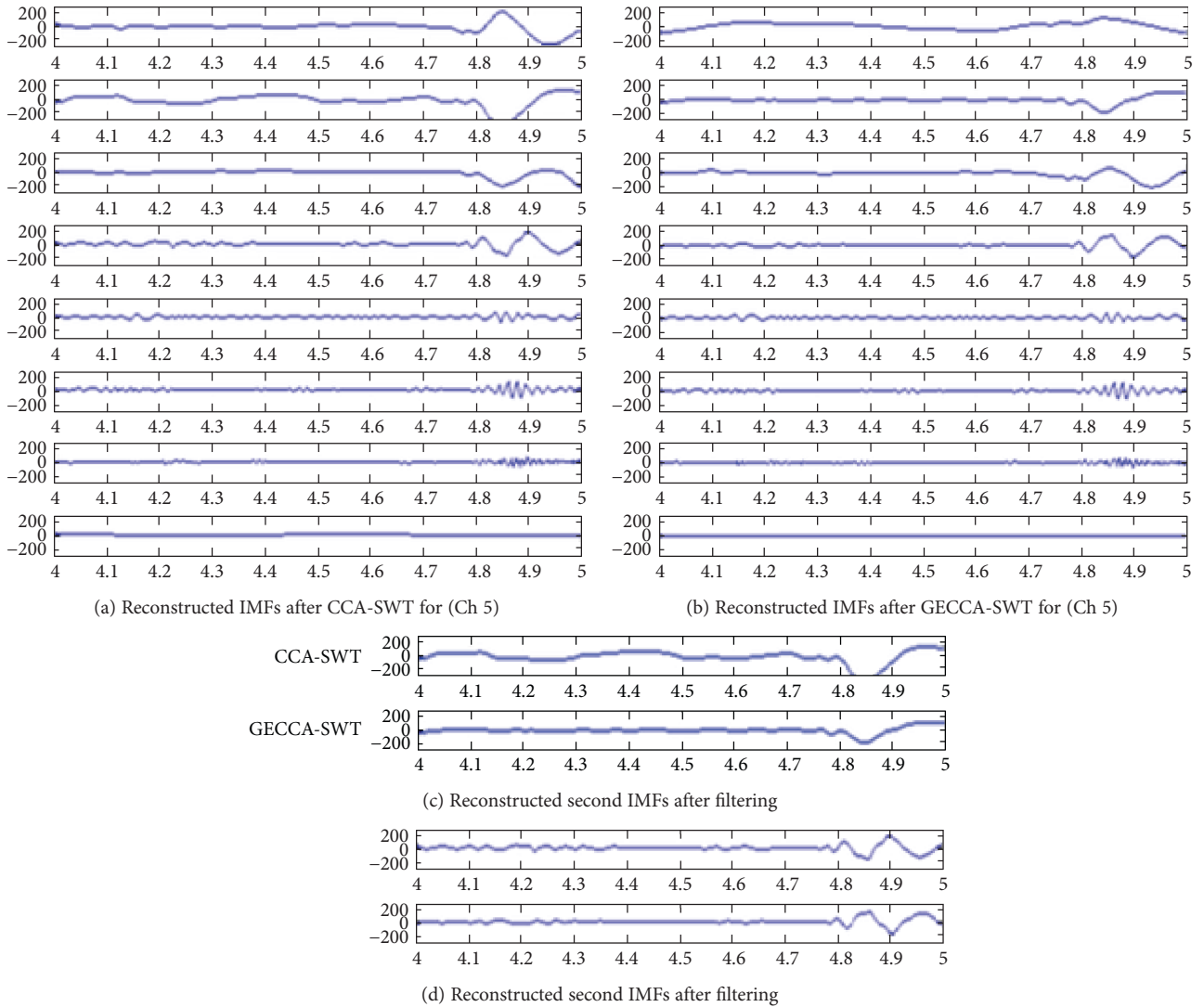


FIGURE 5: Reconstructed IMFs after (a) CCA-SWT, (b) GECCA-SWT, and (c) comparison of CCA and GECCA for the 2nd and 4th reconstructed EEMD (IMFs).

that proposed method of GECCA removes the eye blink artifacts and motion artifacts significantly.

