2018

Real time implementation of multichannel EEG cleaning from muscular artifacts in brain computer interfacing (BCI) paradigm

Ferdous, Mst. Jannatul

University of Rajshahi

UNIVERSITY OF RAJSHAHI

DOCTORAL THESIS

Real time implementation of multichannel EEG cleaning from muscular artifacts in brain computer interfacing (BCI) paradigm

Author:
MST. JANNATUL FERDOUS
Signal Processing & Computational Neuroscience LAB
Department of Computer Science & Engineering
Roll No.: 13712
Registration No.: 10032
Session: 2013-2014

Supervisor:
DR. MD. KHADEMUL ISLAM MOLLA
Cosupervisor:
DR. MD. EKRAMUL HAMID

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Computer Science and Engineering, University of Rajshahi, Rajshahi 6205, Bangladesh.

June 2018

University of Rajshahi

Department of Computer Science and Engineering

Rajshahi, Bangladesh

Certificate

This is certify that the Dissertation entitled **"Real time implementation of multichannel EEG cleaning from muscular artifacts in brain computer interfacing (BCI) paradigm",** submitted by **Mst. Jannatul Ferdous** is a record of original work carried out by him, in the partial fulfilment of the requirement for the Degree of Doctor of Philosophy at the Department of Computer Science and Engineering, University of Rajshahi. This work is done during July 2013 - June 2018, under our guidance. The contents of this report, in full or in parts, have not been submitted to any other Institute or University for the award of any degree.

Prof. Dr. Md. Khademul Islam Molla	Prof. Dr. Ekramul Hamid
Supervisor	**Co-supervisor**

Date:

Declaration of Authorship

I, Mst. Jannatul Ferdous, declare that this thesis titled, "Real time implementation of multichannel EEG cleaning from muscular artifacts in brain computer interfacing (BCI) paradigm" and the work presented in it are my own. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University.
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- where I have consulted the published work of others, this is always clearly attributed.
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- where the thesis is based on work done by myself, I have made clear exactly what I have contributed myself.

Signed: _____

Date: _____

"The brain is a wonderful organ; it starts working the moment you get up in the morning and does not stop until you get into the office."

Robert Frost

Abstract

Electroencephalogram (EEG) can reflect not only the activity which is specific for a cognitive task during an experiment but also the electrical background activity of the brain. The EEG signals recorded from the scalp surface are usually contaminated by different physiological signals which are termed artifacts. Among the artifacts the electrooculography (EOG) makes serious obstacle to many neuroscience experiments including the application for brain-computer interfaces (BCIs). It has noticeable higher energy at lower frequency compared to target EEG signals. This research presents a hybrid wavelet based algorithm to suppress the ocular artifacts from EEG signals. The hybrid wavelet transform (HWT) method is designed by the combination of discrete wavelet decomposition (DWT) and wavelet packet transform (WPT). The artifact suppression is performed by the selection of subbands obtained by HWT. Fractional Gaussian noise (fGn) is used as the reference signal to select the subbands containing the artifacts. The multichannel EEG signal is decomposed by HWT into a finite set of subbands. The energies of the subbands are compared with fGn of the desired subband signals. The EEG signal is reconstructed by the selected subbands consisting of EEG components. The experiments are conducted for both simulated and real EEG signals to study the performance of the proposed algorithm. The results are compared with recently developed artifacts suppression algorithms like stationary subspace analysis (SSA), discrete wavelet transform (DWT) and empirical mode decomposition (EMD). However, the existing EMD method has been successfully used in the field EEG artifact suppression using a data-adaptive subband filtering approach. But the higher computation burden of EMD processing is the main obstacle to online implementation of brain-computer interfacing (BCI). To resolve such limitation, the proposed HWT with higher computation speed is introduced in this study. In this thesis, BCI experiment is conducted to test the cleaning performance followed by the BCI classification with EEG signal. For the motor imagery EEG classification, linear discriminant analysis (LDA) is used. The experimental results prove that the classification performance increases noticeably with the cleaned EEG data, using the proposed algorithm. It is found that the proposed method performs better than the methods compared in terms of perfor-

mance metrics (signal to artifact ratio and mutual information), computational cost and classification accuracy. Therefore, this study shows that the proposed HWT based system performs better than other techniques for elimination of blink contamination from EEG signal.

Acknowledgement

First and foremost, all the praises will be to Almighty Allah for His endless blessings, mercy and guidance upon me. Thank to Allah for bestowing me with wisdom and sustaining me with countless supports throughout this study period.

I would like to express my utmost gratitude to my supervisor Professor Dr. Md. Khademul Islam Molla. Since the beginning of my research study, he has been guiding me with so much skills and care; especially I am quite thankful to him for his invaluable sharing of ideas, inspiration, advices, discussions and above all, great patience.

I am also very grateful to Professor Dr. Md. Ekramul Hamid for his continuous guidance and encouragement throughout the study. His insights, comments, critical assessments and ideas for improvements have always sharpened my views, and helped me to refine my works. I also express my thanks to Professor Dr. Somlal Das and Professor Dr. Bimal Kumar Paramanik for their supports. I would like to acknowledge the entire faculty member of the Department of Computer Science and Engineering for their kind help at various phases of this research.

I am thankful towards our research group members Md. Sujan Ali and others. For this dissertation I would like to thank my reading committee members for their time, interest, and helpful comments. I would also like to thank the other members of my oral defense committee for their time and insightful questions.

I gratefully acknowledge the funding sources that made my Ph.D. work possible. I was funded by the Information and Communication Technology (ICT) division of the ministry of Post, Telecommunication and Information Technology, Bangladesh of ICT fellowship for my first 3 years of Ph.D. tenure.

Lastly, I want to give my deepest thanks to my parents for whom I have come to this stage where I am now and special thanks to my better-half who always encouraged me and stood by my side during my difficult times.

Dedication

Dedicated to my parents especially my father who has always dreamed of my doctoral degree and therefore has provided for me the best education he could afford since beginning. On the other hand, dedication to my mother who has sacrificed everything in her life to raise me, educate me and suffer for me.

Contents

List of Figures

List of Tables

List of Acronyms/ Abbreviations

2T-EMD	Turning Tangent EMD
AAR	Automatic Artifact Removal
ADC	Analog to Digital Converter
ADJUST	Automatic EEG artifact Detector based on the Joint Use of Spatial and Temporal features
AMI	Auto-Mutual Information
ANC	Adaptive Noise Canceler
ANFIS	Adaptive Neural Fuzzy Inference System
ANN	Artificial Neural Network
APF	Adaptive Predictor Filter
ARX	Auto-Regressive eXogenous
AUC	Area Under Curve
BCI	Brain Computer Interface
BMI	Brain Machine Interface
BSS	Blind Source Seperation
CC	Correlation Coefficient
CCA	Canonical Corelation Analysis
CI	Confidene Interval
cICA	constrained Independent Component Analysis
CPU	Central Processing Unit
CRLS	Conventional Re-cursive Least Squares
CSP	Common Spatial Pattern
CWT	Continuous Wavelet Transform
DFT	Discrete Fourier Transform
DWT	Discrete Wavelet Transform

EAS	Ensemble Average Subtraction
ECG	Electrocardiography
ECoG	Electro Cortico Graphy
EEG	Electroenceohalography
EEMD	Enhanced Empirical Mode Decomposition
EMD	Empirical Mode Decomposition
EMG	Electro-myogram
EOG	Electro-oculogram
EP	Evoked Potentials
ERD	Event Related Desynchronisation
ERP	Event Related Potential
ERPLAB	Event Related Potential LAB
FASTER	Fully Automated Statistical Thresholding for EEG artifact Rejection
fBm	fractional Brownian motion
fGN	fractional Gaussian Noise
FLNN	Functional Link Neural Network
fMRI	functional Magnetic Resonance Imaging
FORCe	Fully Online and automated artifact Removal for BCI
FT	Fourier Transform
GUI	Graphical User Interface
HEOG	Horizontal Electro-oculogram
HMM	Hidden Markov Model
HOS	Higher Order Statistics
HWT	Hybrid Wavelet Transform
ICA	Independent Component Analysis
ICA-R	Independent Component Analysis with Reference
IMF	Intrinsic Mode Function
ITR	Information Transfer Rate
kNN	k Nearest Neighbors
LAMIC	Lagged Auto-Mutual Information Clustering
LCL	Lower Confidene Limit

LDA	Linear Discriminant Analysis
LMS	Least Mean Squares
MARA	Multiple Artifact Rejection Algorithm
MATLAB	Matrix Laboratory
MCA	Morphological Component Analysis
MEG	Magneto Encephalo Graphy
MI	Motor Imagery
MuI	Mutual Information
MLP	MultiLayer Perceptron
MSE	Mean Squared Error
MSEC	Multiple Source Eye Correction
MSSA	Multivariate Singular Spectrum Analysis
NIRS	Near Infrared Spectroscopy
NMSE	Normalized Mean Square Eror
NN	Neural Networks
OA	Ocular Artifacts
OSET	Open-Source Electrophysiological Toolbox
PCA	Principal Component Analysis
PSD	Power Spectral Density
RAM	Random Access Memory
REOG	Radial Electro-oculogram
RFLDA	Regularized Fisher's LDA
RLS	Recursive Least Squares
RMSE	Root Mean Square Error
RRMSE	Relative Root Mean Square Error
SAR	Signal to Artifact Ratio
SCA	Sparse Component Analysis
SCICA	Spatially Constrained ICA
SCP	Slow Cortical Potentials
SMC	Sensory Motor Cortex
SNR	Signal to Noise Ratio
SOBI	Second Order Blind Identification

SOS	Second Order Statistics
SP	Spike Potentials
SRLS	Stable Re-cursive Least Squares
SSA	Stationary Subspace Analysis
SSpA	Singular Spectrum Analysis
SSEP	Steady State Evoked Potentials
SST	Synchrosqueezed Transform
SSVEP	Steady State Visual Evoked Potentials
STFT	Short Time Fourier Transform
SURE	Stein's Unbiased Risk Eestimate
SVD	Singular Value Decomposition
SVM	Support Vector Machine
SqWT	Synchrosqueezed Wavelet Transform
SWT	Stationary Wavelet Transform
TDSEP	Temporal Decorrelation Source Separation
TM	Trademark
UCL	Upper Confidene Limit
VEOG	Vertical Electro-oculogram
WNN	Wavelet Neural Network
WPT	Wavelet Packet Transform
WT	Wavelet Transform

List of Symbols

$\boldsymbol{\delta}$: Delta rhythm

$\boldsymbol{\theta}$: Theta rhythm

$\boldsymbol{\alpha}$: Alpha rhythm

$\boldsymbol{\mu}$: Mu rhythm

$\boldsymbol{\beta}$: Beta rhythm

$\boldsymbol{\gamma}$: Gamma rhythm

$\boldsymbol{\sigma}$: Standard deviation

$\boldsymbol{\lambda}$: Multiplication factor

$\mathbf{a}$: Normal vector

$\mathbf{b}$: Intercept

$\mathbf{l}$: Scaling factors

$\mathbf{m}$: Filter pairs

$\mathbf{x}$: Feature vector

$\mathbf{w}$: Projection matrix

$\mathbf{A}$: Weight vector

$\mathbf{C}$: Covariance matrix

$\mathbf{H}$: Hurst exponent

$\mathbf{K}$: Total number of commands

$\mathbf{N}$: Signal length

$\mathbf{W}$: Weight vector

Chapter 1

Introduction

The electrical activity of human brain captured by electroencephalography (EEG) exhibits a significantly complex behavior with strong non-linear and dynamic properties. The communication in the brain cells take place through electrical impulses. It is measured by placing the electrodes on the scalp of the subject. The cortical nerve cell inhibitory and excitatory postsynaptic potentials generate the EEG signals. These postsynaptic potentials summate in the cortex and extend to the scalp surface where they are recorded as EEG. A typical EEG signal, measured from the scalp, will have amplitude of about $10 - 100\mu$V and a frequency in the range of $1 - 100$Hz. Besides, the clinical applications the use of EEG signals as a vector of communication between men and machines represents one of the current challenges in signal theory research. The principal element of such a communication system, more known as "Brain Computer Interface", is the interpretation of the EEG signals related to the characteristic parameters of brain electrical activity.

The identified electrical signals along the scalp by an electroencephalography which are activated from non-cerebral origin are termed as artifacts. EEG data is almost always contaminated by such artifacts. The amplitude of artifacts can be quite bulky relative to the size of amplitude of the cortical signals of interest. This is one of the reasons why it takes considerable skill to properly interpret EEGs clinically. Some of the most common types of biological artifacts include: ocular artifacts (includes eye blinks, eye movements and extra-ocular muscle activity), cardiac ECG (electrocardiography), muscular EMG (electro-myogram) artifacts. The EMG and EOG (electro-oculogram) artifacts are produced by facial muscle movementand, eye movement respectively. The cortical EEG signals are unavoidably combined with the electrical activity of the muscle tissue on the skull. Muscle artifacts are characterized by high frequency and low energy activity relative to

the local background activity. Among different artifacts, the muscular artifacts are more dominating and become potential obstacle to the further processing of the EEG signals. Hence, it is a challenging research issue to develop an efficient method to suppress the muscular artifacts from the raw EEG signal for the application in BCI.

The triumph depends on reliable handling of the noisy, non-linear, and non-stationary brain activity signals for extraction of features and effective grouping of motor imagery (MI) action as well as translation to the corresponding intended actions for MI based brain-computer interfacing (BCI). In this paper, EEG signal processing and classification performances are offered for motor imagery based brain-computer interface. EEG signals have been acquired placing the electrodes following the international 10-20 system. The acquired signals have been pre-processed removing artifacts using hybrid wavelet transform (HWT) leading to better signal to artifact ratio (SAR), mutual information and classification accuracy compared to SSA, DWT and EMD. EEG signals have been decomposed into subbands that are further processed to extract features. The extracted features have been used in linear discriminant analysis to characterize and identify MI activities.

This chapter introduces the background of the current study by describing both invasive and non-invasive EEG signals at first, followed by the motivation of this thesis by describing the challenges faced during preprocessing of these EEG data due to artifacts and the limitations of the existing methods to handle the issues. Then the objectives of this thesis are presented.

1.1 Brain Computer Interface (BCI)

A Brain computer interface is a communication structure of hardware and software that allows cerebral activity alone to restraint computers or peripheral devices. The present immediate aim of BCI study is to provide communications capabilities to severely disabled people who are totally paralyzed or locked in by neurological neuromuscular syndromes, such as amyotrophic lateral sclerosis, brain stem stroke, or spinal cord injury. A brain computer interface (BCI) is also called a brain

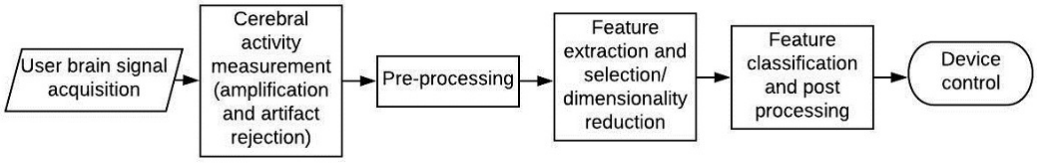

Figure 1.1: Basic BCI outline [3]. BCI system holds four basic components. It contains signal acquisition, signal preprocessing, feature extraction, and classification method.

machine interface (BMI). It is a transmission system that assists individuals to interact with their surroundings, without the involvement of external nerves and muscles, by using control signals generated from EEG activity. BCI creates a fresh non-muscular channel for relaying a persons intentions to external devices such as computers, speech synthesizers, assistive appliances, and neural prostheses. It is particularly attractive for individuals with severe motor disabilities. Such an interface would improve their quality of life and would, at the same time, reduce the cost of intensive care.

A basic BCI layout is illustrated in Figure 1.1. The brain computer interface is an artificial intelligence scheme that can distinguish a definite set of shapes in brain signals ensuing five successive stages: signal acquisition, preprocessing or signal enhancement, feature selection and extraction, classification, and the control interface. The signal acquisition stage captures the brain signals and may also perform noise reduction and artifact processing. The preprocessing stage prepares the signals in a suitable form for further processing. The feature extraction step identifies discriminative features in the brain signals that have been recorded from human brain. Once measured, the signal is mapped onto a vector having effective and discriminant features from the detected signals. The extraction of this attention-grabbing information is a very challenging task. Brain signals are mixed with other signals coming from a finite set of brain activities that overlap in both time and space. Moreover, the signal is not usually stationary and may also be distorted by artifacts such as electromyography (EMG) or electrooculography (EOG). The feature vector must also be of a low dimension, in order to reduce feature extraction part difficulty, but without relevant information loss. The classification stage classifies the signals taking the feature vectors and choice

good discriminative features. It is therefore important to achieve effective pattern recognition, in order to decode the users intentions. Lastly the control interface stage translates the classified signals into meaningful commands for any linked device, such as a wheelchair or a computer [147].

A BCI system can be classified as an invasive or non-invasive BCI according to the way the brain activity is being measured within this BCI [Figure 1.2]. In invasive technology, electrodes are neurosurgically implanted either inside the users brain or over the surface of the brain, while in non-invasive technologies, the brain activity is measured using external sensors [194]. In this manuscript we focus on EEG as the measurement technology.

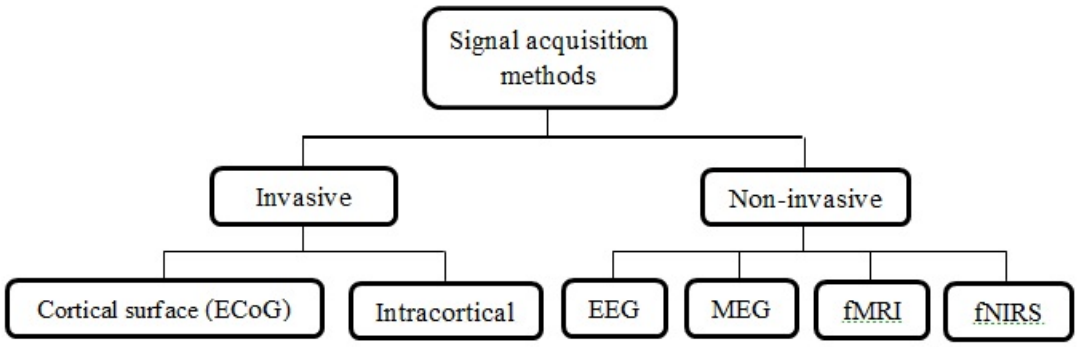

Figure 1.2: Signal acquisition system. If the sensors used for measurement are positioned within the brain, i.e., under the skull, the BCI is said to be invasive. On the contrary, if the measurement sensors are placed outside the head, on the scalp for instance, the BCI is said to be non-invasive.

This whole architecture is summarized in Figure 1.1. The above description defines an online BCI. However, it should be noted that before operating such a BCI, a considerable calibration work is necessary. This work is usually done offline and aims at adjusting the classification algorithm, calibrating and selecting the optimal features, selecting the applicable sensors, etc. In order to do so, a training data set must have been recorded earlier from the user. Certainly, EEG signals are highly subject-specific, and as such, immediate BCI systems must be standardized and adjusted to each user. The training data set should contain EEG signals recorded whereas the subject performed each mental task of interest several times, according to given commands. The recorded EEG signals will be used as mental state examples in order to find the best calibration parameters for this subject.

1.2 Electroencephalography

There is nothing so stable as change

(Bob Dylan)

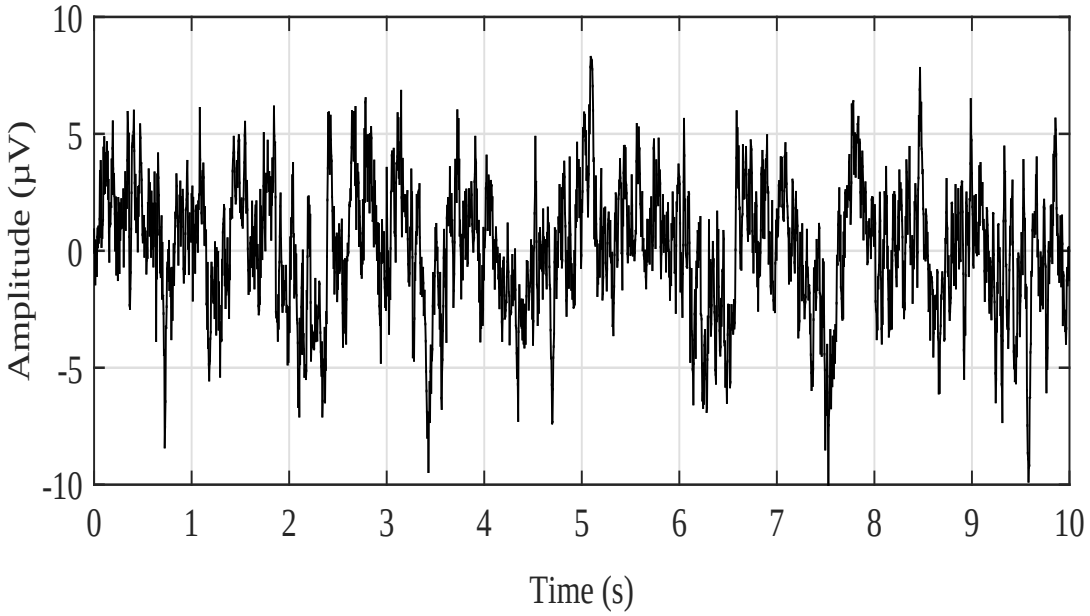

Figure 1.3: Low amplitude and high freuency EEG signal, recorded using electrode. A typical EEG signal, measured from the scalp, will have amplitude of about $10 - 100\mu$V and a frequency in the range of $1 - 100$Hz. EEG recording provides high temporal resolution.

