A New Approach to Denoising EEG Signals - Merger of Translation Invariant Wavelet and ICA

Janett Walters-Williams *jwalters@utech.edu.jm*
School of Computing & Information Technology
Faculty of Engineering & Computing
University of Technology, Jamaica
237 Old Hope Road,Kingston 6, Jamaica W.I.

Yan Li *liyan@usq.edu.au*
Department of Mathematics & Computing
Faculty of Sciences
University of Southern Queensland
Toowoomba, Australia

Abstract

In this paper we present a new algorithm using a merger of Independent Component Analysis and Translation Invariant Wavelet Transform. The efficacy of this algorithm is evaluated by applying contaminated EEG signals. Its performance was compared to three fixed-point ICA algorithms (FastICA, EFICA and Pearson-ICA) using Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR), Signal to Distortion Ratio (SDR), and Amari Performance Index. Experiments reveal that our new technique is the most accurate separation method.

Keywords: Independent Component Analysis, Wavelet Transform, Unscented Kalman Filter, Electroencephalogram (EEG), Cycle Spinning

1. INTRODUCTION

The use of electroencephalogram (EEG) in the field of Medicine has had a great impact on the study of the human brain. The EEG itself represents the brain activity for a subject and gives us an objective mode of recording brain stimulation. It also has been suggested by several studies that EEGs can be used to detect several diseases such as Creutzfeldt-Jakob diseases (CJD), Alzheimer's, Dementia, Epilepsy, and Schizophrenia. The signals received, however, have several origins that lead to the complexity of their identification. This complexity is made of both the pure EEG signal and other non-cerebral signals called artifacts or noise. The artifacts have resulted in the contamination of the EEG signals; hence the removal of these noises has generated a large number of denoising techniques and methods, for example Fourier transform, time-frequency analysis, Wavelet Transform (WT), Neural Networks (NN), and Independent Component Analysis (ICA).

Independent Component Analysis (ICA) originated from the field of Blind Source Separation (BSS) [8]. In the BSS problem, a set of observations is given while the underlying signal information is hidden; the mixing weights of the individual signals are unknown. BSS is aimed at identifying the source signals and/or the mixing weights so as to separate these information sources into signal domain, feature domain or model domain [5]. ICA therefore calls for the separation of the EEG into its constituent independent components (ICs) and then eliminating the ICs that are believed to contribute to the noise. It is subjective, inconvenient and a time consuming process when dealing with large amount of EEG data.

Different types of ICA algorithms were proposed in the last 10 to 12 years. Most of them suppose that the sources are stationary and are based explicitly or implicitly on high order statistics computation. Therefore, Gaussian sources cannot be separated, as they don't have higher than 2 statistic moments. Other types of algorithms do not make the stationarity hypothesis, and use the non stationary structure of the signals (i.e. their time or frequency structure) to separate them. These methods use second order statistics (SOS) only, and they are called SOS algorithms. As EEG signals are highly non-stationary, these type of algorithms are the most widely used.

Like ICA, Wavelet Transform (WT) has been used to study EEG signals [3][20][32][35-36][41][44] successfully because of its good localization properties in time and frequency domain [13]. Here, the EEG signals pass through two complementary filters and emerge as two signals, approximation and details. This is called decomposition or analysis. The components can be assembled back into the original signal without loss of information. This process is called reconstruction or synthesis. The mathematical manipulation, which implies analysis and synthesis, is called discrete wavelet transform (DWT) and inverse discrete wavelet transform (IDWT). There have been many approaches to denoising using WT; those based on shrinkage are the most popular [32] where the EEG signals are decomposed into wavelets and noise removal done using thresholding and shrinkage.

Akin [1] in his research compared WT with fast Fourier transform and found that WT was better in detecting brain diseases. His research was confirmed by Hermann *et al* [16]. Unser *et al* [39] showed that wavelet is good at denoising EEG signals as well as other biomedical signals. WT has therefore emerged as one of the superior technique in analyzing non-stationary signals like EEG. Its capability in transforming a time domain signal into time and frequency localization helps to understand the behaviour of a signal better. WT however has limitations such as Gibbs phenomena [7].

Each of the above methods presents their own limitations. In our opinion a method that aims to fix these limitations should be a more effective denoising method. This is possible as each method is used to overcome the limitation of the other. We present in this paper a new method of extracting noise from EEG signals which aims to remove the limitations of ICA and WT while improving effectiveness – Cycle Spinning Wavelet Transform ICA (CTICA). The performance of CTICA is analyzed and compared with three known fixed-point ICA algorithms – FastICA, EFICA and Pearson-ICA, by using EEG signals contaminated with noise. Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR), Signal to Distortion Ratio (SDR), and Amari Performance Index are used as criteria for testing the quality of denoising.

2. SUPPORTING LITERATURE

EEG Signals

The nervous system sends commands and communicates by trains of electric impulses. When the neurons of the human brain process information they do so by changing the flow of electrical current across their membranes. These changing current (potential) generate electric fields that can be recorded from the scalp. Studies are interested in these electrical potentials but they can only be received by direct measurement. This requires a patient to under-go surgery for electrodes to be placed inside the head. This is not acceptable because of the risk to the patient. Researchers therefore collect recordings from the scalp receiving the global descriptions of the brain activity. Because the same potential is recorded from more than one electrode, signals from the electrodes are supposed to be highly correlated. These are collected by the use of an electroencephalograph and called electroencephalogram (EEG) signals.

Understanding the brain is a huge part of Neuroscience, and the development of EEG was for the elucidation of such a phenomenon. The morphology of the EEG signals has been used by researches and in clinical practice to:

- Diagnose epilepsy and see what type of seizures is occurring.
- Produce the most useful and important test in confirming a diagnosis of epilepsy.
- Check for problems with loss of consciousness or dementia.
- Help find out a person's chance of recovery after a change in consciousness.
- Find out if a person who is in a coma is brain-dead.

- Study sleep disorders, such as narcolepsy.
- Watch brain activity while a person is receiving general anesthesia during brain surgery.
- Help find out if a person has a physical problem (in the brain, spinal cord, or nervous system) or a mental health problem.

The signals must therefore present a true and clear picture about brain activities. Being a physical system, recording electrical potentials, present EEG with problems; all neurons, including those outside the brain, communicate using electrical impulses. These non-cerebral impulses are produced from:

- Eye movements & blinking - Electrooculogram (EOG)
- Cardiac Movements - Cardiograph (ECG/ EKG)
- Muscle Movements - Electromyogram (EMG)
- Chewing & Sucking Movement – Glossokinetic
- The machinery used to record signals
- The power lines.

