

Old Dominion University

## ODU Digital Commons

---

Electrical & Computer Engineering Theses & Dissertations

Electrical & Computer Engineering

---

Fall 2010

# Electroencephalogram Artifact Removal Using a Wavelet Neural Network

Hoang-Anh T. Nguyen  
*Old Dominion University*

Follow this and additional works at: [https://digitalcommons.odu.edu/ece\\_etds](https://digitalcommons.odu.edu/ece_etds)



Part of the [Artificial Intelligence and Robotics Commons](#), [Computational Engineering Commons](#), [Computer and Systems Architecture Commons](#), and the [Theory and Algorithms Commons](#)

---

### Recommended Citation

Nguyen, Hoang-Anh T.. "Electroencephalogram Artifact Removal Using a Wavelet Neural Network" (2010). Master of Science (MS), Thesis, Electrical & Computer Engineering, Old Dominion University, DOI: 10.25777/hthh-w349  
[https://digitalcommons.odu.edu/ece\\_etds/457](https://digitalcommons.odu.edu/ece_etds/457)

This Thesis is brought to you for free and open access by the Electrical & Computer Engineering at ODU Digital Commons. It has been accepted for inclusion in Electrical & Computer Engineering Theses & Dissertations by an authorized administrator of ODU Digital Commons. For more information, please contact [digitalcommons@odu.edu](mailto:digitalcommons@odu.edu).

**ELECTROENCEPHALOGRAPH ARTIFACT REMOVAL USING A WAVELET  
NEURAL NETWORK**

by

Hoang-Anh T. Nguyen

A Thesis Submitted to the Faculty of  
Old Dominion University in Partial Fulfillment of the  
Requirements for the Degree of

MASTER OF SCIENCE

ELECTRICAL ENGINEERING

OLD DOMINION UNIVERSITY

December 2010

Approved by:

---

Jiang Li (Director)

---

Frederic D. McKenzie (Member)

---

Yuzhong Shen (Member)

# **ABSTRACT**

## **ELECTROENCEPHALOGRAM ARTIFACT REMOVAL USING A WAVELET NEURAL NETWORK**

Hoang-Anh T Nguyen

Old Dominion University, December 2010

Director: Dr. Jiang Li

A wavelet neural network (WNN) technique is developed for electroencephalogram (EEG) artifact removal without electrooculographic (EOG) recordings. The algorithm combines the universal approximation characteristics of neural networks and the time/frequency property of wavelet, where the neural network was trained on a simulated dataset with known ground truths. The contribution of this thesis is two-fold. First, many EEG artifact removal algorithms, including regression based methods, require reference EOG signals, which are not always available. To remove EEG artifacts, a WNN tries to learn the characteristics of the artifacts first and does not need reference EOG signals once trained. Second, WNNs are computationally efficient, making them a reliable real time algorithm. A WNN algorithm is then compared with the independent component analysis (ICA) technique and an adaptive wavelet thresholding method is used on both simulated and real datasets. Experimental results show that a WNN can remove EEG artifacts effectively without diminishing useful EEG information even for very noisy datasets.

Copyright, 2010, by Hoang-Anh T. Nguyen, All Rights Reserved.

## ACKNOWLEDGMENTS

I would like to express my deepest gratitude to Prof. Jiang Li for his support, instruction, and patience during my research and study at Old Dominion University. I am honored to have been educated by him.

I would like to give special thanks to Prof. Frederic McKenzie and Prof. Yuzhong Shen. Having each of them serve on my committee is a privilege. Their invaluable instructions and comments are very much appreciated, which helped me improve my work significantly.

I gratefully acknowledge Prof. Rahman, Prof. Popescu and Prof. Gray for bringing out my desire to learn and help me grow tremendously. I would like to thank Prof. Sacharia Albin for his kind help and support during my study at Old Dominion University.

Lastly, I would like to thank my friends for being with me all the time. I would like to thank my family, especially Sonny and Huong, for their love and encouragement. They have been my source of motivation. This thesis would not have been possible without their continued encouragement.

This thesis is dedicated to my parents.

## TABLE OF CONTENTS

TABLE OF CONTENTS.....	vi
LIST OF TABLES.....	viii
LIST OF FIGURES .....	viiiix
CHAPTER 1. INTRODUCTION .....	1
1.1 Electroencephalogram.....	1
1.2 Artifacts.....	1
1.3 EEG Model .....	4
1.4 EOG Artifact Removal Techniques .....	6
1.5 Wavelet Neural Network brief Description .....	7
1.6 Thesis Outline .....	8
CHAPTER 2. RELATED WORK.....	9
2.1 Wavelet Transform and Its Application to EOG Artifact Removal .....	9
2.1.1 Wavelet Transform .....	9
2.1.2 Wavelet Thresholding for EOG Artifact Removal .....	12
2.1.3 Thresholding Function .....	13
2.2 Independent Component Analysis .....	15
2.2.1 ICA in EEG Artifact Removal .....	15
2.2.2 Information Maximization Approach (InfoMax): An ICA Algorithm.....	17
2.3 Artificial Neural Network.....	18
2.3.1 Output Weight Optimization Backpropagation .....	19
2.4 Motivation of the Work .....	21
CHAPTER 3. PROPOSED METHOD.....	23
3.1 EEG Data Simulation.....	24
3.2 Neural Network for EEG Artifact Removal .....	26
3.2.1 Training.....	26
3.2.2 Testing and Artifact Removal.....	27

3.3. Validation Methods.....	27
CHAPTER 4. RESULTS AND DISCUSSIONS.....	29
4.1 Datasets.....	29
4.1.1 Driving Test Dataset.....	29
4.1.2 Visual Selection Task Dataset .....	30
4.1.3 Experimental Settings .....	30
4.2 Results for the Driving Test Dataset.....	31
4.2.1 Results on Simulated Data.....	31
4.2.2 Results on Real Data.....	36
4.3 Results for the Visual Selection Task Dataset .....	39
4.3.1 Results on Simulated Data.....	39
4.3.2 Results on Real Data.....	43
4.4 Discussions .....	48
CHAPTER 5. CONCLUSIONS AND FUTURE WORK.....	51
REFERENCES .....	52



## LIST OF TABLES

Table	Page
<b>Table 1.1.</b> EEG rhythms specifications.....	3
<b>Table 3.1.</b> EEG frequency band specifications. ....	25
<b>Table 4.1.</b> Training and testing (a) MSE and (b) RMSE of signals before and after correction for driving test dataset. ....	33
<b>Table 4.2.</b> Training and testing (a) MSE and (b) RMSE of signal before and after correction for visual selection task dataset. ....	47

## LIST OF FIGURES

Figure	Page
<b>Figure 1.1.</b> EEG data set disturbed by various types of artifacts. ....	2
<b>Figure 1.2.</b> EEG segments contaminated by EOG artifacts. ....	4
<b>Figure 2.1.</b> (a) Wavelet transform and (b) wavelet reconstruction. ....	10
<b>Figure 2.2.</b> Wavelet thresholding method. ....	13
<b>Figure 2.3.</b> General theory of independent component analysis. ....	16
<b>Figure 2.4.</b> EEG artifact removal using ICA procedure. ....	16
<b>Figure 2.5.</b> General artificial neural network structure. ....	18
<b>Figure 2.6.</b> Fully-connected one hidden layer MLP ANN structure. ....	20
<b>Figure 3.1.</b> Proposed wavelet neural network structure. ....	23
<b>Figure 3.2.</b> Simulated EEG generator model. ....	24
<b>Figure 3.3.</b> Neural network training procedure. ....	26
<b>Figure 3.4.</b> EEG corrected by using WNN... ..	27
<b>Figure 4.1.</b> Clean and contaminated simulated signal for (a) training and (b) testing. ....	32
<b>Figure 4.2.</b> Contaminated simulated and WNN corrected singals. ....	33
<b>Figure 4.3.</b> PSD of clean, contaminated and WNN corrected signals for (a) training and (b) testing. ....	34
<b>Figure 4.4.</b> Frequency correlation between (a) contaminated and wavelet thresholding corrected simulated signals and (b) clean and wavelet thresholding corrected simulated signals (c) contaminated and WNN corrected simulated signals and (d) clean and WNN corrected simulated signals, all for testing. ....	35

<b>Figure 4.5.</b> Contaminated and decontaminated EEG (a) contaminated, ICA and WNN corrected EEG (b) contaminated, wavelet thresholding and WNN corrected EEG. ....	37
<b>Figure 4.6.</b> PSD of contaminated and de-contaminated EEG.....	38
<b>Figure 4.7.</b> Frequency correlation between contaminated and decontaminated EEG, (a) by ICA, (b) by wavelet thresholding and (c) by WNN.....	39
<b>Figure 4.8.</b> Contaminated simulated and WNN corrected singals (a) 30 seconds and (b) 7.4219 seconds.....	40
<b>Figure 4.9.</b> PSD of clean, contaminated and WNN corrected signals for (a) training and (b) testing. ....	41
<b>Figure 4.10.</b> Frequency correlation between (a) contaminated and wavelet thresholding corrected simulated signals and (b) clean and wavelet thresholding corrected simulated signals (c) contaminated and WNN corrected simulated signals and (d) clean and WNN corrected simulated signals, all for testing.....	42
<b>Figure 4.11.</b> Contaminated and de-contaminated EEG (a-b) by ICA and WNN and (c-d) by wavelet thresholding and WNN.....	46
<b>Figure 4.12.</b> PSD of contaminated and de-contaminated EEG.....	47
<b>Figure 4.13.</b> Frequency correlation between contaminated and decontaminated EEG by (a) ICA, (b) wavelet thresholding and (c) WNN. ....	48

# CHAPTER 1

## INTRODUCTION

### 1.1 Electroencephalograms

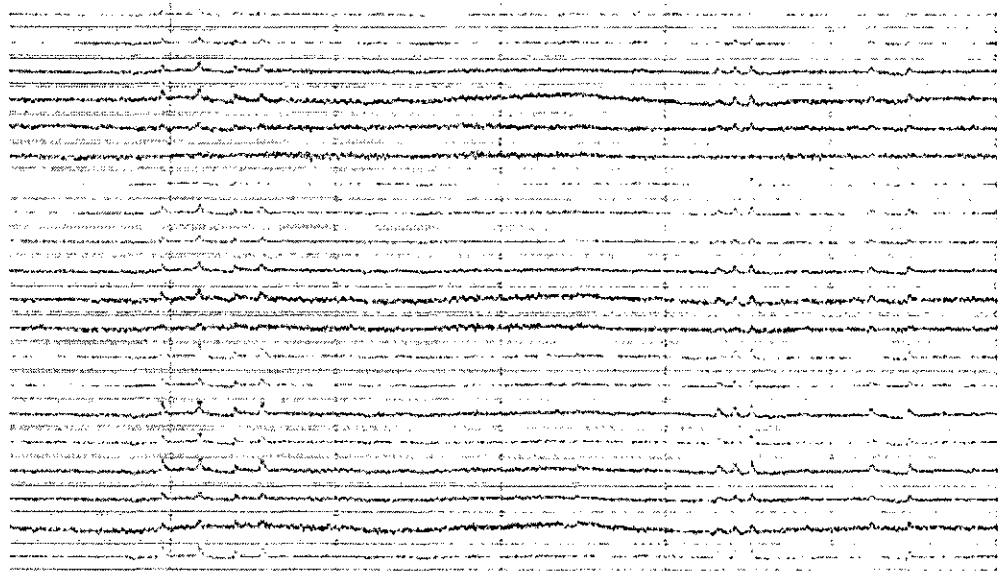
An electroencephalogram (EEG) is the recording of neural electrical activities caused by nerve firings. Typically, EEG signals are recorded using recording systems with electrodes placed across the scalp. EEG waveforms are characterized by three components including shape, frequency, and amplitude. Based on those components, by using visual or non-visual analysis techniques, it is possible to extract useful features carried by brain signals.

An EEG carries information about rhythmic activities at different frequency bandwidths (as shown in Table 1.1):  $\delta$  – delta (1-4 Hz),  $\theta$  – theta (4-8 Hz),  $\alpha$  – alpha (8-13 Hz),  $\beta$  – beta (13 – 30 Hz) and  $\gamma$  – gamma (30 -50 Hz) [24, 49, 52]. EEG rhythms are often followed by external (i.e. some type of stimulation) or internal (e.g., motor preparation, cognitive processing) events. Rhythmic EEG can be extracted by a number of methods including time-frequency analysis [14 – 19], band-pass filtering [11, 54], independent component analysis [13, 37, 50], just to name a few.

### 1.2 Artifacts

EEG recordings are known to be contaminated by physiological artifacts from various sources, such as eye blinking/movement, heart beating and movement of other muscle groups [1]. Such types of artifacts are mixed together with brain signals, making


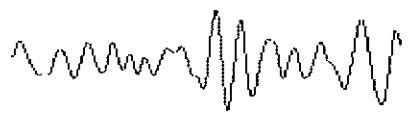




interpretation of EEG signals difficult [2]. An example of highly contaminated EEG dataset is shown in Figure 1.1.



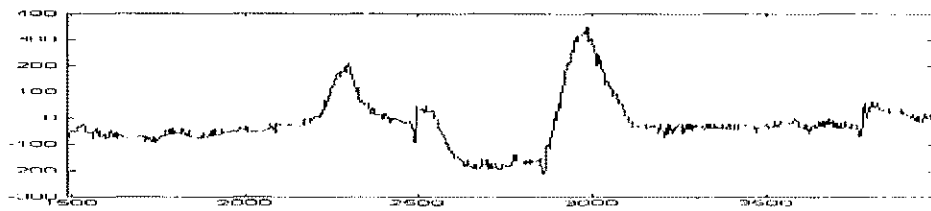
**Figure 1.1.** EEG data set disturbed by various types of artifacts. (Data courtesy of University of Iowa)

Eye movement or blinks, as shown in Figure 1.2, usually produce large electrical potential, which spreads across the scalp and contaminates EEG recordings. This class of potential generates significant electrooculographic (EOG) artifacts in recorded EEG signals. Removal of EOG artifacts is nontrivial because these artifacts contaminate and overlap in frequency and time domains with an EEG. The effect of EOG artifacts on EEG signals is found most significantly in low frequency bands such as  $\delta$ ,  $\theta$  and  $\alpha$  [3]. Eye blinking generates spike-like shaped signal waveforms with their peaks reaching up to 800uV and occurs in a very short period, 200-400 ms [4]. Meanwhile, artifacts generated by eye movement are square-shaped, smaller in amplitude but last longer in time and concentrate in lower frequency bands [5].

Table 1.1. EEG rhythm specifications.

EEG rhythm	Spectral band	Description	Graphical illustration
Delta	1-4 Hz	$\delta$ rhythm is generated in the cortex or thalamic cells and are dominant during the sleep state or in infants [49]. It is often associated with slowed or disabled cognitive processing.	
Theta	4-8 Hz	$\theta$ rhythm is generated in the Hippocampus [52] and is often associated with a drowsy state or occurs during a vigilance decrease.	
Alpha	8-13 Hz	$\alpha$ rhythm often presents while the subject is awake, relaxed with eyes closed [42]. The subject in this state is aware of sensory stimuli without focusing on any specific activity.	
Beta	13 – 30 Hz	$\beta$ , mainly generated by the activation of the cortex, is often associated with a high concentration state. For this reason it is useful in neurofeedback applications to increase concentration of the subject.	
Gamma	30 -50 Hz	$\gamma$ is involved in the formation of percepts and memory, linguistic processing, and other behavioral and perceptual functions [43].	
EEG			

Muscle activity is involved with activities caused by groups of muscles. The frequency and amplitude of muscle activity, which depend strongly on scalp topography, varies widely and overlaps EEG signals. Either asking the subject to be stationary during the course of EEG recording or simply applying a regression technique is not practical [1]. Several works have been proposed to remove the influence of muscle activity to EEG recordings [44-45].