In order to evaluate the performance of the proposed method, parametric evaluation is carried out as shown in Table 1. The performance of artifact removal methods is carried out for CCA and GECCA methods along with DWT and SWT filtering methods. It is observed that proposed method EEMD-GECCA-SWT reduces the mean square error (MSE) by around 15.81% and improves the DSNR performance significantly. Moreover, the higher lambda (λ) value indicates the improved filtering performance [8]. Table 1 shows that GECCA-based artifact removal methods attain improved lambda (λ) value. The vigilant observation of results on Table 1 also presents better correlation improvement of GECCA when compared to CCA, both for DWT and SWT. The backslash operator provides better estimation efficiency. Thus, Gaussian elimination-based approach provides improved correlation by using backslash operation which is more efficient than the matrix inverse method.

The proposed method evaluation is also carried out with the receiver operation characteristic (ROC) curve parameters as shown in Table 2. The positive predictive value (PPV), negative predictive value (NPV), sensitivity (Sen), and specificity (Spe) are calculated by true positive (TP), true negative (TN), false positive (FP), and false negative (FN) [16] as follows:

$$\text{Sen} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (28)$$

$$\text{Spe} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \quad (29)$$

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (30)$$

$$\text{NPV} = \frac{\text{TN}}{\text{TN} + \text{FN}}. \quad (31)$$

The true positive (TP) indicates that the sample is identified as an artifact when it was actually an artifact.

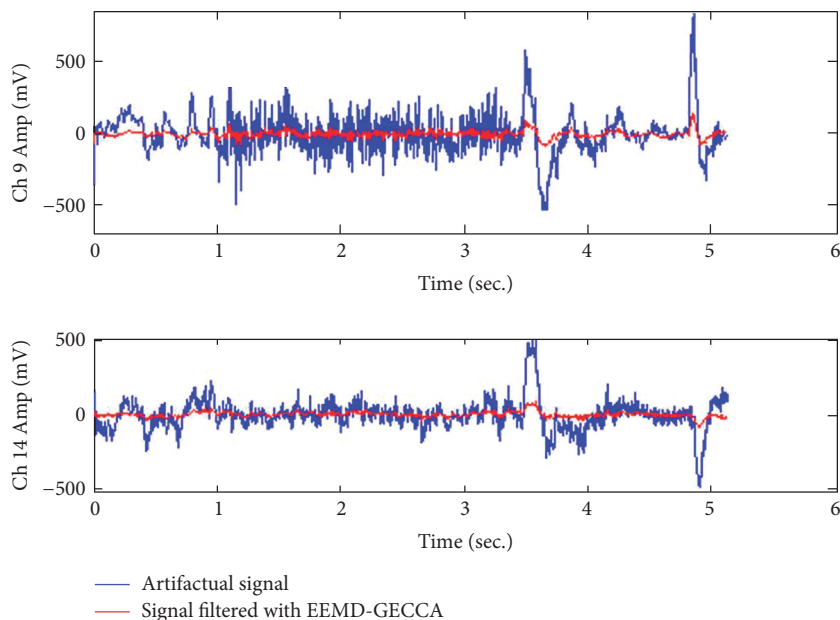


FIGURE 6: Artifact removal consequences with the proposed (EEMD-GECCA-SWT) method.

TABLE 2: Comparison of the ROC parameters with CCA and GECCA for channel 5.

Algorithm Parameters	DWT filtering with		SWT filtering with	
	EEMD_CCA	EEMD_GECCA	EEMD_CCA	EEMD_GECCA
Sensitivity	0.39961	0.45156	0.43438	0.43789
Specificity	0.74062	0.78477	0.65469	0.82656
Accuracy	57.0117%	61.8164%	54.4531%	63.2227%
PPV	60.6402%	67.7211%	55.7114%	71.6294%
NPV	55.2287%	58.8632%	53.6492%	59.5218%

False positive (FP) presents that the sample is identified as an artifact but actually it was not. The false negative (FN) is an indicator for the absence of artifact; however, the artifact was present in the sample. Moreover, true negative (TN) indicates the absence of the artifact when the artifact was actually not present.