EEG recordings are therefore a combination of these signals called artifacts or noise and the pure EEG signal defined mathematically as:

$$E(t) = S(t) + N(t) \qquad (1)$$

where S is pure EEG signal, N is the noise and E represents the recorded signal. The presence of these noises introduces spikes which can be confused with neurological rhythms. They also mimic EEG signals, overlaying these signals resulting in signal distortion (Figure 1). Correct analysis is therefore impossible, resulting in misdiagnosis in the case of some patients. Noise must be eliminated or attenuated.

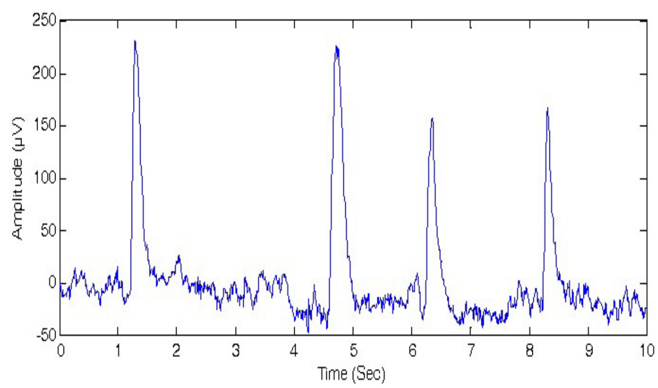


FIGURE 1: EEG contaminated with EOG producing spikes

The method of cancellation of the contaminated segments, although practiced, can lead to considerable information loss thus other methods such as Principal Components Analysis (PCA) [45], the use of a dipole model [18] and more recently ICA and WT have been utilized.

Independent Component Analysis

Independent Component Analysis (ICA) is an approach for the solution of the BSS problem [8]. It can be represented mathematically according to Hyvarinen, Karhunen & Oja [19] as:

$$X = As + n \qquad (2)$$

where X is the observed signal, n is the noise, A is the mixing matrix and s the independent components (ICs) or sources. (It can be seen that mathematically it is similar to Eq. 1). The problem is to determine A and recover s knowing only the measured signal X (equivalent to $E(t)$ in Eq. (1)). This leads to finding the linear transformation W of X, i.e. the inverse of the mixing matrix A, to determine the independent outputs as:

$$u = WX = WAs \qquad (3)$$

where u is the estimated ICs. For this solution to work the assumption is made that the components are statistically independent, while the mixture is not. This is plausible since biological areas are spatially distinct and generate a specific activation; they however correlate in their flow of information [18].

ICA algorithms are suitable for denoising EEG signals because
(i) the signals recorded are the combination of temporal ICs arising from spatially fixed sources and
(ii) the signals tend to be transient (localized in time), restricted to certain ranges of temporal and spatial frequencies (localized in scale) and prominent over certain scalp regions (localized in space) [28].

Wavelet Transform
Wavelet Transform (WT) is a form of time-frequency analysis been used successfully in denoising biomedical signals by decomposing signals in the time-scale space instead of time-frequency space. It is so because it uses a method called wavelet shrinkage proposed by Donoho & Johnstone [9]. Each decomposed signal is called a wavelet (Fig 2).

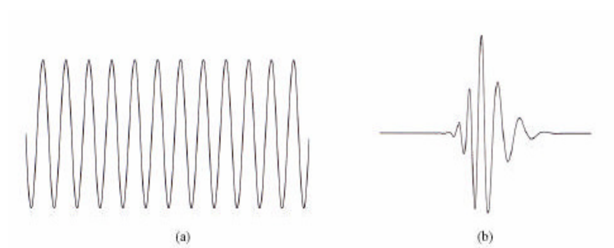


FIGURE 2 : Demonstration of (a) a signal and (b) a wavelet

There are two basic types of WT. One type is designed to be easily reversible (invertible); that means the original signal can be easily recovered after it has been transformed. This kind of WT is used for image compression and cleaning (noise and blur reduction). Typically, the WT of the image is first computed, the wavelet representation is then modified appropriately, and then the WT is reversed (inverted) to obtain a new image.The second type is designed for signal analysis for study of EEG or other biomedical signals. In these cases, a modified form of the original signal is not needed and the WT need not be inverted.

WT decomposes a signal into a set of coefficients called the discrete wavelet transform (DWT) according to:

$$C_{j,k} = \sum_{t \in Z} E(t) g_{j,k}(t) \quad (4)$$

where $C_{j,k}$ is the wavelet coefficient and $g_{j,k}$ is the scaling function defined as:

$$2^{\frac{-j}{2}} g(2^{-j} t - k) \quad (5)$$

The wavelet and scaling functions depend on the chosen wavelet family, such as Haar, Daubechies and Coiflet. Compressed versions of the wavelet function match the high-frequency components, while stretched versions match the low-frequency components. By correlating the original signal with wavelet functions of different sizes, the details of the signal can be obtained at several scales or moments. These correlations with the different wavelet functions can be arranged in a hierarchical scheme called multi-resolution decomposition. The multi-resolution decomposition algorithm separates the signal into "details" at different moments and wavelet coefficients [35-36]. As the moments increase the amplitude of the discrete details become smaller, however the coefficients of the useful signals increase [44-45].

Considering Eq. (1) the wavelet transform of *E(t)* produces wavelet coefficients of the noiseless signal *S(t)* and the coefficients of the noise *N(t).* Researchers found that wavelet denoising is performed by taking the wavelet transform of the noise-corrupted *E(t)* and passing the detail coefficients, of the wavelet transform, through a threshold filter where the details, if small enough, might be omitted without substantially affecting the main signals. There are two main threshold filters – soft and hard. Research has shown that soft-thresholding has better mathematical characteristics [44-45] and provides smoother results [15]. Once discarded these coefficients are replaced with zeroes during reconstruction using an inverse wavelet transform to yield an estimate for the true signal, defined as:

$$\hat{S}(t) = D\left(E(t)]\right) = W^{-1}\left(\Lambda_{th}\left(W\left(E(t)\right)\right)\right) \quad (6)$$

where Λ_{th} is the diagonal thresholding operator that zeroes out wavelet coefficients less than the threshold, *th.* It has been shown that this algorithm offers the advantages of smoothness and adaptation. However, as Coifman and Donoho [7] pointed out, this algorithm exhibits visual artifacts such Gibbs phenomena in the neighbourhoodof discontinuities.

Unscented Kalman Filter

Unscented Kalman Filter (UKF) is a Bayesian filter which uses minimum mean-squared error (MMSE) as the criterion to measure optimality [4][34]. For highly nonlinear systems, the linear estimate of the nonlinear model does not provide a good approximation of the model, and the Extended Kalman Filter (EKF) will not track signals around sharp turning points. Another problem with the EKF is that the estimated covariance matrix tends to underestimate the true covariance matrix and therefore risks becoming inconsistent in the statistical sense without the addition of "stabilising noise". UKF was found to address these flaws. It involves the Unscented Transformation (UT), a method used to calculate the first and second order statistics of the outputs of nonlinear systems with Gaussian. The nonlinear stochastic system used for the algorithm is:

$$\begin{aligned} x_{k+1} &= A x_k + B u_k + v_k \\ y_k &= H x_k + w_k \end{aligned} \quad (7)$$

where *A* and *H* are the known and constant matrices respectively, x_k is the unobserved state of the system, u_k is a known exogenous input, y_k is the observed measurement signal, v_k is the process noise and w_k is the measurement noise.

