International Journal of Bio-Science and Bio-Technology
Vol.8, No.5 (2016), pp. 139-154
http://dx.doi.org/10.14257/ijbsbt.2016.8.5.13

Effective EEG Motion Artifact Removal with KS test Blind Source Separation and Wavelet Transform

Vandana Roy and Shailja Shukla

DoEC, GGITS, Jabalpur, M.P., 482005, INDIA
Professor and Head of DoCSE, JEC, Jabalpur, MP, 482011, INDIA
Email: vandana.roy20@gmail.com, shailja270@gmail.com

Abstract

Artifacts frequently corrupt biomedical signal recording and processing, therefore, removal of these artifacts from physiological signals is an essential step. The acuteness in the performance of healthcare technology has upgraded from the current hospital-centric environment towards portable ubiquitous approaches. The uncertainty in the subsequent performance of these approaches introduced a dedicated research and past few decades have witnessed considerable improvement. In this research work an enhanced empirical approach to model the artifacts of EEG signal are described. The input EEG is a single channel and is converted into multichannel using Ensemble Empirical Mode decomposition (EEMD) operations and further filtered with Independent Component Analysis and Double Density Wavelet Transform to reject any traces of artifacts left at signal. This proposed algorithm is tested with different evaluation parameters and results pronounce the eligibility of the proposed algorithm to stand on top of currently deployed algorithms because significant improvement in results.

***Keywords**: EEG, EMG, EEMD-ICA, ICA, DDWT, EEMD-DDWICA*

1. Introduction

The design of a physiological signal is a long followed approach that illustrates the present condition of the health of an individual. The dynamic research in medical sciences towards superlative health assessment calls for paramount accuracy with low computing cost in signal recordings and imaging. The life of an instrument to function with minimal operation and maintenance cost is one of the primary factors that defines the probability of its selection as the practically acceptable technology. Also, the simplicity of an instrument is directly measured as its capability to function in a standard environment without failure in operation. Further, the complexity of instruments has direct relationship with its cost.

It is evident that measurement of physiological signals in even the surgical environment is accustomed to some noise also referred as artifacts in medical terms. These artifacts are unwanted signals generated due to unregulated sources besides the source under consideration. The artifacts in neural signals have two prominent sources other than the machine and environment noise. The muscular and ocular activities of an individual generate electric pulses of low amplitude and frequency that falls in filter range of sensors and recording equipment. Hence, artifacts rejection is a fundamental subject of research and is well researched [[1][1]]. This paper considers the artifacts caused due to the motion. Till this point, numerous applications of Independent Component Analysis (ICA), wavelets, and adaptive filters are proposed in the same context of research [[2]].

Normally, a common approach is to reject all Electroencephalogram (EEG) epochs containing the signal amplitude larger than some selected value. These methods are inflexible and do not allow for any adaption, which causes in loss of a portion of meaning full data. A component based automated separator of artifacts is required to overcome this

ISSN: 2233-7849 IJBSBT

issue based on the linear decomposition of signals into source components. These source components give the individual nature of information. However artifacts information combines into separate sources. The reconstruction of signals without these artifact sources are claimed as artifact free information. The Fast ICA Algorithm for source has applied by authors [[3]-[4]] using EEGLAB platform. They developed an automated system for artifact removal based on ICA and Bayesian Classification. Authors [[5]] analyzed EEG waves by Discrete Wavelet Transform (DWT) for frequency domain analysis. Authors [[4]] presented a statistical method based on wavelet transform to mimic ocular artifacts in EEG. Authors [[6]] discussed wavelet-based image processing technique with various window sizes. 1-D double density and 1-D double density complex were tested on the window size of 10s, 30s, 60s and 300s for EEG signals. Mijovic *et al.* [[7]] studied two techniques Single-Channel ICA (SCICA) and Wavelet-ICA (WICA) and applied EEMD-ICA algorithm on single channel EEG signal for artifact removal. The EEMD-ICA algorithm was tested on simulated data and then applied on real EEG and EMG data for comparison. The conclusion from results is that the SCICA algorithm has the worst performance with root mean square error (RMSE) as consideration. The WICA algorithm has weaker performance in the simulations and although is comparable to the EEMDICA technique,

The organization of the paper is as follows: In this paper, section 2 describes EEMD followed by ICA in section 3 and Wavelet transform in section 4. The proposed system model is elaborated in section 5 and data acquisition in section 6. The performance evaluation parameters are discussed in section 7 and all results and discussion are in section 8. The results brace the splendid performance of proposed architecture and paper ends with a conclusion.

2. Empirical Mode Decomposition (EEMD)

In 1998 Empirical mode decomposition (EMD) is first defined by Huang et al. [[8]], for nonlinear signal processing and is well appropriate for non-stationary data. A time series signal decomposes into multiple “Intrinsic Mode Functions” (IMFs) by EMD. IMFs must have following property:

(1) Mono-component means all the IMFs should have only one frequency component at a time known as instantaneous frequency.

(2) Zero-mean oscillatory functions define that signals have the same number of local maxima and minima, with positive maxima and negative minima always.

(3) Orthogonal means different IMFs should not have the same frequency.

The EMD technique uses a different approach for decomposition rather than Wavelet analysis. Decomposition of the signal in EMD is a data-driven process, whereas wavelet analysis, decomposition is based on the selection of the appropriate wavelet. Since EMD technique is data driven, hence this approach is more flexible in nature.

The IMFs are functions that must fulfill two conditions:(1) the number of maxima and the number of zero crossings must be the same or differ at most by one over the full length of data (2) the mean value of the envelope defined by the maxima and the envelope defined by the minima must be zero at any point over the data [[8]].

To calculate IMF of a time series, steps are as follows: Time series is $y \in P^L$ where L is the number of samples. EMD is based on using a sifting process that uses only local extreme.

Step 1: All the local maxima and minima will find the full length of the time series. Then, an upper envelope is created by connecting all the maxima using a cubic spline, and the same process is repeated for the all local minima.

Step 2: Average of the two envelopes is calculated and this average is subtracted from the data signal, which produces a new signal $c = p_0 - n$, where $p_0 = y$ and n is average of envelope.

Step 3: Now signal c is considered as a new data signal and above-mentioned steps are repeated till c fulfills all the above-detailed properties of IMFs. Finally, when c, have all the properties of IMF, it is termed as first IMF (f_1).

Step 4: Then all the above mention steps are repeated on the residual signal $p_1 = p_0 - f_1$. As the residual signal p_n becomes a monotonic function this sifting process will stop. Once all the IMFs f_j are calculated, the original data y (or p_0) can be restored by adding them together as shown in equation (1).

$$y = \sum_{j=1}^{l} f_j + p_l \tag{1}$$

Where, f_j are extracted IMF components and p_l is the residual of data [[8]-[10]].