This ROC curve is a comparison plot between true positive rate or sensitivity and false positive rate ($1 - \text{specificity}$) [16] for accurate detection of artifact from EEG signal as shown in Figure 7.

Figure 7(a) shows comparison of CCA and GECCA approaches based on DWT. Blue and red color horizontal line indicates the sensitivity of EEMD-CCA-DWT and EEMD-GECCA-DWT artifact removal methods, respectively. The sensitivity or true artifact detection rate for CCA-based approach is 0.39961, whereas GECCA-based approach presents improved and fast response with increased value of 0.45156. However, Figure 7(b) shows SWT-based approach comparison for CCA and GECCA algorithm with the same blue and red color code. Both approaches present almost similar sensitivity 0.43438 and 0.43789, respectively, as shown in Table 2. Although, GECCA-based approach performs better than CCA artifact removal algorithm. Moreover, the sensitivity is associated

with accuracy of the algorithm too. All the ROC parameters as sensitivity, specificity, PPV, and NPV as mentioned above in (28), (29), (30), and (31) are calculated for EEG signal channel number 5 and tabulated for effective comparison of CCA and GECCA-based approaches as shown in Table 2. It is observed that GECCA-based approach shows improved sensitivity and thus better accuracy than CCA-based approaches with both DWT and SWT filtering. Moreover, specificity, PPV, and NPV also improve by GECCA-based methods.

In addition, the efficiency of GECCA-based artifact removal approach is also evaluated by considering ROC parameters for most artifact contaminated channel numbers (1, 3, 5, 6, 9, and 14) from Figure 1. Table 3 encapsulates all the ROC parameters for filtering approach based on CCA and GECCA. It is observed from Table 3 that GECCA-based methods show improved accuracy in comparison to CCA-based method except channel 1. Similarly, specificity, PPV, and NPV parameters are also improved for most of the channel which presents the success of GECCA-based artifact removal approach.

The use of the left matrix division operator in the GECCA algorithms executes faster than the inverse matrix method used in CCA algorithm. Therefore, the proposed GECCA

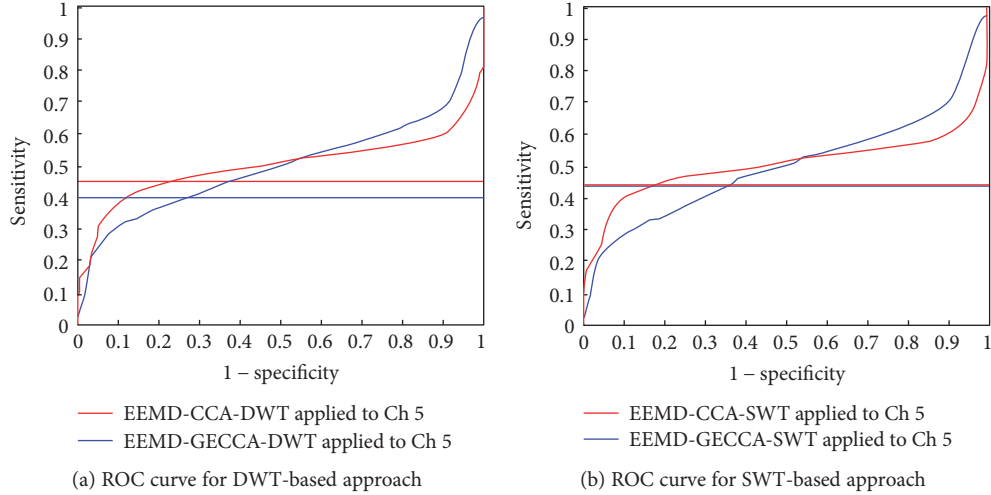


FIGURE 7: ROC plot comparison for CCA and GECCA (a) based on DWT and (b) based on SWT.

TABLE 3: Comparison of ROC parameters for 6 different EEG channels for proposed method.