UKF uses the intuition that it is easier to approximate a probability distribution function rather than to approximate an arbitrary nonlinear function or transformation. Following this intuition, a set of sample points, called sigma points, are generated around the mean, which are thenpropagated through the nonlinear map to get a more accurate estimation of the mean and covariance of the mapping results. In this way, it avoids the need to calculate the Jacobian, which for complex functions can be a difficult task in itself (i.e., requiring complicated derivatives if done analytically or being computationally costly if done numerically).

3. METHODOLOGY

Reasons for Algorithm

Although ICA is popular and for the most part does not result in much data loss; its performance depends on the size of the data set i.e. the number of signals. The larger the set, the higher the probability that the effective number of sources will overcome the number of channels (fixed over time), resulting in an over complete ICA. This algorithm might not be able to separate noise from the signals. Another problem with ICA algorithms has to do with the signals in frequency domain. Although noise has different distinguishing features, once they overlap the EEG signals ICA cannot filter them without discarding the true signals as well. This results in data loss.

WT utilizes the distinguishing features of the noise however. Once wavelet coefficients are created, noise can be identified. Decomposition is done at different levels (L); DWT produces different scale effects (Fig 3). Mallat [2] proved that as scales increase the WT of EEG and noise present different inclination. Noise concentrates on scale 21, decreasing significantly when the scale increases, while EEG concentrates on the 22-25 scales. Elimination of the smaller scales denoise the EEG signals. WT therefore removes any overlapping of noise and EEG signals that ICA cannot filter out.

Recently there has been research comparing the denoising techniques of both ICA and WT. It was found that

(i) if noise and signals are nearly the same or higher amplitude, wavelets had difficultly distinguishing them. ICA, on the other hand, looks at the underlying distributions thus distinguishing each [46] and
(ii) ICA gives high performance when datasets are large. It suffers however from the trade off between a small data set and high performance [20].

Research therefore shows that ICA and wavelets complement each other, removing the limitations of each [35]. Since then research has been done applying a combination of both with ICA as a per- or post- denoising tool. Inuso *et al.* [20] used them where ICA and wavelets are joint. They found that their method outperformed the pre- and post- ICA models.

With or without ICA, conventional wavelet coefficients of 2 signals maybe quite different in many properties as WT is not time invariant, consequently, if the noisy signal is shifted in time, denoised, and then shifted back, the result will, in general, be different from the estimate obtained from denoising without shifting. This result in serious problems such as pseudo-Gibbs phenomenaalternating undershoot and overshoot of a special target level near singularity points of signals [42].

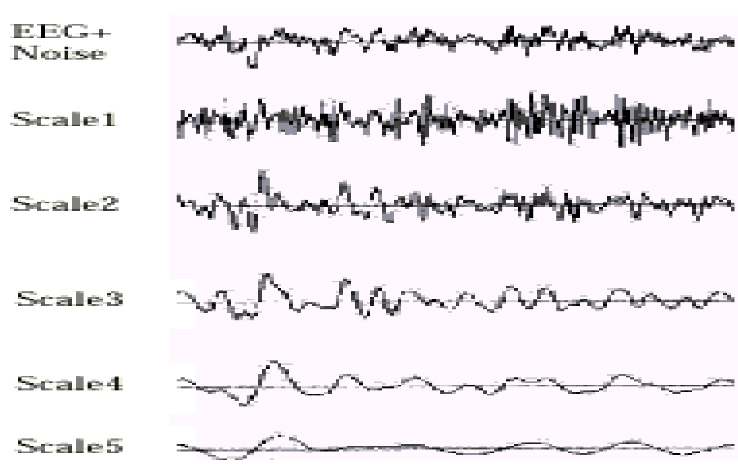


FIGURE 3: Noisy EEG and its Wavelet Transform at different scales

Cycle Spinning (CS) was proposed by Coifman & Donoho [7] as a simple yet efficient method that utilizes the periodic time-invariance of WT in fixing the noise found in wavelet coefficients and defined as:

$$\hat{s} = \frac{1}{k_1 k_2} \sum_{i=1,\, j=1}^{k_1 k_2} S_{-i,-j}\left(T^{-1}\left(\theta\left[T\left(S_{i,j}(x)\right)\right]\right)\right) \quad (8)$$

where (k_1,k_2) are maximum number of shifts, T the shift variant transform, $S_{i,j}$ is the circulant shift, and θ the threshold operator. CS calls for the suppression of these noises by shifting the signals in time and computing the estimate. Using different shifts produce different estimates which are not completely independent; consequently averaging these estimates results in a reduction in the noise generated in each shift. This results in the denoising of all possible unique circularly shifted version of the signal and the creation of the Translation Invariant Wavelet Transform (TIWT) method.

Apart from the use of ICA improvements of WT have been investigated. The idea of Wiener filtering of individual wavelet coefficient arose from the fact that wavelet transforms tend to decorrelate data. An improved wavelet domain denoising technique was therefore proposed that utilizes the Wiener filtering of wavelet coefficients [13]. Research shows that this technique has superior performance over other denoising algorithms using thresholding or shrinkage of wavelet coefficients and has motivated the analysis of many denoising algorithms in terms of optimal filtering of noisy wavelet coefficients. In 2006 the combination of WT and the Kalman Filter (KF) was a novel idea. In the experiments, researchers found that the combination effectively correct overlapped spectra and reduce noise [38]. Mastriani *et al.* [29] created the KalmanShrink for the WT; simulations showed that the threshold had better performance than the most commonly used filters [29]. The use of KF and WT combination therefore improved denoising techniques. Research has also shown that the KF outperforms the Wiener Filter when applied to WT [31]. UKF is advancement on KF.

Each method aims at improving the other in that

(i) WT removes overlapping of noise signals that ICA cannot filter out.
(ii) ICA can distinguish between noise and signals that are nearly the same or higher amplitude which WT has difficultly with.
(iii) WT exhibits serious problems such as pseudo-Gibbs phenomena which CS eliminates and
(iv) Combination of filters and WT effectively correct overlapped spectra

This paper proposes a merger of all four methodologies.