**Figure 1.2.** EEG segments contaminated by EOG artifacts.

A heart beating, or electrocardiogram (ECG), is a rhythmic signal and is unavoidably mixed into the EEG recordings. An ECG artifact, often inter-mixed with other artifacts, makes EEG signals slightly spiky when a heart beats [46]. It is possible to rely on heart beating regularly to recognize the presence of this artifact.

A pulse is another common class of artifact, which appears when the EEG recording electrodes are placed over or close to blood vessels. The recorded electrical potentials are then affected during the expansion or contraction processes of those vessels. The shape of this artifact may be either a sharp spike or a smooth wave [47].

### 1.3 EEG Model

Brain and eye activities are generated by physiologically separate sources. Cerebral signals, recorded by an EEG recording system, result from neural firing

activities. On the other hand, EOG artifacts are non-cerebral activities spreading over the entire recording scalp and then contaminating the EEG electrode recordings. For that reason, an EEG recording can be represented as a superposition of a true EEG and some portions of the artifact signals. When an EOG artifact is present, it is assumed that the model for the contaminated EEG signal as in the following form [14],

$$EEG_{rec}(t) = EEG_{true}(t) + k.EOG(t) \quad (1-1)$$

where  $EEG_{rec}(t)$  is the recorded contaminated EEG,  $EEG_{true}(t)$  denotes the true EEG signal,  $EOG(t)$  represents the original potential changes caused by ocular activity and  $k$  symbolizes the propagated factor and varies between 0 and 1 depending on the location of the recording electrode. Hence,  $k.EOG(t)$  represents the propagated ocular artifact from the eye to the recording site, which directly adulterates the brain signals. Estimating  $EEG_{true}(t)$  from observed  $EEG_{rec}(t)$  is non-trivial and is equivalent to minimizing the effect of ocular artifacts. Similar to other artifact removal techniques, the goal of a proposed wavelet neural network technique is to recover  $EEG_{true}(t)$  from  $EEG_{rec}(t)$ .

As a random signal, a true EEG owns the noise-like (flat) power spectrum. In some cases when a subject performs specific tasks, the biological neural system introduces activities at particular frequencies making the power spectrum deflated. As a major artifactual type, once mixed with  $EEG_{true}(t)$ , the ocular artifact  $k.EOG(t)$  causes proliferation in low frequencies and generates spike-like shape data segments across time domain. These properties are utilized by both wavelet thresholding [14] and the proposed WNN technique for artifact correction.



#### 1.4 EOG Artifact Removal Techniques

In recent years, there has been an increasing interest in applying various techniques to remove ocular artifacts from EEG [1, 2, 5-8, 10, 13, 14-19, 54]. The methods for removing EOG artifacts based on regression have been widely studied [1, 6, 40, 54]. Regression methods often assume that the scalp potential is the linear summation of brain and ocular potentials. By subtracting propagated EOG from EEG recordings, the EEG signals can be recovered [8]. Regression can also be done in frequency domain based on the concept that subtraction in the frequency-domain is equivalent to filtering in the time-domain. By eliminating spectral estimates of EOG from those of EEG recordings, it is possible to recover the non-contaminated EEG [40]. Both types of regression methods are off-line and rely on EOG recordings, which are however, not always available [1, 6, 14].

Berg and Scherg [10, 39] proposed a principle component analysis (PCA) based technique for removing eye movement artifacts. This method assumes that each EEG channel recording is simultaneously generated by multiple sources across the scalp. By decomposing multiple channel EEG data into principle components using PCA, the artifactual and cerebral sources can be identified. The artifacts were removed by eliminating these contaminated PCA components. Their experiments showed that the PCA based method outperformed regression based models. However, PCA models usually failed to completely separate artifacts from cerebral activities [11], and the orthogonal assumption for data components in PCA is hardly satisfied [5]. Independent

component analysis (ICA), which was originally developed for the blind source separation (BSS) problems, has been used as an alternative method for EEG artifact removal [1, 12-14]. ICA usually requires a large amount of data and visual inspection to eliminate noisy independent components, making the method time-consuming and not suitable for real-time applications.

Recently, wavelet analysis has been used as an effective tool for measuring and manipulating non-stationary signals such as EEG. Wavelet-based methods, especially the wavelet-thresholding techniques for EEG artifact removal, have received significant attention [14-19]. For this class of methods, Wavelet coefficients at low-frequency sub-bands are corrected by some thresholding functions before signal reconstruction. As an online artifact removal method, the most important advantage of using this method for EEG correction is that it does not rely on either the reference EOG signal or visual inspection. However, this research's experiments show that wavelet thresholding method is sensitive to the selection of wavelet and threshold functions.

### **1.5 Wavelet Neural Network Brief Description**

This thesis proposes a novel, robust, and efficient Wavelet Neural Network (WNN) technique to remove EEG artifacts by combining the approximation capabilities of both wavelet and neural network methods. Similar to wavelet-based methods, an EOG recording is not required and WNN can perform artifact removal for single channel data where ICA or other BSS methods are not applicable. WNN consists of three steps, it (1) decomposes contaminated EEG signals to a set of wavelet coefficients, (2) corrects low frequency wavelet sub-band coefficients by a trained neural network and (3) reconstructs

EEG signals using the corrected coefficients. The neural network is trained by using a simulated dataset with known ground truths.

## **1.6 Thesis Outline**

The rest of this thesis is organized as follows: Chapter 2 discusses related work. Chapter 3 presents the proposed wavelet neural network technique along with procedures for network training and testing. Chapter 4 describes two EEG datasets from a driving test and a visual selection task and experimental settings. Some of the achieved results on both datasets and discussions are also given in this Chapter. Finally, Chapter 5 concludes the thesis and presents future work.

## CHAPTER 2

### RELATED WORK

#### 2.1 Wavelet Transform and Its Application to EOG Artifact Removal

##### 2.1.1 Wavelet Transform

The wavelet transform [19, 20,25, 28, 29] is an integral transform for which the set of basis functions, known as wavelets, are well localized both in time and frequency domains. Wavelets can be constructed from a single function  $\psi(t)$ , named mother wavelet or analyzing wavelet, by means of translation and dilation,

$$\psi_{a,\tau}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-\tau}{a}\right) \quad (2-1)$$

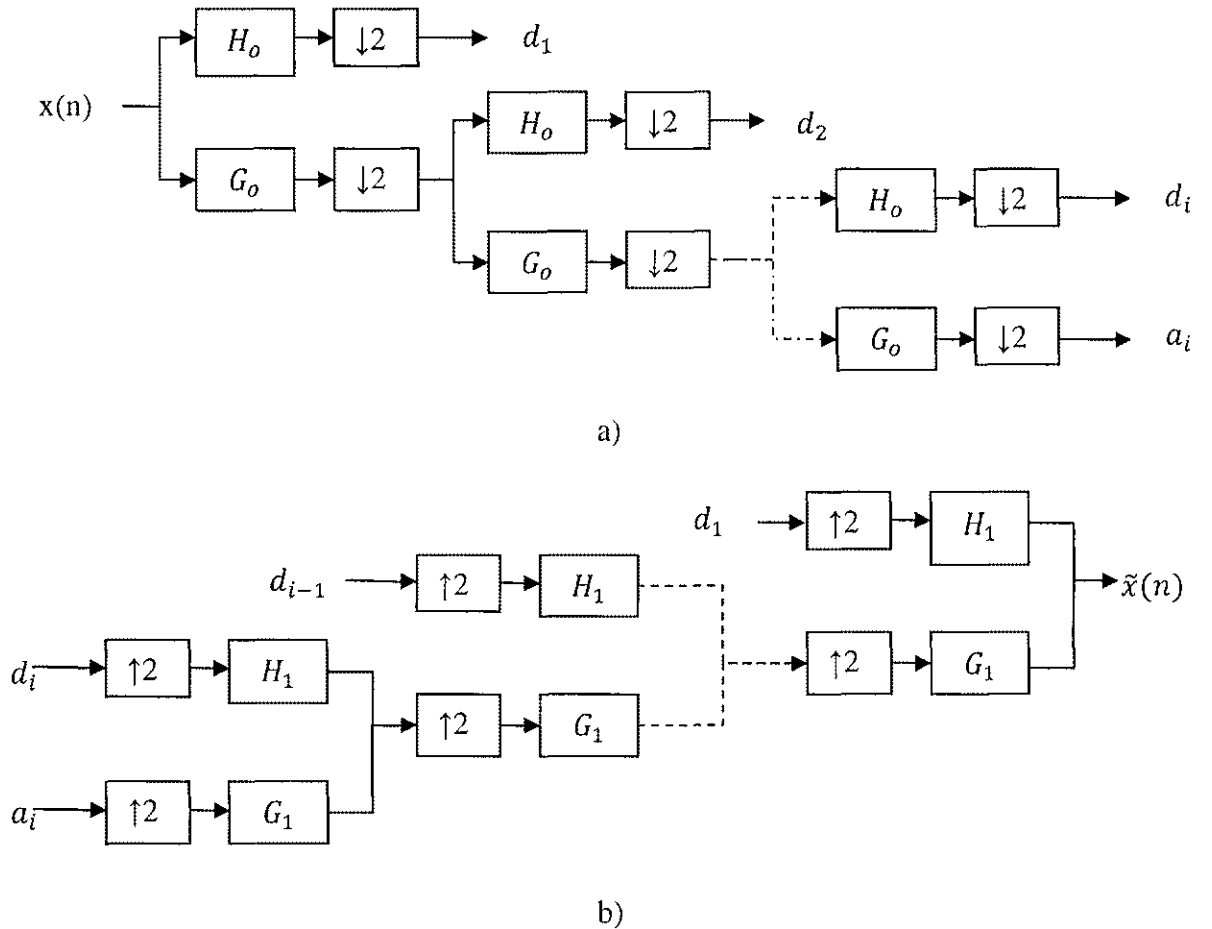
A *continuous wavelet transform* (CWT) of a signal  $x(t)$ , defined as the correlation between the wavelet and the signal itself, can be implemented by the following formula,

$$W(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi_{a,\tau}^*(t) dt \quad (2-2)$$

where  $\psi^*(t)$  denotes the complex conjugate of  $\psi(t)$ . The above equation (2-2) indicates that the wavelet is passed through the analyzed signal and yields a set of coefficients representing an image of an analyzed signal at different scales in time and frequency domains. The translated parameter  $a$  plays a crucial role in wavelet transform. Specifically, for small values of  $a$ , the wavelet is a narrow version of the original function, which corresponds roughly to higher frequencies. For very large values of  $a$ , the wavelet is expanded and owns a low frequency property. Accordingly, the low frequency terms are analyzed with a less sharper time resolution than high frequency components.

This is a desirable property especially in analyzing transient waveforms such as EEGs corrupted with ocular artifacts.

The wavelet transform results in a time-scale decomposition in which scales are related to frequency rather than to a time-frequency representation [26]. The highest scale corresponds to the highest frequencies represented in the signal (less or equal to half of the sampling rate), and the bandwidth of this scale ranges from a half to a quarter of the sampling rate. While the bandwidth is reduced by two, the number of coefficients at lower



**Figure 2.1.** (a) Wavelet transform and (b) wavelet reconstruction.

resolutions decreases approximately by a factor of two compared to that of the higher resolution next to it. A proper selection of coefficients from different scales may be used to suppress artifacts in EEG signals.

The *discrete wavelet transform* (DWT) is the version of wavelet transform applied to discrete time series. DWT can be implemented with a simple recursive filter scheme providing a highly efficient wavelet representation of the original signal. The reconstruction can be realized by using an inverse filtering operation. Then, the parameters  $a$  and  $\tau$  in equation (2-1) and (2-2) can be represented as  $a_i = 2^{-i}$  and  $\tau_{i,j} = 2^{-i}j$ , where  $i$  and  $j$  are positive integers. The property of mother wavelet function  $\psi(t)$  depends on the selection of  $i$  and  $j$ . The family  $\psi_{i,j}(t) = 2^{i/2}\psi(2^i t - j)$  constitutes an orthonormal basis of Hilbert space, consisting of finite energy signals [27]. These coefficients provide full information in a simple way and a direct estimation of local energies at the different scales.

The wavelet transform and wavelet reconstruction can be illustrated as in Figure 2.1. The (forward) wavelet transform is implemented in the following process: the original signal is first passed through the high pass ( $H_o$ ) and low pass ( $G_o$ ) filters, that are designed based on the properties of wavelet basis function, and then downsampled by two. After that, the detail, or high frequency coefficient series and approximation, or low frequency coefficient series are obtained. The approximation term is then continuously used as the input of the next level wavelet transform. Finally, there would be one approximation band ( $a_i$ ) and details ( $d_k, k = 1, \dots, i$ ) representing the signal information at different scales. On the other hand, the reconstruction is the reverse process of

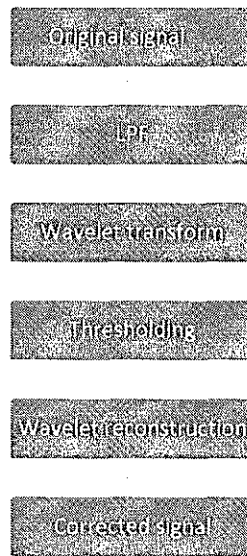
decomposition. The approximation and detail coefficients at each level are upsampled by two and passed through low pass  $G_1$  and high pass  $H_1$  synthetic filters, which satisfy the perfect construction condition [25] and are then summed up. This process is continued  $i$  times, equal to the number of levels of decomposition. The result of this process could be the original or corrected signal. In the later case, the detail and approximation coefficients at some level are modified in desirable ways.

### 2.1.2 Wavelet Thresholding for EOG Artifact Removal

Wavelet thresholding techniques are built on the multiresolution analysis of wavelet transform, a tool that analyses signal at different time and frequency scales [20]. The approximations and details are processed by thresholding before reconstruction [14]-[18]. By selecting a ‘good’ mother wavelet, which resembles the shapes of the artifacts, large-valued coefficients are generated in the areas corresponding to the EEG artifacts at low-frequency sub-bands and are considered as an estimate of the ocular artifacts. Thus, shrinking the amplitude range of these coefficients by nonlinear thresholding functions would remove those artifacts. In this thesis, a wavelet thresholding method was implemented as follows [14],

- Use a Butterworth lowpass filter to smooth the EEG signal before further processing
- Apply wavelet transform to the contaminated EEG signal
- Utilize a thresholding function to automatically corrected high-valued coefficients at low-frequency subbands
- Reconstruct the corrected EEG signal

This process is illustrated by the diagram at Figure 2.2.



**Figure 2.2.** Wavelet thresholding method.

### 2.1.3 Thresholding Function

The core idea of wavelet thresholding method is to use some soft, hard or adaptive functions to threshold the wavelet coefficients and then use that thresholded coefficients to reconstruct the corrected/de-noised signal. The threshold function is selected in the way to suppress the noise and keep the useful information maximally making the threshold function selection become a non-trivial step. A number of schema for thresholding has been presented based on the minimax mean square error [28], Bayesian risk minimization [29] or *Stein's unbiased risk estimate* (SURE) maximization [30-31]. In this thesis, a time-scale adaptive algorithm based on SURE risk estimate along with a soft-like thresholding function [18] was used for EEG ocular artifact removal issues due to its effectiveness illustrated in [14].