The method of detecting IMFs is sensitive to the amalgam of undesired signal components present in surrounding. These noises affect the EMD process. Thus mode mixing is used to overcome the disparate scale oscillations with amplitude in near range of the IMFs peaks and randomly available in the whole dataset [[9]]. Moreover, the momentary spectral components are sometimes misinterpreted as artifact components. An enhanced version of empirical mode decomposition [[10]] minimized this mode mixing quandary. Another version of EMD known as Ensemble-EMD (EEMD) employs the average value of EMD ensembles that filters out the IMFs from signal. Each iteration of EMD process is identical and independent in nature towards the undesired signals.

The IMFs from the EMD process are filtered with ICA to discard the artifacts present in them (details available in the following sections). ICA filtered IMFs are further processed for second stage filtering through wavelet transform for better outputs of evaluation parameters and signal quality.

3. Independent Component Analysis (ICA)

ICA employs statistical and computational techniques for separating the mixture of signals into independent components. The IMF(s) generated by EEMD are further sampled to ICA by equation 2 with input as $C = [c_1, c_2 .. c_j \dots c_n]$ are generated by independent sources $S = [s_1, s_2 .. s_j \dots s_n]$ where A is the $n \times m$ mixing matrix.

$$S = AC \tag{2}$$

Fast ICA algorithm is depicted in figure 1. Here, w_i is a column vector and w_i^+ is temporary variable, $g(.)$ and $g'(.)$ represent first and second derivate of nonlinear and non-quadratic functions. When the convergence is received then signal is reconstructed using equation S=WC with reducing the artifact component to zero.

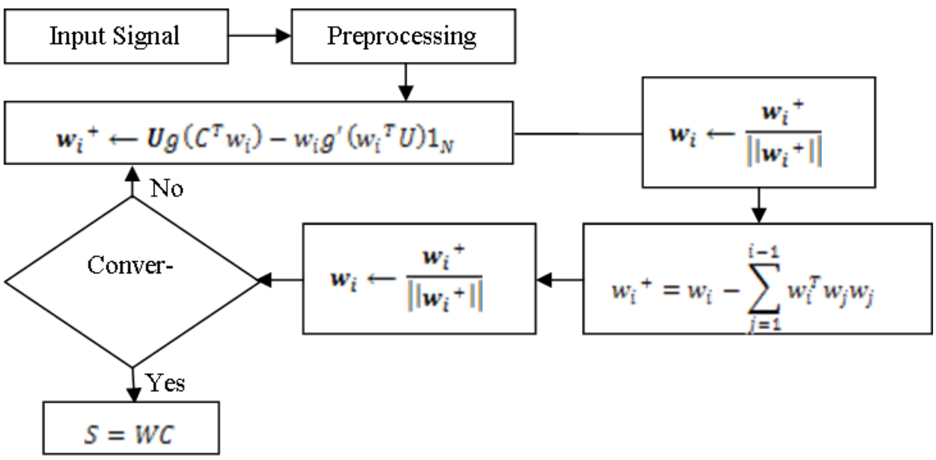

Figure 1. Fast ICA Algorithm

One-sample Kolmogorov-Smirnov test (K-S Test) [[11]] measures the non-Gaussianity in ICA. The output of ICA based on Gaussianity is classified into two types of signals, i.e.

signals either having random behavior or signals with stationary waveforms. The KS test separates both these signals by allocating a '0' and '1' to these signals respectively. The independent source components being artifacts are selected and relative matrix C column components are defined as null.

The KS test is implemented as the function in MATLAB and shown mathematically as equation 3:

$$D^* = \max_C\left(\left|\hat{F}(C) - G(C)\right|\right) \tag{3}$$

Where, $\hat{F}(C)$ is empirical CDF and $G(C)$ is the CDF of hypothesized distribution. The original source signal is thus as in equation 4:

$$S' = (W) * C' \tag{4}$$

The Fast ICA algorithm is a fixed point iterative method that provides the highest value of non-Gaussianity to evaluate the statistical independence. This algorithm looks for the course of weight vector by elevating the value of non-Gaussianity of projection $w_i^T x$ for input data x (figure 1). When the computation of the weight parameters completes, the estimated independent components will be generated by applying the unmixing matrix to EEG source signals.

With this technique, spectral enhancement can be achieved, but at the same time, it's very difficult to estimate the variances of Independent components [[12]]. To make more precise, analysis in time and frequency further wavelet technique is adopted [[4]].

4. Discrete Wavelet Transform (DWT)

The wavelet technique came into existence to overcome the resolution limitations of spectral analysis of Fourier Transform and denoise the corrupted signal. Selesnick and Ivan W [[13]] presented the Double Density Discrete Wavelet Transform (DD-DWT)as the custom-design form of Daubechies orthonormal wavelet transform by manipulating the window size of filter followed by entertaining certain critical polynomial aspects in an oversampled framework. The sensing time of DDWT is twice as that of the conventional DWT due to the more number of samples at each step. This algorithm was further customized as 1-D, 2-D up to Double Density complex [[14]]. The input signals are in the space-time domain and reflects discrete properties because of oversample analysis and respective synthesize filterbank (Figure 2).

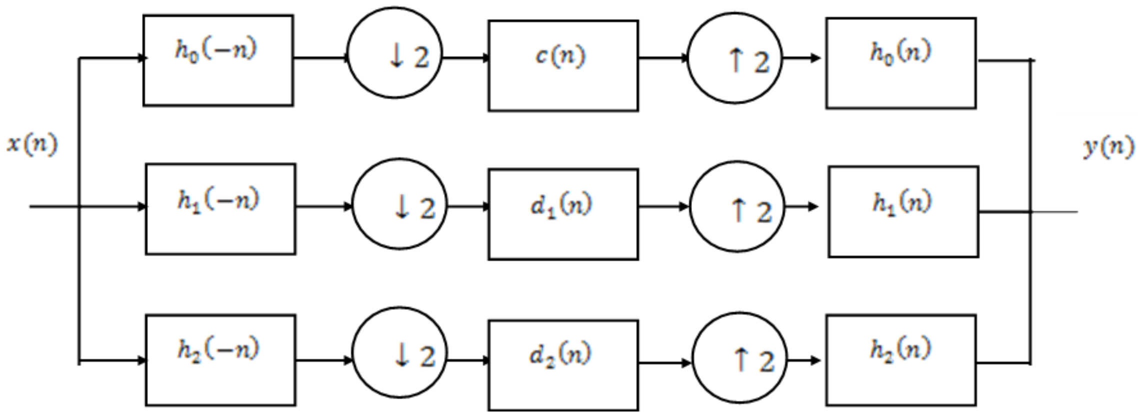

Figure 2. Analysis-Bank Filter Frequency Responses

Figure 2 depicts the frequency response of filters h^0, h^1, h^2 for analysis-bank filters of the defined size 7, 7, 5 respectively. The first is a low pass and remaining two are high pass filters in nature in the constraints of the congruent response of frequency magnitude. Iteration is employed for generation of oversampled filter bank to measure the $h^0$the response of both synthesis filter banks. An application of DD-DWT was illustrated by

Selesnick et al. [[13]] that analyses the unity scaling function $\emptyset$ and multi-resolution along with φ^1, φ^2 wavelets (see figure 2). A "half-delay" property $\varphi^1(t) = \varphi^2\left(t - \frac{1}{2}\right)$ can be satisfied by application of wavelets.