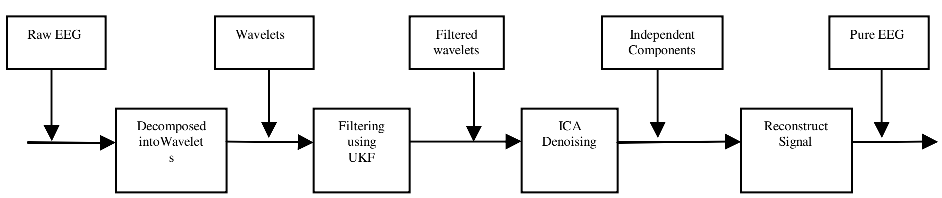


FIGURE 4: Proposed CTICA - Artifacts Removal System

Design

In this paper we are presenting another method to denoising EEG signals using WT and ICA along with smaller methods to improve their performance. Some of the ideas appear in earlier algorithms; however the main differences of CTICA are:

(i)the use of CS, and
(ii) the merger of WT, UKF and ICA into one (this has never been done before).

A block diagram representation of the proposed work is shown in Figure 4. The algorithm can be divided into the following:

1. Signal Collection

This algorithm is designed to denoise both natural and artificially noised EEG signals. They should therefore be mathematically defined based on Eq. (1).

2. Apply CS to signal

The number of time shifts is determined; in so doing signals are forcibly shifted so that their features change positions removing the undesirable oscillations which result in pseudo-Gibbs phenomena. The circulant shift by h is defined as:

$$S_h\left(f(n)\right) = f\left((n+h) \bmod N\right) \quad (9)$$

where *f(n)* is the signal, *S* is time shift operator and *N* is the number of signals. The time-shift operator S is unitary and therefore invertible i.e. $(S_h)^{-1} = S_{-h}$

3. Decomposition of Signal

The signals are decomposed into 5 levels of DWT using the Symmlet family, separating noise and true signals. Symmlets are orthogonal and its regularity increases with the increase in the number of moments [11]. After experiments the number of vanishing moments chosen is 8 (Sym8).

4. Filter Coefficients

Perform UKF on the coefficients to filter out some noise reducing the shrinkage threshold.

5. Choose and Apply Threshold Value

Denoise using the soft-thresholding method discarding all coefficients below the threshold value using VisuShrink based on the universal threshold defined by Donoho & Johnstone [9] given as:

$$T = \sqrt{2\sigma^2 \log N} \quad (10)$$

where N is the number of samples and σ^2 is the noise power.

6. Apply ICA algorithm

Signals and noise may have nearly the same frequency characteristics and overlap in time thus producing noisy coefficients such as beta activity and muscle noise, that WT has not been able to distinguish and remove. ICA is able to look at the underlying distributions thus distinguish noise and remove them. Research has shown that ICA is a robust denoising method where its performance is not affected by the severity of the mixing signals [10]. We implemented a symmetrical fixed-point ICA algorithm based on the Hyvarinen model [19] where the gradient function is:

$$g(y) = \tanh(a, y) \quad (11)$$

A fixed-point algorithm has a cubic or at least a quadratic convergence, is not linear and no parameters have to be chosen for usage which makes it a better choice than other ICA models.

7. Reconstruction of Signals

EEG signals are reconstructed using inverse DWT.

8. Apply CS

Revert signals to their original time shift and average the results obtained to produce the denoised EEG signals.

The proposed algorithm can be expressed as Avg [Shift – Denoise -Unshift] i.e. using Eq. (9) it is defined as:

$$avg_{h \in H}\left(S_{-h} T S_{h}(f)\right) \qquad (12)$$

where H is the range of shifts, T is the wavelet shrinkage denoising operator, h the circular shift and the maximum of H is the length of the signal N from Eq. (9).

Evaluation
There are different means to access the separation quality performed by ICA methods; however the performance measures used throughout this section will be:

(i) the Mean Square Error (MSE),
(ii) the Peak Signal to Noise Ratio (PSNR),
(iii) the Signal to Distortion Ratio (SDR), and
(iv) the Amari Performance Index

We employed fixed point benchmark ICAs with the linearity *g(u) = tanh* and the symmetric orthogonalization for comparison, namely:fixed-point - FastICA[19], EFICA[28] and Pearson-ICA [26]

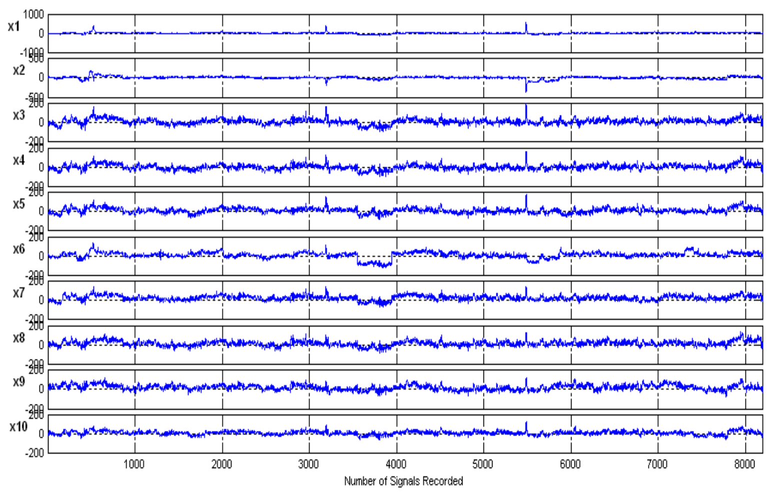


FIGURE 5 : Sample EEG signals from a male subject

4. EXPERIMENTAL DATA

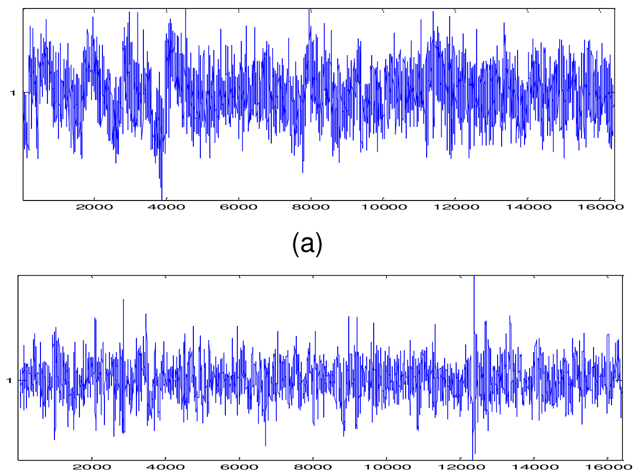
In order to do the study effectively data was collected for analysis from two sites.

(i) http://www.filewatcher.com/b/ftp/ftp.ieee.org/uploads/press/rangayyan.0.0.html [48]. Data was collected at a sampling rate of 100Hz but noise free. These signals were artificially contaminated.

(ii) http://sccn.ucsd.edu/~arno/fam2data/publicly_available_EEG_data.html [47]. All data are real comprised of EEG signals from both human and animals. Data were of different types.
   (a) Data set acquired is a collection of 32-channel data from one male subject who performed a visual task. Fig. 5 shows 10 signals from this dataset.
   (b) Human data based on five disabled and four healthy subjects. The disabled subjects (1-5) were all wheelchair-bound but had varying communication and limb muscle control abilities. The four healthy subjects (6-9) were all male PhD students, ages 30 who had no known neurological deficits. Signals were recorded at 2048 Hz sampling rate from 32 electrodes placed at the standard positions of the 10-20 international system.