In this method, the wavelet coefficients of EEG in specific wavelet sub-bands are corrected as follows,

$$w_k^c(w, t) = \begin{cases} w + t - \frac{t}{2k+1}, w < t \\ \frac{1}{(2k+1)t^{2k}} w^{2k+1}, |w| \leq t \\ w + t + \frac{t}{2k+1}, w > t \end{cases} \quad (2-3)$$

where  $w_k^c$  represents the corrected version of original wavelet coefficients  $w$  using the thresholding value  $t$ . The optimal value of  $t$  can be adjusted under SURE risk using following adaptive steps

$$t(i+1) = t(i) - \nabla t(i) \quad (2-4)$$

where the adjustment of threshold at step  $i$  is defined by

$$\nabla t(i) = \alpha(i) \cdot \frac{\partial R_s(t)}{\partial t} \quad (2-5)$$

where

$$\frac{\partial R_s(t)}{\partial t} = 2 \sum_{i=0}^{N-1} g_i \cdot \frac{\partial g_i}{\partial t} + 2 \sum_{i=0}^{N-1} \frac{\partial^2 g_i}{\partial w_i \partial t} \quad (2-6)$$

and

$$g_i = w_k^c(w_i, t) - w_i \quad (2-7)$$

The SURE risk threshold function is applied to correct the coefficients at approximation and several lowest frequency sub-bands. These corrected coefficients are then used for corrected signal reconstruction.

## 2.2 Independent Component Analysis

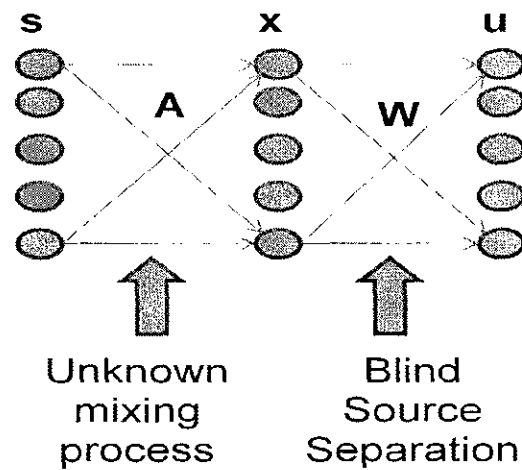
Independent component analysis (ICA) was first proposed by Herault and Jutten at a meeting in Snowbird, Utah in 1986 [1, 11] to solve the blind source separation (BSS) problem. ICA, as shown in Figure 2.3, aims to recover independent source signals  $\mathbf{s} = \{s_1(t), s_2(t), \dots, s_N(t)\}$ , from recorded mixtures  $\mathbf{x} = \{x_1(t), x_2(t), \dots, x_N(t)\}$  by an unknown matrix  $\mathbf{A}$  of full rank. The basic problem of ICA is to estimate the mixing matrix  $\mathbf{A}$  or equivalently, the original independent sources  $\mathbf{s}$  based on the following linear relationship  $\mathbf{x} = \mathbf{A}\mathbf{s}$ , while no knowledge is available about the sources or the mixing matrix. ICA identifies an unmixing matrix,  $\mathbf{W}$ , which decomposes the mixed data into a sum of temporally independent and spatially fixed components. ICA finds  $\mathbf{u} = \mathbf{W}\mathbf{x}$ , where the rows of the output data matrix represent time courses of activation of the ICA components [1, 9, 11].

Several algorithms have been proposed to implement ICA such as information maximization (InfoMax), Fixed-point ICA, Joint approximate diagonalization of eigenmatrices (JADE) and the second-order blind identification (SOBI). In this research, the InfoMax algorithm, as presented in [1, 32], was used to perform EEG artifact removal.

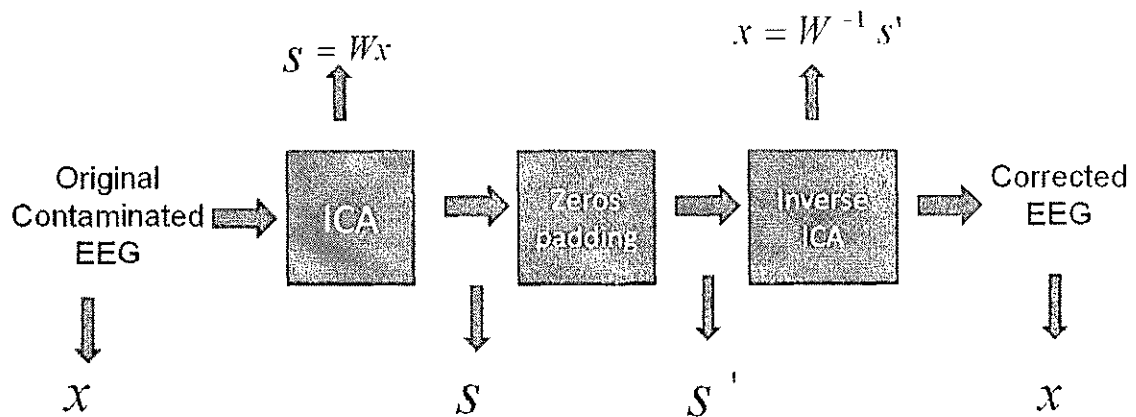
### 2.2.1 ICA in EEG Artifact Removal

EEG artifact removal using ICA can be realized based on several assumptions [1, 13]: (1) neural electrical recording signals are stationary, (2) the number of sources generating EEG is the same as the number of recording channels, (3) the sources are independent, (4) the sources distributions must be non-Gaussian and (5) the mixing

matrix is squared and invertible. Keeping those assumptions in mind, the following procedure, as shown in Figure 2.4, can be used to eliminate EEG artifacts. By performing some sort of ICA technique it is possible to estimate the unmixing matrix  $W$ , which is then utilized to recover the estimated independent components  $\tilde{s}$  from original mixtures  $x$  using the linear relationship  $\tilde{s} = Wx$ . The independent components (ICs) are then categorized into artifactual and non-artifactual components. The artifactual components



**Figure 2.3.** General theory of independent component analysis.



**Figure 2.4.** EEG artifact removal using ICA procedure.

are then replaced by zeros making  $\tilde{\mathbf{s}}$  become  $\tilde{\mathbf{s}}'$ . Multiplying the inverse of unmixing matrix with the new 'clean' set of ICs,  $\tilde{\mathbf{x}} = \mathbf{W}^{-1}\tilde{\mathbf{s}}'$ , it is possible to recover the clean, corrected EEG.

### 2.2.2 Information Maximization Approach (InfoMax): An ICA Algorithm

In [32], Bell and Sejnowski developed a neural network based algorithm named InfoMax that blindly separates mixtures  $\mathbf{x}$ , of independent sources  $\mathbf{s}$ , using information maximization principle. The core idea in InfoMax algorithm is to maximize the joint entropy,  $H[g(\mathbf{s})]$  where  $g()$  is a sigmoid function, by using a stochastic gradient ascent approach [1, 32, 33]. The goal is to minimize the mutual information between the independent component  $s_i$ . InfoMax updates weights of the ICA unmixing matrix  $\mathbf{W}$  by using the gradient of entropy which is defined as:

$$\Delta \mathbf{W} \propto \frac{\partial H(\mathbf{y})}{\partial \mathbf{W}} = E \left[ \frac{\partial \ln |J|}{\partial \mathbf{W}} \right] \quad (2-8)$$

where  $y_i = g(u_i) = \frac{1}{1+e^{-u_i}}$ ,  $E$  denotes the expected value,  $\mathbf{y} = [g(u_1) \dots g(u_N)]^T$  and  $|J|$  is the absolute value of the determinant of the Jacobian matrix:

$$J = \det \left[ \frac{\partial y_i}{\partial x_j} \right]_{ij} \quad (2-9)$$

where  $i, j$  get values between 1 and size of  $\mathbf{W}$ . The learning rule defined by (7) can be defined in the more succinct formula as

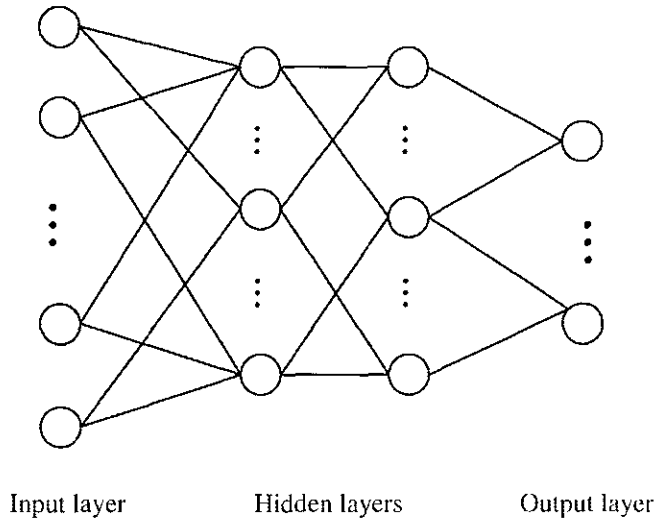
$$\Delta \mathbf{W} \propto [\mathbf{W}^T]^{-1} + \tilde{\mathbf{y}} \mathbf{x}^T \quad (2-10)$$

where  $\tilde{\mathbf{y}} = [\tilde{y}_1 \dots \tilde{y}_N]^T$ , which has each element defined by

$$\tilde{y}_i = \frac{\partial}{\partial y_i} \frac{\partial y_i}{\partial u_i} = \frac{\partial}{\partial u_i} \ln \frac{\partial y_i}{\partial u_i} \quad (2-11)$$

### 2.3 Artificial Neural Network

An artificial neural network (ANN) is often mistaken as a biological term because of its origin and ANN was inspired from the structure of biological neural networks. In the framework of this thesis, ANNs are regarded as mathematical models consisting of neural units called neurons. Those neurons, at different layers, are linked to others via a network of inter-connections. Normally, an ANN contains three main classes of layers: input layer, hidden layers and output layer (Figure 2.5). The number of input and output layer always equals one, although the number of hidden layers may be more than one and may vary based on application. The more hidden layers and number of hidden units the ANN has, the more complex it is.



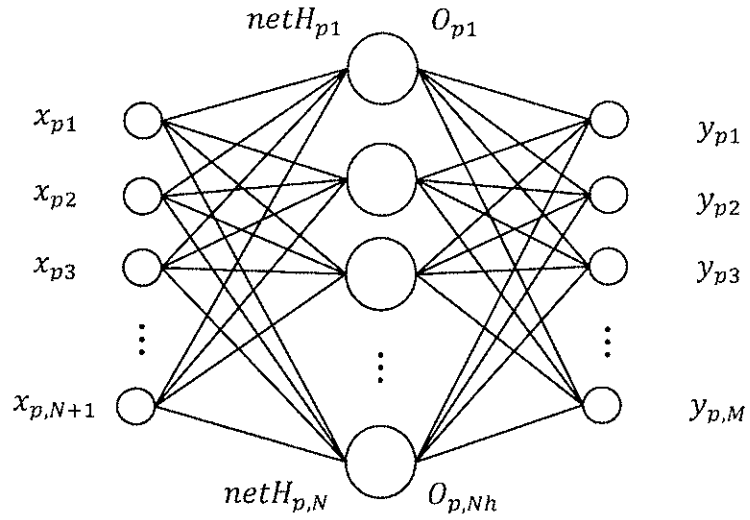
**Figure 2.5.** General artificial neural network structure.

An ANN can be characterized based on neuron properties, network structure, and learning schemes [48]. Neuron properties include the number of input and output units,

the weight associated with each input and output unit, and the activation function used for the hidden layer units. Network structure is the way to connect neurons together in order to establish an approximation net. If there are associations between all the neurons from the input layer to the hidden layer, and from the hidden layer to the output layer, then such an ANN is regarded as fully-connected. Learning schemes, either supervised or unsupervised, are the algorithms used to initialize and adjust weights among individual neurons from different layers. A general structure of ANN is illustrated as in Figure 2.5.

### 2.3.1 Output Weight Optimization Backpropagation

Output weight optimization backpropagation (OWO – BP) [41, 53] is a supervised learning technique, which employs backpropagation (BP) [55-56] and output weight optimization (OWO) to adapt a feed-forward neural network, whose structure is shown in Figure 2.6, to the real model by minimizing the training square error. BP algorithm updates the hidden weights, which connect input units  $x_p$  with hidden units  $netH_p$ . Meanwhile, OWO is capable of solving linear equations for output weights, which connect output activations  $O_p$  with output units  $y_p$ . The output activation is calculated from hidden unit output from a sigmoid function. OWO-BP computes the bypass weights from input layer to output layer as well, yet is not shown here for simplicity.



**Figure 2.6.** Fully-connected one hidden layer MLP structure.

The training square error for the  $i^{th}$  output can be written as,

$$E(i) = \frac{1}{N_v} \sum_{p=1}^{N_v} [\tilde{y}_p(i) - y_p(i)]^2 \quad (2-12)$$

where  $y_p(i)$  and  $\tilde{y}_p(i)$  are the  $i^{th}$  actual and corresponding desired output units.  $N_v$  is the number of training patterns. The actual output is computed as

$$y_p(i) = \sum_{j=1}^{N_o} w_c(i,j) \tilde{x}_p(j) \quad (2-13)$$

where,  $w_c(i,j)$  is the output weight from  $j^{th}$  basis function unit to  $i^{th}$  output unit and  $N_o$  is the number of basis functions. The basis functions are defined as

$$\tilde{x}_p(j) = \begin{cases} x_p(j), & 1 \leq j \leq N \\ O_p(j - N - 1), & N + 2 \leq j \leq N + N_h + 1 \\ 1, & j = N + 1 \end{cases} \quad (2-14)$$

Consequently, the first derivative of  $E(i)$  with respect to  $w_c(i,j)$  is given by

$$\frac{\partial E(i)}{\partial w_c(i,j)} = -2 \left[ \rho(i,j) - \sum_{p=1}^{N_v} w_c(i,j) \cdot \alpha(i,j) \right] \quad (2-15)$$

where,  $N_o = N + N_h + 1$ .  $\alpha(i, j)$  and  $\rho(i, j)$  are the auto-correlation and cross-correlation matrices, respectively. They are defined as

$$\alpha(i, j) = \sum_{p=1}^{N_v} \tilde{x}_p(i) \cdot \tilde{x}_p(j) \quad (2-16)$$

$$\rho(i, j) = \sum_{p=1}^{N_v} \tilde{y}_p(i) \cdot \tilde{x}_p(j) \quad (2-17)$$

The optimized weight  $w_c(i, j)$  is obtained by setting  $\frac{\partial E_i}{\partial w_c(i, j)}$  to zero, or equivalently

$$\sum_{j=1}^{N_o} w_c(i, j) \cdot \alpha(i, j) = \rho(i, j), \quad 1 \leq i \leq M \quad (2-18)$$

where M is the number of outputs. Solving the equation (2-18) is the approach to get the optimized weight  $w_c(i, j)$ , which is equivalent to minimizing  $E(i)$  [53].

## 2.4 Motivation of the Work

Up to this point, a general theory of both wavelet transform and ICA has been presented. Also, the ideas behind using wavelet transform, or more specifically a wavelet thresholding method using adaptive SURE risk threshold function and ICA with the InfoMax as a representative, were summarized. ICA is effective in multi-type artifact removal and should be a good benchmark method for educational purpose in EEG-related research. As previously mentioned, ICA is time-consuming due to its computational complexity. Being a batch algorithm, ICA needs to be performed on the whole data set with at least an adequate number of data points (which is often equivalent to some hours of EEG data recorded by multi-channel EEG recording systems). Instead of using the whole data set, wavelet thresholding can be performed EOG artifact correction on a single channel data. This reduces much time needed for artifact correction. However, the selection of wavelet is sensitive to the time-frequency properties of EEG waves.