EEG is the method of recording electrical activity through the scalp due to firing some of the neurons in the brain. These electrical activities are captured over a short period of time through multiple electrodes set on the scalp directly on the cortex. EEG is measured as the most common scheme for brain signals recording because it has high temporal resolution, easy to use, safe, and reasonably priced. The EEG signal is illustrated in Figure 1.3 where single electrode is used.

Signals recorded by EEG have very weak amplitude, in the order of some microvolts. It is thus necessary to strongly amplify these signals before digitizing and processing them. A 12-bit analog to digital converter (ADC) with sampling frequency ranging from 100Hz to several hundred is used to digitize EEG signal. Typically, EEG signals measurements are performed using a number of electrodes which differs from 1 to about 256, these electrodes being generally attached using a flexible cap. There are two types of electrodes for measuring EEG which are active

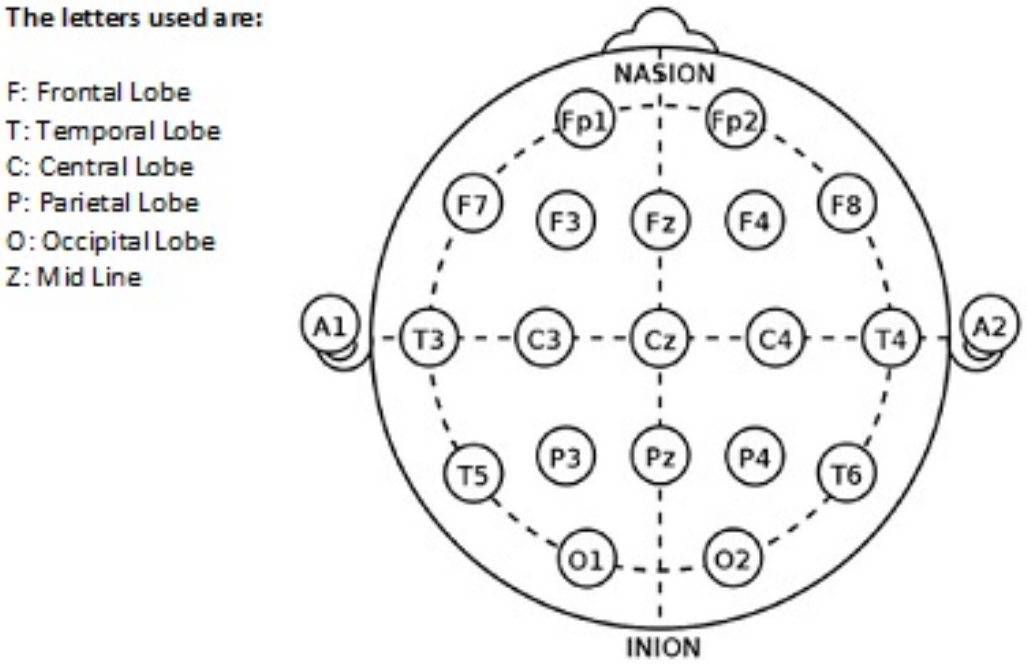

Figure 1.4: 10-20 system and electrode positions [160]. Location and nomenclature of the intermediate 10% electrodes (10/20 electroencephalography system), as standardized by the American Electroencephalographic Society. The letters specify the spatial location as A = ear lobe; Fp = frontal polar, etc.

electrodes or passive electrodes. Passive electrodes require an external amplifier to amplify the measured signals. On the other hand, active electrodes usually have an embedded amplifier. The main purpose of using either embedded or an external amplifier is to reduce the influence of the environment noise as well as the weakness of the signals due to cable movement. One of the major problems of EEG is that it needs gel or saline liquid to reduce the impedance of skin-electrode contact. The problem with the gel or the liquid is that it dries with time. However, currently there are some dry electrodes that are invented and might solve the problem of old electrodes. For reliable signals, it is expected that the distance between the electrodes have to be between 1cm and 2cm for low signal to noise ratio (SNR). In addition, EEG is able to detect changes in the brain signals within millisecond where an action requires almost $0.5 - 130$ milliseconds to propagate across a neuron. The contact between the electrodes and the skin is generally enhanced by the use of a conductive gel or paste. This makes the electrode montage system a generally boring and lengthy process. It is exciting to note that BCI researchers have recently proposed and validated dry

electrodes for BCI, that is, electrodes which do not need conductive gels or pastes for use. However, the performance of the resulting BCI (in terms of maximum information rate) were, on average, 30% lower than the one obtained with a BCI based on electrodes that use conductive gels or pastes. Electrodes are generally placed and named according to a standard model, namely, the 10-20 international system (see Figure 1.4). This system has been initially designed for 19 electrodes; however, extended versions have been proposed in order to deal with a larger number of electrodes.

With more experience with EEG recoding, 10-20 system, shown in Figure 1.4, it is recommended as a standard layout. The 10-20 refers to specific anatomic benchmarks or inter-electrode distance, such that it is 10% to 20% of the front-to-back or right-to-left head perimeters. EEG is usually defined in terms of activity types which are rhythmic activity and transients. The rhythmic activities are divided into certain frequency bands. These bands are verified to have certain biological significant or certain distribution over the scalp. In addition, it turns out to be a kind of nomenclature that EEG signals that falls below 1Hz and above 20Hz are considered as artifactual [160]. EEG measures the sum of the post-synaptic potentials generated by thousands of neurons having the same radial orientation with respect to the scalp (see Figure 1.3). The first EEG measurements on a human subject have been directed in 1924 by Hans Berger. It is at that time that he worked out the name of electroencephalogram. His essential discovery was available in 1929.

1.3 Brain Rhythms

EEG signals are composed of different oscillations named "rhythms". These rhythms have distinct properties in terms of spatial and spectral localization. In healthy adults, the amplitude and frequencies of such signals change with age, race and geographical location. Thus, previous statistical analysis must be done, in order to find mean brain patterns for EEG signal processing. There are six major frequency bands in brain rhythms, from low to high, these are:

- **Delta rhythm:** Delta (δ) waves ranging from 0.5 to 4 Hz. This is a slow

rhythm with a relatively large amplitude, which is mainly found in adults during a deep sleep.

- **Theta rhythm:** Theta (θ) has an amplitude that is greater than 20μV and falls within the range of 4 to 7.5 Hz. This is slightly faster rhythm observed mainly during drowsiness and in young children. It is generated with idling, creative inspiration, unconscious material, drowsiness, and deep meditation.

- **Alpha rhythm:** Alpha (α) is a wave with amplitude of 30-50 μV and change rate between 8 to 13 Hz. It is usually associated with relaxation, concentration, and sometimes in attention. These are oscillations which appear mainly in the posterior regions of the head (occipital lobe) when the subject has closed eyes or is in a relaxation state.

- **Mu rhythm:** Mu (μ) is found in the alpha wave frequency range where the recorded amplitude over motor cortex is maximum. It is usually associated with suppression indicates that motor neurons are working. It is located in the motor and sensorimotor cortex. The amplitude of this rhythm varies when the subject performs movements. Consequently, this rhythm is also known as the sensorimotor rhythm.

- **Beta rhythm:** Beta (β) is associated with alert, thinking and active concentration and falls in the range between 14 to 26 Hz. This is a relatively fast rhythm which is observed in awaken and conscious persons. This rhythm is also affected by the performance of movements, in the motor areas.

- **Gamma rhythm:** Gamma (γ) could be detected at somatosensory cortex with frequency greater than 30 Hz. It is also shown during short term memory matching of recognized objects, sounds, or tactile sensations [160]. This rhythm is sometimes defined as having a maximal frequency around 80 Hz or 100 Hz. It is associated to various cognitive and motor functions.

1.4 Measurements of Brain Activity

It is required to operate a BCI consists in measuring the subjects brain activity. Up to now, about half a dozen different kinds of brain signals have been identified which is suitable for controlling BCI. This section first describes the different available techniques for measuring brain activity i.e., easily observable and controllable for BCI. Then it describes the brain signals that can be used to drive a BCI.

1.4.1 Techniques for Measuring Brain Activity

(A) Overview of measurement techniques used for BCI

Many techniques are available and used, in order to measure brain activity within a BCI. Among these methods, MagnetoEncephaloGraphy (MEG), functional Magnetic Resonance Imaging (fMRI), Near InfraRed Spectroscopy (NIRS), ElectroCorticoGraphy (ECoG) or implanted electrodes, placed under the skull. However, the maximum used method is by far ElectroEncephaloGraphy (EEG). Indeed, this method is relatively cheap, non-invasive, portable and provides a good time resolution. As a result, most current BCI systems are using EEG in order to measure brain action. Thus, in this thesis work, we have focused on EEG-based BCI designs.

(B) Invasive and non-invasive BCI

It should be noted that a large and rapidly growing part of BCI research is dedicated to the use of implanted electrodes which measure the brain activity (most widely used EEG technique) of groups of neurons. At this time, most of this research has focused on the design and evaluation of invasive BCI for primates. However, recent results have exposed the usability of such systems on humans. Implanted electrodes make it possible to obtain signals with a much improved quality and a much better spatial resolution than with non-invasive methods. Indeed, some invasive methods can measure the activity of single neurons while a non-invasive method such as EEG measures the resulting activity of thousands of neurons. As such, it is recommended

that invasive BCI could obtain better results, in terms of performances (information transfer rate, accuracy, fiability, etc.), than non-invasive methods, and especially than EEG.

However, this report still needs to be confirmed and still now it is a topic of debate within the BCI community. Certainly, even if EEG-based BCI are based on much noisier and coarser signals than those of invasive BCI, some studies have reported that they can reach similar information transfer rates. The key difficulty of invasive BCI is exactly the fact that they are invasive and the subject endures a surgery operation in order to use the method. Moreover, limited lifetime of implanted electrodes makes the subject bear regular surgery operations in order to replace the electrodes. At that time, the use of implanted electrodes might be hazardous for the health of the subjects. Finally, numerous ethic problems are occurring when implanting electrodes in a humans brain. From these points of view, non-invasive BCI, and most especially EEG-based BCI becomes the most used and the most popular BCI systems. In the following of this manuscript, we will focus exclusively on non-invasive BCI based on EEG.

1.4.2 Neurophysiological Signals Used to Drive a BCI

In order to associate a command to one or several specific neurophysiological signals, the aim of BCI is Identifying in the brain activity measurements of a given subject. Various signals have been considered and some of them were revealed as relatively easy to identify (automatically), as well as relatively easy to control for the user. In the following of this manuscript we will also denote a neurophysiological signal as a mental state or as a brain activity pattern. These signals can be divided into two main categories:

(A) Evoked potentials

Evoked signals are generated unconsciously by the subject when he observes a specific external stimulus. These signals are also known as evoked potentials (EP). The main advantage of EP is that, different from spontaneous signals, evoked potentials do not require a specific training for the user, as

they are automatically generated by the human brain in response to a stimulus. As such, they can be used efficiently to drive a BCI since the first use. However, as these signals are evoked, they require using external stimulations, which can be uncomfortable, cumbersome or fatiguing for the user. The main signals of evoked potentials are the steady state evoked potentials (SSEP) and the P300. These two potentials are described further in this section.

Steady state evoked potentials:

The SSEP is one kind of brain potentials that appear when the subject perceives a periodic stimulus such as a flickering picture or a sound modulated in amplitude. SSEPs are defined by an increase of the EEG signals power in the frequencies being equal to the stimulation frequency or being equal to its harmonics or sub-harmonics. Various kinds of SSEP are used for BCI, such as steady state visual evoked potentials (SSVEP), which are by far the most used, somatosensory SSEP or auditory SSEP (see Figure 1.5 for an example of SSVEP). These SSEP appear in the brain areas corresponding to the sense which is being stimulated, such as the visual areas when a SSVEP is used.

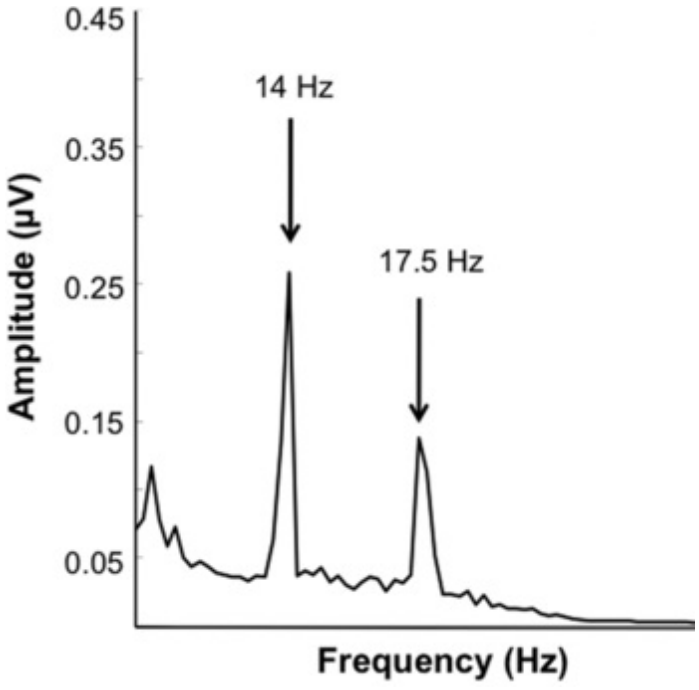

Figure 1.5: The SSVEP ans its harmonics.

An advantage of this kind of signals is that they can be used without training within a BCI. It is possible to use a large number of stimuli in order to obtain and use a large number of mental states for the BCI, as stimuli with different stimulation frequencies will lead to SSEP with different frequencies. As a result, it assists the user to have a large number of commands which makes the whole system more suitable. This explains the collective attention of the BCI community for SSEP, and more especially for SSVEP. For example, a BCI application based on SSVEP can use several flickering buttons displayed on screen, each button having a unique flickering frequency. In such an application, the user should draw his attention on the button he wants to activate. Indeed, it is known that the user draws his attention on this SSVEP corresponding to a given button are enhanced.

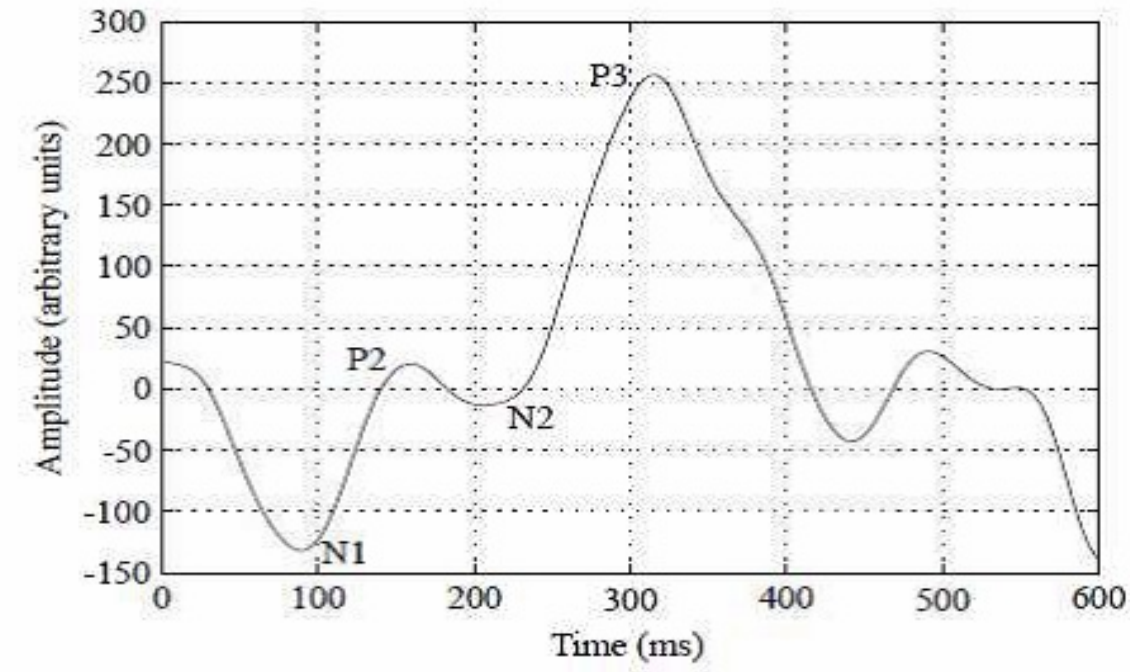

Figure 1.6: A P300 (enhanced by signal averaging) occuring when the desired choice appears.

P300

The P300 consists of a positive waveform appearing approximately 300 ms after a rare and relevant stimulus (see Figure 1.6). Typically, it is generated through the odd-ball paradigm, in which the user is requested to attend to a random sequence composed of two kinds of stimuli with one of these stimuli being less frequent than the other. If the rare stimulus is relevant to the

user, its actual appearance activates a P300 observable in the users EEG. This potential is mainly found in the parietal regions.

As like as SSVEP-based BCI, several buttons or objects are displayed on the users screen in P300-based BCI applications. These buttons or objects are randomly highlighted on the screen. The user is instructed to count the number of times that the button he wants to activate is highlighted over a finite time period. The desired button is highlighted by a rare and relevant stimulus only when the trigger is applied at the appearance of the P300. Thus, when a P300 is detected in the EEG signals, the system identifies the button desired by the user was highlighted 300 ms earlier. The P300 is generally used in a kind of virtual keyboard which is known as the P300 speller. No training is required for the subject in order to be used P300. But P300-based BCI applications need the user to continuously focus on fast and repetitive visual stimuli, which can be fatiguing and inconvenient.

(B) Spontaneous signals

Spontaneous signals are voluntarily generated signals by the user following an internal cognitive process without external stimulation. There are two types of spontaneous signals: the most used signals are sensorimotor rhythms and other called neurophysiological signals (such as slow cortical potentials or non-motor cognitive signals.

Motor and sensorimotor rhythms:

Sensorimotor rhythms are brain rhythms related to motor actions, such as arm movements, for instance. These rhythms are mainly found in the α ($\simeq$ 8-13 Hz) and β ($\simeq$ 13-30 Hz) frequency bands, over the motor cortex, which can be voluntarily controlled in amplitude by a user. Considering these sensorimotor rhythms will be used in BCI then, two different strategies have been proposed in order to make the signal user control:

- Operant conditioning: A subject can change his sensorimotor rhythms amplitude willingly through a long training known as "operant conditioning" (see Figure 1.8). In order to reach this goal, the user is free to select the mental strategy which is comfortable for him. In this case,

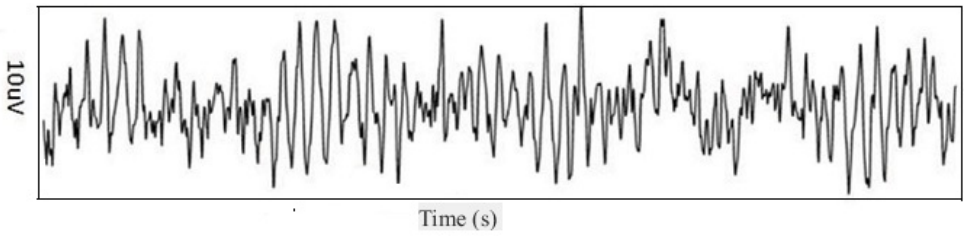

Figure 1.7: Time series of a typical μ component.