(c) Data set is a collection of 32-channel data from 14 subjects (7 males, 7 females) who performed a go-nogo categorization task and a go-no recognition task on natural photographs presented very briefly (20 ms). Each subject responded to a total of 2500 trials. The data is CZ referenced and is sampled at 1000 Hz.
(d) Five data sets containing quasi-stationary, noise-free EEG signals both in normal and epileptic subjects. Each data set contains 100 single channel EEG segments of 23.6 sec duration.

These two sites produce real signals of different sizes as well as 1D and 2D signals. A total of 1,383 signals were tested.



(b)

FIGURE 6 : (a) EEG Signal with EOG (b) Denoised EEG Signal

5. RESULTS & DISCUSSION

We conducted experiments, using the above mentioned signals, in Matlab 7.8.0 (R2009) on a laptop with AMD Athlon 64x2 Dual-core Processor 1.80GHz. Noisy signals were generated by adding noise to the original noise-free signals and the length of all signals, *N*, were truncated to lengths of power of twos i.e. 2^x.

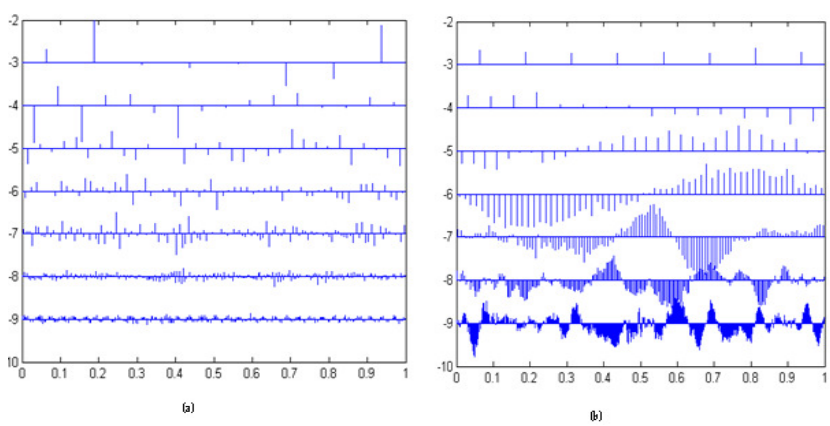


FIGURE 7: (a) Wave Coefficient before denoising
(b) Wave Coefficient after denoising

Figure 6 shows the results of the above algorithm on one EEG signal contaminated with EOG. Investigations on the wavelet coefficients (Figure 7) also show that there are major changes in the wavelets - some wavelets have been zeroed because of their identification to noise.

CTICA	FASTICA	EFICA	PEARSON
1411.00	1489.70	1412.40	1337.20
950.09	1014.80	951.75	890.67
714.51	769.91	715.31	662.73
693.71	749.04	694.70	642.40
757.94	815.16	758.81	704.47
858.63	920.96	859.34	799.85
849.22	923.80	849.61	778.53
788.22	850.93	788.77	728.96
875.44	941.76	876.23	813.07
369.86	410.65	370.30	332.04
297.61	333.63	298.10	264.57
526.54	574.16	527.08	482.01
636.14	696.30	636.44	579.28
1031.70	1116.70	1032.00	950.71
236.44	269.07	236.93	206.84
383.87	424.57	384.31	346.05
650.56	598.48	651.06	598.48
694.03	752.93	694.55	638.52
577.78	646.60	577.90	513.06
449.89	504.49	450.01	398.64

TABLE 1: MSE for 15 EEG signals with EOG noise

Noise/Signal Measures

The MSE measures the average of the square of the "error" which is the amount by which the estimator differs from the quantity to be estimated. Mathematically it is defined as:

$$MSE = \frac{1}{MN}\sum_{y=1}^{M}\sum_{x=1}^{N}[I(x,y) - I'(x,y)]^2 \quad (13)$$

The error is the amount by which the estimator differs from the quantity to be estimated. The difference occurs because of randomness or because the estimator doesn't account for information that could produce a more accurate estimate. . For a perfect fit, *I(x,y) = I'(x,y)* and MSE = 0; so, the MSE index ranges from 0 to infinity, with 0 corresponding to the ideal. The smaller the
MSE therefore the closer the estimator is to the actual data and the less the error on the signal; CTICA was compared in both Table 1 and Table 2. Examination shows that on average our method had the second lowest MSE next to Pearson-ICA.

PSNR is the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Mathematically it is defined as:

$$PSNR = 10 \times \log_{10}\left(\frac{MAX^2}{MSE}\right) \quad (14)$$

CTICA	FASTICA	EFICA	PEARSON
26351.00	26678.00	26352.00	26028.00
12824.00	13051.00	12823.00	12597.00
6449.30	6610.00	6447.80	6287.70
5493.40	5642.20	5492.60	5345.20
6221.40	6379.90	6220.60	6063.50
106.51	87.34	107.03	87.35
2481.90	2583.60	2482.60	2383.80
6457.50	6618.90	6456.80	6296.90
6451.60	6612.90	6450.70	6290.70
12811.00	13038.00	12810.00	12584.00
5508.10	5657.20	5507.50	5359.80
16004.00	16265.00	16008.00	15755.00
3839.70	3968.90	3842.10	3717.90
1999.40	20921.00	2000.20	1910.70
451.69	494.80	451.31	409.98
470.52	517.86	471.58	427.64
13454.00	13694.00	13458.00	13255.00
13089.00	13324.00	13092.00	12863.00
3850.60	3979.80	3853.00	3728.60

TABLE 2: MSE for 19 EEG signals with artificially added noise

Because many signals have a very wide dynamic range, PSNR is usually expressed in terms of the logarithmic decibel scale.In this research MAX takes the value of 255. Unlike MSE which represents the cumulative squared error between the denoised and mixed signal, PSNR represents a measure of the peak error i.e. when the two signals are identical the MSE will be equal to zero, resulting in an infinite PSNR.The higher the PSNR therefore, the better the quality of the reconstructed signal i.e. a higher PSNR indicates that the reconstruction is of a higher qualityand therefore the algorithm is considered good. Table 3 shows the PSNR for EOG contaminated signals and Table 4 shows those with artificially contaminated noise.

Examination of Table 3 and Table 4 show that Pearson-ICA is the algorithm that has the highest PSNR with CTICA been second. It can also be seen in Table 4 that CTICA and EFICA both have similar to MSE when signals have artificial noise added for 83%. The other 17% CTICA performed better. CTICA however in both cases presents more signal than noise in its denoised results than FastICA. CTICA is the second best in performance therefore because it outperformed FastICA and it never produces a PSNR lower than EFICA; in fact it sometimes performed better than EFICA. Algorithms therefore follow the same behavior as seen with the MSE investigations.