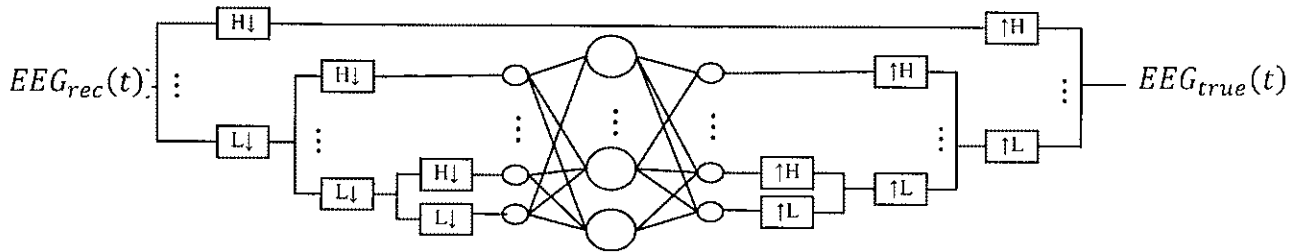


Specifically, results from the experiments that were conducted during this work show that using only one specific wavelet or even various wavelets from one mother wavelet for various EEG data segments usually gives out unacceptable correction results. Also, the threshold function selection still needs more investigation to improve this method. Obviously, ICA is not a desirable method for real-time processing applications. In this thesis, a solution is being sought that is applicable to an online EOG artifact removal task. Recognizing the universal approximation of both neural network and wavelet-based methods, the idea of combining them should be very attractive. However, the large questions are how to combine them properly and what should be the proper manner to apply the new method on EOG artifact removal. These issues will be addressed in the thesis.

## CHAPTER 3

### PROPOSED METHOD

In this thesis, a novel algorithm, Wavelet Neural Network, for EEG artifact removal is presented. In this method, the WNN is trained with simulated data resembling the properties in both time and frequency domains of an EEG signal. The trained WNN is then used as the corrector for contaminated data. In both testing and training processes, the original signal is decomposed first with a wavelet to get different frequency components. The low frequency subband coefficients are then interpolated to maintain same lengths. A trained artificial neural network (ANN) is fed with such interpolated inputs to yield the corrected coefficients at its outputs. Finally, the corrected coefficients are downsampled for the wavelet construction to get the corrected signal of original contaminated signal. The structure of the proposed wavelet neural network is illustrated in Figure 3.1.



**Figure 3.1.** Proposed wavelet neural network structure.

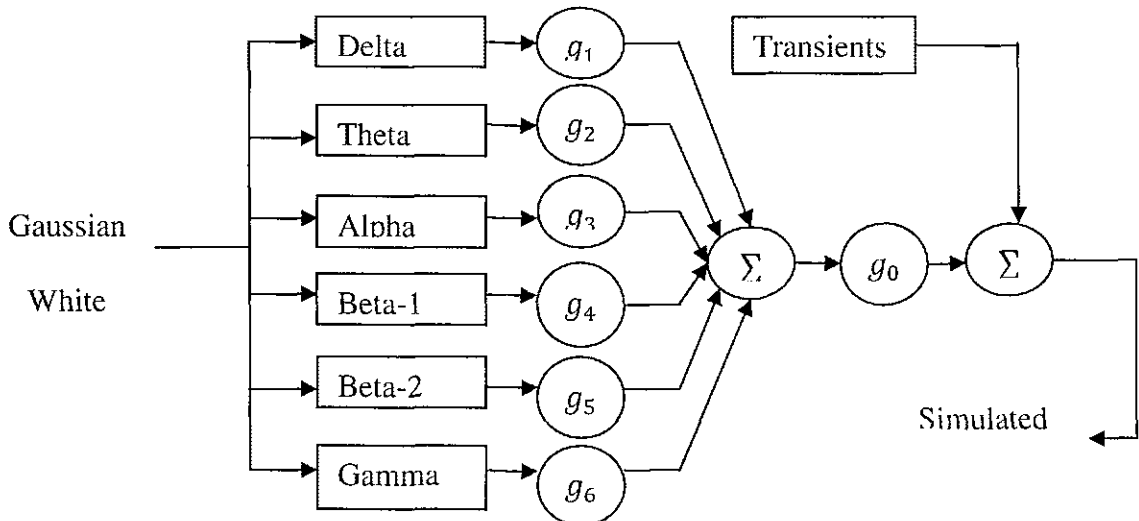
The core idea of the method, decomposing the signal in both time and frequency domains with wavelet and using an ANN to correct them, can be viewed as follows. By

combining the time/frequency properties of wavelet and the universal approximation capability of neural network, useful information can be retained that is related to cognitive activities while eliminating artifacts in EEG.

### 3.1 EEG Data Simulation

It has been summarized in [24] that a desirable EEG model should (1) give a better understanding of brain function, (2) provide testing tool for novel EEG-related method, and (3) be useful for medical education practice and training. There have been a number of EEG simulating models [34, 36, 37, 38]. The utilized model [24] was originally designed for medical research in pharmacodynamics, the psychological science of effects caused by drug on humans [24, 34]. This EEG simulated model is used for the purpose of training and testing the proposed WNN method as well as testing other algorithms like wavelet thresholding on EEG ocular artifact removal [35].

As described in [24], an EEG signal can be simulated based on three assumptions,



**Figure 3.2.** Simulated EEG generator model.

(1) short segments of the spontaneous EEGs can be described as linearly filtered, (2) non-stationary components in the spontaneous EEG can be simulated by changing the characteristics of this filtering process and (3) the spectral property of the simulated EEG data resembles that of actual signal. As shown in Figure 3.2, a set of Gaussian noises (GN) were generated and then filtered by a number of lowpass and bandpass filters with different cut-off frequencies (Table 3.1) that are similar to the spectral property of EEG frequency bands. The filtered signals are multiplied with different gains ( $g_i, i = 1, \dots, 6$ ) corresponding to its EEG rhythmic properties. These EEG simulated rhythmic signals are then summed. That synthetic clean simulated signal is multiplied with an overall gain  $g_0$  to get an amplitude corresponding to that of actual EEG. After all, transients such as blinking and other eye movement, collected from real signals were then filtered by lowpass filters and added to contaminate the simulated data.

**Table 3.1.** EEG frequency band specifications.

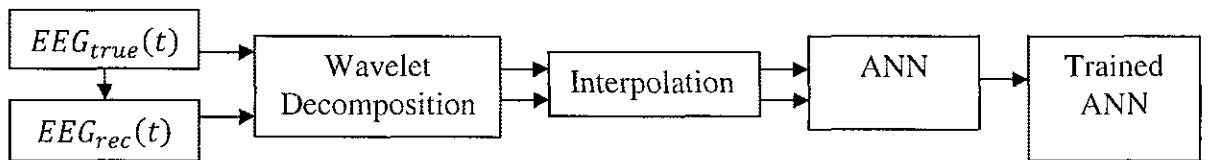
Frequency bands	Lower (Hz)	Upper (Hz)
Delta	0.5	4
Theta	4	8
Alpha	8	13
Beta-1	13	20
Beta-2	20	30
Gamma	30	50

The model (Figure 3.2) slightly differs from the model presented in [24] in the last step. Instead of adding the transients before multiplying with the overall gain  $g_0$ , it is done right after obtaining the clean simulated EEG signal.

## 3.2 Neural Network for EEG Artifact Removal

### 3.2.1 Training

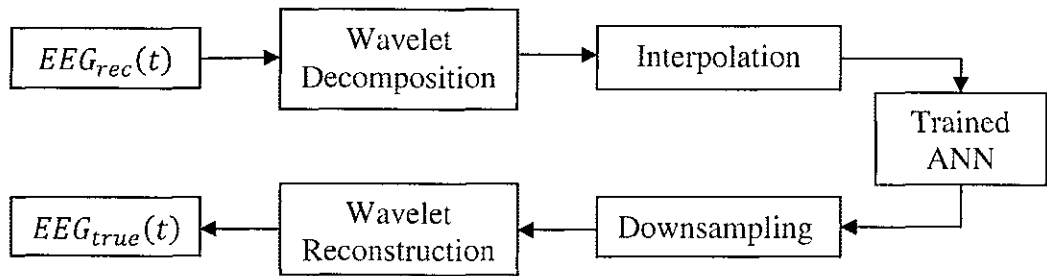
For training purposes, first, the clean simulated EEG is generated using an EEG generator model without adding transients. That clean EEG is then contaminated by adding real processed artifacts (transients). By doing that, two separate EEG signals can be collected that own almost all identical time-frequency properties. The only difference is with the contamination issue. Looking back at the EEG model described earlier in formula (1-1), clean and contaminated signals can be considered as  $EEG_{true}(t)$  and  $EEG_{rec}(t)$  respectively. The network training problem is equivalent to teach the NN how to recover the  $EEG_{true}(t)$  from  $EEG_{rec}(t)$ . Both  $EEG_{true}(t)$  and  $EEG_{rec}(t)$  are decomposed with wavelet decomposition. The approximation and detail coefficients at several low-frequency sub-bands are interpolated in order to obtain equal length series. These modified coefficient series are then combined as a training dataset. The training dataset then is passed through an ANN using an iterative machine learning technique for training, as described in Section 2.3.1. Finally, the trained ANN is obtained. A diagram of the presented process is given at Figure 3.3.



**Figure 3.3.** Neural network training procedure.

### 3.2.2 Testing and Artifact Removal

For testing and real data artifact removal purposes, the beginning steps are overlapped with those described in ANN training. The  $EEG_{rec}(t)$  is decomposed by a wavelet transform. Coefficients at a low frequency sub-band are then interpolated and combined together as testing data. The testing data are then passed over the trained ANN. The corrected data after ANN correction are then down-sampled to the original length and returned to original wavelet coefficient series for wavelet reconstruction. Figure 3.4 illustrates the testing/correcting procedure.



**Figure 3.4.** EEG corrected by using WNN.

### 3.3. Validation Methods

Three metrics are used, power spectrum density (PSD), MSE/RMSE and frequency correlation, to assess the proposed method. The PSD is a popular metric used to show information about the power spectrum of EEG signal at specific frequencies. Calculation of the correlation in frequency domain before and after artifact removal is equivalent to the correlation in time domain after filtering the time series with the corresponding

frequency filter [14]-[24]. The frequency correlation between  $\tilde{x}$  and  $\tilde{y}$  is computed as in the following formula,

$$c = \frac{\frac{1}{2} \sum_{w_1}^{w_2} (\tilde{x}^* \tilde{y} + \tilde{y}^* \tilde{x})}{\sqrt{\sum_{w_1}^{w_2} \tilde{x} \tilde{x}^* \sum_{w_1}^{w_2} \tilde{y} \tilde{y}^*}} \quad (3-1)$$

where  $w_1$  and  $w_2$  are the lower and upper limits of the interested power spectrum region to be calculated,  $c$  is the correlation value that will be assigned to the frequency of  $(w_1+w_2)/2$ . If  $\tilde{x}$  and  $\tilde{y}$  are identical,  $c$  gets 1, otherwise,  $c$  obtains a value between 0 and 1. In this thesis, the ‘window size’,  $w_1-w_2$ , is selected equal to 2.

The value of MSE/RMSE shows the difference between the corrected and original non-contaminated signal. This value is proportional to the accuracy of the method used for correction.

Mean square error (MSE) is defined by

$$MSE = \frac{1}{N_v} \sum_{t=1}^{N_v} (EEG_{rec}(t) - EEG_{true}(t))^2 \quad (3-2)$$

where  $N_v$  is the length of the contaminated EEG segment. Root mean square error (RMSE) is defined by

$$RMSE = \sqrt{\frac{1}{N_v} \sum_{t=1}^{N_v} (EEG_{rec}(t) - EEG_{true}(t))^2} \quad (3-3)$$

## CHAPTER 4

### RESULTS AND DISCUSSIONS

In this chapter, the experimental results obtained using the proposed algorithm are shown and the performance of the algorithm was assessed by comparing it with other algorithms.

#### 4.1 Datasets

The method was validated on two datasets recorded in a driving test [35] and a visual selection task experiment (for more information, see [http://sccn.ucsd.edu/~arno/fam2data/publicly\\_available\\_EEG\\_data.html](http://sccn.ucsd.edu/~arno/fam2data/publicly_available_EEG_data.html)).

##### 4.1.1 Driving Test Dataset

This dataset was collected when participants were performing a driving test [35]. The EEG information was collected by a 128-channel recording system at the sampling rate of 1000 Hz along with other information including description of the task, system dynamics related information, performance measures, physiological signals (ECG, respiration, etc.), and eye tracking. The workload was also analyzed according to the driving conditions (city-driving, stopped, highway passing, etc.). Due to the recording condition, the subject eye movements and blinks happen at high frequencies making the data, especially at frontal recording channels, highly contaminated by ocular artifacts.



#### 4.1.2 Visual Selection Task Dataset

This dataset was recorded by a 32-channel recording system at sampling rate of 128 Hz during the course of 238.3125 seconds while the subject participated in a selective attention task, where the subject was asked to attend to circles flashed in random order at one of five displayed locations [1]. Even the design of the task helps restrict the eye movements and blinks, but generally, the data are still highly contaminated by the ocular artifacts, which are dominant at EEG recorded from the electrodes F3, Fz, F4, etc, located in the frontal sites of the subject head.

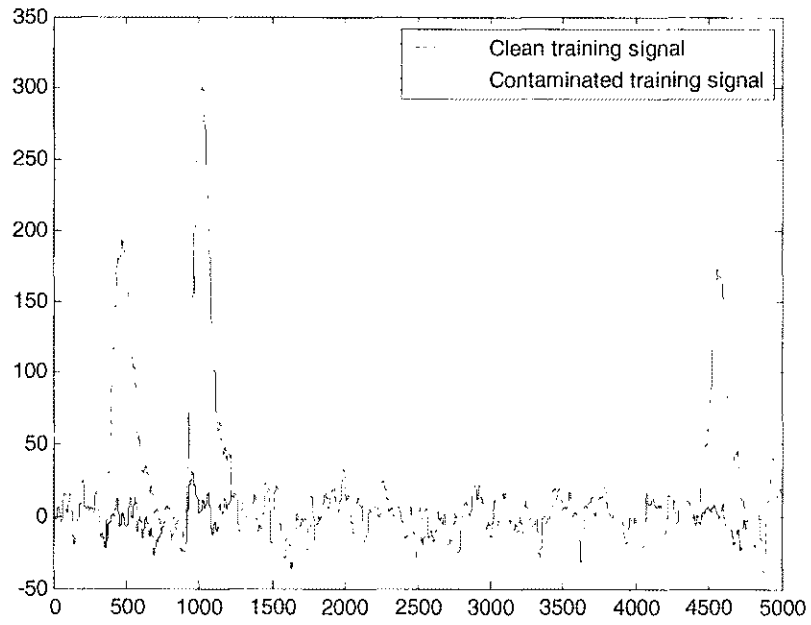
#### 4.1.3 Experimental Settings

For each dataset, three artifact removal methods were implemented for comparison: the ICA method, the wavelet thresholding algorithm and the proposed WNN technique. For each algorithm, a PSD and frequency correlation were computed before and after artifact removal to illustrate the effectiveness of each algorithm. The MSE/RMSE was used for method accuracy comparison between WNN and wavelet thresholding. For the proposed method, an EEG signal was first simulated to train an ANN and the trained WNN was tested on each of the data sets. The wavelet thresholding method was implemented by following the instructions in [14] and for the ICA, EEGLAB software [21] downloaded from <http://sccn.ucsd.edu/eeqlab/downloadtoolbox.html> was used.

