

# Image Denoising by Data Adaptive and Non-Data Adaptive Transform Domain Denoising Method Using EEG Signal

Vandana Roy and Shailja Shukla

**Abstract** This chapter proposes an automatic method for artifact removal and noise elimination from scalp electroencephalogram recordings (EEG). The method is based on transform domain method having combination of data adaptive and non-data adaptive transform domain image denoising method to improve artifact elimination (ocular, high frequency muscle, and electrocardiogram (ECG) artifacts). The elimination of artifact from scalp EEGs is of substantial significance for both the automated and visual examination of underlying brainwave actions. These noise sources increase the difficulty in analyzing the EEG and obtaining clinical information related to pathology. Hence it is crucial to design a procedure to decrease such artifacts in EEG records. The role of a data adaptive transform domain, i.e., ICA to separate the signal from multichannel sources, then non-data adaptive transform, i.e., wavelet is applied to denoise the signal. The proposed methodology successfully rejected a good percentage of artifacts and noise, while preserving almost all the cerebral activity. The “denoised artifact-free” EEG presents a very good improvement compared with recorded raw EEG.

**Keywords** Artifact removal • Electroencephalogram (EEG) • Wavelet denoising • Independent component analysis (ICA)

---

V. Roy (✉)

Department of Electronics and Communication, Jabalpur Engineering College, Jabalpur, Madhya Pradesh, India  
e-mail: vandana.roy20@gmail.com

S. Shukla

Department of Computer Science Engineering, Jabalpur Engineering College, Jabalpur, Madhya Pradesh, India

## Introduction

The electrical activity produced by the brain is recorded by the electroencephalogram (EEG) using several electrodes placed on the scalp. Signal characteristics vary from one state to another, such as wakefulness/sleep or normal/pathological. EEG is the multivariate time series data measured using multiple sensors positioned on the scalp that imitates electrical potential produced by behaviors of brain and is a record of the electrical potentials created by the cerebral cortex nerve cells. There are two categories of EEG based on where the signal is obtained in the head: scalp or intracranial. Scalp EEG being the main focus of the research, uses small metal discs, also called electrodes, which are kept on the scalp with good mechanical and electrical touch. Intracranial EEG is obtained by special electrodes placed in the brain during a surgery. The electrodes should be of low impedance in order to record the exact voltage of the brain neuron. The variations in the voltage difference among electrodes are sensed and amplified before being transmitted to a computer program [1]. Classically, five major brain waves can be distinguished by their frequency ranges: delta ( $\delta$ ) 0.5–4 Hz, theta ( $\theta$ ) 4–8 Hz, alpha ( $\alpha$ ) 8–13 Hz, beta ( $\beta$ ) 13–30 Hz and gamma ( $\gamma$ ) 30–128 Hz. The informative cortically generated signals are contaminated by extra-cerebral artifact sources: ocular movements, eye blinks, electrocardiogram (ECG), muscular artifacts. Generally, the mixture between brain signals and artifactual signals is present in all sensors, although not necessarily in the same proportions (depending on the spatial distribution). Moreover, the EEG recordings are also affected by other unknown basically random signals (instrumentation noise, other physiological generators, external electromagnetic activity, etc.) which can be modeled as additive random noise. These phenomena make difficult the analysis and interpretation of the EEGs, and a first important processing step would be the elimination of the artifacts and noise. Several methods for artifact elimination were proposed. Most of them consist of two main steps: artifact extraction from the multichannel recorded signals, generally using some signal separation methods, followed by signal classification. Our goal is to contribute to EEG artifact rejection by proposing an original and more complete automatic methodology consisting of an optimized combination of several signal processing and data analysis techniques [2].

This chapter is organized as follows: [Supporting Literature](#) briefs the supporting literature, [Data Adaptive Transform Domain Image Denoising Method: ICA](#) states the data adaptive transform domain method to separate the signals from multichannel sources, then [Non Data Adaptive Transform Domain Based Denoising \(Wavelet denoising\)](#) gives details of non-data adaptive transform domain method to denoise the signal to remove artifacts. This method assumes that EEG contains two classes namely, artifact and non-artifact signal, and then it calculates the optimum threshold separating these two classes. [Proposed Method](#) is dedicated to our present approach to denoise the signal and [Experimental Results](#) presents the main results in Sect. 6.