CTICA	FASTICA	EFICA	PEARSON
16.63	16.40	16.63	16.87
18.35	18.07	18.35	18.63
19.59	19.27	19.59	19.92
19.72	19.39	19.71	20.05
19.33	19.02	19.33	19.65
18.79	18.49	18.79	19.10
18.84	18.48	18.84	19.22
19.16	18.83	19.16	19.50
18.71	18.39	18.70	19.03
22.45	22.00	22.45	22.92
23.39	22.90	23.39	23.91
20.92	20.54	20.91	21.30
20.10	19.70	20.09	20.50
18.00	17.65	17.99	18.35
24.39	23.83	24.38	24.97
22.29	21.85	22.28	22.74
20.00	20.36	19.99	20.36
19.72	19.36	19.71	20.08
20.51	20.02	20.51	21.03
21.60	21.10	21.60	22.12

TABLE 3 : PSNR for 20 EEG signals with EOG noise

Separation Accuracy Measures

How accurate the separation of an ICA algorithm in terms of the signals can be calculated by the total SDR which is defined as:

$$SDR(x_i, y_i) = \frac{\sum_{n=1}^{L} x_i(n)^2}{\sum_{n=1}^{L} \left(y_i(n) - x_i(n) \right)^2} \quad i = 1, \ldots m \tag{15}$$

CTICA	FASTICA	EFICA	PEARSON
3.92	3.87	3.92	3.98
7.05	6.97	7.05	7.13
10.04	9.93	10.04	10.15
10.73	10.62	10.73	10.85
10.19	10.08	10.19	10.30
27.86	28.72	27.84	28.72
14.18	14.01	14.18	14.36
10.03	9.92	10.03	10.14
10.03	9.93	10.03	10.14
10.72	10.60	10.72	10.84
6.09	6.02	6.09	6.16
12.29	12.14	12.29	12.43
15.12	14.93	15.12	15.32
21.41	20.99	21.40	21.82
6.84	6.77	6.84	6.92
6.96	6.88	6.96	7.04
12.28	12.13	12.27	12.42
11.81	11.72	11.81	11.99

TABLE 4 : PSNR for 18 EEG signals with artificially added noise

where $x_i(n)$ is the original source signal and $y_i(n)$ is the reconstructed signal. When the SDR is calculated if it is found to be below 8-10dB the algorithm is considered to have failed separation. Fig. 8 shows all four algorithms having SDR above 8dB; there is not much differentiation in the graph for the algorithms however. Where there were difference in the SDR calculations CTICA had the most consistent. Table 5 showed that CTICA produced the largest SDR on average. This shows that on average CTICA had best separation of signal from noise than the other algorithms.

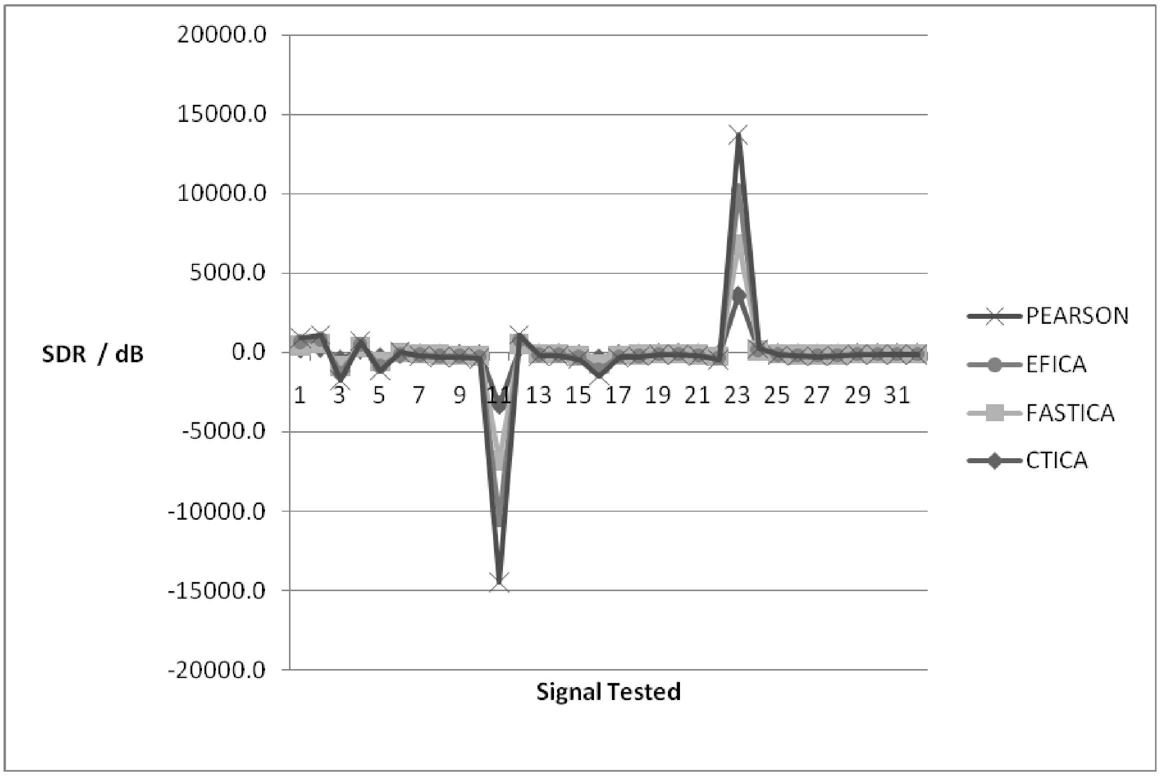


FIGURE 8 : SDR for 32 signals with EOG

The most widely used measure for assessing the accuracy of the estimated mixing matrix is the Amari performance index defined as:

$$P_{err} = \frac{1}{2m} \sum_{i,j=1}^{m} \left(\frac{|p_{ij}|}{\max_k |p_{ik}|} + \frac{|p_{ij}|}{\max_k |p_{kj}|} \right) - 1 \quad (16)$$

where $p_{ij} = (BA)_{ij}$. It assesses the quality of the de-mixing matrix *W* for separating observations generated by the mixing matrix *A***.** When the separation is perfect, the Amari index is equal to zero. In the worst case, i.e. when the estimated sources contain the same proportion of each original source signal, the Amari index is equal to *m*/2–1. However, the most common situation is between both. The lower the Amari index therefore, the more accurate the separation is. The Amari indexes obtained for the different algorithms and for different sample sizes are presented in Figure 9.