The double density DWT is applying the oversampled filter bank on a low-pass subband signal $c(n)$. In this paper the double density algorithm is used along with the filter banks for sampling purposes.

Figure 2 represents the formation of DD-DWT along with FIR filters as an oversampled filter bank. The filters employed are the low and high pass h^0, h^1, h^2 respectively as discussed above.

To develop the perfect reconstruction condition, standard multi-rate identities are used to write $Y(z)$ in terms of $X(z)$ as in equation 5.

$$Y(z) = \frac{1}{2}\left[H_0(z)H_0\left(\frac{1}{z}\right) + H_1(z)H_1\left(\frac{1}{z}\right) + H_2(z)H_2\left(\frac{1}{z}\right)\right]S'(z) \tag{5}$$

$$= \frac{1}{2}\left[H_0(z)H_0\left(-\frac{1}{z}\right) + H_1(z)H_1\left(-\frac{1}{z}\right) + H_2(z)H_2\left(-\frac{1}{z}\right)\right]S'(-z) \tag{6}$$

The filter bank employed to structure the double-density discrete wavelet transform resembles the wavelet frame with $\phi(t)$ as scaling function and $\psi_1(t)$ and $\psi_2(t)$ as wavelet functions. According to the documents of dyadic wavelet space

$$V_j = \underset{n\epsilon Z}{Span}\{\phi(2^j t - n)\} \tag{7}$$

$$W_{i,j} = \underset{n\epsilon Z}{Span}\{\psi_i(2^j t - n)\}, \quad i = 1,2 \ldots n. \tag{8}$$

Where, V_j is the single level decomposition and $W_{i,j}$ is group of V_j presented mathematically by equation 7 and 8.

The single wavelet ψ and scaling function ϕ defines the structure of a dyadic wavelet. The wavelets and scaling function satisfy the dilation equations.

$$\emptyset(t) = \sqrt{2}\sum_n h_0(n)\,\emptyset(2t - n) \tag{9}$$

$$\psi_i(t) = \sqrt{2}\sum_n h_i(n)\,\emptyset(2t - n) \qquad i = 1,2 \ldots n. \tag{10}$$

$h_0(n), h_1(n)$ and $h_2(n)$ from above section illustrates about wavelets $\psi_1(t), \psi_2(t)$ and scaling function $\emptyset(t)$ of equation 9 and 10.

A filtering chain for DWT decomposition of Single Channel Source which generates wavelets through z-transform is shown in figure 3. Figure 3 details the single channel source signal decomposition which is already processed through ICA.

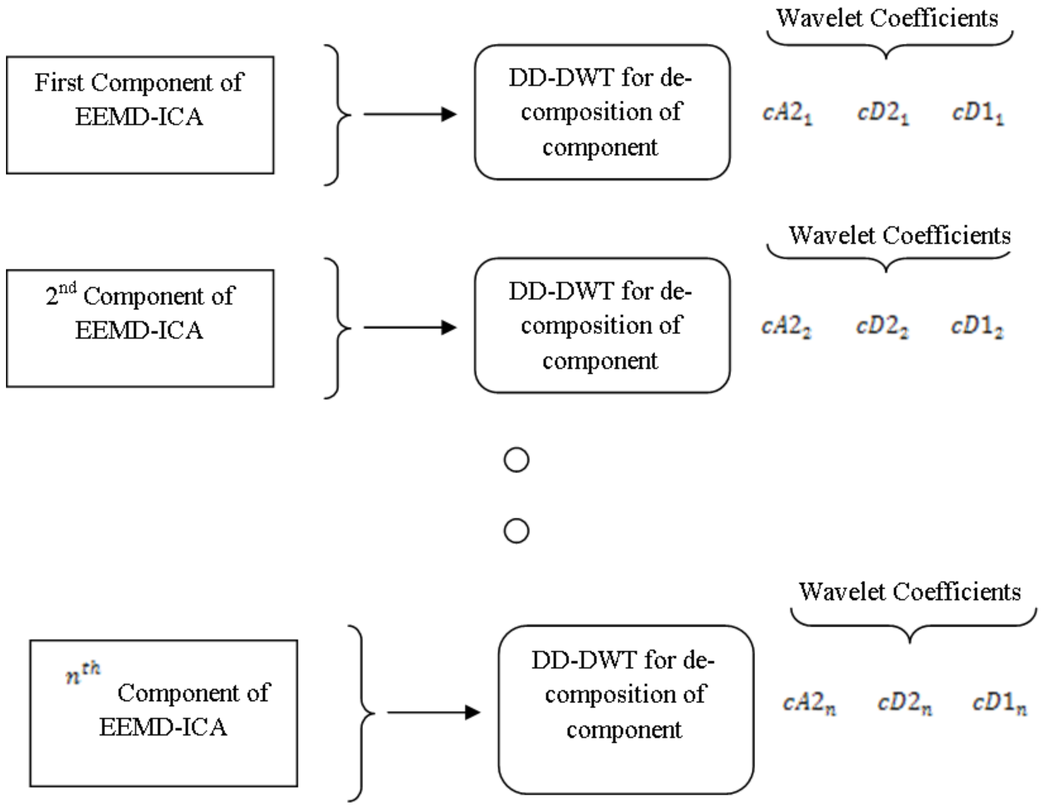

Figure 3. DWT Decomposition of Single Channel Source Signal Obtained from ICA

Figure 3 represents the wavelet coefficients for the output of ICA components. Here, n components are considered as the figure is the generalized notation of wavelet step. The coefficients of the wavelet (for example: $cA2_1$, $cD2_1$, $cD1_1$) are further optimized for the error free signal.