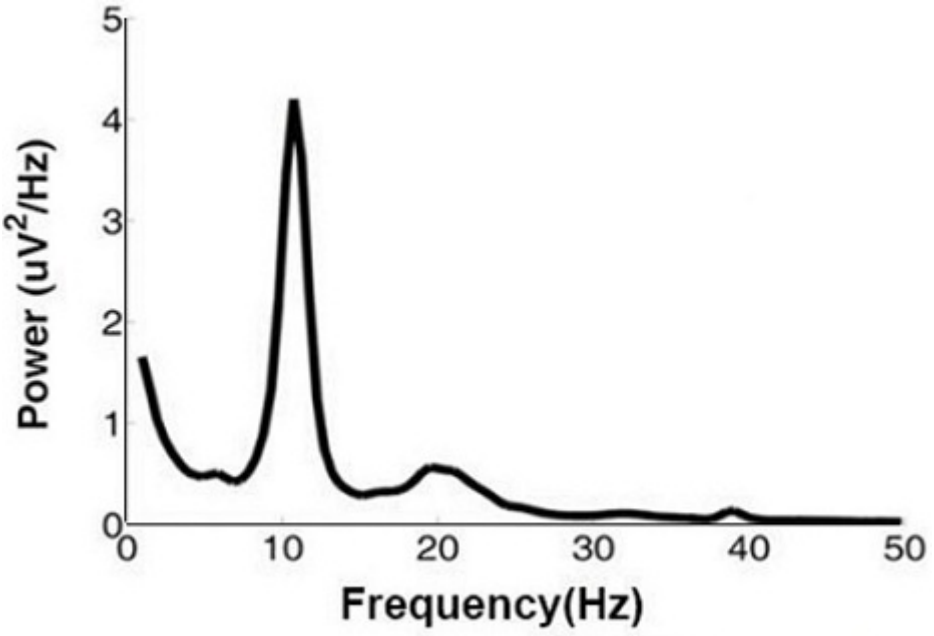

Figure 1.8: Power spectrum of the μ component shown in Figure 1.7. [225].

motor imagery (see below) is one possible approach which is often used. The role of feedback is essential for operant condition. Using the operant conditioning, the user is able to understand how he should modify his brain activity in order to control the system. When operant conditioning based BCI is used, the power of the μ and β rhythms in different electrode locations are linearly combined in order to build a control signal which will be used to perform 1D, 2D or 3D cursor control. The main problem of this method is that it needs very long training time (training time may be several weeks or even several months). However, very good performances (in terms of information transfer rate) are obtained when this training is completed.

- Motor imagery (MI): We are able to imagine almost everything. Sitting at my desk and overlooking a canal in down-town Dhaka, I can at the

same time easily imagine that I walk outside the city on a country road. I can imagine scenes or objects that are not there, or no longer there. I can mentally perform actions which I can not perform in reality. I can imagine myself as a perfect dancer although reality has taught me that my motor networks are very reluctant to learn the repetitive and rhythmic motor patterns needed for that type of skill. Although this imaginative power can be used in the visual, auditory, and tactile domains. I will give attention solely on the imagination of motor imagery movements.

Motor imagery is an intellectual development in which a subject imagines that he or she performs a movement of various body parts resulting in sensorimotor cortex activation without actually performing the movement and without even tensing the muscles. It is a dynamic state during which motor action is internally activated without any motor output. In other words, the conscious activation of brain regions is required by motor imagery. The imagery is also involved in movement preparation and execution, attended by a voluntary inhibition of the actual movement [141].

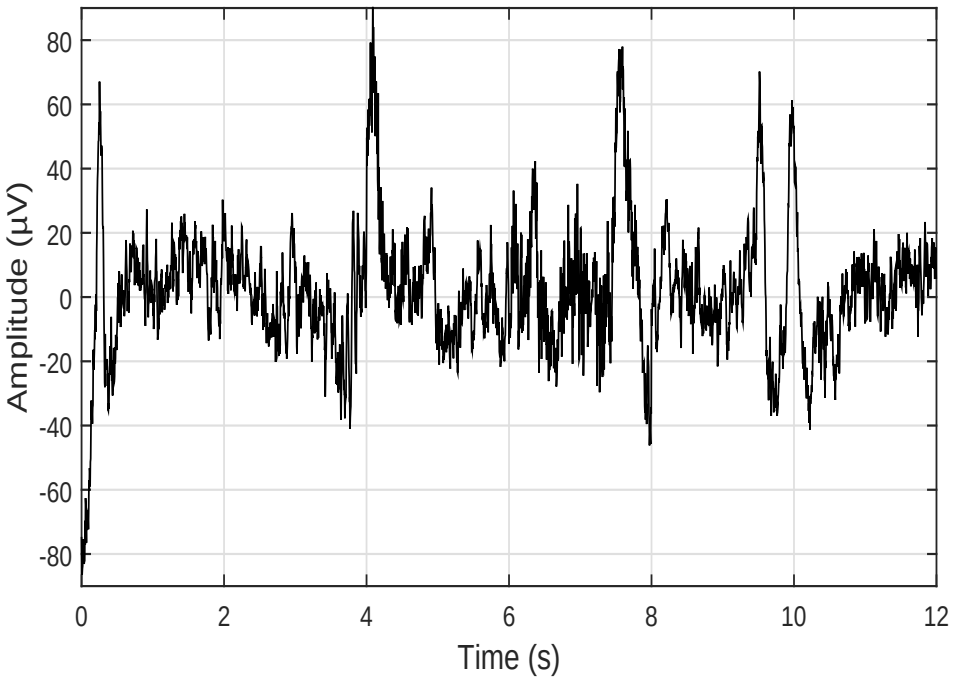

Figure 1.9: The motor imagery EEG signals [189].

For a subject, implementation of motor imagery consists in imagining movements of his own limbs (hands or feet for example). The signals resulting from performing or imagining a limb movement have very specific temporal, frequential and spatial features, which makes them relatively easy to recognize automatically. For instance, imagining a left hand movement is known to trigger a decrease of power (event related desynchronisation (ERD)) in the μ and β rhythms, over the right motor cortex. The motor imagery EEG signals are illustrated in Figure 1.9. Advance machine learning methods were recently developed to compute a subject-specific model for detecting the performance of motor imagery. The top performing algorithm is the filter bank common spatial pattern from BCI Competition IV dataset 2 for motor imagery.

Slow cortical potentials: The characteristics of slow cortical potentials (SCP) are based on the slow event-related direct-current shifts of the electroencephalogram. The shifts in the electrical negative direction redirect the depolarization of large cortical cell assemblies, reducing their excitation threshold. They are very slow variations of the cortical activity, which can last from hundreds of milliseconds to several seconds. Based on the operant conditioning, it is possible to learn to make these variations positive or negative (see Figure 1.10). According to the positivity or negativity of the potential, SCP can be used to generate a binary command in BCI. If the operant condition is used to control SCP then the signal requires generally a very long training time. The training time of operant conditioning of SCP is longer than motor rhythms. However, SCP would be considered a more stable signal.

Non-motor cognitive tasks: There are large number of non-motor cognitive processing tasks that are used to drive a BCI. For example, these tasks are-mathematical computations, mental rotation of geometric figures, visual counting, mental generation of words, music imagination, etc. All these mental tasks generate specific EEG signal variations, in specific cortical areas and frequency bands, which make them relatively easy to identify.

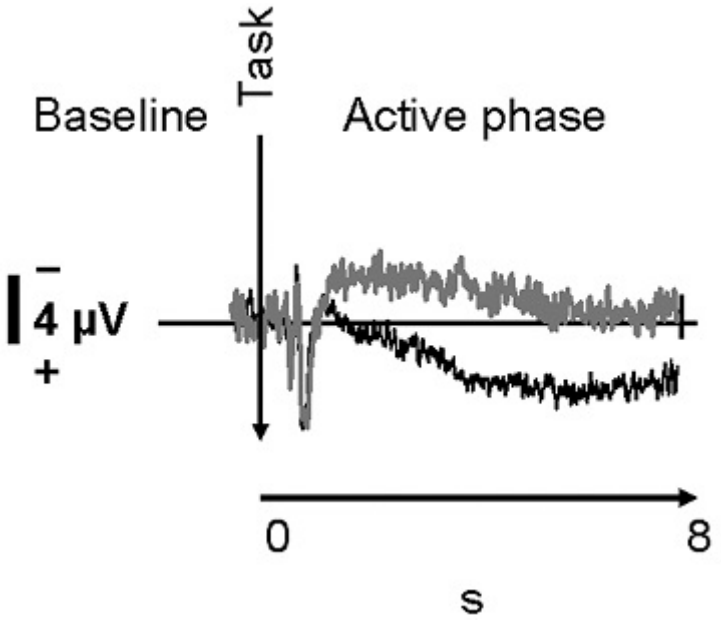

Figure 1.10: The upper line indicates negative shifts; lower line, positive shifts.

1.5 Artifacts

Electrical signals are recorded along the scalp by an EEG, but that originated from non-cerebral origin are called artifacts. It is observed that EEG data is almost always contaminated by such artifacts. The amplitude of artifacts is quite large relative to the size of amplitude of the cortical signals. The most common types of biological artifacts are ocular artifacts (includes eye blinks, eye movements and extra-ocular muscle activity), cardiac ECG, muscular EMG artifacts. Figure 1.11, Figure 1.12, and Figure 1.13 shows the most common artifacts. The EMG and EOG artifacts are produced by facial muscle movement and eye movement respectively.

1.5.1 Sources of Artifacts

EEG recordings are often contaminated by different forms of artifacts. In EEG recording, the artifacts are of several types that come from different sources. In broad sense, artifacts in EEG can be originated from internal and external sources and contaminate the recordings in both temporal and spectral domains with wide frequency band. Internal source of artifacts are due to physiological activities of the subject (e.g. ECG, EMG/muscle artifacts, EOG) and its movement. External

Table 1.1: Different types of artifacts and their origins.

Artifacts types and sources						
Physiological/Internal				Extra-physiological/External		
Ocular	Cardiac	Muscular	Others	Instrumental	Interference	Movement
Eyeblink, Eye movement, Eye flutter, REM sleep, etc.	ECG, Pulse	Chewing, Swallowing, Clenching, Sniffing, Talking, Scalp contraction, etc.	Glossokinetic, Skin Respiration, etc.	Electrode displacement, Cable motion, Poor ground, etc.	Electrical, Magnetic, Sound, Optical, EM waves etc.	Head movement, Body movement, Limbs movement, Tremors, Other movements, etc.

source of artifacts are environmental interferences, recording equipment, electrode pop-up and cable movement. Also some artifacts may be presented in several adjacent channels (global) while some of them can be found only in single-channel (local). In addition, some artifacts look as regular periodic events such as ECG or pulse artifacts (regular/periodic) while some others may be extremely irregular. A summary of different artifact types and their origins is presented in Table 1.1. [86].

1.5.2 Artifact Characterization in EEG

In order to reduce the influence of artifacts on the EEG signal, it is very essential to know what are the most common types that may be recorded during EEG recording. Artifacts can be strongly separated from other possible categorizations are physiological and non-physiological or technical origin. The artifacts during EEG recording are of various types that come from different sources. In broad sense, artifacts in EEG can be originated from internal and external sources. The

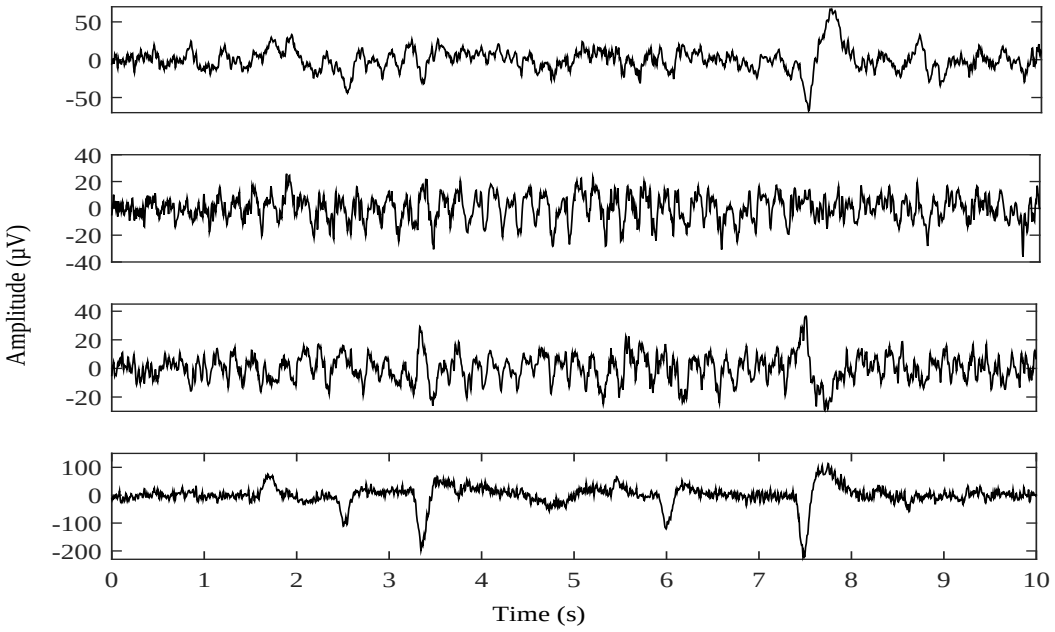

Figure 1.11: EEG recording corrupted by ocular artifacts [105].

sources and types of artifacts are discussed detail in next chapter.

1.6 Artifact Suppression

Artifact removal is the process of recognizing and eliminating artifacts from brain signals. An artifact removal method should be able to remove the artifacts as well as keep the information of neurological signal intact. There are a wide range of techniques for detecting and/or removing artifacts in EEG. Among them, the simplest techniques based on thresholding of the various parameters calculated from the EEG signal, fully automated techniques. There are, for example, standard and adaptive filters regression methods (both in the time domain and in the frequency domain), PCA-based approaches, blind source separation approaches, algorithms based on ICA, and various hybrid methods. Recently, independent component analysis (ICA) techniques have been used to correct or remove EEG impurities. These techniques attempt to "unmix" the EEG signals into some number of underlying components. Depending on the behaviors or natures of EEG, there are many source separation algorithms. Fully automated artifact rejection methods, which use ICA, have also been developed [92]. EEG contamination by muscle has been shown to be far more prevalent than last few years. By compar-

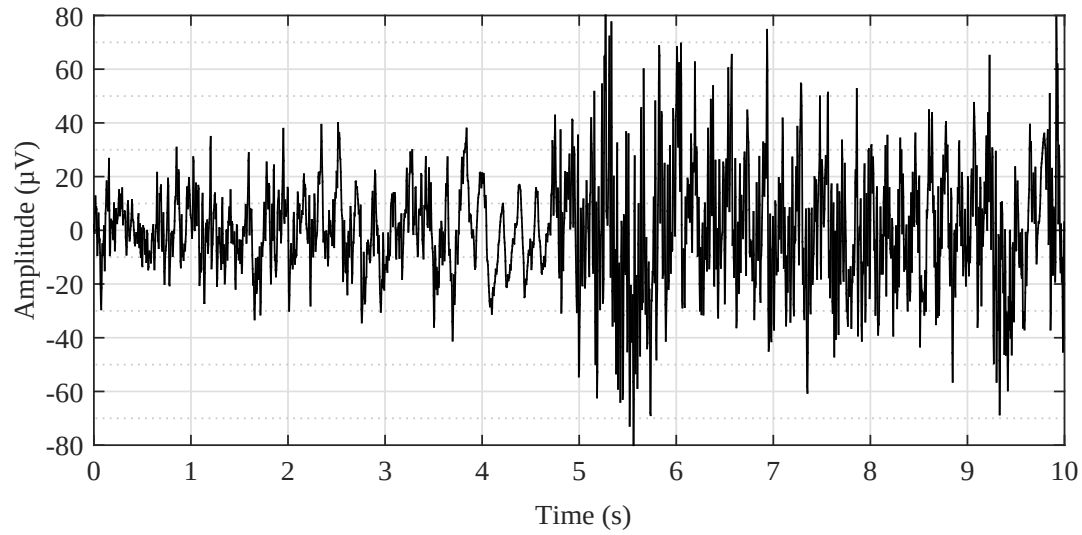

Figure 1.12: The contaminated EEG data by muscle activity [22].

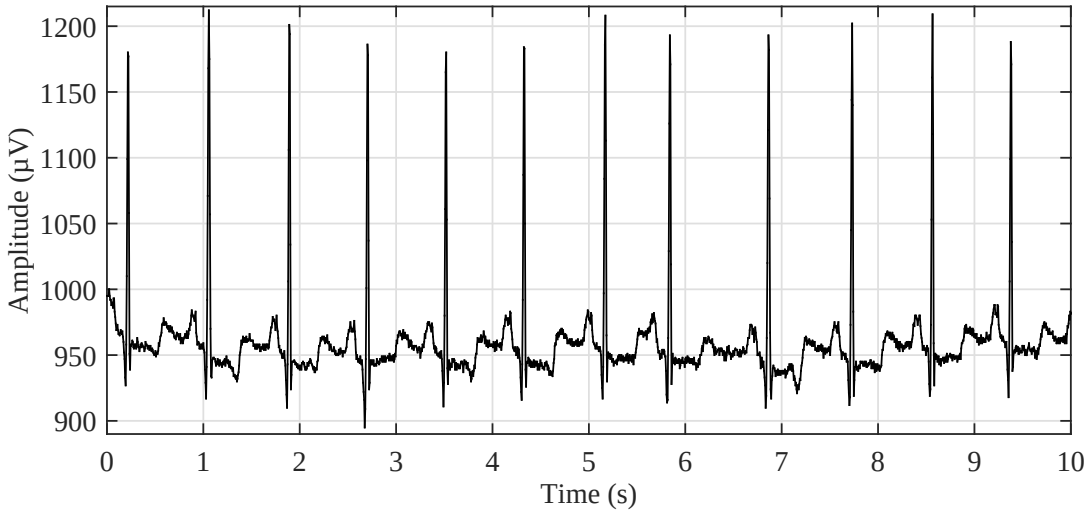

Figure 1.13: ECG signal

ing data from paralysed and unparalysed subjects, muscle artifacts are dominant particularly in the gamma range above 20 Hz. However, Surface Laplacian has been shown to be effective in reducing muscle artifacts, particularly for central electrodes. The arrangement of Surface Laplacian with automated techniques are known as hybrid method that is used for reducing muscle components using ICA proved particularly effective in a follow up study [52].

1.7 Existing Methods for Artifact Removal

In this section we give an overview of the measured EEG signal, the different kinds of artifacts that may affect the recordings, how these sources combine in the scalp EEG and finally, existing ways for validating artifact removal methods.

Unfortunately, EEG signal is often contaminated by different biological measures of non-interest. Among them, muscular or myogenic activity arising strongly opaque EEG signals from the contraction of head muscles. It is particularly difficult to correct this type of artifact which is induced by perturbation because myogenic activity is of high amplitude (possibly several times larger than the EEG signal), wide spectral distribution, and variable topographical distribution. Therefore, it is a challenging preprocessing step prior to qualitative or quantitative EEG analysis for denoising of EEG. A simple algorithm is introduced that compares high frequency activity in each epoch with the activity level in a local window to detect bursts of myogenic activity. If it exceeded the local background activity by a certain factor, the window is considered artifactual. The artifact detection algorithms are empirically adjusted by applying different factors such as artifact thresholds, sensitivity and specificity. Without a priori (or very little) knowledge about the sources themselves and about the mixing process, it is very difficult to minimize the disturbances due to muscular activity in EEG signals can be considered as a blind source separation (BSS) problem, which consists in estimating the original sources underlying the multi-channel EEG signals. BSS techniques have been applied in various domains including the denoising of EEG as found in the recent reviews. Among the existing BSS algorithms, four main ones have emerged to specifically tackle the difficult problem of EEG data denoising. The ICA was successfully applied to EEG denoising for muscular activity. It has been introduced for removing a wide variety of artifacts from multichannel EEG signals. However, the method requires visual inspection of ICA components and manual classification of the interference components. This can be time-consuming and is not desirable for real-time artifact suppression. In particular, these techniques require multichannel EEG and cannot be applied to a single-channel recording. More recently, another BSS approach called canonical correlation analysis (CCA)

is also proposed to remove muscle artifacts from EEG and improve the interpretation of EEG channels. For the detection of epileptic shapes, elimination of different kinds of noises and for the removal of particular electrophysiological artifacts, such as ocular movement, and heart signal, the wavelet transforms (WT) has been used with respect to the EEG background. Finally, a fourth method called empirical mode decomposition (EMD) has appeared as a promising tool in the particular field of EEG data denoising which is not suitable for real time implementation. Most of the existing methods of muscular artifact suppression suffer from high computational load. An efficient model based approach is demanding to implement the BCI system in real time applications.