CTICA	FASTICA	EFICA	PEARSON
10089.00	9916.50	9977.60	10039.00
1900.10	1867.70	1884.20	1901.00
930.91	911.92	923.32	935.00
1141.00	1112.70	1127.00	1143.10
966.00	945.51	957.53	969.86
600.63	503.94	455.22	503.89
-663.61	-660.57	-673.86	-687.69
1266.60	1237.20	1252.60	1268.40
961.88	942.02	965.84	965.84
1371.90	1351.60	1363.60	1375.80
1074.20	1048.20	1062.40	1076.90
901.30	888.19	895.28	902.46
371.50	361.34	367.26	373.35
258.60	249.01	254.67	260.57
-2441.10	-6579.90	-6889.70	-7228.60
80.17	74.99	78.58	82.52
771.02	759.16	765.79	772.50
848.85	834.97	842.33	849.80
384.01	373.40	379.50	385.77
1095.42	849.36	841.53	836.29

TABLE 5: SDR for 20 with artificially added noise

Observations show that the Amari indexes for our method is lower for sample sizes greater than 2^7 i.e. it clearly outperforms the other algorithms with sample size greater than 128. Figure 9 also shows that unlike the other algorithms, the Amari index for CTICA is inversely proportional to sample size.

6. CONCLUSION & FUTURE WORK

In recent years researchers have used both ICA algorithms and WT to denoise EEG signals. In this paper we propose a new method – Cycle Spinning Wavelet Transform ICA (CTICA). From the experiments we can conclude the following for CTICA

(i) It has outperformed FastICA and EFICA as far as MSE and PSNR were concerned.
(ii) It has the best SDR and
(iii) It has the best Amari index for signals greater than 2^7 in size which decreases as sample size increases

Based on these results it can be concluded that CTICA is the most accurate separation method and is the second best at signal/noise measurement.

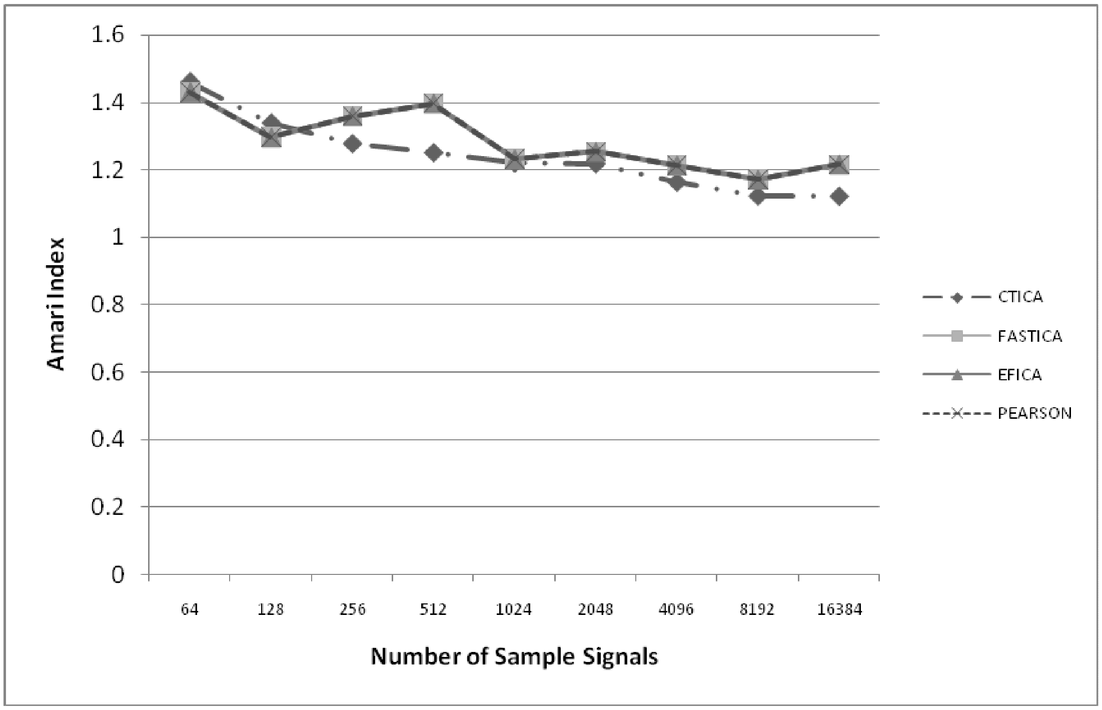


FIGURE 9: Amari Results for the four algorithms

7. REFERENCES

[1] M. Akin, "Comparison of Wavelet Transform and FFT Methods in the Analysis of EEG Signals", Journal of Medical Systems 26(3), 241-247, 2002.

[2] M. Alfaouri and K. Daqrouq, "ECGSignal Denoising By Wavelet Transform Thresholding", American Journal of Applied Sciences 5 (3), 276-281, 2008.

[3] M.I. Bhatti, A. Pervaiz and M.H. Baig, "EEG Signal Decomposition and Improved Spectral Analysis Using Wavelet Transform", In Proceedings of the 23rd Engineering in Medicine and Biology Society 2, 2001, 1862-1864.

[4] Z. Chen, "Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond", Adaptive Systems Lab., McMaster University., Hamilton, Ontario, Canada, 2003, Retrieved June 20, 2009, from: http://users.isr.ist.utl.pt/~jpg/tfc0607/chen_bayesian.pdf.

[5] J. Chien; H. Hsin-Lung and S. Furui, "A new mutual information measure for independent component analysis", In the Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2008 (ICASSP 2008), 2008, 1817 – 1820

[6] S. Choi, A. Cichocki, L. Zhang, and S. Amari, "Approximate Maximum Likelihood Source Separation Using the Natural Gradient",In the Proceedings of theIEICE Transaction Fundamental E84A(12), 2002.

[7] R.R. Coifman, and D.L. Donoho, "Translation Invariant Denoising",Lecture Notes in Statistics: Wavelets and Statistics, 125-150, 1995.

[8] P. Comon, "Independent Component Analysis, a new concept?", Signal Processing, Elsevier, 36(3), 287-314, 1994.

[9] D.L. Donoho and I. Johnstone, “Adapting to unknown smoothness via wavelet shrinkage”, Journal of American Statistical Association, 90,1200—1224, 1995.

[10] U.E. Emuir, A. Akin, A. Ertuzun, B. Sankur and K. Harmanci, “Wavelet Denoising vs ICA Denoising for Functional Optical Imaging”, In the Proceedings of the 1st International IEEE EMBS Conference on Neural Engineering, 2003, 384-387.

[11] B. Ferguson, D. Abbott, “Denoising Techniques for Terahertz Response of Biological Samples”, Microelectronics Journal 32, 943-953, 2001.

[12] N. Gadhok, and W. Kinsner, “Robust ICA for Cognitive Informatics”. International Journal of Cognitive Informatics and Natural Intelligence (IJCINI) 2(4),86-92, 2008.