The wavelet coefficients for a single level (equation 7) and for multi-level (equation 8) are used to derive coefficients of ICA. The equation for multi-level decomposition will be as in equation 10:

$$\mathrm{W_{I,j}} = \underset{\mathrm{n\epsilon Z}}{\mathrm{Span}}\{\psi_\mathrm{i}(2^\mathrm{j}\mathrm{t}-\mathrm{n})\}\mathrm{S_i}, \quad \mathrm{i} = 1,2 \ldots \mathrm{n}. \tag{10}$$

The coefficients are thus separate units of original signal and artifacts. To differentiate among them and discard the Gaussian components thresholding is applied. There are two common thresholding processes in this field. The first is hard-thresholding [[15]] presented mathematically by equation 11:

$$\widehat{\mathrm{W}}_{\mathrm{ij}}(\mathrm{t}) = \begin{cases} \mathrm{W_{ij}}(\mathrm{t}) \text{for } |\mathrm{W_{ij}}(\mathrm{t})| \geq \mathrm{T} \\ 0 \qquad \text{otherwise} \end{cases} \tag{11}$$

The other is soft thresholding [[16]-[17]] depicted by equation 12.

$$\widehat{\mathrm{W}}_{\mathrm{ij}}(\mathrm{t}) = \begin{cases} \mathrm{sgn}\left(\mathrm{W_{ij}}(\mathrm{t})\right)\left(|\mathrm{W_{ij}}(\mathrm{t})| - \mathrm{T}\right) \text{for} |\mathrm{W_{ij}}(\mathrm{t})| \geq \mathrm{T} \\ 0 \qquad \text{otherwise} \end{cases} \tag{12}$$

The IMF(s) reconstructed from wavelet coefficients (eq. 12) have the comparatively lesser amount of artifacts compared with conventional ICA techniques and obtained using equation as

$$\hat{S}_n = \sum_{i=1}^{k} w_{ij} \tag{13}$$

Where, $\hat{S}_n$ is number of IMF(s) reconstructed, k is the level of decompositions and w_{ij} are wavelet coefficients. The original signal is thus the sum of all the m number of IMF(s) reconstructed from EMD as by equation 14.

$$y = \sum_{i=1}^{m} \hat{S}_i \quad (14)$$

5. Proposed System Model

To obtain better motion artifact removal from a single channel input signal, the following approach is used whose functional description of the prototype is depicted by the model as shown in Figure 4.

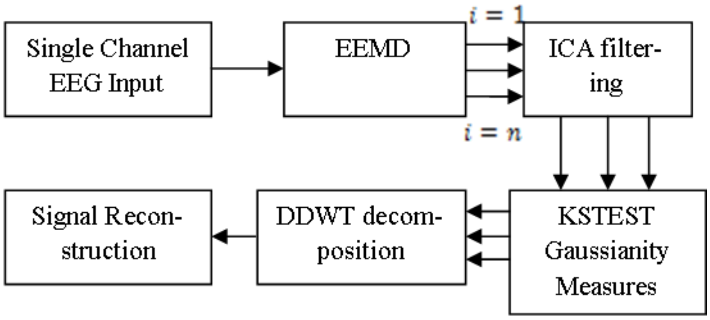

Figure 4. Proposed Architecture for EEG Artifacts Removal in Single Channel Signal

First [[7]] proposed the ICA and EEMD in combined form for separation of source and artifacts from the input signal in a given channel. Authors employed EEMD method to segment the input single channel signal in multi-channel signal and IMFs were generated. The output of EEMD was sourced to Fast ICA algorithm to modify the signals according to their respective sources S. The ICA separates the Independent components from input IMF(s). Finally, noisy components were removed to find an artifact-free signal.

The author used the system to eradicate artifacts of muscle movement and eye blinks primarily from EEG signals as shown in figure 5. As figure 5 presents that randomness in the EEG signal is introduced due to muscle movement and eye blinks artifacts.

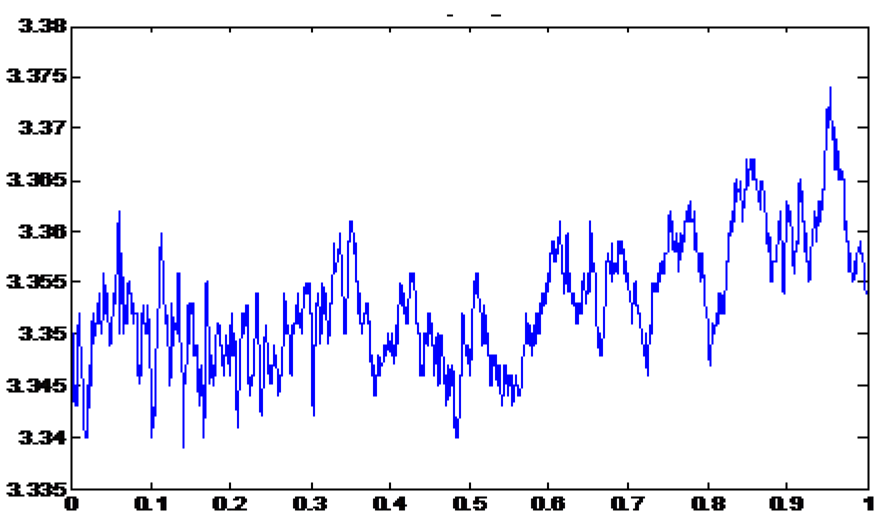

Figure 5. Input Noisy EEG Signal with Muscle Artifacts (X axis is Time and y-axisis Amplitude (Microvolt))

In the application of human health diagnosis, motion artifacts play a vital role because long-term monitoring of the patients causes discomfort and hence motion in subjects introduces motion artifact in captured EEG signal. These artifacts have higher amplitude than EEG signal and having wider spectral distribution. Whereas Electrooculogram

(EOG) artifacts are easy to remove through adaptive filters and blind source separation. To improve the motion artifacts separation from EEG signal the proposed methodology employed the combination of EEMD, Fast ICA, and double density DWT. In this research paper, EEMD is applied to convert single channel input signal to multichannel signal and subsequently, Fast ICA is applied on IMFs from EEMD process to find independent components. Out of these components, a single component has Gaussian amplitude and random behavior. The KS test allocates '0' to this component to filter it as an artifact. At last as an improvement, the double density wavelet transform is applied on filtered and processed signal to remove the randomness of artifacts, if any left in the signal and finally reconstructed to find a better artifact-free signal.