Regression analysis is a method using a multi-modal linear model between observed and a reference signal. It is a traditional way of identifying artifactual data and consequently removing such data that do not belong to the model. There are common methods for removing some physiological artifacts such as ocular and cardiac artifacts from measured artifact-contaminated EEG signal and an artifact reference signal. As a result, regression analysis is widely used artifact removing method to remove ocular artifacts from EEG signals. It observes that the measured EEG signals are a linear superposition of EEG and EOG signals. The part of any EOG component which is present in the EEG signal is estimated and then removed using the least squares criterion. It has a drawback that the method requires the recording of source signals from the EOG channels to remove ocular artifacts. For this case, the muscle artifacts which originated from different muscle groups are difficult to identify from the source EEG signals. For this reason, different reference channels from multiple muscle groups are required. This in turn can greatly increase the complexity of the algorithm [226]. In addition, EEG signal being non-linear and non-stationary process, linear regression is not the best choice for analysis in such applications. Moreover, it can only be used to treat few particular types of artifact; not all types.

Another popular approach for artifact removal method is blind source separation (BSS) [67, 92], including ICA or maximally uncorrelated (for CCA) algorithms. Blind source separation is a method to estimate original signals from observed signals which consists of combination of original signals and noise. The

hidden factors such as sets of random variables, measurements, or signals are revealing using another blind source separation technique ICA which is a statistical and computational technique. It is assumed that the sources are linearly independent. ICA defines a generative model for the observed multivariate data, which is typically given as a large database of samples. The main problem with ICA-based artifact detection and removal method is that, it is not automatic. In this method, manual intervention is needed to reject ICs with visually detected artifacts after decomposition. By uniting ICA with another balancing method such as wavelet transform or EMD or using a classifier like SVM or even with a help of reference channel, it (i.e. artifact detection and removal) can be work automatically [230]. However, even in such case, the independent component with artifacts may also contain some residual neural signals. Also it introduces distortion to the neural signal during signal reconstruction after completely rejecting that particular IC. Another problem is that it cannot operate on single channel data, since the number of recording channels must be at least equal to the number of independent sources. Another factor that limits the choice of ICA for artifact removal in applications that require online/real-time implementation of the algorithm is the computational complexity. Finally the involvement of iterative process in computing ICA algorithm makes it difficult to perform robustly. E.g. ICA may be useful to remove global artifacts such as ocular artifacts [55, 65, 14, 108, 101] or sometimes other physiological artifacts, but not external artifacts. There are few works reported the use of modified [108] or constrained ICA [167, 178, 2, 89] for automated and better performance in artifact detection and removal.

Constrained ICA (cICA) [90] or ICA with Reference (ICA-R) incorporates prior knowledge about the source signals, making a semi-blind source separation problem. This is done by imposing temporal or spatial constraints on the source mixture model. The temporal, spectral or time-frequency information of the biomedical measurements such as heart beat morphology, frequency bands of EEG activity and temporal dynamics of certain high amplitude artifacts are available [90].

Correlation analysis or CCA is another BSS method for separating a number of mixed or contaminated signals that uses second-order statistics (SOS) to generate

components derived from their uncorrelated nature. By looking for uncorrelated components, the approach uses a weaker condition than statistical independence sought by the ICA algorithm. Temporal correlations is absent in ICA into account while CCA addresses this point by being capable of finding uncorrelated components. In addition, have maximum spatial or temporal correlation within each component [175]. Both methods provide qualitatively the same results, but CCA method is more computationally efficient.

Morphological component analysis (MCA) decomposes the recorded signal into components that have different morphological characteristics where each component is sparsely represented in an over-complete dictionary [175]. If the wave shape or morphology of MCA is known then it is stored in a database and applicable to certain known artifacts. The effectiveness of this method greatly depends on the available artifact-template database. In [129, 227, 228], MCA is used to remove ocular artifacts for real time and some of the muscle artifacts originating from jaw clenching, swallowing, and eye-brow raising and in [180], it is also used to remove ocular artifacts (eye movement and blink artifacts) from the EEG raw signal. The raw EEG signal was derived from signal morphology characterized as slow and smooth change of EOG time series.

To address the weakness of ICA-based artifact correction methods another effective BSS method, stationary subspace analysis (SSA) is used [210, 70]. The SSA is a time sequence study method, can be used to identify high-dimensional EEG measurements that are the underlying stationary and non-stationary brain sources. The stationary sources found by SSA can significantly increase the restricted BCI performance [211]. Frequently used SSA algorithms are based on exploiting higher-order statistics (HOS)[152]. The SSA reveals their spatial characteristics using topographic maps corresponding to stationary and non-stationary brain sources. The EOG signals are considered as low frequency and high energy trend in the recorded EEG signals, the SSA is performed better to suppress it from mixing. The components that are determined to be related to the ocular artifacts are projected back to the EOG data which is subtracted from EEG signals. Hence it is very evident to employ SSA for EOG identification and suppression. But it is used for multiple channel data to partition the whole space into stationary and

non-stationary subspaces [229]. The stationary channels are considered as EEG.

Other component based techniques that have appeared in the literature are multiple source eye correction (MSEC) [8, 9], multivariate singular spectrum analysis (MSSA) [32] and sparse component analysis (SCA) [64, 59], include among others for artifact removal from the EEG. The MSEC determines signal topographies in the presence of predefined artifact topographies [82] by spatio temporal dipole source analysis. Eye activity in EEG and event-related response data can be estimated and corrected in the presence of brain activity using the resulting spatial vectors together with a brain model [9].

Based on embedding a signal into a higher dimensional space, singular spectrum analysis (SSpA) is a nonparametric spectral estimation method that later represented by the signal trajectory matrix formed with time-delayed versions of the signals. By clustering the columns of this trajectory matrix, artifacts are removed using singular value decomposition. The trajectory matrix is constructed by concatenating the trajectory matrices from each individual signal in the multivariate signal set [32, 199] for MSSA. In order to solve the BSS problem [109], SCA assumes sparsity of the sources in some (transformed) domain .

It is quite popular to analyze time-frequency of non-stationary time-series data in biomedical signal processing, e.g. in EEG signal processing. The reason of using simultaneous time and frequency domain analysis is that the frequency and statistics of time-series data vary with time because of the non-stationary properties of this type of signal. Therefore, any quick change in frequency values can be captured in a particular temporal window for any signal components (e.g., either artifact or seizure [161, 186]. A time-frequency representation of ocular artifacts (OAs) including blinks and saccades found in EOG and observed that frequencies up to 181Hz can be present in a subject's EOG for certain tasks which is proposed in [149]. It is suggested that if EOG is used for ocular artifact removal from EEG, then EOG should be sampled at least 362 Hz to avoid aliasing.

Short-window or short-time Fourier transform (STFT) is the common time-frequency representation. However, this method is not as effective as it has identical time-frequency resolution at all frequency values. The bandwidth of EEG signal is around 0.5-120 Hz and most of the artifacts appear in the lower frequency

region (< 10 Hz). Therefore, STFT cannot provide high frequency resolution in lower frequency region which is required. The good solution of this problem is to use wavelet transform since it provides a decent time-frequency resolution for EEG signals [86].

It is assumed that the signal and artifacts are uncorrelated in adaptive filtering. A signal correlated with the artifact which is generated from the filter using a reference signal and then the estimate is subtracted from the acquired EEG [193]. The proper functioning of the algorithm depends on the choice of the artifact reference and it may be obtained from EOG recordings for the removal of eye movements or blinks [29], or from EMG recordings for the removal of muscle artifacts [32].

As like as an adaptive filter, Wiener filter doesn't require the use of an external reference signal, however, it assumes that both the signal and artifact are stationary linear random processes with known spectral characteristics and also the signal and artifact are uncorrelated. The neural signal shows non-stationary characteristics and it is believed to originate from a complicated non-linear stochastic process. Due to the uncertainty of different types of artifact sources, the spectral characteristics cannot be determined accurately, although the spectral characteristics of EEG rhythms are known. In addition, the Wiener filter cannot be implemented in real-time application. So it is not a good choice using wiener filter for artifact removal in applications that require real-time processing.

Bayes filtering based on the system is Markov [193] and assuming it is a probabilistic system estimation method from noisy observations. These filters are capable of working without a reference signal and operate in real time that overcome some of the limitations of the aforementioned techniques in that way. Bayes filters are similar to Kalman filters and particle filters and it is not directly implementable due to their complexity. However, Kalman filters and particle filters have been used for EEG artifact removal for instance in [140] and in a nonlinear fashion in [171]. Bayes filters first estimate the state at a given time and then obtain a feedback in the form of noisy measurements [193], which is used to predict a new a priori estimate. The algorithm needs to be calibrated.

The spatial filter PCA reduces data dimension and highlights specific features

of data which is usually difficult to identify in the spatially unfiltered data as the new components are created by weighted combinations of all EEG channels. Based on PCA, two recent articles have proposed artifact removal method. The paper [202] reported the use of robust PCA after preprocessing is done based on wavelet denoising and band-pass-filtering; while [203] compared PCA with ICA for artifact removal and originate ICA performs better than PCA. Both these articles have evaluated their method qualitatively; therefore, it is not possible to comment exclusively on the efficacy of PCA in detecting and removing artifacts. One important limitation of PCA (or SVD) is that it fails to identify ocular or similar artifacts when amplitudes are comparable since PCA depends on the higher order statistical property [167] from EEG. In [221], spatial filters-two common spatial pattern (CSP) filters by making use of fuzzy sets reduced the root mean square estimation error and significantly increased the EEG signal quality.

In the recent time, literature is an option that has gained attention for multi channel [2, 17, 60] and single channel processing [134, 206, 192] using a combination of algorithms to remove artifacts from the EEG. It has become very popular in recent years that researchers are very keen to use the advantages of different methods by combining them into a single method for artifact detection and removal, i.e. hybrid method which has two or more stages.

This hybrid method is formed by combining two popular methods: wavelet transform and blind source separation. This hybrid is mainly encouraged from the fact that only BSS based separation of artifactual components (e.g. ICs) is often inaccurate since the separated artifactual component also contains residual neural information. Therefore, such component will introduce significant distortion in reconstructed EEG signal which are completely rejecting. Hence, the multi-channel datasets are converted into ICs or CCs and then possible artifactual part is decomposed to different frequency bands of detail coefficients by wavelet transform. After that, the artifactual coefficients of higher amplitude are denoised by thresholding which eventually preserve the residual neural signals of low amplitude. The related articles are [17, 119, 69] for wavelet-ICA, [159, 231] for wavelet-CCA. On the other hand, there are related hybrid methods that can be applied to single-channel EEG data by reversing the order of wavelet transform and BSS blocks.

E.g. [125, 15] reported that first signal is decomposing into wavelet coefficients then artifactual coefficients are passed through BSS block to separate artifacts from neural signal. However, typically the former way is more known to the research communal as wavelet enhanced ICA or wavelet enhanced CCA. Please note that the type of wavelet transform can be DWT, CWT, SWT or sometimes WPT [12].

The combination of BSS and EMD is another hybrid method instead of wavelet transform. The difference is that usually the first stage is to decompose the signal into IMFs by EMD or EEMD and then apply BSS (ICA or CCA) on the IMFs to identify artifactual component followed by rejecting the artifactual IC or CC. Such methods are reported in [192, 230, 20].

It is reported in [178], a hybrid BSS-SVM algorithm where certain carefully chosen features are extracted from separated source components and then fed into a SVM classifier to identify artifact components followed by removing them for eye blink and ECG artifact removal. Lastly, to reconstructing artifact-free EEG, the rest of the source components are re-projected.

In [101], it is reported that a hybrid approach is made by combining BSS and regression-based adaptive filtering (with VEOG and HEOG as reference channels) for elimination of ocular artifacts. Similar methods have been used in [65] to remove ocular artifacts by combining ICA and adaptive filtering.

To remove ocular artifacts, another hybrid approach is made by combining ICA and auto-regressive exogenous (ARX) was proposed in [214]. In this method, ARX is used to reduce the negative effect induced by ICA by building the ARX multi-models based on the ICA correlated signals and the reference EEG that are selected prior to the artifact-contamination.

In [145], it is reported that EOG artifact removal used a hybrid method named as wavelet neural network (WNN) which is the combination of wavelet decomposition and artificial neural network. In this method, the reference EOG channel is only required for ANN classifier throughout training session.

Another method combining DWT and ANC (adaptive noise canceler) is proposed in [153] to remove eye blink artifacts. In this method, the ocular artifact reference is resulting from DWT decomposition and then used in the adaptive fil-

ter as reference. On the other hand, Navarro et al. [143] used the combination of EMD and adaptive filter (with RLS algorithm) to remove ECG artifacts not EOG from EEG recordings. To remove EOG and EMG artifacts from EEG, the authors in [75] presented a new way by using a hybrid combination of adaptive neural fuzzy inference system (ANFIS) and functional link neural network (FLNN). The ANFIS generally has two parts: a non-linear antecedent and a linear consequent; however, in their proposed system to enhance the non-linear approximation ability, the second part is replaced with the FLNN. Then an adaptive filtering system adjusts the parameters of both ANFIS and FLNN.

It is impossible to explain all existing methods on artifact removal from the EEG for lack of time and space even if we try to be comprehensive in literature. Otherwise, we have explained the most usual methods to remove EOG, EMG and ECG impurity from the EEG.

Depending on the need for human intervention or not, artifact removal algorithms can be classified as semi-automated or automated. Semi-automatic methods can only be used for off-line applications because they require visual inspection of the measured signal or of the components obtained by the artifact removal method. In general, an automated execution is usually preferred. So, to be able to run an algorithm in an online application only when to avoid introducing subjectivity in the process.

In order to get automatic identification, multiple procedures and artifacts correction have been developed, however no single one views out among the others. In real life measurements, there are multiple types of signals and contaminants that are mixed in undetermined ways. As a result, automating of existing algorithms is not easy, as mentioned throughout the text. In this way, some standard methods have limited application unless they can be readily adapted to specific scenarios. Regression and filtering approaches need a reference channel if they are meant to operate automatically [193, 173]. On the other hand, based on values for typical characteristics of the EEG signal or artifacts and deviations from normality [33, 42, 136], component based methods are more versatile in that they can be made automatic via a reference signal.

An automation may be performed by computing the correlation of certain

ICs with either reference [213, 82], or by combining signal features and correlation [166, 165], when a reference waveform exists or a prototype signal can be generated. Mostly, the EEG signal is decomposed and the derived components are based on the spatial and temporal probabilistic characteristics which are compared to standard values for background EEG and based on the combination of thresholds, artifacts may aid to automate an algorithm. A completely automatic ICA-based algorithm for identification of artifact related components in EEG recordings, ADJUST [136], works quite consistently following this idea. Similar approaches are presented in [33, 42]. Another alternative, consists in performing a training phase followed by a clustering step, implemented in [32, 146]. In those papers, a reference channel is needed for the training phase that is either measured or achieved from clean epochs of the same recording or from epochs of another recording that holds similar types signal or artifact components to those of interest. Clustering is performed based on the likeness of the temporal dynamics of ICs termed by their auto mutual information.

To conclude, WT and EMD can be automated more easily and in a common manner using thresholds [168]. For the WT, the SURE [44] shrinkage rule and a soft thresholding strategy seem to provide good results [168]. Inspired on wavelet thresholding and with respect to EMD, the writers of [103] use an algorithm that assesses the noise level and filters each IMF.

Most of the EEG based applications require automated information processing; specially it is an online/real-time implementation. In addition, manual identification of artifactual component or epoch is very time-consuming and laborious for multi-channel long-term data sequences. Therefore, many signal processing techniques have been proposed along with some useful apriori signal or artifact statistics/characteristics have been utilized. Among them, BSS-based techniques sometimes act as semi-automated when it requires some training or parameter selection for identification of artifactual component. Although there are few papers presented that propose automated identification of ICs after ICA [217]; however, they have the need of training samples for supervised classification and also need an additional information in the form of contact impedance measurement [10]. In order to make the whole process automated, there are some methods that includes

ICA method for an automatic artifacts detection, then there has to be another stage.

Online/Real-time implementation demands the algorithm to be fast enough and to have low-enough complexity for such application. Here, online implementation refers to the algorithms implemented in software platform capable of online/real-time processing, not in hardware platform. However, some EEG-based applications such as wireless ambulatory EEG monitoring may require on-chip implementation of the artifact detection/removal algorithm. In such case, the computational complexity has to be minimum which is a great challenge and so far according to the best of our knowledge no real-time hardware implementation has been performed [86].

BSS-based methods require multi-channels to function. Therefore, such methods cannot be used in low-channel (e.g. 4-6) or single-channel based applications (e.g. in ambulatory monitoring of epilepsy patient). On the other hand, wavelet transform and EMD-based techniques can work with single-channel analysis by decomposing a single data sequence into multiple components (approx./detail coefficient for wavelet decomposition and IMF for EMD) [86].

Generally, used artifact removing methods need a dedicated artifact channel to remove artifact. In order to remove ocular or cardiac artifacts, the reference channel often provides satisfactory complementary information to identify ECG/EOG artifacts. Besides, real-time contact impedance measurement can provide the complementary information about artifacts due to electrode pop, movement or loose connection. Some movement tracking devices such as motion captured camera, accelerometer and/or gyroscope can help to detect motion artifacts [86].

It is an important matter in developing any artifact removal algorithm is robustness. Different types of artifacts can contaminate the EEG differently in different recording environments. Some factors that should be considered for robustness are artifact-SNR, type of artifact, duration of artifacts, subject-variability, environmental variability, application-specificity, etc.

Spatial distribution or topographic mapping helps to identify the origin of many artifacts (e.g. ocular artifacts are dominant in frontal EEG channels) [190]. In addition, some artifacts may appear in several nearby channels (global artifacts

such as eye blink) where some appear only in one channel (i.e. local artifacts). Therefore, spatial features along with their spectral content are important to identify artifacts from EEG signals [136, 181].

Still now, it is an active area of research to develop methods for artifact detection and removal in EEG applications and for this significant amount of efforts have been made. Most of the artifact handling method removes single type artifact, some of them cannot work for single-channel EEG, few methods require training data, some need a dedicated reference channel, some are designed for general purpose applications that often leads to overcorrection of data and some of them are not fully automated. Currently available some of the major software plug-in GUIs are FORCe, FASTER, LAMIC, PureEEG, OSET, MARA, AAR, ADJUST and ERPLAB.

Fully online and automated artifact Removal for brain-Computer interfacing or FORCe is the most recent method reported in [34] that is based on a unique combination of WT, ICA and thresholding. Compared with two other state of-the-art methods namely LAMIC and FASTER, FORCe has been shown to be better than other and is capable of removing different types of artifacts including eyeblink, EOG and EMG. One of noticeable features of FORCe is that it does not require any reference channel and can operate on fewer numbers of channels which makes it suitable for ambulatory EEG applications.

On the contrary, FASTER is an unsupervised algorithm for parameter estimation in both EEG time series and in the ICs of EEG [148]. In this method, the achieved sensitivity and specificity is $> 90\%$ for finding of EOG and EMG artifacts, linear trends and white noise in the contaminated channels.

Another method LAMIC is a clustering algorithm developed for automatic artifact removal from EEG [146]. It includes data decomposition by a BSS algorithm called TDSEP, which is a temporal extension of ICA. Then the decomposed components are clustered using the similarity of their lagged auto-mutual information (AMI). This is inspired from the fact that EEG and artifacts are different from their temporal dynamics point of view. The clustering procedure follows the usual steps of hierarchical clustering.

An automatic EEG artifact removal algorithm for epilepsy monitoring is PureEEG

by utilizing an iterative Bayesian estimation scheme [71] that is based on a neurophysiological model. It targets to remove all types of artifact without requiring any manual intervention. It is reported by the authors that the performance of PureEEG method significantly improves the readability of EEG recordings after artifact elimination.

An open-source electrophysiological toolbox named as OSET is used for biomedical signal generation, modeling, processing, and filtering [170]. Using this method cardiac artifacts are detached from any bioelectrical signal including EEG. To detect and remove EOG artifacts from multi-channel EEG, semi-blind source separation techniques is used.

Multiple artifact rejection algorithm is called MARA that is actually an open source MATLAB-based EEGLAB^2 plug-in which automatically identify the artifact-contaminated ICs for artifact removal [218, 217]. A supervised machine knowledge based algorithm is MARA that learns from labeled components by experts and utilizes six features based on spatial, spectral and temporal domain. It is used for all types of artifact. EEGLAB^2 is an open-source MATLAB-based interactive GUI toolbox for analyzing and processing continuous and event-related EEG, MEG and other electrophysiological signals. It uses ICA, time-frequency analysis, artifact rejection, event-related statistics and different modes for visualizing the averaged or single-trial EEG data [40].