[13] S. P. Ghael, A. M. Sayeed and R. G. Baraniuk, “Improved Wavelet Denoising via Empirical Wiener Filtering”, In the Proceedings of theSPIEvol. 3169, 389-399, 1997.

[14] A. Graps, An Introduction to Wavelets. IEEE Journal of Computational Science and Engineering 2(2),1-17, 1995.

[15] Y.M. Hawwar, A.M. Reza, and R.D. Turney, Filtering(Denoising) in the Wavelet Transform Domain,, Unpublished, Department of Electrical Engineering And Computer Science, University of Wisconsin-Milwaukee, 2002.

[16] C.S. Herrmann, M. Grigutsch and N.A. Busch, (2005). EEG oscillations and wavelet analysis. Event-related potentials: A methods handbook,MIT Press, 229-259, 2005.

[17] G.G. Herrero, and K. Egiazarian, “Independent Component Analysis by a Resampling Strategy”, Technical Report 2005, Retrieved September 18, 2009, fromhttp://www.cs.tut.fi/~gomezher/projects/bss/rica/rica.pdf

[18] S. Hoffman, and M. Falkenstien , “The Correction of Eye Blink Artefacts in the EEG: A Comparison of a Two Prominent Methods”, PLoS One 3(8):e3004, 2008

[19] A. Hyvärinen and E. Oja, “A Fast Fixed-Point Algorithm for Independent Component Analysis”, Neural Computation,9(7), 1483-1492, 1997

[20] A. Hyvarinen, J. Karhunen and E. Oja , Independent Component Analysis, eds. Wiley & Sons, 2001

[21] G. Inuso, F. La Foresta, N. Mammone, and F.C. Morabito, “Wavelet-ICA methodology for efficient artifact removal from Electroencephalographic recordings”, In the Proceedings of the International Conference on Neural Networks,1524-1529, 2007.

[22] N. Jacob, and A. Martin, “Image Denoising in the Wavelet Domain Using Wiener Filtering”, Unpublished course project, University of Wisconsin, Madison, Wisconsin, USA, 2004.

[23] S. Julier, and J.K. Uhlmann, “Unscented Filtering and Nonlinear Estimation”,Proceedings of the. IEEE 92(3), 401-421, 2004.

[24] S. Julier, and J.K. Uhlmann, “A New Extension of the Kalman Filter to Nonlinear Systems” In the Proceedingd of AeroSense: 11th Int. Symp. Aerospace/Defense Sensing, Simulation and Controls,182-193, 1997.

[25] A. Kallapur, S. Anavatti, and M. Garratt, “Extended and Unscented Kalman Filters for Attitude Estimation of an Unmanned Aerial Vehicle,In the Proceedings of the 27th IASTED Int. Conf. Modelling, Identification, and Control (MIC 2008) , 2008

[26] J. Karvanen, J.,Eriksson, and K.V. Pearson, “System Based Method for Blind Separation", In the Proceedings of Second International Workshop on Independent Component Analysis and Blind Signal Separation, Helsinki 2000, 585—590, 2000.

[27] L. Kaur, S. Gupta, and R.C. Chauhan, “Image Denoising using Wavelet Thresholding”, In the Proceedings of the 3rd Indian Conf. Computer Vision, Graphics & Image Processing (ICVGIP 2002), 22(14), 2002.

[28] Z. Koldovský and P. Tichavský, P., "Time-Domain Blind Audio Source Separation Using Advanced ICA Methods", In the Proceedings of the 8th Annual Conference of the International Speech Communication Association (Interspeech 2007), pp. 846-849, 2007.

[29] V. Krishnaveni, S. Jayaraman, A. Gunasekaran, and K. Ramadoss, Automatic Removal of Ocular Artifacts using JADE Algorithm and Neural Network, International Journal of Intelligent Systems and Technologies 1(4), 322-333, 2006.

[30] V. Krishnaveni, S. Jayaraman, S. Aravind, V. Hariharasudhan, and K. Ramadoss, “Automatic Identification and Removal of Ocular Artifacts from EEG using Wavelet Transform”,Measurement Science Review 6(2, 4), 45-57, 2006.

[31] S. Makeig, J. Anthony, A. J. Bell, T. Jung, and T.J. Sejnowski,. “Independent Component Analysis of Electroencephalographic data”, Advances in Neural Information Processing Systems 8, 1996.

[32] M. Mastriani, and A.E. Giraldez, “Kalman’s Shrinkage for Wavelet-Based Despeckling of SAR Images”, International Journal of Intelligent Systems and Technologies 1(3),190-196, 2006.

[33] N. Nikolaev, A. Gotchev, “ECG Signal Denoising using Wavelet Domain Wiener Filtering” In the Proceedings of the 10thEuropean Signal Processing Conference (EUSIPCO 2000), 2000

[34] S. Postalcioglu, K. Erkan, E.D. Bolat, “Comparison of Kalman Filter and Wavelet Filter for Denoising”, In the Proceedings of the International Conference onNeural Networks and Brain 2005 Vol. 2, 951 – 954, 13-15 Oct. 2005

[35] V.V.K.D.V. Prasad, P. Siddaiah, and B. Prabhaksrs Rao, “A New Wavelet Based Method for Denoising of Biological Signals”, International Journal of Computer Science and Network Security 8(1), 2008, 238-244, 2008.

[36] R. Romo-Vazquez, R., Ranta, V. Louis-Dorr, and D. Maquin, “Ocular Artifacts Removal in Scalp EEG: Combining ICA and Wavelet Denoising”, In the Proceedings of Physics in Signal and Image Processing (PSISP 07), 2007

[37] R. Sameni, M.B. Shamsollahi, and C. Jutten, “Filtering Electrocardiogram Signals using the Extended Kalman Filter”, In the Proceedings of the 27th IEEE Engineering in Medicine and Biology (EMBS) Annual Conference, 5639-5642, 2005.

[38] P. Senthil Kumar, R. Arumuganathan, K. Sivakumar, and C. Vimal, "A Wavelet based Statistical Method for De-noising of Ocular Artifacts in EEG Signals",IJCSNS International Journal of Computer Science and Network Security. 8(9), 87-92, 2008.

[39] P. Senthil 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 in Computer Science & Mathematics 1(3). 2008

[40] P. Shui, and Y. Zhao, Image Denoising Algorithm using Doubling Local Wiener Filtering with Block Adaptive Windows in Wavelet Domain, Signal Processing 87(7), 1721-1734, 2007.

[41] L. SuWen, L, WenQing, X. PinHua, Z. YuJui, "Application of Kalman Filtering and Wavelet Transform in DOAS", In the Proceedings of the 2006 IEEE International Conference on Information Acquisition, 748-753, 2006.

[42] M. Unser, and A. Aldroubi, "A Review of Wavelets in Biomedical Applications", In the Proceedings of the IEEE 84(4), 626-638, 1996.