Method ↓	Parameters in %	Ch 1	Ch 3	Ch 5	Ch 6	Ch 9	Ch 14
With GECCA	Accuracy	57.3047	59.4141	63.2227	65.3125	63.0859	63.5352
	PPV	57.18	57.0674	71.6294	75.0319	71.9241	59.802
	NPV	55.3368	64.0936	59.5218	61.0298	59.3263	71.8612
With CCA	Accuracy	57.6953	58.8672	54.4531	60.2734	62.2461	61.2109
	PPV	56.0578	62.6251	55.7114	64.7422	69.3399	58.2566
	NPV	60.546	56.8332	53.6492	57.8837	58.9597	67.4574

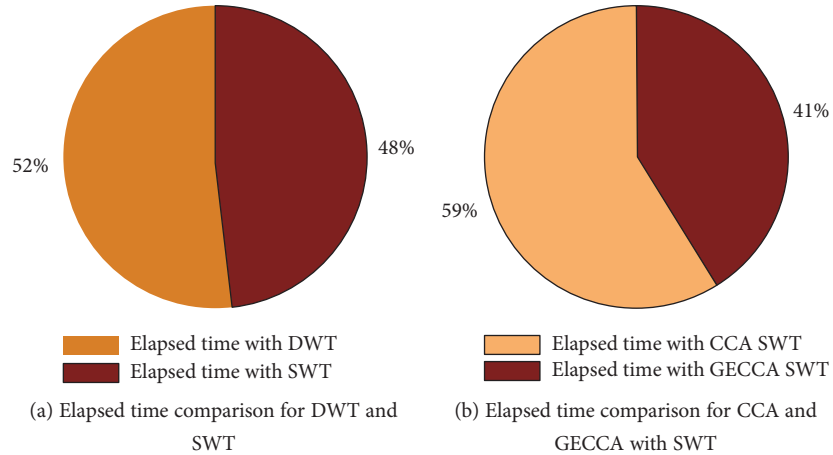


FIGURE 8: (a) Elapsed time comparison for DWT and SWT approach and (b) elapsed time comparison for proposed GECCA method with CCA method.

method is faster than the CCA. Moreover, left matrix division allows better estimates of the matrix inversion which improves the SWT filtering efficiency thus improves the overall efficiency of the Pearson’s correlation-based artifact removal.

This execution or computation speed is evaluated by elapsed time. The elapsed time comparison of DWT and SWT is shown in Figure 8(a) suggesting that SWT is faster than DWT. Figure 8(b) compares the elapsed time for

GECCA and CCA with SWT. This presents that GECCA with SWT algorithm computation time is faster than CCA with SWT approach.

Figure 8(a) suggests that SWT is faster than DWT approach for artifact removal. Thus, SWT elapsed time is evaluated with CCA and GECCA approaches and concluded that GECCA approach with SWT is faster and efficient for EEG artifact removal.