Automatic artifact removal (AAR) is a MATLAB toolbox which joined as a plug-in into EEGLAB, includes various types of artifact removal methods for eliminating only EOG and EMG artifacts [61]. In order to remove simply EOG artifacts, regression-based methods such as conventional re-cursive least squares (CRLS), least mean squares (LMS), stable re-cursive least squares (SRLS) are used. Spatial filters based techniques have been adopted for removing both EOG and EMG artifacts.

ADJUST, reported by Mognon et al. [136], is an EEGLab supported plug-in for automated EEG artifact detection. This algorithm is based on the combination of specific spatial and temporal features of stereotyped artifact to identify automatically the artifactual ICs after ICA is executed. Four different artifact types (i.e. eye blink, vertical eye movement, horizontal eye movement and generic discontinu-

ities) are chosen for extracting features such as temporal kurtosis, spatial average and variance difference, maximum epoch variance, spatial eye difference. The main feature of ADJUST is that it is completely automated and unsupervised with reported accuracy of 95.2% in classifying all of the four artifacts. The clean ERP topographies are successfully reconstructed from heavy artifact-contamination.

ERPLAB is also EEGLAB-compatible open-source toolbox for analyzing ERP data, which has artifact rejection capability in both manual and automated manner [112].

1.8 Recommended Scheme

In order to choose the appropriate artifact handling method, required specification to be satisfied given the computational resources and recording environment are available to consider the particular application. In EEG applications, there are one or two types of artifacts that affect the later stage information decoding or processing, thus it is not wise to attempt to identify and remove all the artifacts as other artifacts may not (or minimally) harm for a particular signal processing purpose. If any reference channel is available in the targeted application, then regression or adaptive filtering technique may be a favorite solution. In the case of EEG monitoring, when number of channels are fewer, no reference channel is available and wireless EEG transfer preferred, in such case it is recommended to use computationally cheaper method that can work without reference and on single or few channels, e.g. wavelet-based methods. Since BSS-based methods may not perform satisfactorily with less number of channels because it needs multichannel. In some applications, if it is possible to have some a priori knowledge about artifacts and some training data are available, and the application only requires to identify artifacts not to remove them, then machine learning based classifiers can be good choice. When EEG recording holds high-volume channels, then principal component analysis (PCA) may be preferred to reduce the dimensionality before applying any artifact reduction methods, as for example BSS-based methods. If the application is based on offline analysis, then we can afford some computational expensive techniques such as ICA or EMD.

The existing artifact detection and removal methods have been presented broadly with their comparison, advantages and limitations. Still now it is an active area of research to handle artifacts present in the typical EEG recordings and none of the existing methods can be considered as the perfect solution. Most of the solutions and adjustment do not consider for the particular application. Although, maximum elimination algorithms provide good performance and they are only suitable for offline analysis because of their high computational complexity and unsupervised nature. It is not possible for particular applications which method requires a dedicated reference channel. To observe the effects of artifact contamination to the desired signal processing method and further studies are required to characterize the properties of commonly encountered artifacts. Some applications e.g. classification/identification of two classes are required to identify artifacts and not to remove them. In this case, a more accurate mathematical model of the preferred event(s) to be identified is essential in order to easily ignore other non-brain signals (i.e. artifacts or interferences). Finally, application-specific solutions with reasonable complexity, optimized performance and most importantly with feasible solutions will need for future plan.

1.9 Thesis Motivation

1.9.1 Artifact Reduction from Scalp EEG

Nowadays, it is an active area of research to develop methods for artifact detection and removal in EEG applications where significant amount of efforts have been made. Most of them handle single type of artifact, many of them cannot work for single-channel EEG, some of them require training data, some require a dedicated reference channel, some are designed for general purpose applications that often leads to overcorrection of data and some of them are not fully automated. The proposed artifact reduction method for clean EEG application is automatic, independent of artifact types, does not require a reference channel, can work for both single and multi-channels. The method uses hybrid wavelet transform with adaptive threshold parameters which relies on some spectral information of pure EEG. In order to separate artifacts from contaminated EEG, it utilizes adaptive

threshold method. The proposed method is to preserve the clean EEG at all time and then to reduce artifacts as much as possible.

1.9.2 Artifact Reduction in BCI Experiments

Brain computer interface (BCI) is establishing a direct link between human brain and an external computerized device to allow communications for the person suffering from brain/spinal cord related injury/disease. EEG is the most popular brain signal recording techniques in BCI research. It is widely used because of its non-invasive nature along with other attractive features such as fine temporal resolution, simple to use, portability and lower cost. In EEG-based BCI applications, the EEG recording is often contaminated by different types of artifacts that can misinterpret the BCI output. Recently, many efforts/algorithms have been made to develop suitable methods for artifact detection and removal from EEG in the field of signal processing technique. However, there has no universal and complete solution been proposed yet (e.g. some of them can only remove ocular artifacts, some of them require an additional reference channel whereas some of them cannot work on single-channel, etc.) which indicates the future need for the removal of algorithms to be application-specific such as BCI applications [85].

This thesis also presents a distinctive artifact detection method for BCI application based on adaptive threshold method. Successively a removal method is also proposed which is based on hybrid wavelet transform based denoising and relies on the subband approach and EEG signal enhancement for improved motor imagery classification particularly for BCI applications. The proposed method is demonstrated with both real and synthesized database and extensive quantitative measures show the efficacy of this method with obtained satisfactory results. Moreover, the proposed method is also compared with some of state-of-the-art methods and proves its superiority over others both in terms of performance metrics and computational time. The artifact removal effect on real BCI database has also been proved that it can substantially improve the classifier accuracy in BCI experiments.

From above discussion, it is clearly shown that the available EEG artifact removal methods have some practical limitations for real time applications in

BCI experiments. The present research is very important in signal preprocessing platform to reliably detect and eliminate artifacts in non-invasive EEG recordings that mainly motivate the study presented in the thesis.

1.10 Thesis Objectives

Based on the problem definition mentioned above, the aim of this thesis is to provide methods for improving the quality of EEG signals which are widely used in basic neuroscience studies, diagnosis of different neurological diseases and modern BCIs (both invasive and non-invasive). The target of this research project is the development of the novel online methods to preprocess and clean EEG from multichannel signals. The model based approach e.g. hybrid wavelet transform can be employed to achieve the goal of this research. The work presented in this PhD manuscript belongs to the framework of BCI research. More precisely, it focused on the study of EEG signal processing and classification techniques in order to design and use BCI for interacting with virtual reality applications. Indeed, despite the valuable and promising achievements already obtained in the literature, the BCI field is still a relatively young research field and there is still much to do in order to make BCI become a mature technology. Accordingly, the specific objectives of this thesis are as follows:

- Study of artifacts present at EEG recordings and characterize them in a systematic way.
- Analysis of the effects of artifacts on EEG signals.
- Propose an artifact detection and removal algorithm for cleaning EEG.
- Observe the after-effects of artifact removal on later-stage signal processing applications (e.g. BCI performance).
- Analysis of the effects for proper quantitative evaluation of both artifact removal and the after effect of removal to neural information processing.

This is done in terms of EEG artifact detection performance and motor imagery EEG signal classification accuracy. Facilitated and encouraged by the new

understanding of brain functions and by the low-cost computer equipment, these programs have concentrated mainly in developing new communication and control technologies for people with severe neuromuscular disorders. The immediate goal of BCI is to provide communication capabilities without using the brain's normal output pathways of peripheral nerves and muscles. As a result any subject can control the external world with BCI. The implementation of online (real time) BCI has multiple uses e.g., biomedical engineering applications-assisting devices for disabled people; human subject monitoring-sleep disorders, neurological diseases, attention monitoring, and/or overall mental state"; neuroscience research-real time methods for correlating observable behavior with recorded neural signals; brain machine interaction (BMI) - interfacing devices between human brain with other machines.

Chapter 2

Artifacts with EEG

The electroencephalogram (EEG) as a multi-channel signal of neuronal brain activity can reflect brain states linked to the mental condition of a person. Due to its temporal resolution, it is an excellent and widely used technique for investigating human brain functioning. Unfortunately, EEG signals are often contaminated with artifacts. Artifacts change the quality of EEG signals and subsequently degrade the BCIs performance. So, the recorded EEG samples are not equally reliable due to the artifacts at the time of recording. Different biological and non-biological artifacts such as eye movements, blinks, muscle activity, heartbeat, high electrode impedance, line noise, and interference from electric devices are responsible for the contamination of the EEG signal. The discarding of entire EEG segments due to noise is a widely applied method in research settings and results in the loss of experimental data. If only a few epochs are available and artifacts such as blinks or movements occur too frequently, then it becomes especially problematic. Moreover, this approach is inappropriate when working with the continuous EEG and non-event-locked activity, long-range temporal correlations, real-time brain-computer interface (BCI) applications, and online mental state monitoring [158]. This chapter describes the study on types of artifacts and effects of artifacts on electroencephalography.

2.1 Types of Artifacts

Different types of artifact in EEG typically have different characteristics. For example, the amplitudes of ocular or body movement artifacts are usually much higher than those of the EEG activities of interest. High-frequency and low-amplitude activities accompany muscle artifacts. The removal of undesired arti-

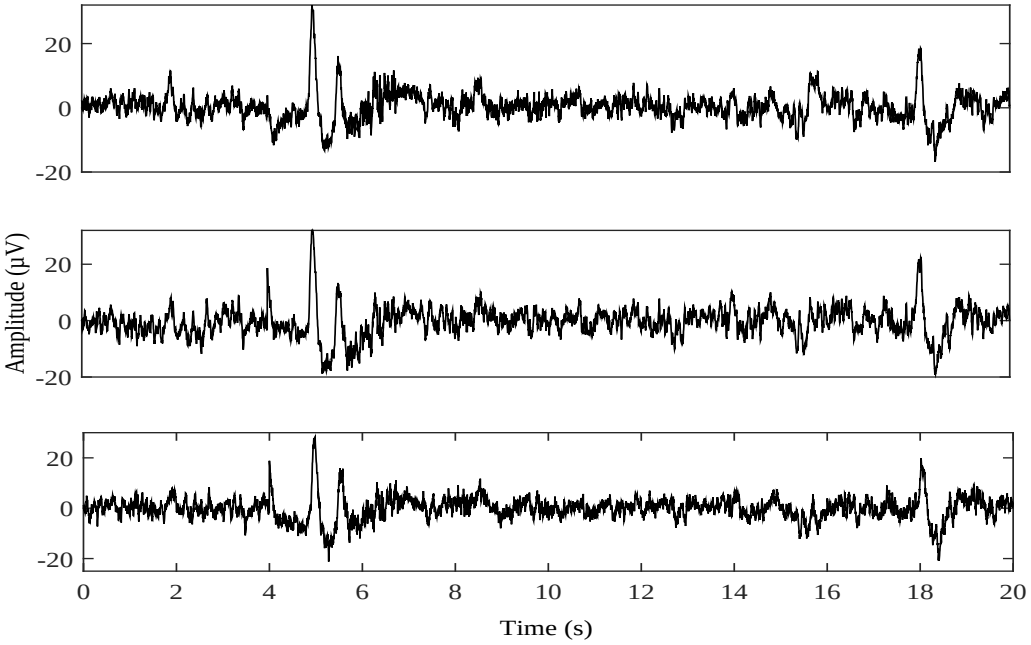

Figure 2.1: EEG recording corrupted by eyeblink artifacts.

facts from the EEG is a major preprocessing step for most EEG analysis. Such artifacts stem from eye and muscle movement, the heart beat or external technical sources. In this paper, we are concerned with the removal of muscle artifacts. These are typically caused by muscle activity near the head, such as swallowing or head movements, and are characterized by high-frequency activity ($> 20Hz$). Because muscle activity arises from different type of muscle groups, muscle artifacts are harder to stereotype than eye artifacts [62].

2.1.1 Internal/Physiological Artifacts

Most of the algorithms developed for EEG artifact processing are intended for the reduction of physiological artifacts. The physiological artifacts are:

- **Ocular Artifacts:** The electrooculogram (EOG) measures the electrical activity produced by eye movement, which is normally strong enough to be recorded along with the EEG. These movements are primarily picked up by the frontal electrodes, although they also extend further. The strength of the interference depends on the proximity of the electrodes to the eyes and the direction in which the eye is moving. Blinking also causes the EEG recording to become contaminated, usually with a more abrupt change than that

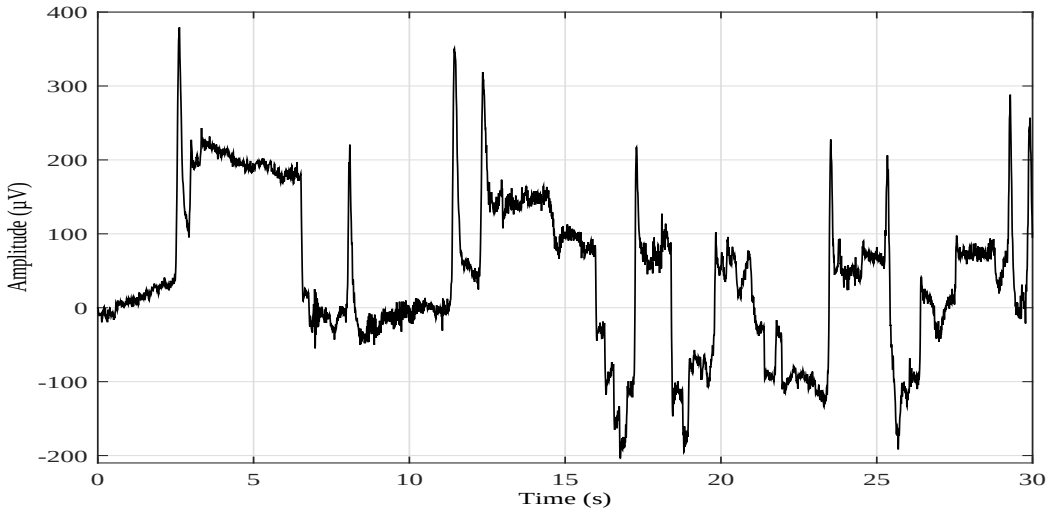

Figure 2.2: EEG signal with horizontal and vertical eye movement artifacts.

produced by eye movement, which is associated to higher frequency interference. Furthermore, the amplitude of the background EEG activity is smaller than the blinking artifact. Ocular artifacts in the literature are referred to as OAs or EOG artifacts even though we only use the latter in the text. The eyeball acts as an electrical dipole and therefore any movement in eyeball generates large-amplitude artifacts in EEG recordings. Ocular artifacts include eye blink (Figure 2.1), both horizontal and vertical eye movement (Figure 2.2), eye flatter, eye movement during REM sleep, eye saccade, etc. When the electrode in frontal position then eye blinks is typically generated abrupt amplitude. As blink activities are notably different from neuronal activities, it is possible to detect them using a suitable threshold for HWT. And these blinks are typically faster than horizontal eye movements which also generate large amplitude fluctuations in frontal channels. Vertical eye movements are caused by the time course of artifacts and it is similar to the one generated by horizontal eye movements [127]. It is very advantageous for ocular artifact elimination from a practical point of view, having reference EOG waveforms, measured simultaneously to the EEG. It is advisable for the measurements of vertical (VEOG), horizontal (HEOG) and radial (REOG) EOG since these signals are propagated differently across the scalp [206].

Source of ocular artifacts: Generally, the cerebral cortex potentials is relatively large compared to the most noticeable eye-induced artifacts produced by the potential difference between the cornea and retina. This corneo-retinal dipole does not affect EEG when the eyes and eyelids are completely static. However, the eyes movements occur several times per second and eye blinks occur several times per minute. Eyelid movements, elicit a large potential seen mostly in the difference between the EOG channels above and below the eyes and occurring generally during blinking or vertical eye movements. A well-known description of this potential that short circuit the positively charged cornea to the extra-ocular skin regards the eyelids considered as sliding electrodes. Increases the potential in one side electrodes towards which the eyes are rotated, and decrease the potentials in the opposing electrodes during rotation of the eyeballs, and consequently of the corneo-retinal dipole. Generally, saccades are caused from eye movements and also generate transient electromyographic potentials, known as saccadic spike potentials (SPs). It is requiring tailored artifact correction approaches when the spectrum of these SPs overlaps the gamma-band and seriously confounds analysis of induced gamma-band responses. There is reflexive movement of the eyeball during blinking that gives a characteristic artifactual appearance of the EEG and the purposeful or reflexive eye blinking also generates electromyographic potentials.

Kappa rhythm (or Kappa wave) is characterized by the characteristic of Eyelid fluttering artifacts. It is usually seen just over the eyes that means in the prefrontal leads. Sometimes, they are seen with mental activity. They are usually in the Theta (47 Hz) or Alpha (714 Hz) range. It is believed that they were originating from the brain. Sometimes, it was so minute that it was difficult to see but later study revealed they were generated by rapid fluttering of the eyelids. It should not technically be called a rhythm or wave because they are in fact noise in the EEG reading.

Some of these artifacts are useful in many applications. For example, the EOG signals are used to identify and track eye-movements, and are also in conventional EEG for evaluating possible changes in alertness, drowsiness or

sleep [96] and very essential in polysomnography.

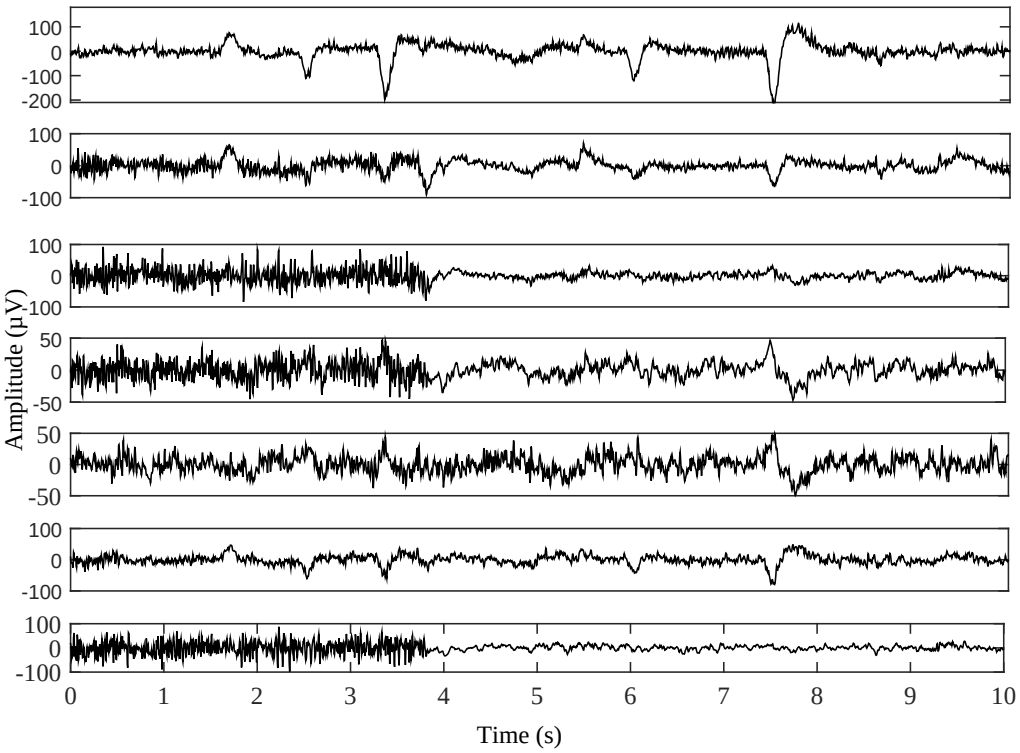

Figure 2.3: Raw EEG signals with strong EOG and EMG artifacts.

- **Muscle Artifacts:** The electromyogram (EMG) measures the electrical activity on the body surface caused by contracting muscles. The muscle artifacts occur when the patient swallows, talks, walks, etc., being more detrimental in uncontrolled environments. This artifact is typical of awake patients. They are hard to stereotype and the shapes and amplitudes of the interferences depend on the degree of muscle contraction and on the type of muscle contracted. The artifacts are termed MAs or EMG artifacts. EMG are responsible for its adverse effects on the EEG background activity and make it more difficult to correct than other types of artifacts because it has number of cranial properties. In exact it noticeably overlaps with beta activity in the 15-30 Hz range but may be as low as 20 Hz, making the widely used alpha band also susceptible to muscle artifacts. Although, EMG presents a wide spectral distribution thus perturbing all classic EEG bands. Due to volume conduction of myogenic activity independently generated by muscles across the head, face and neck, EMG can often be detected across the entire scalp. Figure 2.3 shows the different EEG channels which are

contaminated by EOG and EMG artifacts.