6. Data Acquisition

The data for experimentation is contributed by Kevin Sweeney at the National University of Ireland at Maynooth. The data are available on an online open source interface [[18]] and description of recordings is acquired from the same source only. The recorded samples are EEG samples contaminated by motion artifacts. Every recording is a single pair of similar psychological signals recorded in close proximity via transducers. Keeping the single transducer stationary, the second transducer is manipulated to create artifacts related to motion and of variable duration within each 2-minute interval of recording. To mimic the real time motion artifact effect on original EEG data the second transducer has moved randomly. The EEG signals sampling frequency was 2048 Hz. Comparison of the artifact-free EEG having a high correlation during motion-free intervals and lower correlation during artifact-contaminated intervals defines the efficiency of proposed work. The slightly increased time cost is quite acceptable given the improvement in artifact removal. This trade-off is especially beneficial in ambulatory systems in which clean information is important for diagnosis. All the simulations and implementations have been done in MATLAB (MathWorks) run under Microsoft Windows 8.1 x64 OS on the computer with Intel(R) core(TM) i-5-4200U, 2.30 GHz CPU, and 8.00 GB RAM. The synthetic EEG data band pass filtered at 0.5-50 Hz and notch filtered at 50 Hz.

7. Performance Evaluation Parameters

7.1. Signal to Noise Ratio (SNR)

The SNR is a measure of low magnitude waves stimulated by some sinusoid approach. The input wave in space-time domain is studied via periodogram optimized with Kaiser Window to attenuate large side lobes. The algorithm looks for the non-zero spectral component to result in fundamental frequency. The central moment of each subject is computed for all adjacent bins in a decreasing order (i.e. from maximum to minimum). These frequencies are detectable in second bin and further frequencies are the replica of these steps. The power of a signal is selected as the larger harmonic in case if the signal shows the monotonically decreasing behavior compared with the neighboring signal. The function is the ratio of noise intensity in the noise contaminated region derived via median power. To calculate the performance, the DC component is rejected and noise at every step could be either ordinate of a point or estimated level. This noise is eliminated for an artifact-free signal.

7.2. POWER spectral Density (PSD)

The supplemented readings of Electroencephalography signals are sourced by Gaussian noise. The obtained variation in signals with reference to theoretical model leads to restricted access in the accuracy of the power spectrum of signals as in equation 15.

$$P_{\overline{xx}}(w) = P_{rr}(w) + \Delta P(w) \quad (15)$$

Where,

P_{rr} = Reference Power Spectrum of Artifact free signal

$P_{\overline{xx}}$ = Power Spectrum of ICA modified Signal

P = Distortion Spectrum by side effects of methods (Should be 0 ideally)

$$\Delta P_j(w) = -m_{j1}^2 P_{nn}(w) \quad (16)$$

Here,

m_{j1} Denotes weight from Matrix M

P_{nn} Component of brain signals in ICA

Equation 16 represents the minimization of ICA-EEG based on spectral function (P_{nn}) with factors m_{j1}^2. As there exists a direct relationship among decrease in $j$and m_{j1}^2, front end experience high distorted signals of the spectrum. DDWICA along with reduction of residual EEG signals in artifact components simultaneously minimize P_{nn} on a serious note that leads to the improved approximated value of pure EEG power spectrum.

8. Results and Discussion

EEG signal with motion artifacts are acquired from [[18]]. Signals having motion artifacts have a low SNR due to the harmonics created by random frequency artifacts as shown in figure6. As the motion artifacts have been introduced, the original EEG signal will have high amplitude randomness due to artifacts.

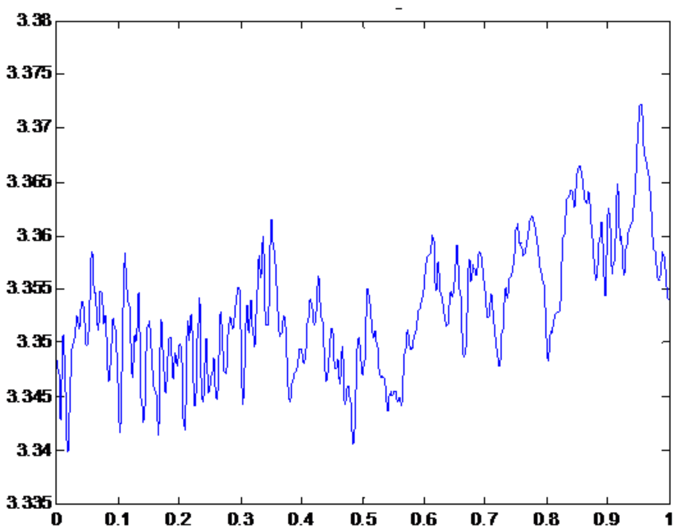

Figure 6. EEG Signal with Artifacts (X axis is Time and y-Axisis Amplitude (Microvolt))

The signal is decomposed to IMFs using EEMD to convert single channel input signal into multichannel signal, so as to provide a monotonic component of the signal as a separate source for ICA decomposition. After EEMD decomposition, IMFs are shown in figure 7. X axis is the time and y-axisis the amplitude (microvolt).