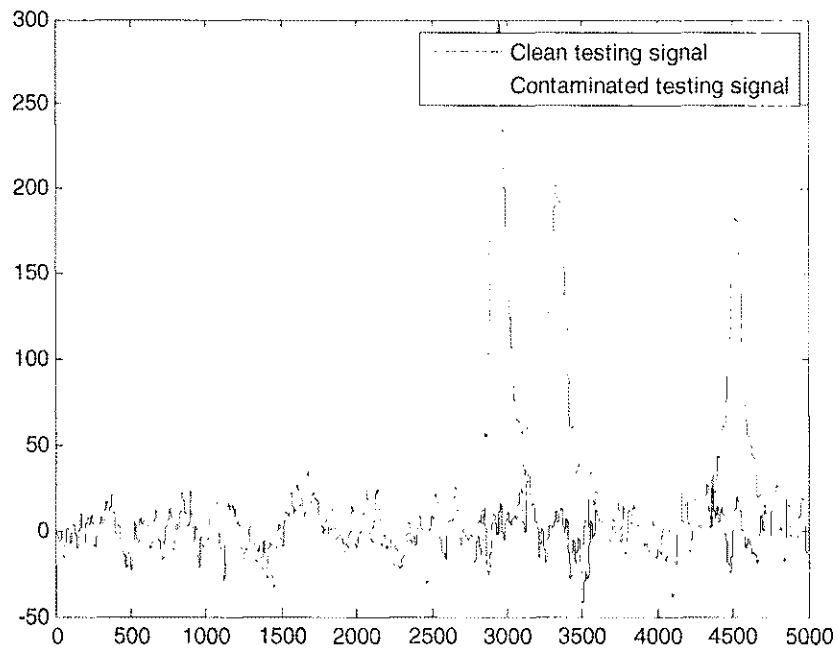
## 4.2 Results for the Driving Test Dataset

### 4.2.1 Results on Simulated Data

For the proposed WNN algorithm, two simulated segments with a length of 5 seconds, created by the simulation model described in Section 3.1, for training and testing at sampling rate of 1000 Hz are displayed in Figure 4.1 a) and b), respectively. The artifacts were taken from the driving test data set and added to the simulated data segments. Data in Figure 4.1 a) were then used to train the neural network in the proposed WNN algorithm. The trained WNN model was applied to the testing data segment (Figure 4.1 b)). The simulated signal was decomposed with wavelet Coif3 up to 8 levels. The three lowest frequency sub-band coefficients were then corrected by the trained ANN and used for wavelet reconstruction. The corrected EEG signal is shown in Figure 4.2. Figure 4.3 shows PSD of the contaminated, corrected and clean EEG signals both for neural network training and testing. Figure 4.4 shows frequency correlations among those simulated signals. As mentioned previously, the ICA needs to be performed on the whole dataset so in this case that only single channel data is available and ICA is not applicable.

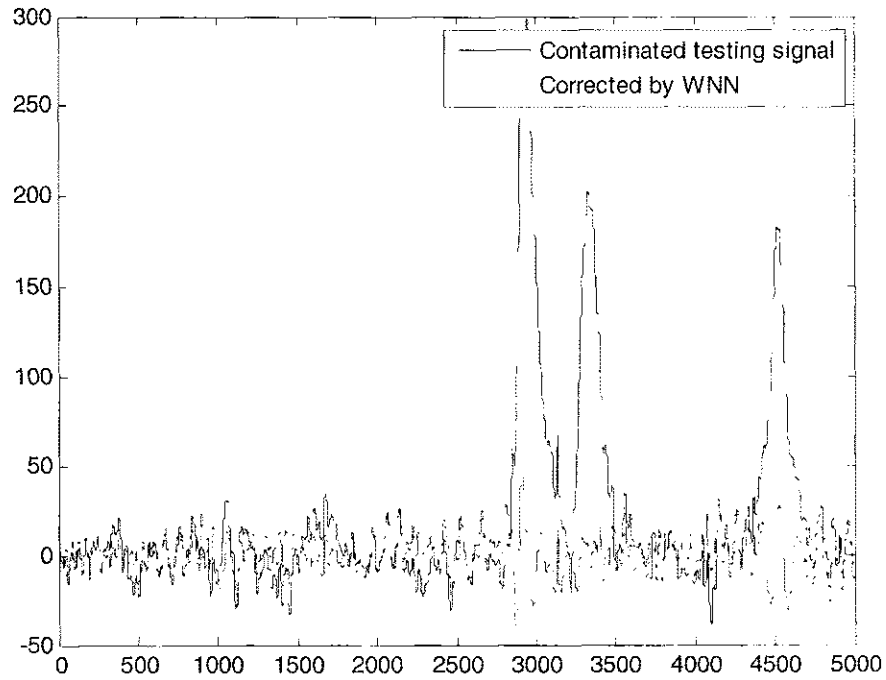


(a)



(b)

**Figure 4.1.** Clean and contaminated simulated signal for (a) training and (b) testing.



**Figure 4.2.** Contaminated simulated and WNN corrected singals.

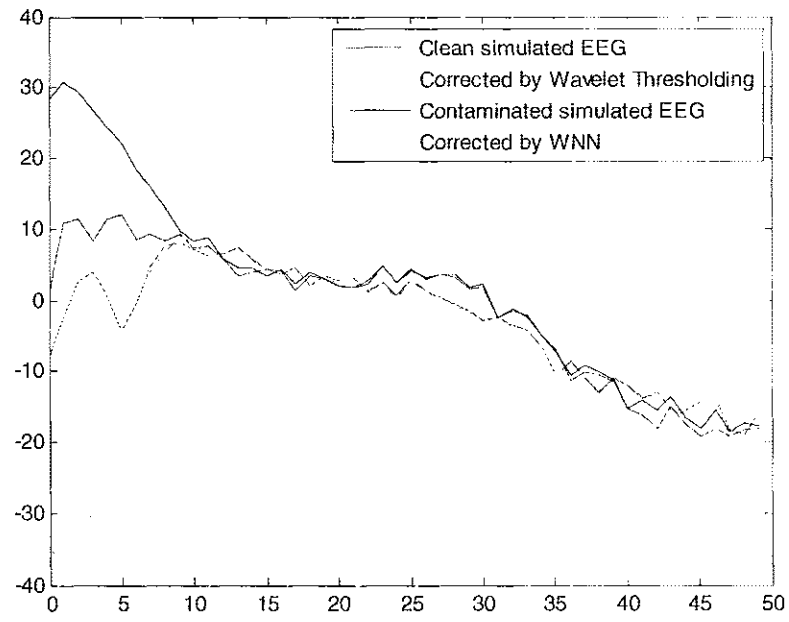
**Table 4.1.** Training and testing (a) MSE and (b) RMSE of signals before and after correction for driving test dataset.

MSE	Training	Testing
WNN	126.3122	149.9971
Wavelet Thresholding	262.2498	272.0597

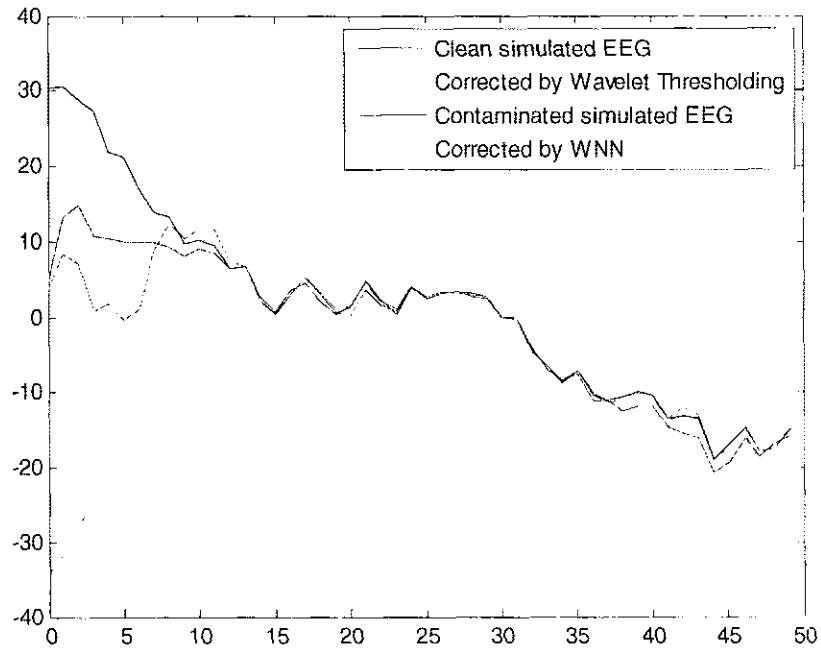
(a)

RMSE	Training	Testing
WNN	11.2389	12.2473
Wavelet Thresholding	16.1941	16.4942

(b)

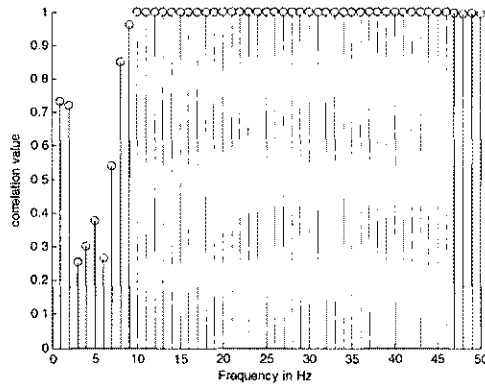


(a)

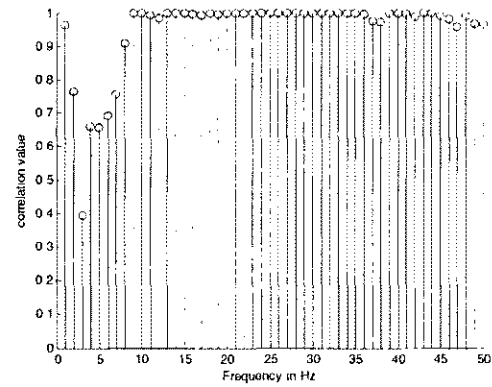


(b)

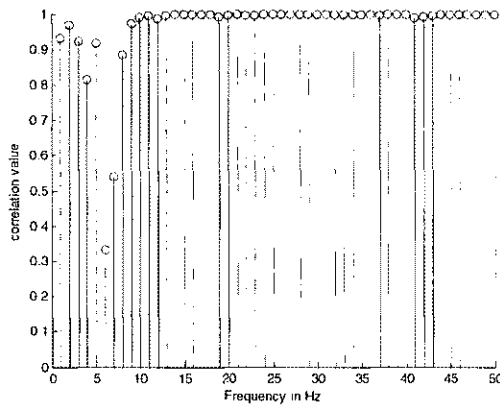
**Figure 4.3.** PSD of clean, contaminated and WNN corrected signals for (a) training and (b) testing.



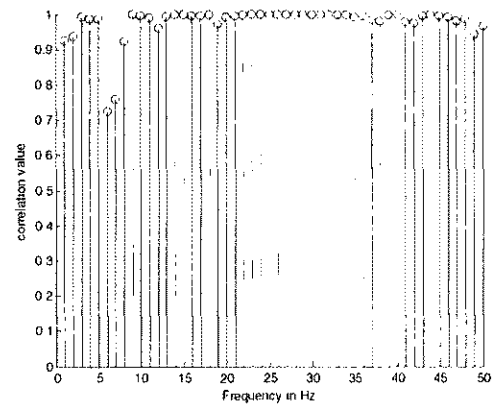
(a)



(b)



(c)



(d)

**Figure 4.4.** Frequency correlation between (a) contaminated and wavelet thresholding corrected simulated signals and (b) clean and wavelet thresholding corrected simulated signals (c) contaminated and WNN corrected simulated signals and (d) clean and WNN corrected simulated signals, all for testing.

Table 4.1 shows the difference between the original non-contaminated and decontaminated signals by WNN and wavelet thresholding via the metric MSE/RMSE.

#### 4.2.2 Results on Real Data

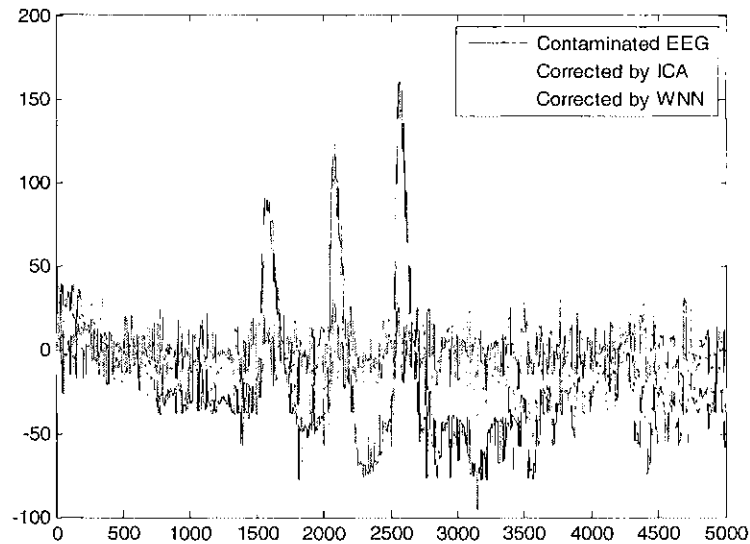
Due to the recording condition, signals from some channels with bad-connections to the scalp were deteriorated and not usable. Data for those channels (25, 47, 48, 112, 125 and 126) were removed from the original dataset. The resized dataset contains data from 122 channels, each of 60 second or equivalent to the total of 7,320,000 data points. Three methods were applied, WNN, ICA and wavelet thresholding, on this resized dataset.

For the WNN technique, the EEG signals were decomposed to 8 levels with wavelet Coif3. Then, three low frequency sub-band coefficients were corrected. Those corrected coefficients were then used to reconstruct the decontaminated signals.

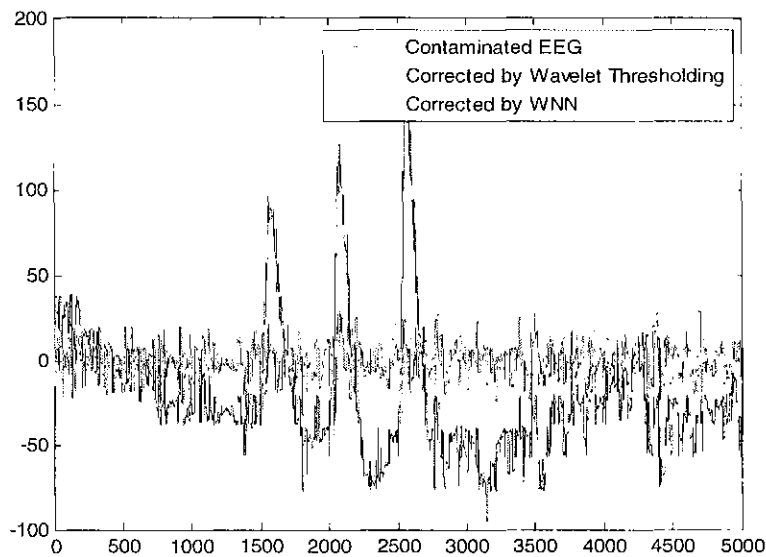
For Infomax ICA, it took a computer, equipped with Intel(R) Core(TM) 2 CPU 6400 @ 2.13 GHz and RAM 2.00 GB, 27 minutes with 382 steps to yield the unmixing matrix. After that, the independent components (ICs) were obtained by multiplying the unmixing matrix with the original mixtures. After visually inspection, the artifactual ICs: 1→7, 10, 11, 13, 14, 16→20, 23, 27, 29, 33, 40, 44, 45, 49 and 57 were made zero. The corrected signal was obtained by multiplying the inverse version of unmixing matrix with the 'new' IC set.