A variety of experimental manipulations like cognitive load and vocalization are responsible for EMG which is temporally mixed with those. Finally, EMG is more difficult to characterize and also exhibits less repetition than other biological artifacts. Since it arises from functionally independent muscle groups, the activity is spatially distributed with distinct topographic and spectral signatures.

One of the most prominent physiological artifacts comes from muscle activity of the subject (EMG). Usually muscle artifacts are of high frequency range (e.g. from 20 Hz to 40 Hz) and are generated from activities like chewing, swallowing, clenching, sniffing, talking, scalp contraction, eyebrows raising, etc. Figure 1.12 illustrated the clean EEG is contaminated by muscle activity.

- **Cardiac Artifacts:** Cardiac artifacts are due to the electromagnetic field produced by heart and are of two types: ECG and pulse artifacts.
  Electrocardiogram Artifacts: The electrocardiogram (ECG) artifacts are rhythmic regular activities while the pulsation sometimes can cause slow waves which might mimic the EEG activity. Figure 1.13 shows the ECG waveform. ECG artifacts are common in EEG and it is deliberated act as like as spike activity. Modern EEG acquisition commonly includes a one-channel ECG from the extremities. This also sanctions the EEG to recognize cardiac arrhythmias that is an important differential diagnosis to syncope or other episodic/attack disorders.

  The ECG measures the electrical activity of the heart. Depending on the electrode positions and differs from certain body types, the amplitude of the cardiac activity on the scalp is usually of low amplitude. The ECG has a very characteristic repetitive and regular pattern, which unfortunately may be sometimes mistaken for epileptic form activity when the ECG is barely visible in the EEG. The ECG is routinely measured aside cerebral activity which making this artifact easier to correct. In ECG signal, reference waveform is usually available.

Pulse artifacts: It occurs when an EEG electrode is generating slow periodic waves that may resemble EEG activity and it places over a pulsating vessel such a scalp artery. The pulse activity only occurs on one electrode and can be minimized by proper sensor positioning. Thus, it is much harder to correct the ECG since it may be similar in time and frequency content to the measured EEG itself. In fact pulse waves are easily recognizable due to their regular occurrence and since they precede the ECG by a constant interval [206] though there is a direct relationship that exists between ECG and the pulse activity.

- **Glossokinetic Artifacts:** Glossokinetic artifacts are caused by the potential difference between the base and the tip of the tongue. Minor tongue movements, mainly occurs in parkinsonian and tremor disorders and the EEG signal can be contaminated by this artifacts.

- **Respiration Artifacts:** Respiration artifacts originate from the movement of an electrode with inhalation or exhalation and can take the form of slow, rhythmic EEG activity.

- **Sweat Artifacts:** Sweat artifacts or electrodermal takes the shape of a long, slow baseline drift in the spectral band of 0.25 -0.5 Hz [85] and originated from changes in electrolyte concentration of electrode due to sweat secretion on the scalp.

- **Less common physiological artifacts:** In addition to the artifacts described before, two interferences may arise from skin potential: perspiration artifacts and to a smaller extent. The perspiration artifacts which are slow waves produced by shifts of the electrical baseline of certain electrodes; and to a smaller extent, the sympathetic skin response, which also holds of slow waves and is an autonomic response produce by sweat gland and skin potentials. Other possible artifacts include, dental restorations with different metals, movements of the tongue, electrodermal interferences due to sweating, or breathing artifacts in the lower part of the spectrum, chest movements, etc [206]. We do not meet with these artifacts further since they are hardly treated in the literature that we have described.

2.1.2 External Artifacts

- **Movement Artifacts:** Movement of patient especially in an ambulatory EEG monitoring system [85], generates a lot of motion artifacts. This artifact often has extremely high amplitude such that it can saturate the recordings. Head movement, body movement, limbs movement, tremor, walking, running, browsing PC, and many other movements in daily activities are responsible for this type of artifact.

- **Environmental Artifacts:**

  - **Interferences:** This type of artifacts is due to the interferences coming from the surrounding electrical/electronic devices/machines that produce EM waves. Also any sound or optical interference may also be picked up by the EEG electrodes as artifacts.

  - **Mains voltage:** One of most common source of artifacts in any biomedical signal acquisition is the 50/60 Hz main voltage and its harmonics.

  - **Loose electrode:** Loose electrode contact with scalp can lead to impedance change on the tissue-electrode interface and results in prolonged EEG spike-like artifact.

  - **Electrode pop and movement:** Another common source of artifact is due to electrode pop which produces sudden change in impedance in the electrode-tissue interface and results in high amplitude sharp waveform-shaped artifacts. Electrode movement occurs when it moves with respect to the scalp and produces high-amplitude detection in EEG generally in the low frequency range of 1-10 Hz.

2.2 Artifact Avoidance

Artifact avoidance is a protective and defensive way to minimize or avoide artifacts and it tries to avoid unnecessary eye blinks, eye/body movements by instructing the subject to remain static. Also one can reduce the supply mains interference by proper grounding of the EEG recorder. Artifact avoidance is minimizing artifacts,

it can reduce both the data loss and the computational complexity but it is not the best way to get rid of artifacts completely. Sometimes this is a very unrealistic solution based on applications, e.g. in an EEG monitoring or BCI applications. Moreover, some of the physiological artifacts (e.g. ECG) are involuntary and therefore cannot be avoided. So, there are several limitations to employ such approach in artifact removing. In addition, the subject cannot limit eye blinking or movement for a long period of time, especially if the subject is neonatal or children. Therefore, those should be handled in the digital signal processing domain because there will always be some artifacts present in the EEG recording.

2.3 Artifact Detection

Identifying artifacts is the first and most important step for handling artifacts. It becomes difficult to use simple filtering or straight forward signal processing technique in both spectral and temporal domain when the artifacts overlap with EEG signals. In order to identify the applications faster, it is required to identify or separate artifacts in real-time. Therefore, it is really necessary to know both the artifact and signal characteristics. Detection of artifacts may refer to detect an independent component or detect a particular epoch to be artifactual. Whether it should be detected in time domain or frequency domain or even in both by utilizing time-frequency analysis, this decision depends on the type of artifacts and/or type of applications. Depending on the various factors, the detection method also varies whether a reference artifact source is available or not, whether we want to remove the artifacts after detection, whether the no. of channels are enough, and so on.

To detect artifacts caused by eye-blinks, saccades, and head movements, the eye-gaze information obtained from the eyetracker is employed. The images of the right and left pupil were used to detect eye blinks. A blink was declared present whenever both pupils were not detected by the eye-tracker. In addition, a saccade was declared as present whenever a change in the point of gaze happened. To detect head movements, the change in the distance (from the screen to the pupils) between that of where the pupils were calibrated and their present position was used.

2.4 Artifact Rejection

Artifact rejection denotes to the practice of eliminating the trials or epoch from EEG affected by artifacts. Perhaps it is the simplest way of cleaning the brain signals which is polluted with artifacts. It has several essential benefits over the artifact avoidance method. For instance, it would be relaxed for those subjects with motor disabilities and the subjects to contribute in the experiments and execute the necessary tasks. The cognitive task resulting from a subject annoying to escape producing a specific artifact, will not be existent in the EEG signal [48]. This artifacts removeing process not only removes artifact but also removes important EEG information which results in the loss of data. This is the early days way of handling artifacts, but nowadays with the introduction of recent signal processing techniques, the preference is more on the techniques for artifact removal or correcting them instead of rejecting the data epoch. However, in certain applications, this technique can still work reasonably well, e.g. offline analysis or during training of any classifier.

2.5 Existing Artifact Cleaning Methods for Different Artifacts

The process of recognizing and eliminating artifacts from brain signals is called artifact removal. An artifact subtraction technique should be able to eliminate the artifacts as well as to keep the related neurological phenomenon intact. There are a wide-ranging techniques for detecting and/or removing artifacts (e.g., EOG, EMG, ECG) in EEG. Some of the prominent artifacts correction is described below:

2.5.1 Electrooculography

The EOG is one the best methods for removing blinks and saccades from EEG measurements, as confirmed in [30, 156]. It is possible, if we go to publications that use real EEGs, the aligned-artifact average solution which is described in [29].

Depending on taking a reference channel for reducing the ocular activity from the EEG (up to vertical, horizontal and radial EOG references for optimal performance [29]), regression algorithm is used. The other way, artifacts in experimental states [41, 74, 209, 89, 94], of which undoubtedly that of [94], using extended InfoMax, is the most thoroughly acceptable. Recently, the similar research group showing that ICA is able to produce a denoised EEG of quality equivalent to that from the EOG correction methods and confirmed that RAAA has presented initial results [206]. According to simulations and visual inspection of results on attained EEGs, SOBI stands out as the best performing method for removing ocular artifacts from background EEG activity and ERPs in multiple publications [166, 68, 165, 99, 92]. Moreover, it is easily used to facilitate source analysis from EEGs [198] and magnetoencephalographic (MEG) data [197, 196, 196], since it has been shown to be robust across subjects [198] and over large time intervals [195].

Many articles reported to remove EOG artifacts by the use of EOG reference channel [55, 101, 232]. In [232], it has been stated that it combines discrete wavelet transformation (DWT) and an adaptive predictor filter (APF) which is termed as a hybrid denoising method. To accomplish lower MSE and higher correlation among cleaned and original EEG, APF is establish to better performance than existing methods such as wavelet packet transform (WPT) and independent component analysis (ICA), DWT and adaptive noise cancellation (ANC). The APF is performing well for automatic identification and removal of ocular artifacts for portable EEG applications. Another article [101] stated an automated ocular (eyeblink) artifact elimination method through adaptive filtering and ICA with the help of horizontal (HEOG) and vertical (VEOG) EOG channel as reference. On the other hand, Flexer et al. [55] offered an ICA-based ocular artifact elimination technique from blind subjects EEG using both vertical and horizontal EOG references.

2.5.2 Electromyography

The existence of muscle artifacts in EEG signal stances a more challenging situation than that of ocular artifacts, subsequently a reference waveform is rarely available [94] and because spatial topographies need not to be as well defined as

with contamination coming from the eyes. There is no existing EMG correction methods unlike ocular artifacts that can be removed by the aforementioned EOG correction methods, as well as by other more general methods such as BSS, i.e. authors have not developed algorithms in the literature intended specifically to cancel out muscular interference. Newly, a literature review in [142] presented on several techniques for high frequency interference from EEG and MEG activity that help suppress muscle artifacts and others. According to his survey, disagreement exists in the literature about the effectiveness of ICA in removing EMG activity from data, a fact that is also reported elsewhere [131, 130, 176, 132, 150]. The writer concludes that even though there is a number of methods are available for the reduction of EMG artifacts, at present none can guarantee processed data free of high frequency artifacts. Considering the earlier statements, EMG has been reduced quite successfully from contaminated EEGs. The muscular artifacts are cancelled in [212, 208] by using CCA, in [39, 66] by using InfoMax, in [28, 32] by employing SOBI and in [168] by using EMD for low SNR, CoM2 (and ICA algorithm) and CCA for intermediate SNR and DWT for high SNR. Among them, the cases in [168, 28, 208, 32] are mainly based on simulations. Then, in addition to the previous papers, [212, 93, 39, 32, 66, 84] analyze the results visually and [131] validates ICA on the basis of its sensitivity and specificity. Obviously, the results verified by visual inspection and the performances obtained with simulated data need to be properly validated, as in [131] but present readings do not yet offer a unified method for doing this. The EMG artifact reduction is not as straightforward as EOG suppression because reference signals are rarely available. Even though CCA is widely used in the literature, according to our practical experience it does not outperform ICA.

2.5.3 Electrocardiography

Writers in [38] proposed elimination/reduction of ECG/cardiac artifacts from EEG using a separate ECG reference channel. In [65], an automatic technique based on a modified ICA system has been proposed that works for a single-channel EEG and the ECG (as reference) which gives promising results when compared with two popular methods that use a reference channel namely ensemble average subtrac-

tion (EAS) and adaptive filtering. The other two articles proposed their methods for application in neonatal EEG monitoring. Another paper [143] proposed a combination of EMD and adaptive filtering based approach for ECG artifact removal in preterm EEG and reported up to 17% improvement in correlation coefficient between original and cleaned datasets compared with removal by only adaptive filtering.

2.6 Target Artifact for EEG Cleaning

The canceling or correcting the artifacts without distorting the signal of interest is the great challenge for artifact removal methods. This is primarily done in two ways: either by filtering and regression or by separating/decomposing the EEG data into other domains. In this dissertation, the low frequency and high energy signal components representing electrooculogram artifacts are suppressed from the recorded electroencephalography. It concentrated mainly on removing eye blink ocular artifacts.

Chapter 3

Subband Decomposition Methods

The electroencephogram (EEG) signal is decomposed into multiple subbands (multiband) signals to extract the brain waves. Such multiband decomposition can be achieved by applying the filter bank technique on the EEG signals. In processing of a signal, a group of bandpass filters are called a filter bank that separates the input signal into multiple components and component carrying a frequency subband of the original signal. The analysis of signal in terms of each subband is the process of decomposition of a signal performed by the filter bank. The output of analysis signal is denoted as a subband signal. Synthesis is the reconstruction process which is the reconstitution of a complete signal resulting from the filtering process. To separate the brain waves [137], the both analysis and synthesis of the EEG signal are required. The EEG contains information about brain hence the subband decomposition of EEG is used for analyzing many brain diseases. To extract various brain waves with different frequency bands such as alpha, beta, delta, theta and gamma from EEG signal to get more information from [177] it is called the subband decomposition. Using EMD, discrete wavelet transform and the proposed hybrid wavelet transform (HWT) method, the EEG signal is decomposed here. The EEG signal is decomposed into five sub-bands alpha, beta, gamma, theta, delta using daubechies wavelet. Based on application, these decomposed brain waves can be given to any network as input for further analysis. The decomposed signal was further reconstructed to obtain the original signal. Original signal was compared with the reconstructed signal and mean square error (MSE) was calculated. The work carried out shows that the MSE for HWT is less as compared to that of the other methods. So, HWT is the best suited wavelet for subband decomposition.

3.1 Fourier Transform (FT)

The Fourier transform widely used to implement filter bank is a mathematical operation with various physics and engineering applications. It expresses as a frequency spectrum which established a relation as a function of time and as a function of frequency. The function of time and frequency is often called time domain and frequency domain representation, respectively. The opposite Fourier transform is called inverse Fourier transform states a frequency domain function in the time domain. Each value of the function is usually expressed as a complex number (called complex amplitude) that can be interpreted as a magnitude and a phase component. The transform is called Fourier transform for the transform operation and to the complex-valued function production. In the case of a sinusoidal or non-sinusoidal periodic function, the Fourier transform is used to simplify the calculation of a discrete set of complex amplitudes. Still now it is possible to recreate discrete-time Fourier transform which is a version of the original Fourier transform according to the Poisson summation formula. It is essential when a time-domain function is sampled to facilitate storage or computer processing.

The conversions of a time-domain signal of infinite duration into a continuous spectrum composed of an infinite number of sinusoids are termed as the continuous Fourier transform. It is observed, we deal with the signal that are discretely sampled, and of finite duration or periodic, usually at constant intervals. For this data, the discrete Fourier transform (DFT) is appropriate only a finite number of sinusoids are needed. There is a related theorem or property for the DFT, for almost every Fourier transform theorem or property. The DFT of N uniformly sampled data points x_n (where $n = 0, ..., N-1$) and its inverse are given by:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-2\pi jnk/N} \quad and \quad x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k e^{2\pi jnk/N} \tag{3.1}$$

where, the signal $x[n] \approx x_n \approx s(t)$. Sign and normalization agreements may vary, but the definition is most common over and over. The DFT, the outcome of an N-point input time series is an N-point frequency spectrum, with Fourier frequencies k ranging from $-(N/2-1)$, through the 0-frequency or the so-called DC component, and up to the highest Fourier frequency $N/2$. Each bin number denotes the

integer number of sinusoidal periods existing in the time series. The amplitudes and phases represent the amplitudes A_k and phases ϕ_k of those sinusoids. In summary, each bin can be described by $X_k = A_k e^{j\phi k}$. A set of bandpass filters are the digital filter bank is with either a common input or a summed output. An M-band analysis filter bank is shown in Figure 3.1. In this Figure, the signal $x[n]$ is decomposed into multiple band.

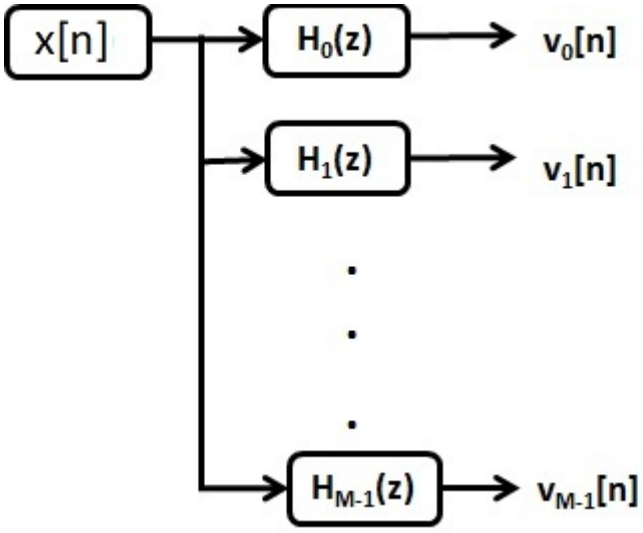

Figure 3.1: M-band analysis filter bank.

The subfilters $H_k(z)$ in the analysis filter bank are known as analysis filters. The analysis filter bank is used to decompose the input signal $x[n]$ into a set of subband signals with each subband signal $v_k[n]$ occupying a portion of the original frequency band. The mentioned M filters can be used as the analysis filters in the analysis filter bank or as the synthesis filters in the synthesis filter bank. Since the magnitude responses of all M filters are uniformly shifted version of that of the prototype filter, the filter bank obtained is called a uniform filter bank. The signal is synthesized by summing up all the subbands. The use of Fourier transform-based filter bank introduces a noticeable reconstruction error during synthesis.

It is well known from Fourier theory that a signal can be expressed as the sum of a possibly infinite, series of sines and cosines. This sum is also referred to as a Fourier expansion. The principal disadvantage of a Fourier transform is that it has only frequency resolution and no time resolution. This means that although we might be able to determine all the frequencies present in a signal, we do not know

when they are present. To overcome this problem in the past decades, several solutions have been developed which are more or less able to represent a signal in the time and frequency domain at the same time.

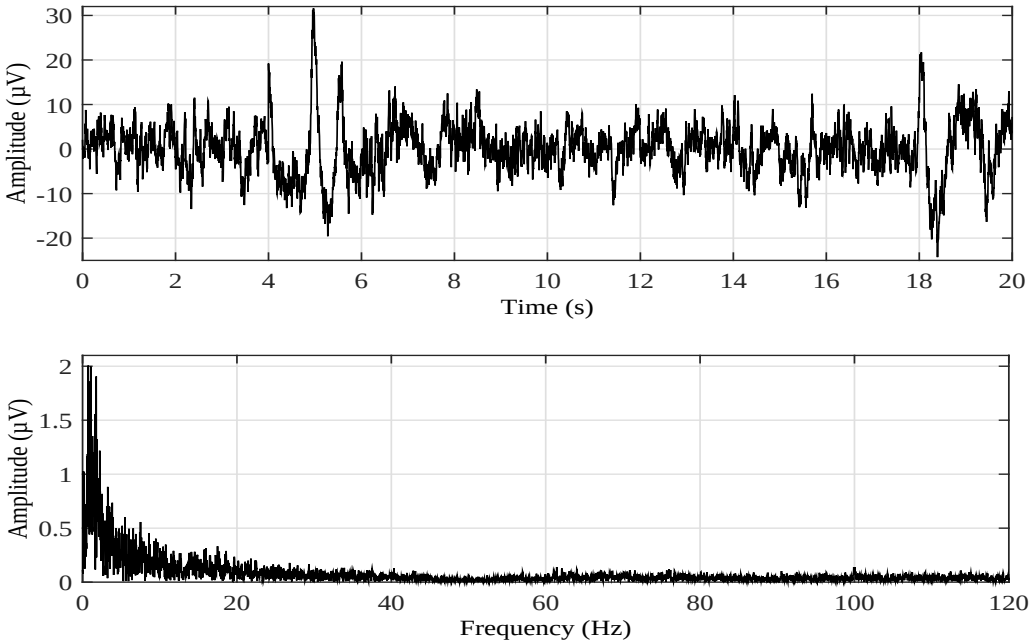

Figure 3.2: Contaminated EEG (top) in time domain and EEG signal (bottom) in frequency domain using Fourier transform.