## 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 currents (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 undergo 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 are called EEG signals. Understanding the brain is a huge part of Neuroscience, and the development of EEG was for the explanation of such a phenomenon. The morphology of 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 conforming a diagnosis of epilepsy.
- Check for problem with loss of consciousness or dementia.
- Help find out a person's chance of recovery after change in consciousness.
- Find out whether a person is in coma or is brain dead.
- Study sleep disorder such as narcolepsy.
- What brain activity occurs while a person is receiving general anesthesia during brain surgery.
- Help find out whether 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 of 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:

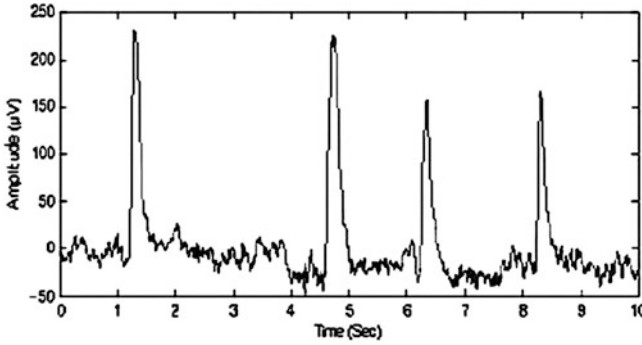
- Eye movement and blinking—Electrooculogram (EOG)
- Cardiac movement—Cardiograph (ECG/EKG)
- Muscle movement—Electromyogram (EMG)
- Chewing and sucking movement—Glassokinetic
- The power lines.

EEG recordings are therefore a combination of these signals called artifacts or noise and the pure EEG signal defined mathematically as in Eq. (1):

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

where,

$S$  is pure EEG signal,



**Fig. 1** EEG contaminated with EOG producing spikes

$N$  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 overlying these signals resulting in signal distortion (Fig. 1). Correct analysis is therefore impossible, resulting in misdiagnosis in the case of some patients. Noise must be eliminated or attenuated.

The method of cancellation of the contaminated segments, although practice can lead to considerable information loss, thus other methods such as principal component analysis (PCA) and more recently ICA and WT have been utilized [3].

## Data Adaptive Transform Domain Image Denoising Method: ICA

**Definitions of ICA:** We can define the ICA as it is a random vector  $X$  which consists of finding a linear transform as in Eq. (2):

$$X = AS \quad (2)$$

so that the components  $s_i$  are as independent as possible w.r.t. some maximum function that measures independence. This definition is known as a general definition where no assumptions on the data are made. Independent component analysis (ICA) is the decomposition of a random vector in linear components which are “as independent as possible”. Here, ‘independence’ should be understood in its strong statistical sense: it goes beyond second order decorrelation and thus involves the non-gaussianity of the data. The ideal measure of independence is the higher order cumulants like kurtosis and mutual information.

In addition to the basic assumption of statistically independence, by imposing the following fundamental restrictions, the noise free ICA model can be defined if:

1. All the independent components  $S_i$ , with the possible exception of one component, must be non-Gaussian
2. The number of observed linear mixtures  $m$  must be at least as large as the number of independent components  $n$ , i.e.,  $m > n$
3. The matrix  $A$  must be of full column rank.

We can invert the mixing matrix as in Eq. (3):

$$S = A^{-1} X \quad (3)$$

Thus, to estimate one of the independent components, we can consider a linear combination of  $X_i$ . Let us denote this by Eq. (4):

$$Y = b^T X = b^T A S \quad (4)$$

Hence, if  $b$  were one of the rows of  $A^{-1}$ , this linear combination  $b^T X$  would actually equal one of the independent components. But in practice we cannot determine such ' $b$ ' exactly because we have no knowledge of matrix  $A$ , but we can find an estimator that gives a good approximation. In practice there are two different measures of non-Gaussianity.