The wavelet thresholding method was used to adaptively correct four low frequency sub-band coefficients. For specific data segments, the corrections were repeated a number of times with various wavelets and at different levels of decomposition in order to make the corrected data most acceptable. The wavelets from

the Coiflet and Daubechies family were chosen because experiments show that they could extract the features of artifacts efficiently.



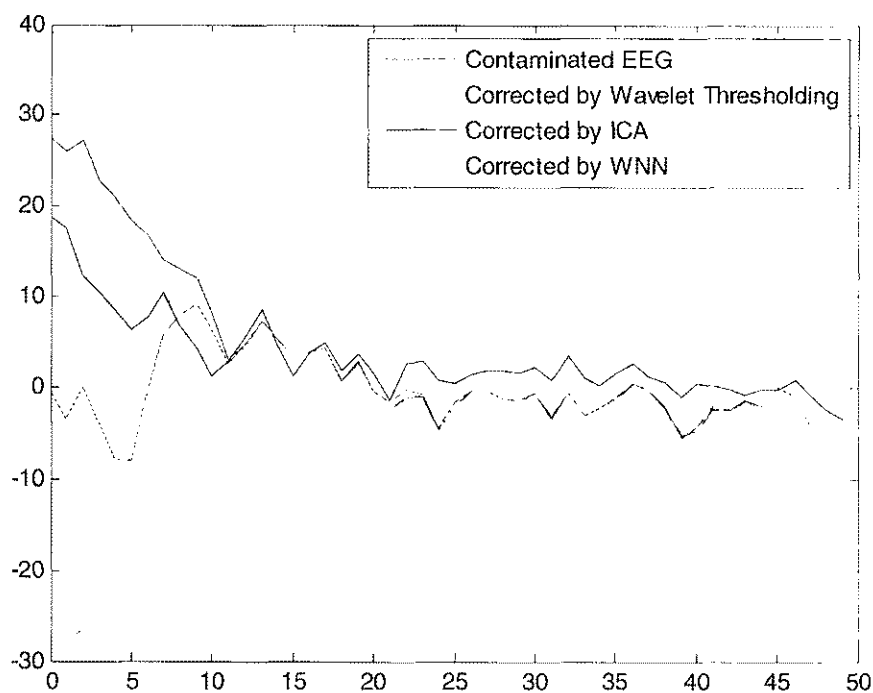
(a)



(b)

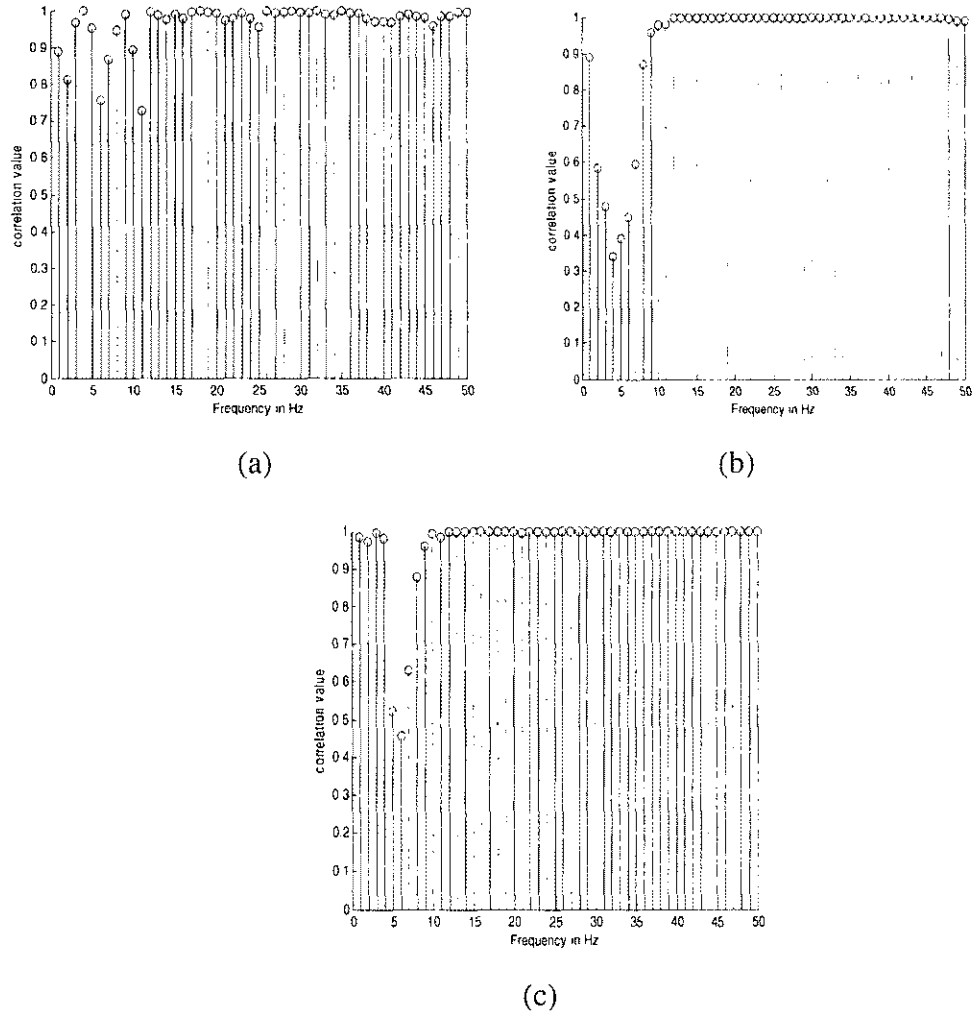
**Figure 4.5.** Contaminated and decontaminated EEG (a) contaminated, ICA and WNN corrected EEG (b) contaminated, wavelet thresholding and WNN corrected EEG.





**Figure 4.6.** PSD of contaminated and de-contaminated EEG.

Figure 4.5 shows a segment with three spike artifacts and its corrected versions made by various methods in time domain. The results given by three methods confirm their ability on artifact removal. Figure 4.6 shows PSD plots for one sample artifact removed segment in the driving test data by the three algorithms. Figure 4.7 shows frequency correlations between the contaminated and corrected segments.



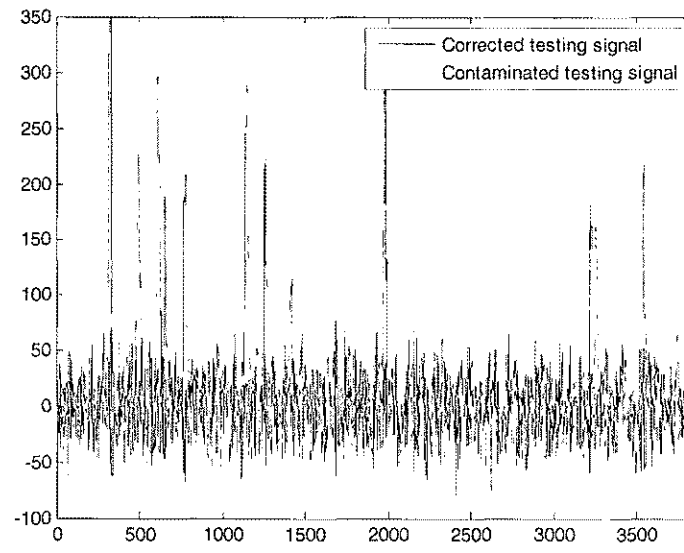
**Figure 4.7.** Frequency correlation between contaminated and decontaminated EEG, (a) by ICA, (b) by wavelet thresholding and (c) by WNN.

### 4.3 Results for the Visual Selection Task Dataset

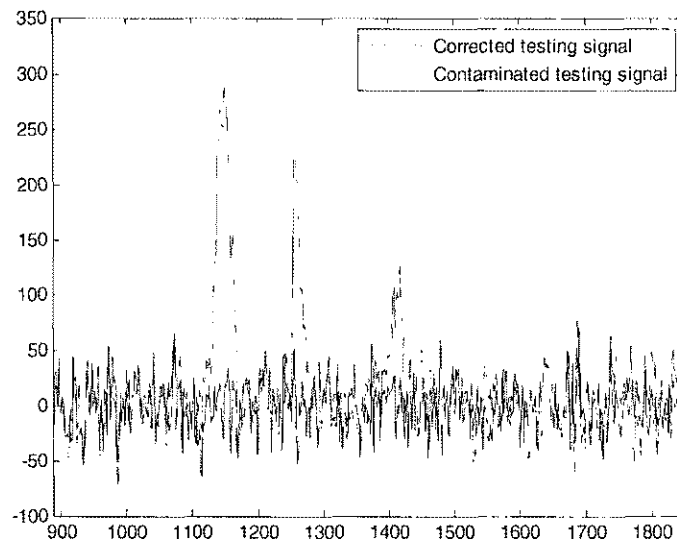
#### 4.3.1 Results on Simulated Data

For the WNN training and testing, the simulated data segments of length 30 seconds, equivalently 3,840 samples, were generated at the sampling rate of 128 Hz. Then the simulated data were contaminated with artifacts taken from real data at channels number 1 and 3, or FP1 and FP3 equivalently. After being decomposed by wavelet

transform with wavelet Coif3 to six levels, the four sub-band low frequency coefficients were passed through the ANN of structure 4-6-4 for training purposes.

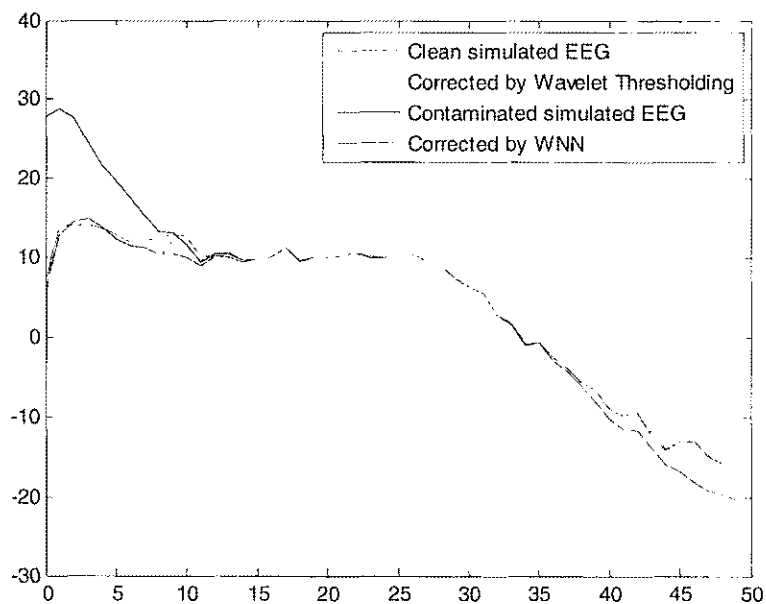


(a)

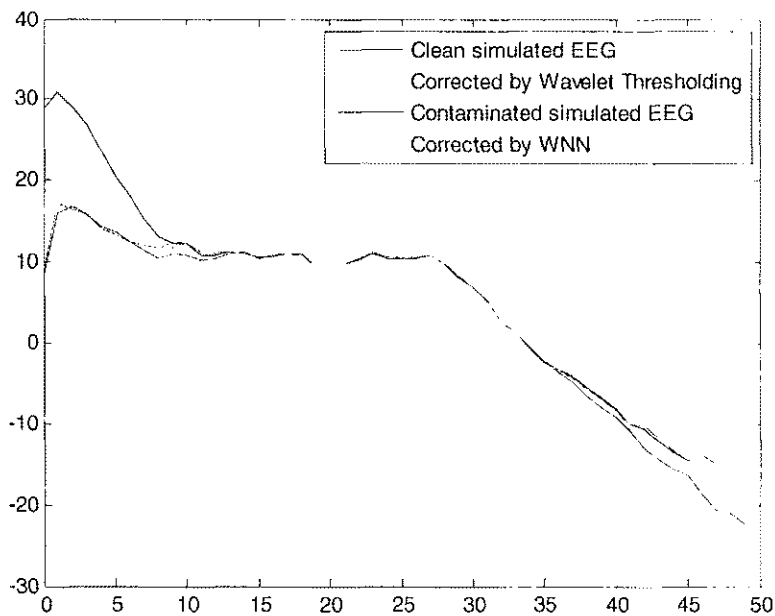


(b)

**Figure 4.8.** Contaminated simulated and WNN corrected singals (a) 30 seconds and (b) 7.4219 seconds.

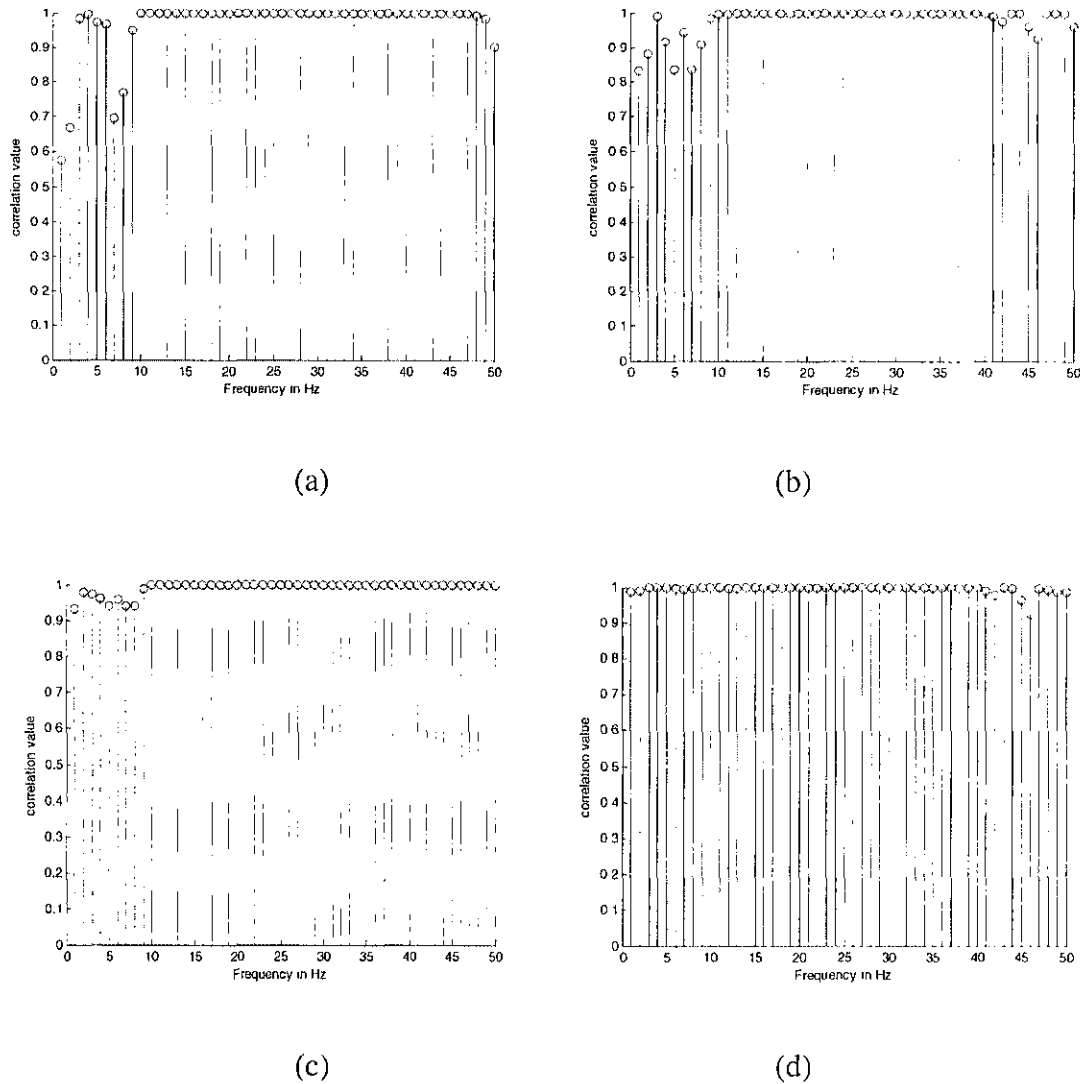


(a)



(b)

**Figure 4.9.** PSD of clean, contaminated and WNN corrected signals for (a) training and (b) testing.



**Figure 4.10.** Frequency correlation between (a) contaminated and wavelet thresholding corrected simulated signals and (b) clean and wavelet thresholding corrected simulated signals (c) contaminated and WNN corrected simulated signals and (d) clean and WNN corrected simulated signals, all for testing.