Figure 3.2 is illustrated the contaminated EEG (top) in time domain and EEG signal (bottom) in frequency domain using Fourier transform. Also, Figure 3.3 shows Fourier log spectrum (bottom) using Fourier transform. From Figure 3.2, it is clear that the Fourier transform-based filter bank is not suitable for non-stationary EEG signals [137].

3.2 Multiband Decomposition Methods

Successful algorithms are based on the point that some sources (either signal or artifacts) can be exemplified by a single decomposition unit, such as an intrinsic mode function (IMF) for empirical mode decomposition, or resemble certain wavelet basis for the wavelet transform.

Wavelets are perfect for biomedical application in that they permit to design methods that are robust and work in most circumstances, for their finely tunable time-frequency trade-off, and for their versatility such that they can accommodate

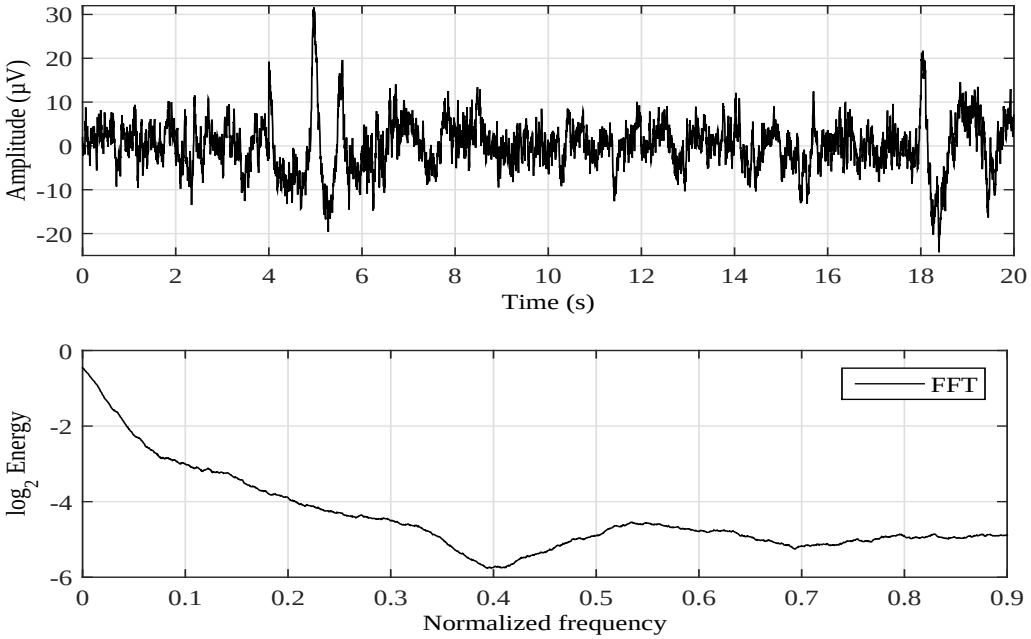

Figure 3.3: Contaminated EEG (top) and Fourier log spectrum (bottom) using Fourier transform.

biomedical signals that generally combine features with good time or frequency localization [204]. The WT has been widely used possibly commencing in the early 90s in the context of EEG denoising. The subject was evolving very rapidly by the time the review [204] was published.

It is observed that the standard Fourier transform is only localized in frequency whereas wavelets are localized in both time and frequency. Although the short-time Fourier transform (STFT) is more similar to the wavelet transform, in a sense that it is also time and frequency localized, but wavelets afford a better multi-resolution analysis of signal representation, with stable resolution at any time and frequency. The wavelet transform is a time-scale representation method that decomposes signal f(t) into basis functions of time and scale which are dilated and translated versions of a basis function $\psi(t)$ which is called mother wavelet [123]. Translation is accomplished by considering all possible integer translations of $\psi(t)$ and dilation is obtained by multiplying t by a scaling factor which is usually factors of 2.

The WT is defined as the inner product of the signal $f(t)$ with the time scaled and shifted version of the wavelet function $\psi(t)$ [204]. The WT decomposes the signal into a set of coefficients, for various scales, which represent the similarity

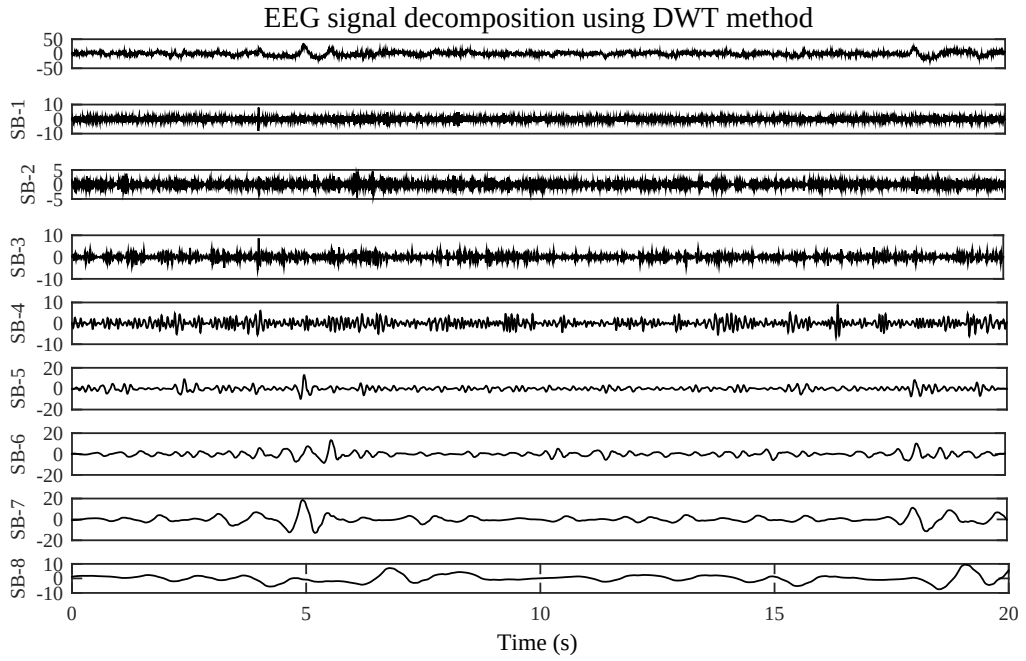

Figure 3.4: The subband decomposition of the contaminated EEG signal using DWT.

of the signal with the wavelet at that scale. There are many techniques based on wavelet theory, such as wavelet packets, wavelet approximation and decomposition, discrete and continuous wavelet transform, and so forth. The most commonly used technique is DWT. The DWT is developed from continuous wavelet transform with discrete input, but it is simplified mathematical derivation.

The DWT is any wavelet transform for which the wavelets are discretely sampled in time, that is $t + mT$. A common form of the DWT employs a dyadic grid where the amplitude and time scales are $a + 2j$ and the time shift is $b + k2j$, with both j and k integers. Briefly, discrete wavelet transform is entering a signal into a low pass filter to get the low frequency component and into a high pass filter to get the high frequency component. The DWT is usually calculated by filtering the input vector through a series of low-pass and high-pass filters that provide one approximation and D detail coefficients respectively. Thresholding is applied on the decomposed coefficients to denoise the signal from artifacts when the signal is divided into detail and approximate coefficients. Then the new set of coefficients (all detail with final level approximate coefficients) are added up to reconstruct back the artifact-reduced signal.

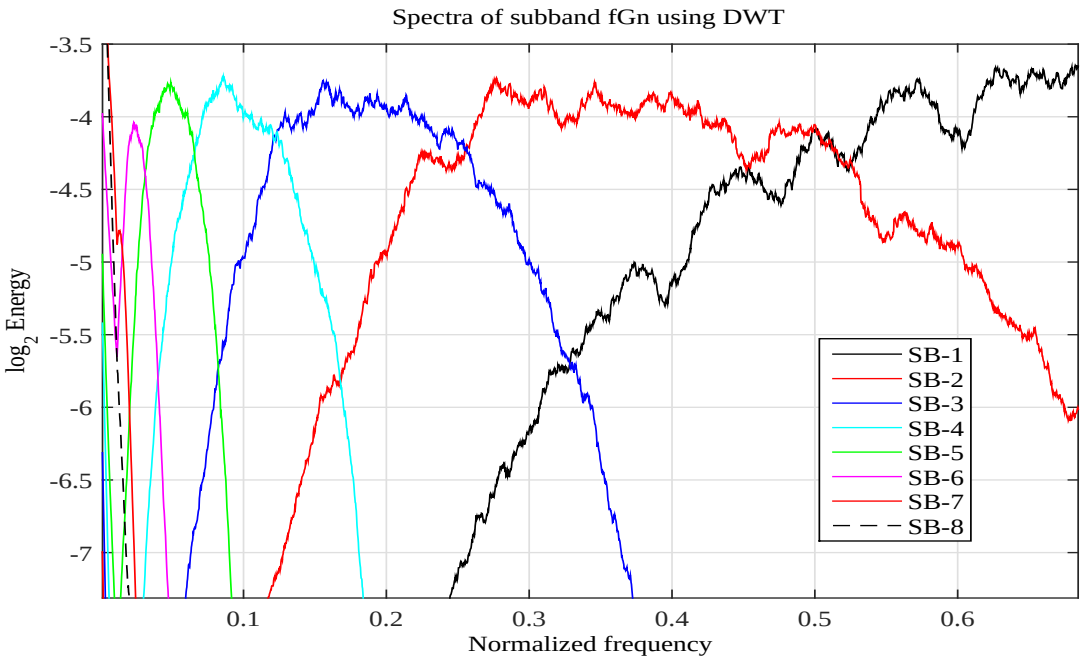

Figure 3.5: Fourier log spectrum of subbands of fGn using DWT.

3.2.1 Discrete Wavelet Transform (DWT)

In this section, we briefly review the DWT, which is the wavelet counterpart to the discrete Fourier transform. One of the superior technique in analyzing non-stationary signals like EEG is wavelet transform. To understand more the behavior of a signal, its capability helps in transforming a time domain signal into time and frequency localization. The details of the signal can be obtained at several scales by correlating the original signal with wavelet functions of different sizes. The multi-resolution decomposition using DWT is observed when these correlations with the different wavelet functions can be arranged in a hierarchical scheme. In such decomposition the signal is represented as finite set band-passed signals of different frequency bands [138]. The DWT is also widely used to perform the multiband representation of EEG data. Reconstruction of the original signal from the details obtained by wavelet filterbank is more convenient than Fourier-based method. It is also possible to design perfect reconstruction filterbank in Fourier domain, whereas wavelet filter bank automatically satisfies such requirement. It is well known in the subband filtering community that symmetry and exact reconstruction are incompatible if the same FIR filters are used for reconstruction and decomposition. Therefore, Duabachies wavelet is used here for decomposition and reconstruction

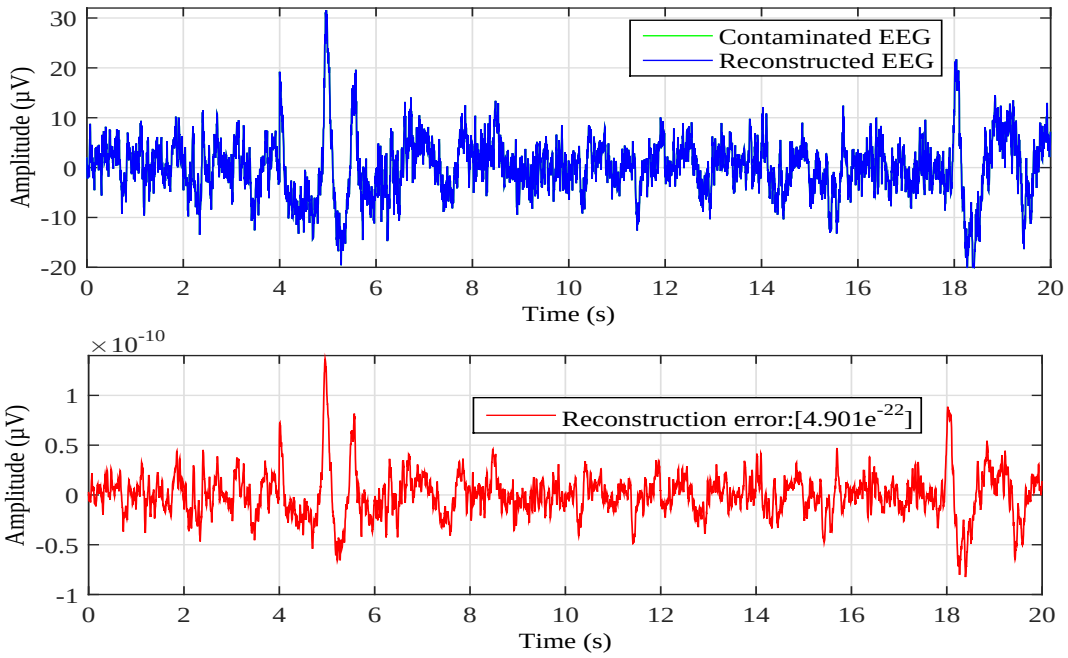

Figure 3.6: Reconstruction of contaminated EEG (top) and reconstruction error (bottom) using DWT.

signal. Figure 3.4 shows subband decomposition of the contaminated EEG signal and Figure 3.5 is represented Fourier log spectrum of subbands of fGn using DWT. In this section,the reconstruction of contaminated EEG (top) and reconstruction error (bottom) are illustrated in Figure 3.6.

3.2.2 Wavelet Packet Transform (WPT)

The WPT is a generalization of the wavelet decomposition process that offers a better performance compared to the ordinary wavelet methods. In the wavelet analysis, a signal is split into an approximation and a detail coefficients. The approximation is then itself split into a second-level approximation and detail and the process is repeated. On the other hand, WPT is applied in both the detail and the approximation coefficients are divided to get all nodes for the decomposed levels and generates the full decomposition tree. In Figure 3.7, a low (l) and high (h) pass filter is frequently applied to generate a complete subband tree to some desired depth. The low-pass and high-pass filters are generated using orthogonal basis functions [91]. The wavelet packet coefficients $C^i_{j,k}$ succeeding to the signal s(t) can be obtain as,

$$C^i_{j,k} = \int_{-\infty}^{\infty} s(t) W^i_{j,k}(t) dt \tag{3.2}$$

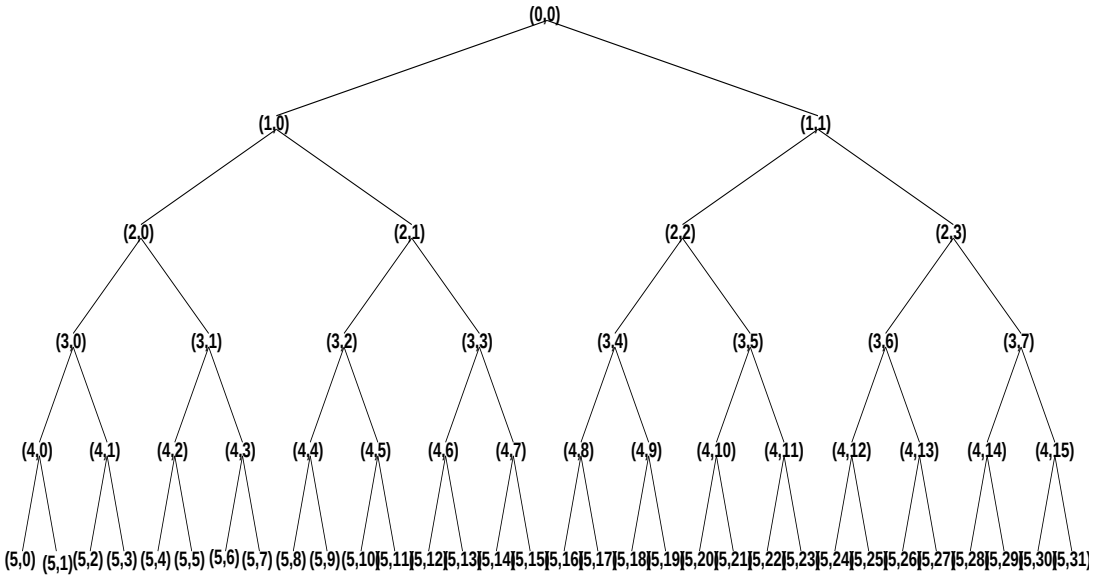

Figure 3.7: Wavelet packet tree

where, i is the modulation parameter, j is the dilation parameter, k is the translation parameter, $i = 1, 2, ..., j^L$, and L is the level of decomposition in wavelet packet tree.

The discrete filters (k) and $h(k)$ are quadrature mirror filters coupled with the scaling function and the mother wavelet function. The mother wavelet $W^i_{j,k}(t)$ is used for reconstruction and the reconstructed signal is obtained as

$$S^i_j(t) = \sum_{k=-\infty}^{\infty} C^i_{j,k} W^i_{j,k}(t)dt \tag{3.3}$$

Implementing wavelet packet decomposition up to j^{th} level, the original signal can be characterized as a summation of all wavelet packet components at j^{th} level as shown in following equation:

$$\tilde{S}(t) = \sum_{i=1}^{2i} S^i_j(t) \tag{3.4}$$

At the end of the WPT, the reconstructed subband is represented as:

$$\hat{S}(t) = \sum_{b=1}^{2^L} C_b(t) \tag{3.5}$$

where,$\tilde{S}(t) \approx \hat{S}(t)$ and $C_b(t)$ is the reconstructed subband of WPT.

By applying WPT on each channel, it produce 2^L sub band wavelet packets, where L is the number of levels. The structure of WPT decomposition, the lower

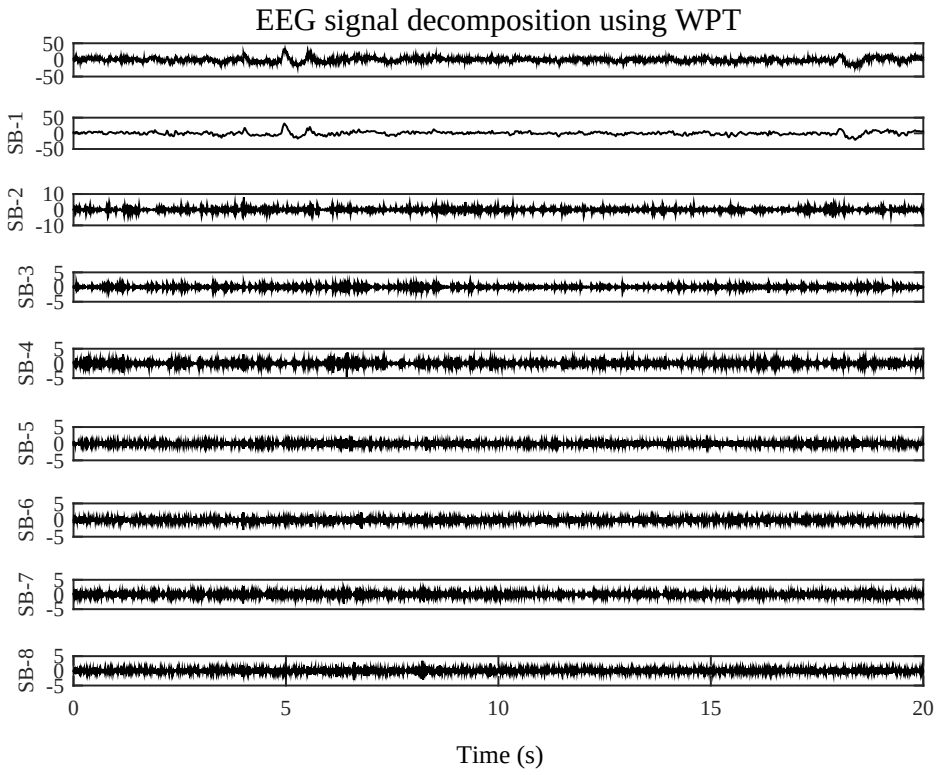

Figure 3.8: The subband decomposition of the contaminated EEG signal using WPT.

and the higher frequency bands are decomposed giving a balanced binary tree structure in Figure 3.8, the subband decomposition of the contaminated EEG signal using WPT is illustrated. In this present work, three levels generated 8 subspaces ($2^L = 2^3$) and wavelet frequency interval of each subspace is calculated by

$$\left\{ \frac{(b-1)f_s}{2^{L+1}}, \frac{bf_s}{2^{L+1}} \right\}$$

where, the frequency factor, b=1,2,3,4,5,.....2^L, f_s is the sampling frequency of the EEG signal. In this study f_s=250 Hz, $s(t)$ is the original signal with the frequency $[0 \sim \frac{f_s}{2}]$ [223].