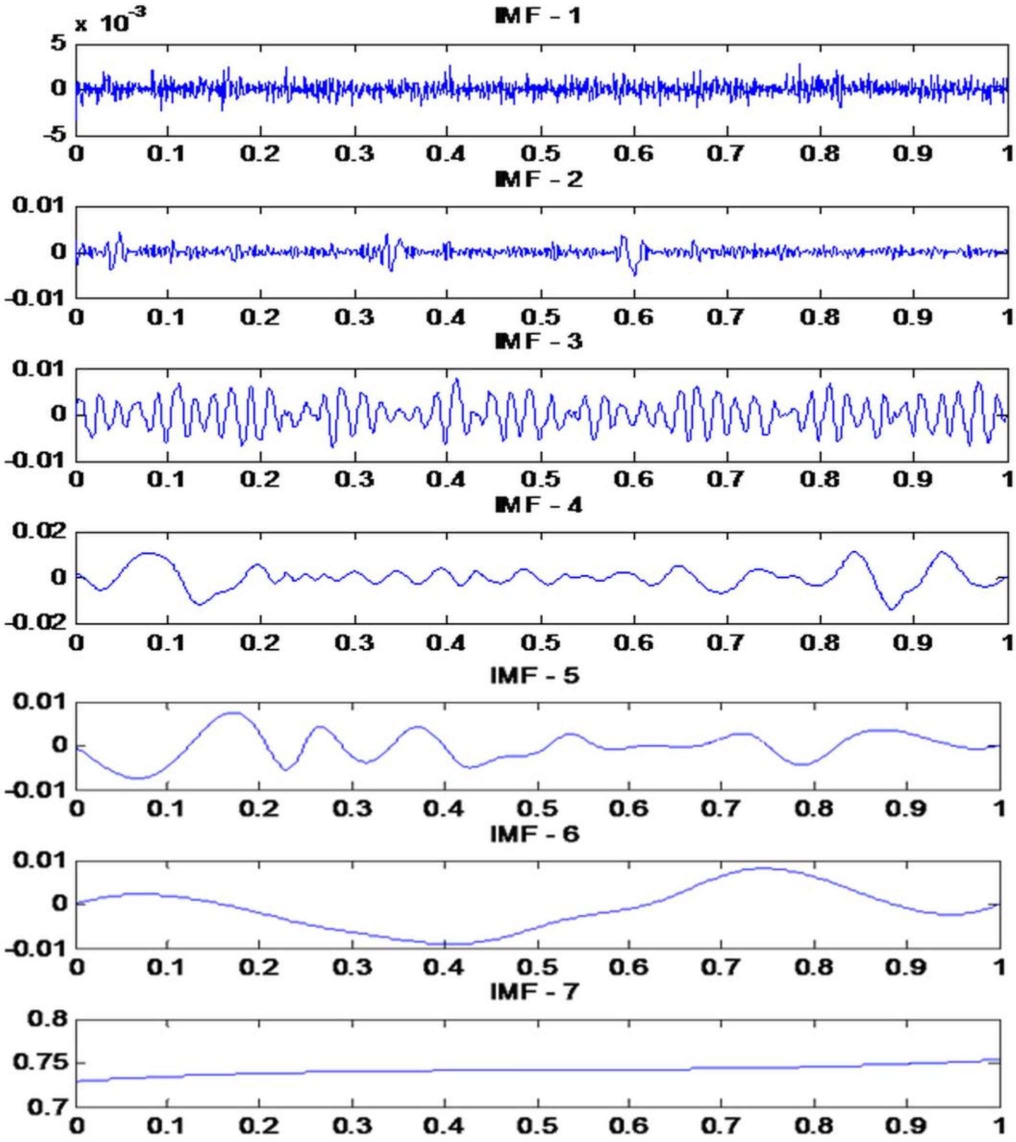

Figure 7. EEMD Extracted IMFs from EEG Signal (X axis is Time and y-Axisis Amplitude (Microvolt))

A blind source separation with Fast ICA is performed over these IMFS so that all the signals from different sources, following different distribution will be separated as shown below in figure 8. Blind source separation will decompose the signal according to various EEG electrode resources. This will facilitate the easy identification and separation of artifact source.

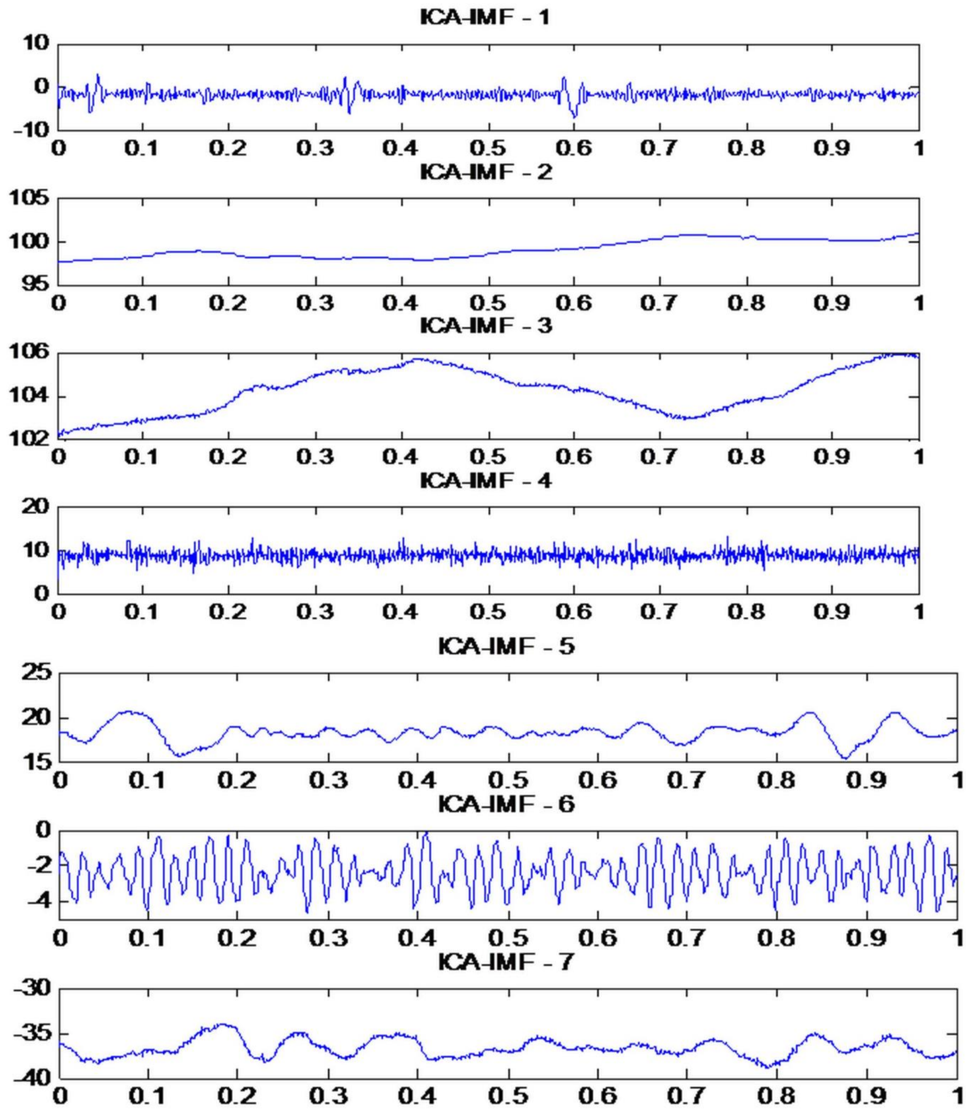

Figure 8. ICA Extracted Independent Components from IMFs (X Axis is Time and y-Axisis Amplitude (Microvolt))

KS test is performed after Fast ICA operation, and analyzed IC4 having a random distribution, thus it has the noise component. Hence, this independent component (IC4) has been reduced to zero and reconstructing IMFs with weighting matrix W. This process removed a big part of the noise from the signal as the SNR for EEG signal increased. Finally three levels Double Density Wavelet Transform with soft thresholding is applied on reconstructed IMFs to remove artifact traces left in the signal. Filtered IMFs are reconstructed to achieve error-free signal as shown in figure 9. The results after artifact removal show the smoothened effect with original information of EEG signal preserved. To evaluate the performance of proposed algorithm SNR is calculated after filtering and compared with existing algorithm as in table 1. SNR of the EEG signal with proposed artifact removal algorithm has been improved. The signal quality after artifact removal can also be evaluated by comparing the power spectral density of signal before and after artifact removal.