## ***Kurtosis***

The classical measure of non-Gaussianity is kurtosis or the fourth order cumulant. It is defined by Eq. (5)

$$\text{Kurt}(y) = E\{y^4\} - 3(E\{y^2\})^2 \quad (5)$$

As the variable  $y$  is assumed to be standardized we can say in Eq. (6) as:

$$\text{Kurt}(y) = E\{y^4\} - 3 \quad (6)$$

Hence the kurtosis is simply a normalized version of the fourth moment  $E\{y^4\}$ . For the Gaussian case the fourth moment is equal to 3 and hence  $\text{kurt}(y) = 0$ . Thus, for Gaussian variable kurtosis is zero but for non-Gaussian random variable it is nonzero [4, 5].

## Non-Data Adaptive Transform Domain-Based Denoising (Wavelet Denoising)

We know that Fast ICA are expected to correspond to artifacts only, on the other hand, some brain action might escape to these gathered signals. The purpose of conventional filtering is to process raw EEG data  $x(t)$  to eliminate 50 Hz line noise, baseline values, artifacts inhabiting very low frequencies and high frequency sensor noise  $v(t)$ , and this phase may include mixture of different existing notch, lowpass, and/or highpass filters. As artifacts have a frequency overlap with the brain signals, the conventional filtering technique cannot be utilized, and therefore this paper focuses on using Wavelet Denoising to explore brain activity from gathered independent components [1].

Image signal and noise signal by the wavelet transform have different characteristics:

1. In the wavelet transform, the noise energy reduces rapidly as scale increases, but the image signal does not reduce rapidly.
2. Noise is not highly relevant at different scales of the wavelet transform. But the wavelet transform of image signal generally has a strong correlation, the scale of the adjacent local maxima almost appear in the same position and have the same symbol.

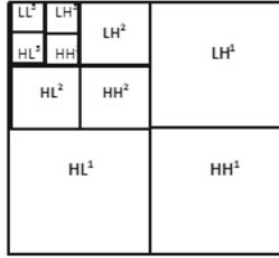
The two above-mentioned points will separate image signal and noise signal, that is to say they are the base of image denoising [6].

### *Wavelet Domain-Based Denoising Algorithm*

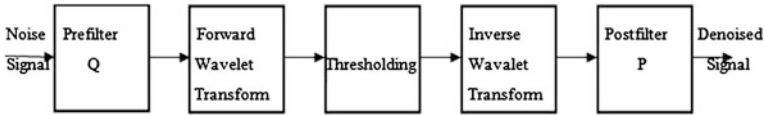
An image is often corrupted by noise in its acquisition or transmission. Wavelet provides an appropriate basis for separating noisy signal from image signal. The motivation is that as the wavelet transform is good at energy compaction, small coefficient is more likely due to noise and large coefficient due to important signal features. These small coefficients can be threshold without affecting the significant features of the images.

The problem that arises is how to find an optimal threshold such that the mean squared error between the signal and its estimate is minimized. The wavelet decomposition of an image is done as the image is split into 4 subbands, namely the HH, HL, LH, and LL subbands as shown in Fig. 2. The HH subband gives the diagonal details of the image; the HL subband gives the horizontal features while the LH subband represents the vertical structures. The LL subband is the low resolution residual consisting of low frequency components and it is this subband which is further split at a higher level of decomposition as shown in Fig. 2 [7].

The low pass filters represent the “approximation” of the signal or its dc component and the high pass filters represent the “details” or its high frequency



**Fig. 2** Wavelet decomposition: 1, 2, 3 decomposition levels, *H* high frequency bands, *L* low frequency bands



**Fig. 3** Denoising by wavelet domain

components. The successive analysis of the low pass component only is called wavelet decomposition, (Fig. 1b), whereas the analysis of both the low and high pass components is called wavelet packet decomposition; the existence of small coefficients is more likely to be due to the noise contamination, whereas the large coefficients contain significant image details. Hence, the small magnitude coefficients may be thresholded without affecting the large ones and therefore the quality of the image [8].

The investigations show that the method for denoising differs only in the selection of the wavelets and their decomposition levels [6].

The algorithm has the following steps:

1. Calculate the DWT of the image.
2. Threshold the wavelet coefficients (Threshold may be universal or subband adaptive).
3. Compute the IDWT to get the denoised estimate.

Wavelet transform of noisy signal should be taken first and then thresholding function is applied on it. Finally the output should be undergone inverse wavelet transformation to obtain the estimate  $x$  as shown in Fig. 3.