Figure 4.8 a) shows the visual appearance of simulated EEG data (30 seconds), which is contaminated by numerous artifacts, before and after it has been corrected. The zoom-in result displayed in Figure 4.8 b) indicates that the WNN just focuses on removing

artifacts across contaminated data segments; meanwhile, the non-contaminated information is very well-preserved. Figure 4.9 shows PSD of clean, contaminated, wavelet-thesholding-corrected and WNN-corrected simulated signals used for both training and testing the ANN. The PSDs given by WNN are corrected and the clean simulated EEG signals are almost overlapped. That is promising because it demonstrates that the WNN correction performed very well in terms of frequency property. Figure 4.10 shows the frequency correlation of the contaminated and clean simulated signals with their corrected versions by using WNN and wavelet thresholding. The frequency correlation (Figure 4.10 d)) shows negligible differences between clean and WNN corrected simulated signals, which affirms the previous discussion drawn from the PSD plot.

#### **4.3.2 Results on Real Data**

The real EEG data segments of 30 seconds, or 3840 samples, were decomposed with wavelet Coif3 to six levels. Then, the four low frequency sub-band coefficients were interpolated and fed to the trained WNN using the Neural Network. The WNN outputs corrected coefficients, which were down-sampled and then used to reconstruct the corrected EEG signal.

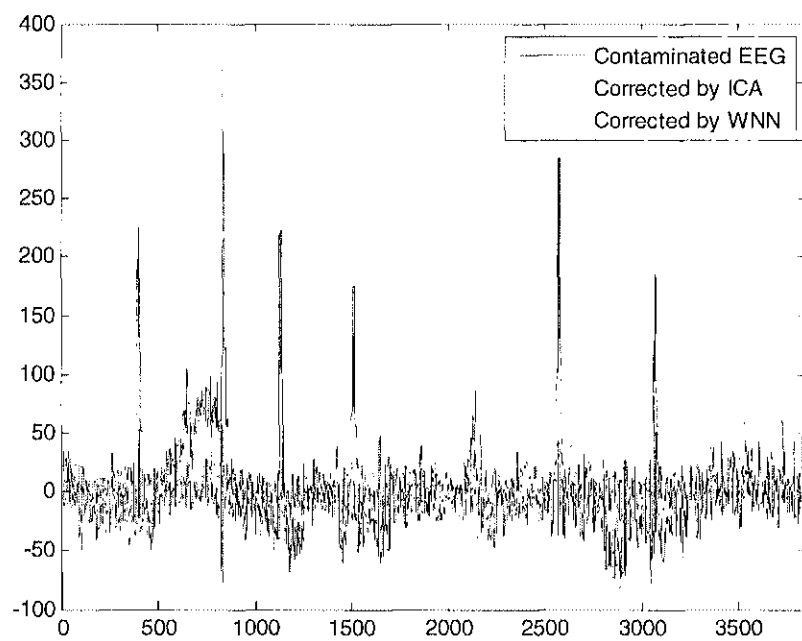
The ICA correction was realized by using the same computer as in the driving test experiment. Infomax ICA was implemented with EEGLAB. Almost three minutes with 275 steps were needed to obtain the unmixing matrix of size 32 x 32. By careful visual inspection combined with IC topographic mapping method, which was realized by using

the recording system descriptions given along with the dataset, the ICs number 1, 2, 4→7 and 22 were made zero. After multiplying the new zero-valued and other non-zero valued ICs with the mixing matrix the corrected EEG data were obtained.

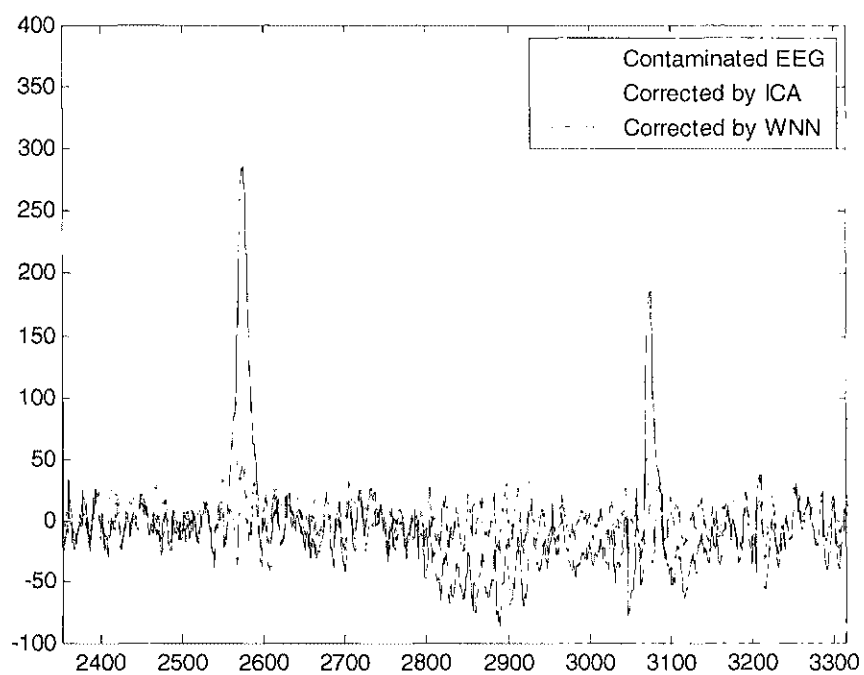
The Wavelet thresholding method was implemented with wavelet Coif4 with 6 levels of composition. Coefficients at four low frequency sub-bands: 0-2, 2-4, 4-8 and 8-16 Hz were corrected by the adaptive threshold function.

Figure 4.11 a-c) shows the visual appearance of a real EEG data segment (30 seconds), which was contaminated by numerous artifacts, before and after correction. Specifically, the artifacts were removed efficiently by WNN (Figure 4.11 b-d)). Figure 4.12 shows the PSD of signal before and after correction by all three methods. The frequency correlation was shown in Figure 4.13. WNN and wavelet thresholding correct the contaminated signal in the low frequency range. Meanwhile, ICA performed the corrections in the entire useful EEG frequency range.

MSE/RMSE , as displayed in Table 4.2, shows the difference between the original non-contaminated and corrected signals made by WNN and wavelet thresholding for both the training and testing data.

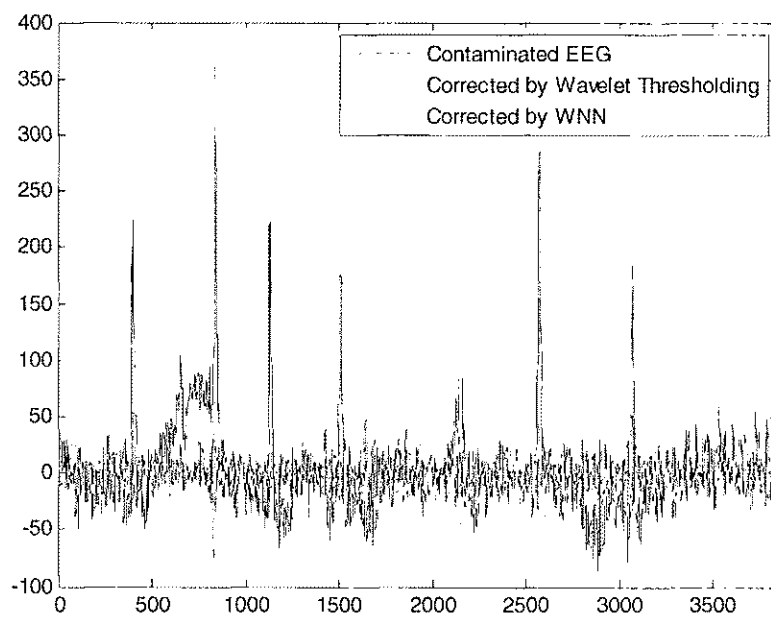


(a)

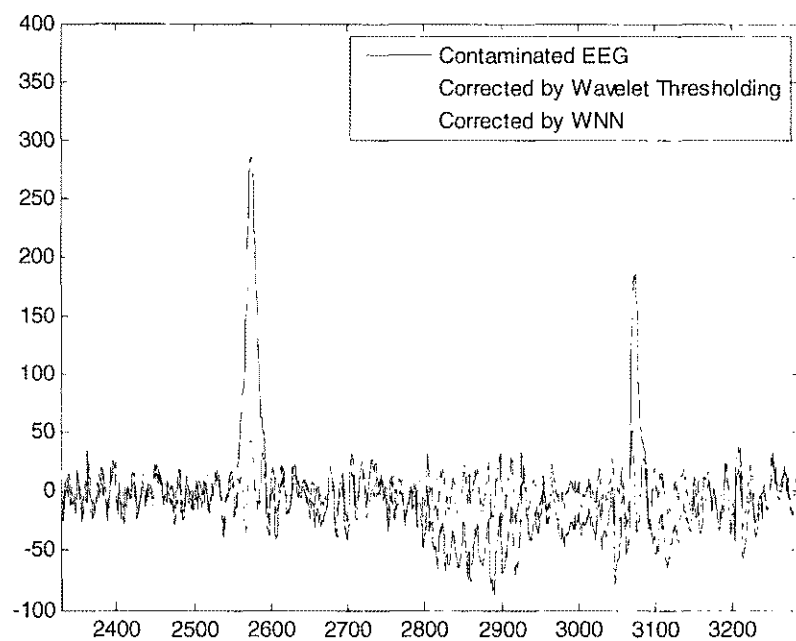


(b)



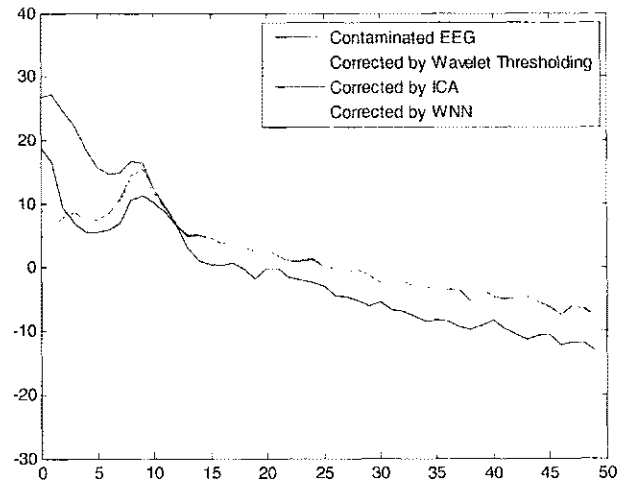


(c)



(d)

**Figure 4.11.** Contaminated and de-contaminated EEG (a-b) by ICA and WNN and (c-d) by wavelet thresholding and WNN.



**Figure 4.12.** PSD of contaminated and de-contaminated EEG.

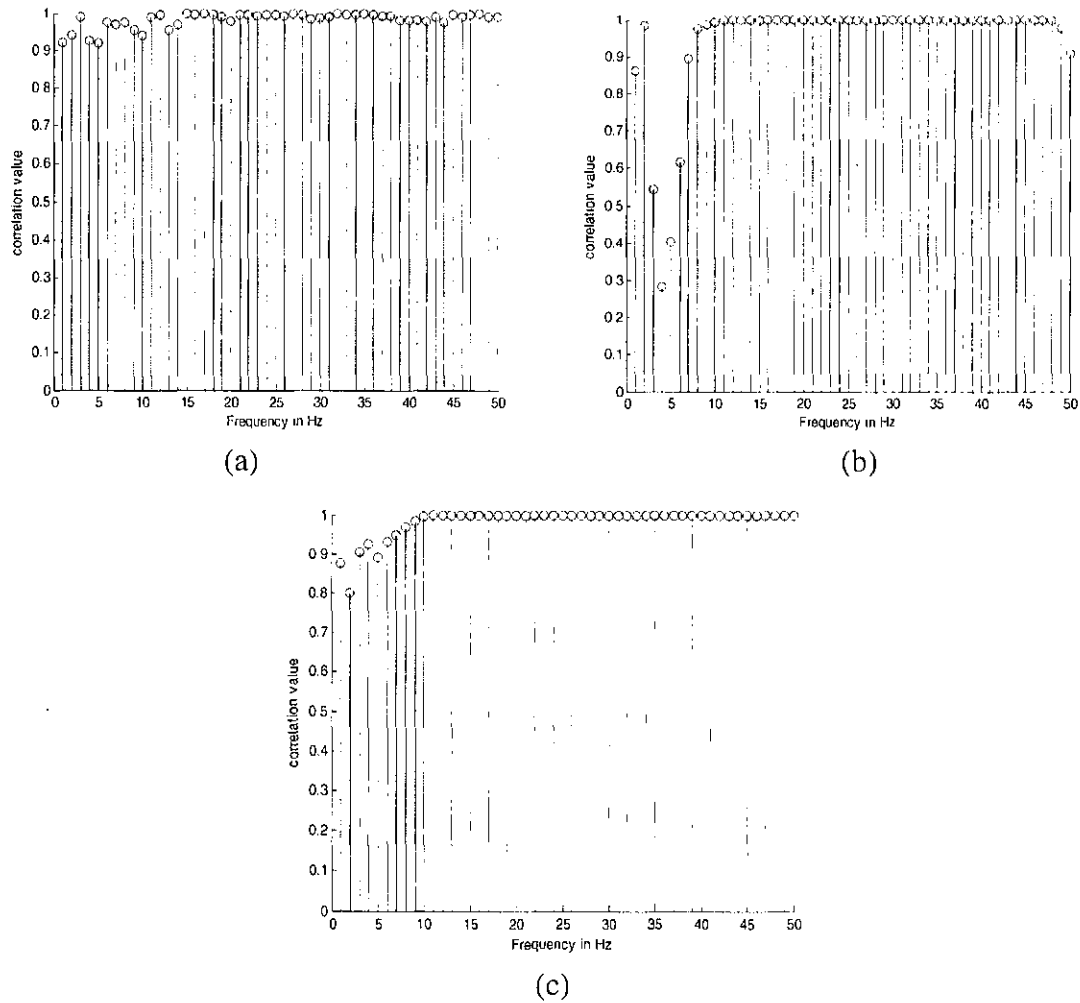
**Table 4.2.** Training and testing (a) MSE and (b) RMSE of signal before and after correction for visual selection task dataset.

MSE	Training	Testing
WNN	204.0771	369.2298
Wavelet Thresholding	571.0926	830.4674

(a)

MSE	Training	Testing
WNN	14.2856	19.2154
Wavelet Thresholding	23.8975	28.8178

(b)



**Figure 4.13.** Frequency correlation between contaminated and decontaminated EEG by (a) ICA, (b) wavelet thresholding and (c) WNN.

#### 4.4 Discussions

It has been observed from various results that the WNN algorithm removed ocular artifacts efficiently while keeping cerebral background information. Like wavelet thresholding, WNN just needs one single channel data to perform correction which makes more advantageous than ICA, which needs to perform on the whole dataset. Furthermore, the method was proved through repeated experiments on various data

segments for its effectiveness and stability, which is not true for the wavelet thresholding algorithm.