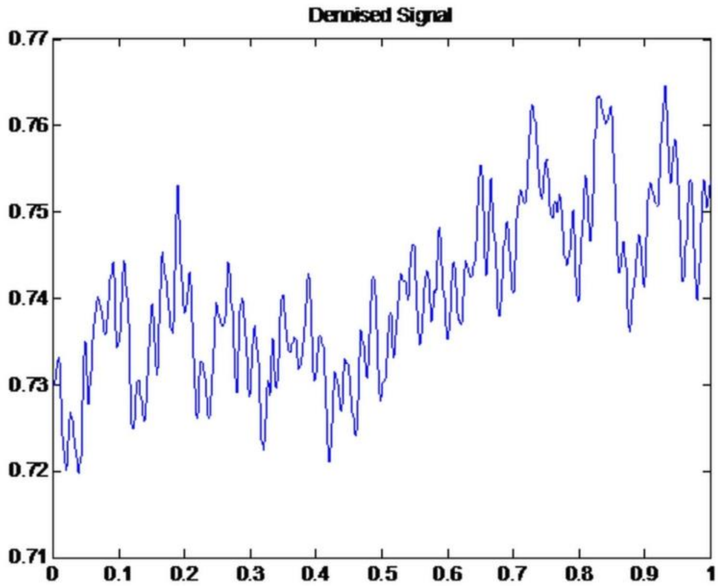

Figure 9. EEG with Removed Artifacts (X axis is Time and y-Axisis Amplitude (Microvolt))

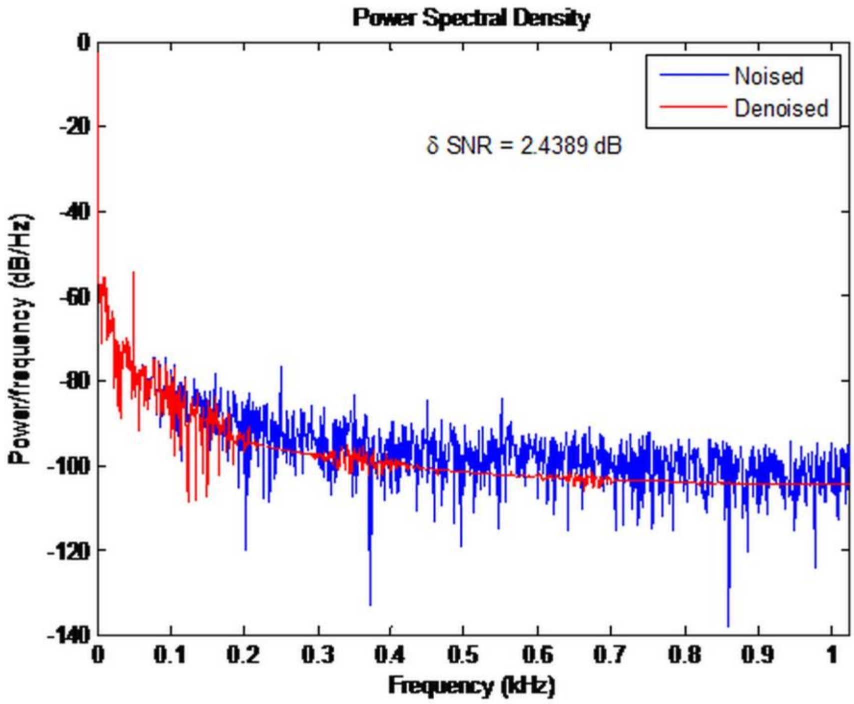

Figure 10. PSD Comparison of Noisy and Denoised EEG Signals

Figure 10 shows the Power Spectral Density for EEG with artifacts (blue) and without artifacts (Red), as observed from figure power of high-frequency components has been lowered, as in EEG signal frequency components above a particular level is considered as noise. When these high-frequency components were removed then the power of that frequency region will be reduced, which gives a rise in SNR of 2.4389 dB.

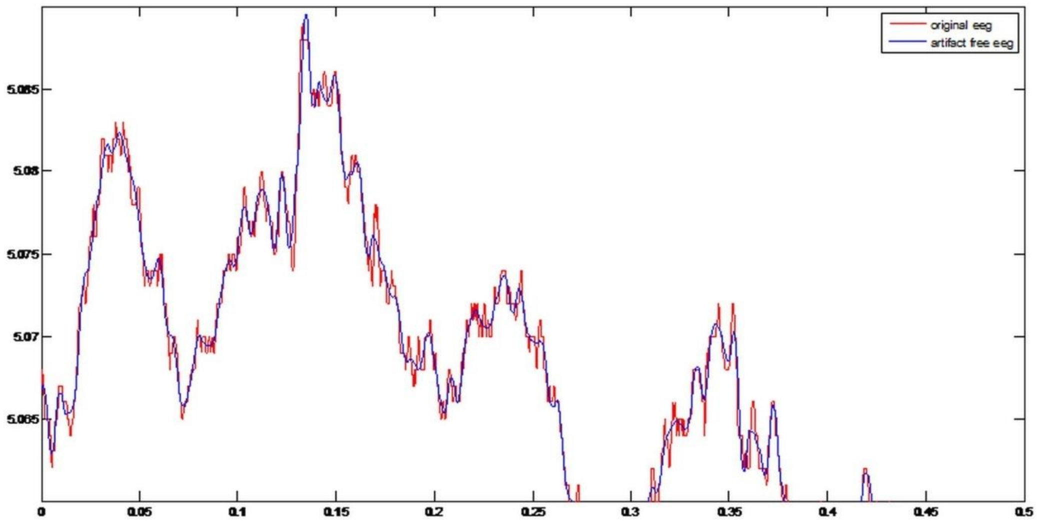

Figure 11. Noisy and Denoised Signals (X axis is Time and y-Axisis Amplitude (Microvolt))

Figure 11 shows the plot comparison of the signal before and after artifact removal and shows that after artifact removal, the signal smoothened and peaks due to motion artifacts have been removed to a great extent.

Table 1shows the SNR comparison for 7 EEG signals by two methods. One method is to remove artifacts by a combination of EEMD and ICA and another proposed approach is remove artifacts with a combination of EEMD and Fast ICA followed by KS test and double density wavelet transform.