The DWT of any signal sample is given by Eq. (7)

$$S_{\text{DWT}}(j, k) = \sum_{n=0}^{N-1} S_{nj,k} w_n, \quad a = 2^j, \tau = k2^j \quad (7)$$

$S_n$	s(nT) Signal samples
$j, k$	$n$ th sample of $k$ th sifted version of a $2^j$ scaled discrete wavelet
$N$	number of signal samples.

There are four thresholds frequently used, i.e., hard threshold, soft threshold, semi-soft threshold, and semi-hard threshold. The hard-thresholding function keeps the input if it is larger than the threshold, otherwise, it is set to zero. It is described as in Eq. (8)

$$\begin{aligned} f_h(x) &= x \quad \text{if } x \geq \lambda \\ &= 0 \quad \text{otherwise} \end{aligned} \quad (8)$$

The hard-thresholding function chooses all wavelet coefficients that are greater than the given threshold  $\lambda$  and sets the others to zero. The threshold  $\lambda$  is chosen according to the signal energy and the noise variance  $2\sigma$ . If a wavelet coefficient is greater than  $\lambda$ , we assume that as significant and attribute it to the original signal. Otherwise, we consider it to be due to the additive noise and discard the value. The soft-thresholding function has a somewhat different rule from the hard-thresholding function. It shrinks the wavelet coefficients by  $\lambda$  toward zero, which is the reason why it is also called the wavelet shrinkage function. It is explained in Eq. (9) as:

$$\begin{aligned} f_s(x) &= x - \lambda \quad \text{if } x > \lambda \\ &= 0 \quad \text{if } x < \lambda \\ &= x + \lambda \quad \text{if } x \leq -\lambda \end{aligned} \quad (9)$$

The soft-thresholding rule is chosen over hard-thresholding, as the soft-thresholding method yields more visually pleasant images over hard-thresholding. Then finally IDWT is calculated by Eq. (10)

$$s_n = \frac{1}{N} \sum_{j=0}^{\log_2 N} \sum_{k=0}^{\text{int}(N/2^{j+1})} S_{j,k} j, kw_n \quad (10)$$

where

$j, kw_n$	$n$ th sample of $k$ th sifted version of a $2^j$ scaled discrete wavelet
$J, k$	row index
$n$	column index [8].

In our work we have used OTSU'S thresholding to denoise the image which chooses the threshold in such a way that all variances available in black and white pixels in the same signal are minimized.

level = graythresh (I) computes a global threshold (level) that can be used to convert an intensity image into a binary image with im2bw. level is a normalized intensity value that lies in the range [0, 1].



The `graythresh` function uses Otsu's method, which chooses the threshold to minimize the intraclass variance of the black and white pixels.

Multidimensional arrays are converted automatically in to 2D arrays using `reshape`. The `graythresh` function ignores any nonzero imaginary part of `I`.

`[level EM] = graythresh (I)` returns the effectiveness metric, `EM`, as the second output argument. The effectiveness metric is a value in the range `[0 1]` that indicates the effectiveness of the thresholding of the input image. The lower bound is attainable only by images having a single gray level, and the upper bound is attainable only by two-valued images.

## Proposed Method

In this chapter we have taken the pure EEG signals having four samples that are then mixed with noise. The signal is processed with ICA and then further with wavelet denoise. ICA is applied so as to separate the signals from a multichannel source of signals and then wavelet denoising to remove noise from an independent component of the signal; we find that the final signal shows better artifacts removal as compared to simple filtering methods. The complete process is explained in the following algorithm:

Algorithm

1. Plot the EEG signal that is mixed with noise with respect to  $j$  and  $k$ .
2. Apply conventional filtering through Kurtosis that is defined by Eq. (5):

$$\text{Kurt}(y) = E\{y^4\} - 3(E\{y^2\})^2$$

3. Let the original signal be defined by  $X$ , then basic ICA model is expressed by Eq. (11)

$$X = AS + N \quad (11)$$

where,

$A$     mixing matrix  
 $S$     independent component  
 $N$     noise added

4. To denoise the image using ICA we have to process the image. The preprocessing consists of two steps:
  - a. Centering: the signal is first centered, i.e., we subtract the image mean from the noisy image. It is expressed mathematically as in Eq. (12):

$$X \leftarrow X - E\{X\} \quad (12)$$

- b. Whitening: in whitening we remove the second order statistical dependence in the data, i.e., the whitened data have unit variance and they are uncorrelated. Let  $Z$  be zero mean random vector, then in terms of covariance matrix we can write

$$E \{ZZ^T\} = I \quad (13)$$

where  $I$  is identity matrix.