The PSD plot shows that the low frequency components were reduced significantly in the corrected signal. That is more evident when looking at the frequency correlation metric plot between contaminated and corrected signals. There are slight differences in the range of low frequency components while in other ranges, the useful information is well-preserved.

The frequency correlation plots also show that the correction made by ICA spreads over the entire frequency range and the power of low frequency components are reduced not significantly. Meanwhile, the low frequency components in the signal were derogated by wavelet thresholding and WNN while high frequency components are well preserved by both.

Besides, the MSE/RMSE of the correction to the simulated data made by WNN is smaller than that of the correction by wavelet thresholding, which implies the WNN is more accurate.

ICA requires a lot more computing power and multiple channel data sources for artifact removal. It also requires either an automatic or a manual step to determine which independent component is artifact, making an online implementation of ICA difficult. Meanwhile, as mentioned previously, WNN can be performed on signals recorded from only a single channel without reference to EOG recording. In addition, the processing time for each data segment correction made by WNN is negligible. Thus, by using a

sliding window across the entire contaminated EEG data, it is possible for a WNN to perform the artifact removal in real time.

In the experiment for WNN correction, the Coiflet 3 basis functions were used throughout all the implementations. The number of wavelet stages (eight for driving data set and six for visual selection task data set) were also fixed for each dataset even when the method is implemented on a large number of data segments with various shapes and time-frequency properties. The selection of the wavelet stage can be determined on the base of sampling rate without reference to any other specific characteristics of the contaminated EEG segments themselves.

## **CHAPTER 5**

### **CONCLUSIONS AND FUTURE WORK**

A novel algorithm was proposed: a wavelet neural network capable of removing EEG artifacts. The algorithm combines the time/frequency property of wavelet and the approximating capability of neural network to locate and eliminate artifacts. Experimental results on driving and visual task selection datasets show that WNN can effectively remove artifact and achieve better results than the wavelet thresholding algorithm. WNN is also much computationally efficient than the ICA algorithm making it possible an automatic online algorithm.

Future work includes (1) using more data for training, (2) enhancing the EEG data simulating techniques, (3) finding the best-fitted wavelet and number of sub-bands to feed the neural network and (4) applying more advanced machine learning techniques. A WNN can be extended to solve the problem of removing other EEG artifacts like muscle or cardiac artifacts if the time-frequency properties of these artifacts are known. Finally, WNN is applicable to other 1-D and 2-D signal processing issues like noise cancellation in speech, medical image processing and digital communication.

## REFERENCES

- [1] T. Jung, S. Makeig, C. Humphries, T. Lee, M. Mckeown, V. Iragui, and T. J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, Vol. 37, pp. 163-178, 2000.
- [2] E. Urretarazu, J. Iriarte, M. Alegre, M. Valencia, C. Vireri, and J. Artieda, "Independent component analysis removing artifacts in ictal recordings," *Epilepsia*, Vol. 45, pp. 1071-1078, 2004.
- [3] D. A. Pizzagalli, "Electroencephalography and High-Density Electrophysiological Source Localization", in Cacioppo, J.T. Tassinary, G.G. Bernston, *Handbook of Psychophysiology* (3<sup>rd</sup> Edition), Cambridge University Press, Cambridge, U. K., pp. 56 – 84, 2007.
- [4] D. Hagemann, and E. Naumann, "The effects of ocular artifacts on (lateralized) broadband power in the EEG," *Clinical Neurophysiols*, Vol. 112, pp. 215-231, 2001.
- [5] C. Joyce, I. Gorodnitsky and M. Kutas, "Automatic removal of eye movement and blink artifacts from EEG data using blind component separation," *Psychophysiology*, Vol. 41, pp. 313-325, 2004.
- [6] R. J. Croft and R. J. Barry, "Removal of ocular artifact from the EEG: a review," *Clinical Neurophysiols*, Vol. 30, pp. 5-19, 2000.
- [7] L. Shoker, S. Sanei and M. A. Latif, "Artifact removal from electroencephalograms using a hybrid BSS-SVM algorithm," *IEEE signal Processing Lett*, Vol. 12, no. 10, pp. 721-724, 2005.

- [8] G. Gratton, M. G. Coles and E. Donchin, "A new method for off-line removal of ocular artifact," *Electroencephalogram Clin Neurophysiol.* Vol. 55, pp. 84-468, 1983.
- [9] J. C. Woestenburg, M. N. Verbaten and J. L. Slangen, "The removal of the eye movement artifact from the EEG by regression analysis in the frequency domain," *Biological Psychology*, Vol. 16, pp. 47-127, 1983.
- [10] P. Berg, and M. Scherg, "Dipole models of eye activity and its application to the removal of eye artifacts from the EEG ad MEG," *Clinical Physics and Physiological Measurements*, vol.12 (Supplement A), pp. 49–54, 1991.
- [11] T. D. Lagerlund, F. W. Sharbrough, and N. E. Busacker, "Spatial filtering of multichannel electroencephalographic recordings through principal component analysis by singular value decomposition," *Journal of Clinical Neurophysiology*, Vol. 14, No. 1, pp. 73-82, 1997.
- [12] P. Common, "Independent Component Analysis, A new concept?," *Signal Processing*, Vol. 36, pp. 287-314, 1994.
- [13] A. Delorme, S. Makeig and T. Sejnowski, "Automatic artifact rejection for EEG data using high-order statistics and independent component analysis," in *Proceedings of the Third International ICA Conference*, Dec 9-12, San Diego, 2001.
- [14] V. Krishnaveni, S. Jayaraman, L. Anitha and K. Ramadoss, "Removal of ocular artifacts from EEG using adaptive thresholding of wavelet coefficients," *Journal of Neural Engineering*, Vol. 3, pp. 338-346, 2006.
- [15] K. P. Indiradevi, E. Elias, P. S. Sathidevi, S. D. Nayak and K. Radhakrishnan, "A multi-level wavelet approach for automatic detection of epileptic spikes in the



- electroencephalogram,” *Journal of Computers in Biology and Medicine*, Vol. 38, No. 7, pp. 805-816, 2008.
- [16] L. Senhadji, J-L Dillenseger, F. Wendling, C. Rocha, and A. Kinie, “Wavelet analysis of EEG for three-dimensional mapping of epileptic events,” *Ann Biomedical Engineering*, Vol. 23(5), pp. 543-552, 1995.
- [17] W. Zhou and J. Gotman, “Removal of EMG and ECG artefacts from EEG based on wavelet transform and ICA,” in *Proceedings of 26th Annual International Conference IEEE Engineering in Medicine and Biology Society (EMBS'04)*, San Francisco, CA, Sep. 1–5, pp. 392–395, 2004.
- [18] X. P. Zhang and M. D. Desai “Adaptive denoising based on SURE risk,” *IEEE Signal Processing Lett*, Vol. 10, No. 5, pp. 265 – 267, Oct. 1998
- [19] V. J. Samar, A. Bopardikar, R. Rao, and K. Swartz, “Wavelet analysis of neuroelectric waveforms: a conceptual tutorial,” *Brain and Language*, Vol. 66, pp. 7-60, 1999.
- [20] Daubechies Ingrid, *Ten lectures on wavelets*, Philadelphia: Soc. Ind. Applied Math., 1992.
- [21] A. Delorme and S. Makeig, “EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis,” *Journal of Neuroscience Methods*, Vol. 134 , pp. 9-21, 2004.
- [22] A. Oonsivilai and M. E. El-Hawary, “Wavelet neural network based short term load forecasting of electric power system commercial load,” in *Proceedings of*

- IEEE Canadian Conference in Electrical and Computer Engineering*, Vol. 3, pp. 1223-1228, May 1999.
- [23] N. A. de Beer, W. L. van Meurs, M. B. Grit, M. L. Good and D. Gravenstein, "Educational simulation of the electroencephalogram (EEG)," *Technol Health Care* 2001, Vol. 9, pp. 237-256, 2001.
- [24] A. Jung, *Statistical analysis of biomedical data*, Ph.D. thesis, University of Regensburg, Dec. 2003.
- [25] O. Rioul and M. Vetterli, "Wavelet and signal processing," *IEEE Signal Processing Magazine*, Vol. 8, No. 4, pp. 14 – 38, 2002.
- [26] T. Kalayci and O. Ozdamar, "Wavelet preprocessing for automated neural network detection of EEG spikes," *IEEE Engineering in Medicine and Biology Magazine*, Vol. 14, No. 2, pp. 160-166, 1995.
- [27] O. A. Rosso, M. T. Martin, A. Figliola, K. Kelle and A. Plastino, "EEG analysis using wavelet-based information tools," *Journal of Neuroscience methods*, Vol. 153, No. 2, pp 163-182, 2006.
- [28] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transaction on Information Theory*, Vol. 41, pp. 613-627, 1995.
- [29] S. G. Chang, B. Yu and M. Vetterli, "Adaptive Wavelet thresholding for image denoising and compression," *IEEE Transaction on Image Processing*, Vol. 9, No. 9, pp. 1532-1546, 2000.
- [30] C. M. Stein, "Estimation of the mean of a multivariate normal distribution," *The Annals of Statistics*, Vol. 9, No. 6, pp. 1135-1151, 1996.

- [31] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *Journal of the American Statistical Association*, Vol. 90, pp. 1200-1224, 1995.
- [32] J. Bell, and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation*, Vol. 7, pp. 1129-1159, 1995.
- [33] J. Bell, and T. J. Sejnowski, "the independent components of natural scenes are edge filters," *Vision Research*, Vol. 37, pp. 3327-3338, 1997.
- [34] P. Lees, F. M. Cunningham and J. Elliott, "Principles of pharmacodynamics and their applications in veterinary pharmacology," *Journal of Veterinary Pharmacology & Therapeutics*, Vol. 27, No. 6, pp. 397-414, 2004.
- [35] H.-A. T. Nguyen, J. Musson, J. Li, F. McKenzie, G. Zhang, R. Zu, C. Richey and T. Schell, "EEG artifact removal using a Wavelet Neural Network," in *Proceedings of ModSim World*, 2010.
- [36] O. Bai, M. Nakamura, A. Ikeda and H. Shibasaki "Nonlinear Markov Process Amplitude EEG model for nonlinear coupling interaction of spontaneous EEG," *IEEE Transaction on Biomedical Engineering*, Vol. 47, No. 9, pp. 1141-1146, 2000.
- [37] D. Ghahremani, S. Makeig, T. P. Jung, A. J. Bell and T. J. Sejnowski, "Independent component analysis of simulated EEG using a three-shell spherical head model," *Technical report series*, UCSD, 1996.
- [38] J. P. Kaipio and P. A. Karjainen, "Simulation of nonstationary EEG," *Biological Cybernetics*, Vol. 76, No. 5, pp. 349-356, 1994.

- [39] P. Berg and M. Scherg, "A multiple source approach to the correction of eye artifacts," *Electroencephalography and clinical Neurophysiology*, Vol. 90, No. 3, pp. 229-241, 1994.
- [40] J. L. Whitton, F. Lue and H. Moldofsky, "A spectral method for removing eye movement artifacts from the EEG," *Electroencephalography and Clinical Neurophysiology*, Vol. 44, No. 6, pp. 735 – 741, 1978.
- [41] C. Yu, M. T. Manry, J. Li and P. L. Narasimha "An efficient hidden layer training method for the multilayer perceptron," *Neurocomputing*, Vol. 70, No. 1-3, pp. 525–535, 2006.
- [42] R. I. Goldman, J. Stern, J. Engel and M. Cohen, "Simultaneous EEG and fMRI of the alpha rhythm," *Neuroreport*, Vol. 13, No. 18, pp. 2487-2492, 2002.
- [43] W. H. R. Miltner, C. Braun, M. A. H. Witte and E. Taub, "Coherence of gamma-band EEG activity as a basis for associative learning," *Nature*, Vol. 397, pp. 434 – 436, 1999.
- [44] W. D. Clercq, A. Vergult, B. Vanrumste, W. V. Paesschen, and S. V. Huffel, "Canonical correlation analysis applied to remove muscle artifacts from the Electroencephalogram," *IEEE Transactions on Biomedical Engineering*, Vol. 53, No. 12, pp. 2583 – 2587, 2006.
- [45] S. Romero, M. A. Mailanas, S. Clos, S. Gimenez, and M. J. Barbanoj, "Reduction of EEG Artifacts by ICA in Different Sleep Stages," in *Proceedings of the 25<sup>th</sup> Annual International Conference of the IEEE EMBS*, Vol. 3, pp. 2675-2678, 2003.

- [46] J. Sijbers, J. V. Audekerke, M. Verhoye, A. V. Der Linden, and D. V. Dyck, "Reduction of ECG and gradient related artifacts in simultaneously recorded human EEG/MRI data," *Magnetic Resonance Imaging*, Vol. 18, pp. 881-886, 2000.
- [47] J. N. Knight, *Signal fraction analysis and artifact removal in EEG*, master thesis, Colorado State University, 2003.
- [48] D. J. Livingstone, *Artificial Neural Networks Methods and Applications: Methods and Application*, Humana Press, 2009.
- [49] L. Sherlin, *Diagnosing and treating brain function through the use of Low Resolution Electromagnetic Tomography (LORETA)*, Book Chapter in Introduction to quantitative EEG and neurofeedback, advanced theory and applications(2nd ed.). New York: Elsevier.
- [50] K. S. Sri and J. C. Rajapakse, "Extracting EEG rhythms using ICA-R," in *Proceedings of IEEE International Joint Conference On Neural Networks*, pp. 2133-2138, 2008.
- [51] H. Bland and S. D. Oddie, "Theta band oscillation and synchrony in the hippocampal formation and associated structures: the case for its role in sensorimotor integration," *Behavioural Brain Research*, Vol. 127, No. 1-2, pp. 119-136, 2001.
- [52] J. P. Rosenfeld, A. M. Reinhart and S. Srivastava, "The Effects of Alpha (10-Hz) and Beta (22-Hz) "Entrainment" Simulation on the Alpha and Beta EEG Bands: Individual differences are critical to prediction of effects," *Applied Psychophysiology and Biofeedback*, Vol. 22, No. 1, pp. 1023 - 1034, 1997.

- [53] J. Li, M. T. Manry, L.-M. Liu, C. Yu and J. Wei, "Iterative improvement of neural classifiers," in *Proceedings of the Seventeenth International Conference of the Florida AI Research Society*, pp. 700-705, 2004.
- [54] J. C. Woestenburg, M. N. Verbaten and J. L. Slangen, "The removal of the eye movement artifact from the EEG by regression analysis in the frequency domain," *Biological Psychology*, Vol. 14, pp. 73-82, 1997.
- [55] D. E. Rumelhart, G. E. Hinton and R. J. Williams, "Learning internal representation by error propagation," *Parallel Distributed Processing*, Vol. 1, pp. 318-362, 1986.
- [56] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

**HOANG-ANH T NGUYEN**

Department of Electrical and Computer Engineering  
Old Dominion University  
757-275-5624  
hnguy025@odu.edu

---

**Education:**

**Master of Science in Electrical and Computer Engineering**  
Old Dominion University, Norfolk, VA. (December 2010)

**Bachelor of Science in Electronics and Computer Engineering**  
Vietnam National University. (May 2008)

**Research Publications:**

- Hoang-Anh T. Nguyen, John Musson, Guangfan Zhang, Roger Xu, Carl Richey, Tom Schnell, Frederic McKenzie and Jiang Li, 'EEG artifact removal using a wavelet neural network,' MODSIM WORLD, Hampton, VA, 2010.