Table 1. Comparative table between ICA and Proposed Double Density DWT of SNR for 7 EEG signals

EEG Signals	EEMD-ICA-DD_DWT (dB)	EEMD-ICA(dB)
EEG 1	3.3008	2.869
EEG 2	4.0558	0.20838
EEG 3	5.6965	2.9495
EEG 4	3.0611	2.1422
EEG 5	2.4398	0.59004
EEG 6	2.9307	0.51126
EEG 7	1.3885	2.2284

Table 1 suggests that proposed approach shows improvement in SNR over existing approach to remove motion artifacts from EEG signal. One important point is that as motion artifacts are random in behavior. Therefore, the performances of algorithm do not follow any specific trend with respect to evaluation parameter. This is the reason for EEG dataset 7 the proposed algorithm unable to show an improved performance with respect to existing artifact method. One more possible reason is as the synthetic artifacts have been randomly created which can limit the outcomes of artifact removal methods occasionally.

9. Conclusion

An improved technique for motion artifact removal in EEG signal has been proposed. Moreover, results have been compared and proved better than results of method proposed in [[7]]. EEG Signals have specified a range of frequency with different amplitude for particular points of the EEG acquisition system, but when artifacts are introduced, the range of frequency becomes higher with random high-frequency components shows significant power in the power spectrum. Removing these artifacts directly with filters may lose the necessary information because of amplitude variation, so EEMD is applied first to decompose the signal into IMFs for converting single time samples signal to multiple sources with different frequencies. Each IMF has a small frequency range, but if the signal has artifacts than IMFs will be having a low amplitude of high-frequency components. To separate this Fast ICA is performed over these IMFs to get independent components of the signal so that the component of artifacts will be treated as one IC with a random distribution, which can easily be eliminated with the check of Gaussianity with KS test and IMFs can be reconstructed from the rest ICs with inverse ICA. To get better and smooth performance, before reconstructing the signal from IMFs, three Levels Double Density DWT with soft thresholding method over each IMF have been applied to remove the high-frequency energy components form signal. Finally reconstructed the signal and compared with existing results pronounce the eligibility of the proposed algorithm to stand on top of currently deployed algorithms on account of significant improvement in results.

Future research: In future research based on optimal wavelets with an optimal number of levels can be applied to remove motion artifacts from EEG big data efficiently. Further work may be done to optimize our method for real-time applications.

References

[1] K. Sweeney, T. Ward and S. McLoone, "Artifact Removal in Physio-logical Signals -Practices and Possibilities", Information Technology in Biomedicine, vol. 16, no. 3, **(2012)**, pp. 588-600.

[2] F. C. Robertson, T. S. Douglas and E. M. Meintjes, "Motion Artifact Removal for Functional Near Infrared Spectroscopy: A Comparison of Methods", IEEE Transactions Biomedical Engineering, vol. 57, no. 6, **(2010)**, pp. 1377-1387.

[3] A. Hyvarinen and O. Erkki, "A fast fixed-point algorithm for independent component analysis. Neural computation, vol. 9, no. 7, **(1997)**, pp. 1483-1492.

[4] P. S. Kumar, R. Arumuganathan, K. Sivakumar and C. Vimal, "Removal of Ocular Artifacts in the EEG through Wavelet Transform without using an EOG Reference Channel", International Journal of Open Problems Compt. Math, vol. 1, no. 3, **(2008)**, pp. 188-200.

[5] H. K. Sawant and Z. A. Jalali, "Detection and classification of eeg waves", Oriental Journal of Computer Science and Technology, vol. 3, no. 1, **(2010)**, pp. 207-213.

[6] H. Abdullah and D. Cvetkovic, "Double Density Wavelet for EEG Signal Denoising", International Conference on Machine Learning and Computer Science, **(2013)**, pp. 51-53.

[7] B. Mijovic, M. De Vos, I. Gligorijevic, J. Taelman and S. V. Huffel, "Source Separation From Single-Channel Recordings by Combining Empirical-Mode Decomposition and Independent Component Analysis", IEEE Transactions Biomedical Engineering, vol. 57, no. 9, **(2010)**, pp. 2188-2196.

[8] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih and Q. Zheng, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis", Proceedings of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences, vol. 454, no. 19, **(1998)**, pp. 903-955

[9] Z. Shen, Q. Wang, Y. Shen, J. Jin and Y. Lin, "Accent extraction of emotional speech based on modified ensemble empirical mode decomposition", Instrumentation and Measurement Technology Conference, vol. 20, **(2010)**, pp. 600-604.

[10] Z. Wu and N. E. Huang, "Ensemble Empirical Mode Decomposition: a Noise-Assisted Data Analysis Method", Advances in Adaptive Data Analysis, vol. 1, no. 1, **(2009)**, pp. 1-41.

[11] A. P. Bradley, "ROC curve equivalence using the Kolmogorov–Smirnov test", Pattern Recognition Letters, vol. 34, no. 5, **(2013)**, pp. 470-475.

[12] S. Arora, R. Ge, A. Moitra and S. Sachdeva, "Provable ICA with unknown Gaussian noise, with implications for Gaussian mixtures and autoencoders", Advances in Neural Information Processing Systems, **(2012)**, 2375-2383.

[13] I. W. Selesnick, "Hilbert Transform Pairs of Wavelet Bases", IEEE SIGNAL PROCESSING LETTERS, vol. 8, no. 6, **(2001)**, pp. 170-173.

[14] S. Chitchian, M. A. Mayer, A. R. Boretsky, F. J. Van Kuijk and M. Motamedi, "Retinal optical coherence tomography image enhancement via shrinkage denoising using double-density dual-tree complex wavelet transform", Journal of biomedical optics, vol. 17, no. 11, **(2012)**, 116009.

[15] P. Quan, Z. Lei, D. Guanzhong and Z. Hongai, "Two denoising methods by wavelet transform", IEEE Trans. on Acoustics, Speech, and Signal Processing, vol. 47, no. 12, **(1999)**, pp. 3401-3406.

[16] Y. Chen and C. Han, "Adaptive wavelet threshold for image denoising", Electronics Letters, vol. 41, no. 10, **(2005)**, pp. 586-587.
[17] M. L. Hilton and R. T. Ogden, "Date analytic wavelet threshold selection in 2-D signal denoising", IEEE Trans. on Signal Processing, vol. 45, no. 2, **(1997)**, pp. 496-500.
[18] Physionet. (n.d.). Retrieved from PhysioNet-motion artifact contaminated fNIRS and EEG data (motion artifact): http://physionet.org/cgi-bin/atm/ATM
[19] P. C. Nazareth and A. M. Valeri, "Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis", Journal of Neuroscience Methods, **(2006)**, pp. 300-312.