Finally whitened data  $Z$  is expressed by Eq. (14):

$$Z = D^{-1/2}.E^T.X \quad (14)$$

where  $E$  is matrix whose columns are unit norm eigenvectors of covariance matrix.

$$C_X = E\{XX^T\} \quad (15)$$

$D$  is diagonal matrix of eigenvalues of  $C_X$

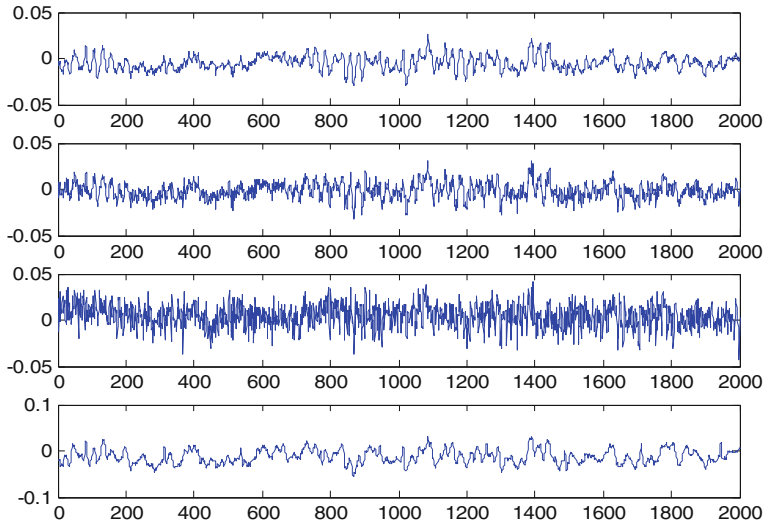
5. Using the demixing matrix obtained above we obtain the mixing matrix  $A$  by the equation

$$A = \text{pseudoinverse}(w) \quad (16)$$

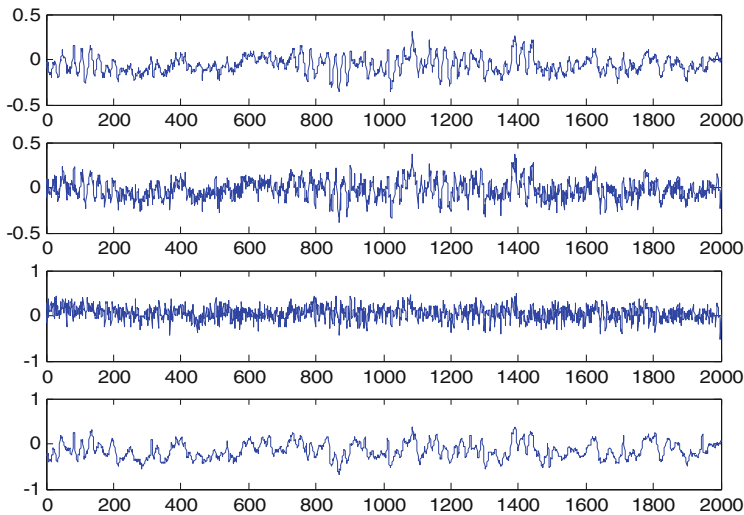
6. Finally denoised image is obtained by the mixing matrix  $A$  and the independent component.
7. Then Otsu's thresholding is done and values below a certain threshold are set to zero.
8. This gives the whitened denoised image. To reconstruct the image from this we add the mean to this image which was subtracted earlier and multiply the whitening matrix to obtain the final denoised image.