6. Conclusion

The measurement and processing of EEG signal result in the probability of signal contamination prominently through motion artifacts which can obstruct the important features and information quality existing in the original EEG signal. To diagnose the human neurological diseases like epilepsy, tumors, and problems associated with trauma [15], these artifacts must be properly pruned assuring that there is no loss of the main attributes of EEG signals. Thus, a novel algorithm GECCA is introduced in cascade with EEMD and SWT for fast and effective suppression of motion artifacts from single-channel EEG signal. The proposed GECCA method uses backslash operation to solve the linear equations. This improves the computation efficiency of the methods. The application of SWT instead of DWT improves the SNR performance of the method and faster than the DWT method too. However, the proposed algorithm may give over smoothing if not properly designed. The proposed method based on GECCA is 18% faster than the conventional CCA. The various evaluation parameters as del signal to noise ratio (DSNR), lambda (λ), root mean square error (RMSE), and ROC parameters are employed to compare the performance of the proposed artifact removal algorithm. The ROC parameter comparison suggested that the improved accuracy is attained with GECCA algorithm except for Ch 1. The optimum evaluation result shows the success of proposed motion artifact removal method.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] B. W. McMenamin, A. J. Shackman, L. L. Greischar, and R. J. Davidson, "Electromyogenic artifacts and electroencephalographic inferences revisited," *NeuroImage*, vol. 54, no. 1, pp. 4–9, 2011.
- [2] K. T. Sweeney, T. E. Ward, and S. F. McLoone, "Artifact removal in physiological signals—practices and possibilities," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 3, pp. 488–500, 2012.
- [3] H. Zeng, A. Song, R. Yan, and H. Quin, "EOG artifact correction from EEG recording using stationary subspace analysis and empirical mode decomposition," *Sensors*, vol. 13, no. 11, pp. 14839–14859, 2013.
- [4] M. H. Soomro, N. Badruddin, and M. Z. Yusoff, "Comparison of blind source separation methods for removal of eye blink artifacts from EEG," in *2014 5th International Conference on Intelligent and Advanced Systems (ICIAS)*, pp. 1–6, Kuala Lumpur, 2014.
- [5] J. Gao, C. Zheng, and P. Wang, "Online removal of muscles artefact from electroencephalogram signal based on canonical correlation analysis," *Clinical EEG and Neuroscience*, vol. 41, no. 1, pp. 53–59, 2010.
- [6] M. Anastasiadou, A. Hadjipapas, M. Christodoulakis, E. S. Papathanasiou, S. S. Papacostas, and G. D. Mitsis, "Detection and removal of muscle artifacts from scalp EEG recordings in patients with epilepsy," in *2014 IEEE International Conference on Bioinformatics and Bioengineering*, pp. 291–296, Boca Raton, FL, 2014.
- [7] M. Anastasiadou, M. Christodoulakis, E. S. Papathanasiou, S. S. Papacostas, and G. D. Mitsis, "Automatic detection and removal of muscle artifacts from scalp EEG recordings in patients with epilepsy," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 1946–1950, Milan, 2015.
- [8] K. T. Sweeney, S. F. McLoone, and T. E. Ward, "The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique," *IEEE Transactions Biomedical Engineering*, vol. 60, no. 1, pp. 97–105, 2013.
- [9] M. H. Soomro, N. Badruddin, M. Z. Yusoff, and M. A. Jatoi, "Automatic eye-blink artifact removal method based on EMD-CCA," in *2013 ICME International Conference on Complex Medical Engineering*, pp. 186–190, Beijing, 2013.
- [10] X. Chen, C. He, and H. Peng, "Removal of muscle artifacts from single-channel EEG based on ensemble empirical mode decomposition and multi-set canonical correlation analysis," *Journal of Applied Mathematics*, vol. 2014, Article ID 261347, 10 pages, 2014.
- [11] C. Zhao and T. Qiu, "An automatic ocular artifacts removal method based on wavelet-enhanced canonical correlation analysis," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4191–4194, Boston, MA, 2011.
- [12] D. Safieddine, A. Kachenoura, L. Albera et al., "Removal of muscle artifact from EEG data: comparison between stochastic (ICA and CCA) and deterministic (EMD and wavelet-based) approaches," *EURASIP Journal on Advances in Signal Processing*, vol. 127, pp. 1–15, 2012.
- [13] V. Roy and S. Shukla, "A methodical health-care model to eliminate motion artifacts from big EEG data," *JOEUC, Big Data Analytics in Business, Healthcare & Governance*, vol. 29, 2016.
- [14] S. Khatun, R. Mahajan, and B. I. Morshed, "Comparative analysis of wavelet based approaches for reliable removal of ocular artifacts from single channel EEG," in *2015 IEEE International Conference on Electro/Information Technology (EIT)*, pp. 335–340, Dekalb, IL, 2015.
- [15] A. H. Shoeb, *Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment*, [Ph.D. thesis], Massachusetts Institute of Technology, 2009.
- [16] R. Mahajan and B. I. Morshed, "Unsupervised eye blink artifact denoising of EEG data with modified multi-scale sample entropy, kurtosis and wavelet-ICA," *Biomedical and Health Informatics*, vol. 19, no. 1, pp. 158–165, 2015.