## Experimental Results

This section presents the evaluation of the proposed artifact removal technique. Initially, EEG signals are captured with occurrence of artifacts. Figure 4 shows the four samples of EEG signal that is effected through noise and artifacts. We defined  $n$  as the number of iterations and to plot our data to number of values to number of iteration we defined  $j$  and also  $j$  as the original matrix value of data and  $k$  defines the number of blocks available in data. Then Fig. 5 shows the result using ICA so as to find the independent component. Figure 6 is a result of signal after implementation of wavelet denoising on the signals of Fig. 5. From these figures, it has been observed that the proposed artifact removal technique results in better removal of artifacts. This will help in improving the performance of the further processing of this EEG signal.

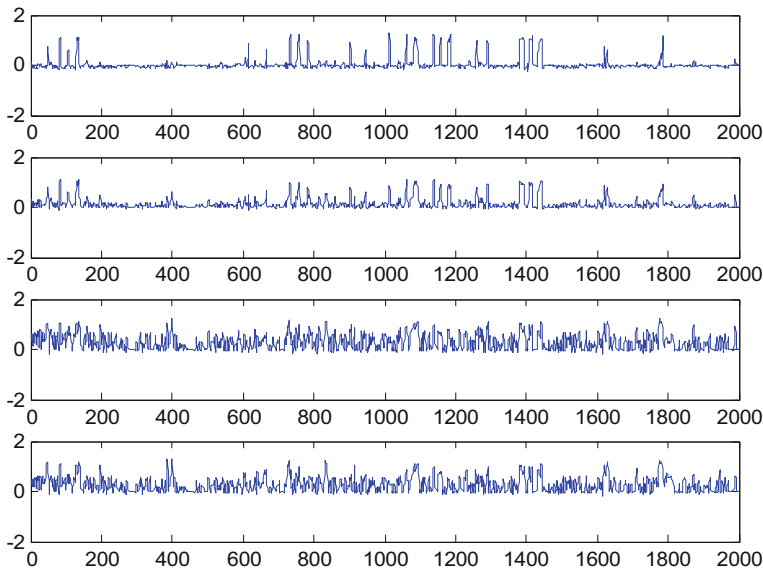


**Fig. 4** EEG signal mixed with noise



**Fig. 5** EEG signal after applying independent component analysis

Here, from Fig. 6 we can conclude that artifacts and noise is removed from the original EEG signal to a great extent because of use of both methods ICA and wavelet denoising, which use Otsu's method for thresholding.



**Fig. 6** EEG signal after applying ICA and wavelet denoising

## References

1. Geethaa G, Geethalakshmia SN (2012) "Artifact removal from EEG using Spatially constrained independent component analysis and wavelet denoising with otsu's thresholding technique", International conference on communication technology and system design 2011, Procedia Engineering 30 (2012) 1064–1071
2. Romo Vázquez R, Vélez-Pérez H, Rantab R, Louis Dorrb V, Maquinb D, Maillard L (2011) Blind source separation, wavelet denoising and discriminant analysis for EEG artifacts and noise cancelling, Biomedical Signal Processing and Control
3. Walters-Williams J, Li A Y (2011) New Approach to Denoising EEG Signals: Merger of Translation Invariant Wavelet and ICA. Int J Biometrics Bioinform 5(2):130–149
4. Khaparde A, Madhavilath M, Manasa MBL, Anil Babu P, Pradeep Kumar S (2008) FastICA Algorithm for the Separation of Mixed Images. Wseas Trans Sign Proces 4(5):271–277 ISSN: 1790-5052
5. Hyvarinen A, Oja E (1997) A fast fixed point algorithm for independent component analysis. Neural Comput 9(7):1483–1492
6. Hongqiao1 L, Shengqian W (2009) A new image denoising method using wavelet transform International forum on information technology and applications, vol 1. pp 111–114
7. Gonzalez RC, Elwood's R (1993) Digital Image Processing. Addison-Wesley, MA
8. Ellinas JN, Mandadelis T, Tzortzis A, Aslanoglou L (2004) "Image denoising using wavelets", T.E.I. of Piraeus Applied Research Review, vol IX. no.1 pp 97–109

Proceedings of All India Seminar on Biomedical  
Engineering 2012 (AISOB 2012)

Kumar, V.; Bhatele, M. (Eds.)

2013, XVI, 296 p., Hardcover

ISBN: 978-81-322-0969-0