The WPT decomposition up to level three preserves the frequency granularity at the leaf nodes and also helps preserving most of the useful information of the original EEG. After that the wavelet packet coefficients are reconstructed. The reconstructed each wavelet packet decomposition subband with its spectrum is shown in Figure 3.9. From Figure 3.9, it can be found that the natural orders of wavelet packet tree node were discrepancies with the frequency order. From the literature [179], it is seen that any wavelet packet frequency sequences would

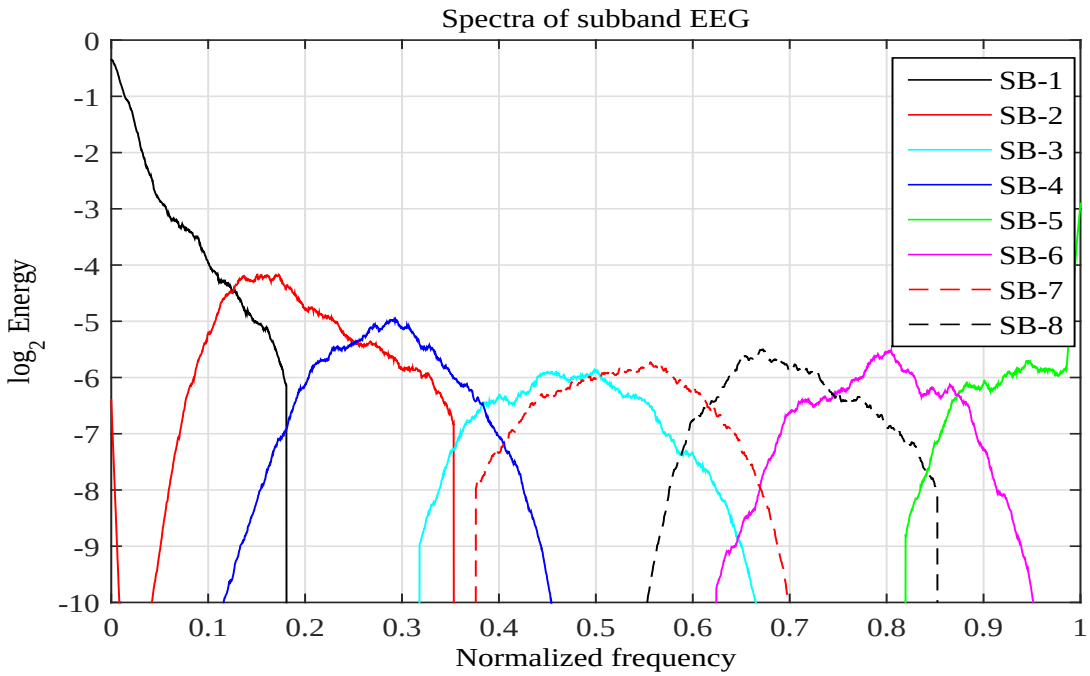

Figure 3.9: Spectrum of subbands using WPT in the form of frequency disorder.

produce dislocation, and the orders of frequency are different with the natural order. Each layer of wavelet packet decomposition, the low-frequency decomposition parts are sorted by frequency, high frequency part in descending order according to the frequency. In WPT, high-pass filter would conduct a "flip" operation due to the properties. By the nature of wavelet packet, it can prove that any wavelet packet decomposition is produced the natural order and the order of frequency discrepancy and the discrepancy of the situation is the same. Hence, it must be needed the subbands in frequency order to remove the artifact from EEG signal.

Frequency Sequences:

Whenever high-pass filtered signals (in WPT) are down sampled, the origin of the frequency axis is translated to. So, the frequency order is corrected in sequence using the following rule and then wrong frequency orders difficulty are solved.

$$f_b = (N_b - 1)\frac{f_s}{2n_b} \tag{3.6}$$

where, f_b =dominant frequency of b^{th} subbands, N_b =maximum energy bin of b^{th} subbands, n_b=length of b^{th} subbands.

Using Eq.(3.6), wavelet packet decomposition b^{th} subbands, are arranged in the frequency order according to the magnitude of the frequency. The orderly sequenced subbands are shown in Figure 3.10. From Figure 3.10, it is clearly

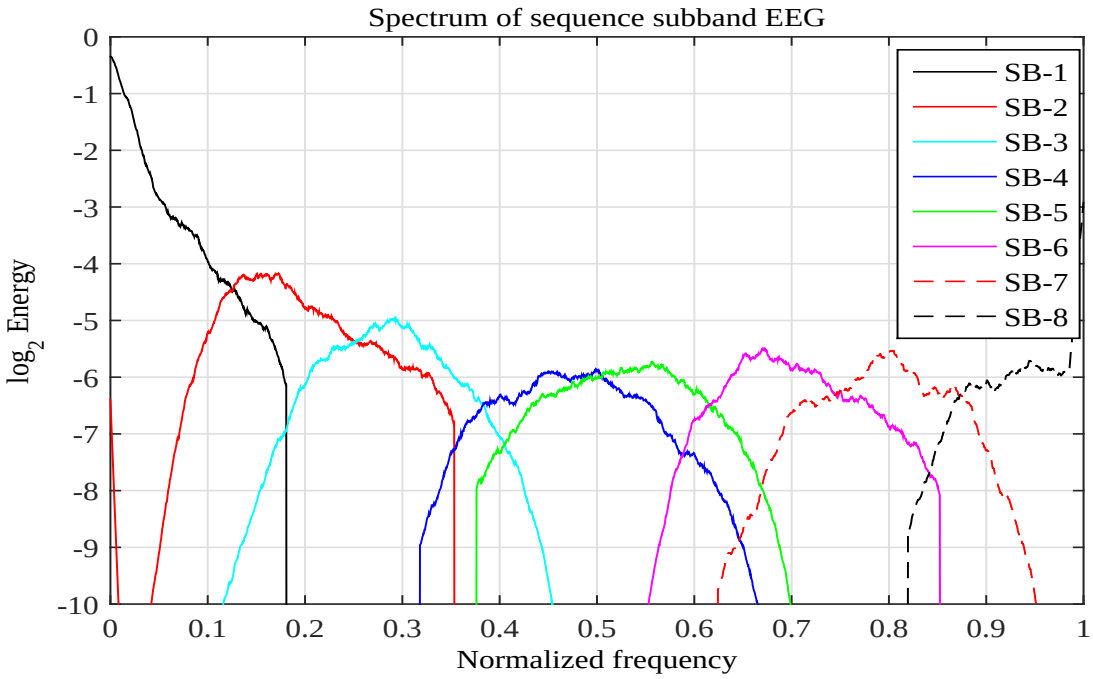

Figure 3.10: Fourier log spectrum of sequence subbands in WPT using Eq.(3.6).

visualized that the subbands of WPT act as band-pass filter. It is spectacled that the eight spectra and also showed how they complement each other to cover the signal bandwidth.

The reconstruction of contaminated EEG (top) and reconstruction error (bottom) of wavelet packet transform is illustrated in Figure 3.11. From this Figure, it is seen than the signal reconstruction error of WPT is very small as compared to DWT.

3.2.3 Empirical Mode Decomposition (EMD)

The EMD is an empirical and data-driven method developed to perform on non-stationary, non-linear, stochastic processes and therefore it is ideally suitable for EEG signal analysis and processing. However, the computational complexity of EMD is quite heavy, so it may not be suitable for online applications. Moreover, the theory behind EMD is still not complete and so far used in empirical studies. Therefore, it is difficult to predict its robustness in all EEG recordings.

EMD algorithm decomposes a signal, $s(n)$ into a sum of the band limited components/functions, $d_m(n)$ called intrinsic mode functions (IMF) with well defined

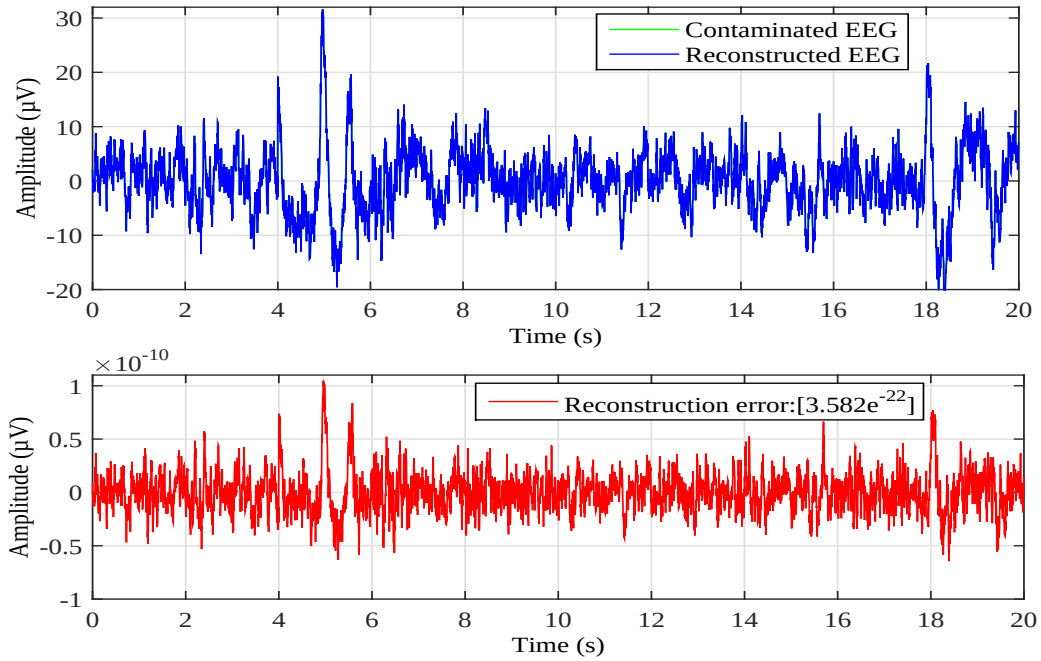

Figure 3.11: Reconstruction of contaminated EEG (top) and reconstruction error (bottom) using WPT.

instantaneous frequencies [138]. There are two basic conditions to be an IMF: (i) the number of extrema must be equal (or at most may differ by one) to the number of zero crossings (ii) any point, the mean value of the two envelopes defined by the local maxima and the local minima has to be zero [163].

To describe how EMD works, it can be said that it extracts out the highest frequency oscillation that remains in the signal. Locally, in this way, each IMF contains lower frequency oscillation than the one extracted just before this one. Being data adaptive, the basis usually offers a physically meaningful representation of the underlying processes. There is no need of considering the signal as a stack of harmonics and therefore, EMD is ideal for analyzing nonstationary and nonlinear data. Each IMF is considered as a monocomponent contribution such that the derivation of instantaneous amplitude and frequency provides a physical significance. The advantage of this time-space filtering is that the resulting band passed signals preserve the full nonstationary property in physical space. This filtering method is intuitive and direct and its basis is a posteriori and data adaptive. The completeness of the EMD of a EEG data is given in Figure 3.12. It is also observed that higher-order IMFs contain lower frequency oscillations than that of lower-order IMFs. The EMD on fGn acts as dyadic filter-bank [53]. The dyadic

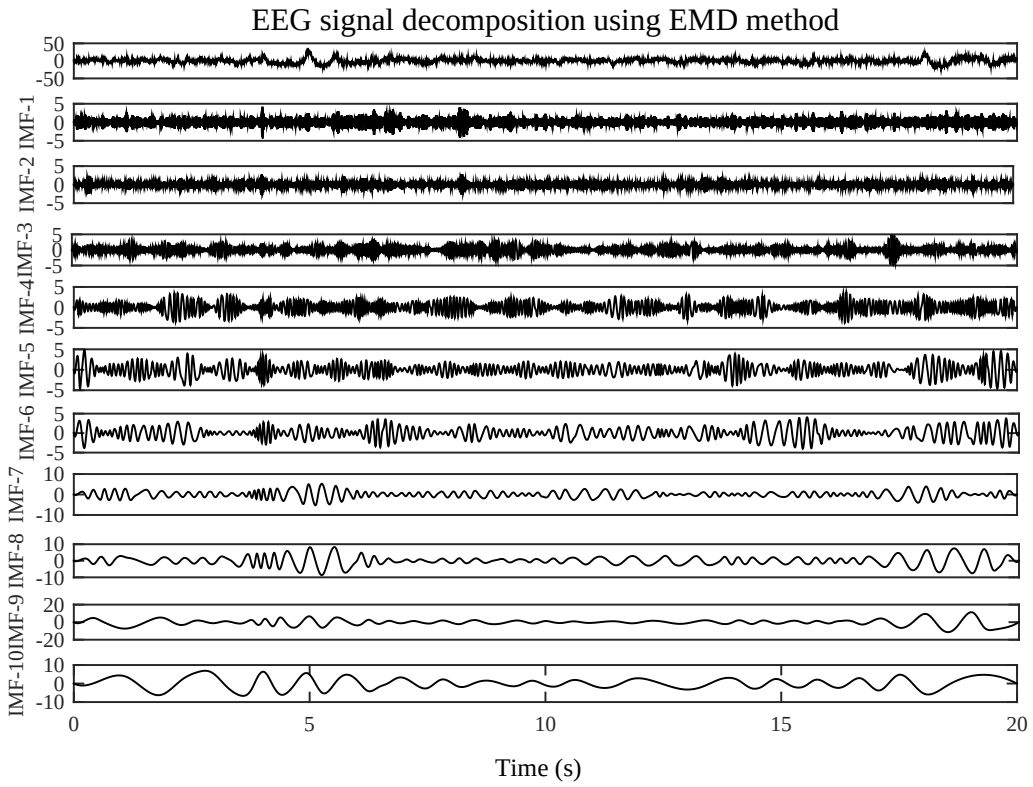

Figure 3.12: The intrinsic mode functions (subband signals) obtained by empirical mode decomposition.

property of any filter bank structure refers that the bandwidth of any subband is the half of its just previous (high frequency) subband. When the EEG signal is decomposed together with fGn using EMD, the overall decomposition acts like a dyadic filter bank. The Fourier log spectrum of the IMFs of fGn is shown in Figure 3.13. It is observed that the spectra of fGns IMF represent the dyadic characteristics. The original signal can easily be reconstructed by simply adding the bases (including the residue) with negligible error term. The original signal and reconstructed signal and error are shown in Figure 3.14. From this Figure, it is recommended that the reconstruction error of EMD is very small as compared to DWT and WPT, but it needs more time to decompose a signal as compared to other decomposed method.

3.3 Proposed Hybrid Wavelet Transform (HWT)

Recently, for artifact detection and removal, the researchers are keen to utilize the benefits of different methods by combining them into a single, i.e. hybrid method

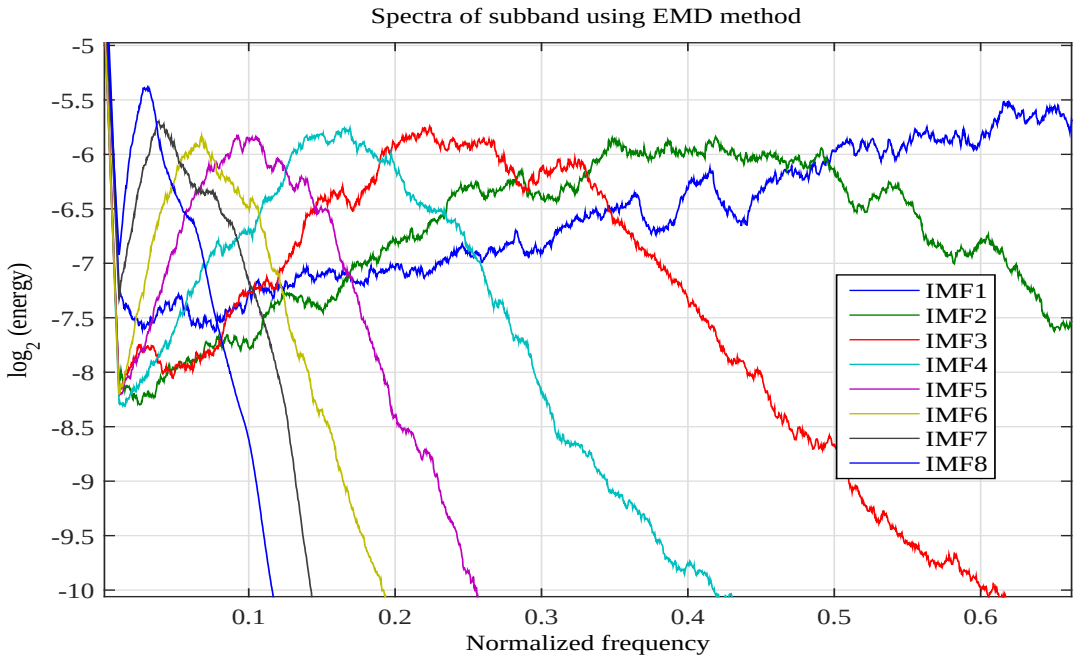

Figure 3.13: Fourier log spectrum of first 8 IMFs of fGn using EMD. The frequency bands of the IMFs is ovelapped each other.

which has two or more stages. Some of these methods are discussed in previous chapter. In the implementation, the hybrid wavelet transform (HWT) method is used which based on a novel combination of discrete wavelet decomposition and wavelet packet tree. In this method, the EEG signal is decomposed using DWT and WPT method and in this way the original signal can easily be reconstructed by simply adding all subbands. The original signal and reconstructed signal are depicted in the uppermost part and reconstruction error are shown in lowest of Figure 3.15. From the top of this Figure, it is clearly presented that the contaminated EEG signal and the recreated signal using HWT method are overlapping each other and the reconstruction error is very small and it is $4.63e^{-22}$. Thus, the value of reconstruction error is used to evaluate the similarity between the original signal and rebuilt signal.

In order to compare the spectral separation properties, it can be observed that the modes spectra obtained by the WPT decomposition are less overlapped than those obtained from DWT and EMD, showing a clearer separation of the frequency content between the bands. In summary, it can be appreciated that in the case of delta like signals for WPT, the proposed method provides here a clearer net decomposition than the others.

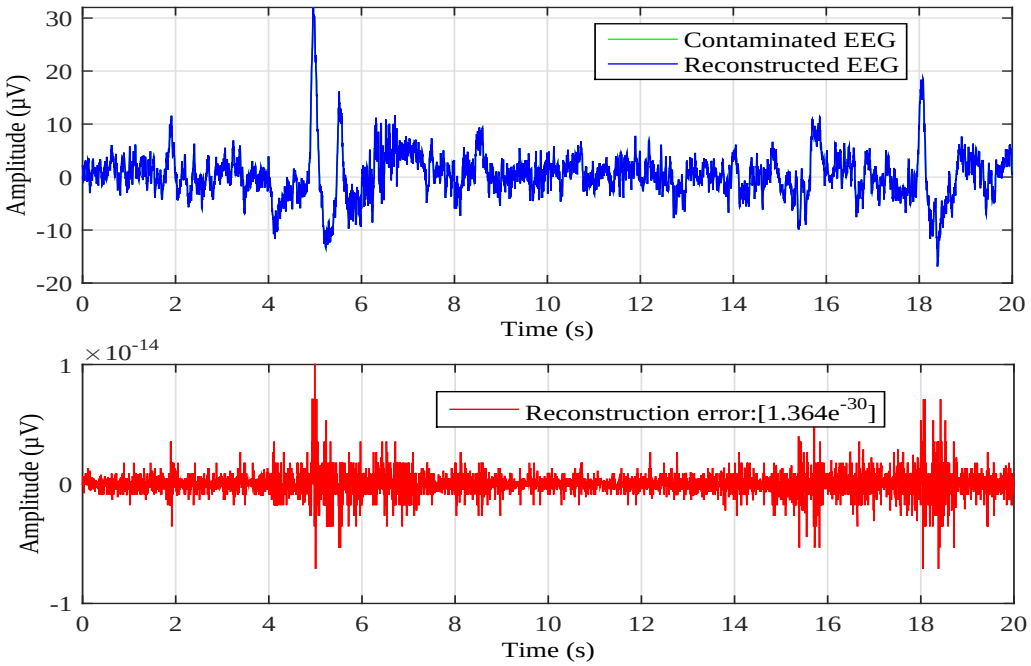

Figure 3.14: The results of EMD (different IMFs) application for contaminated EEG data. Original contaminated EEG signal and reconstructed data by simply adding the IMFs (top) and reconstruction error (bottom)

In this work, we have presented a new algorithm for analyzing and processing non-linear and non-stationary signals. The new method was successfully tested on artificial and real signals. The method here proposed has the advantages of requiring less decompose iterations that DWT and EMD does, and that the original signal can be exactly reconstructed by summing the bands. Decomposition completeness was theoretically demonstrated and numerically verified in the case of EEG signal. Because of that, a smaller level is needed resulting in a significant computational cost saving. In that sense, the novel method recovers some of the EEG properties lost by DWT and EMD, such as completeness and a fully data-driven process.

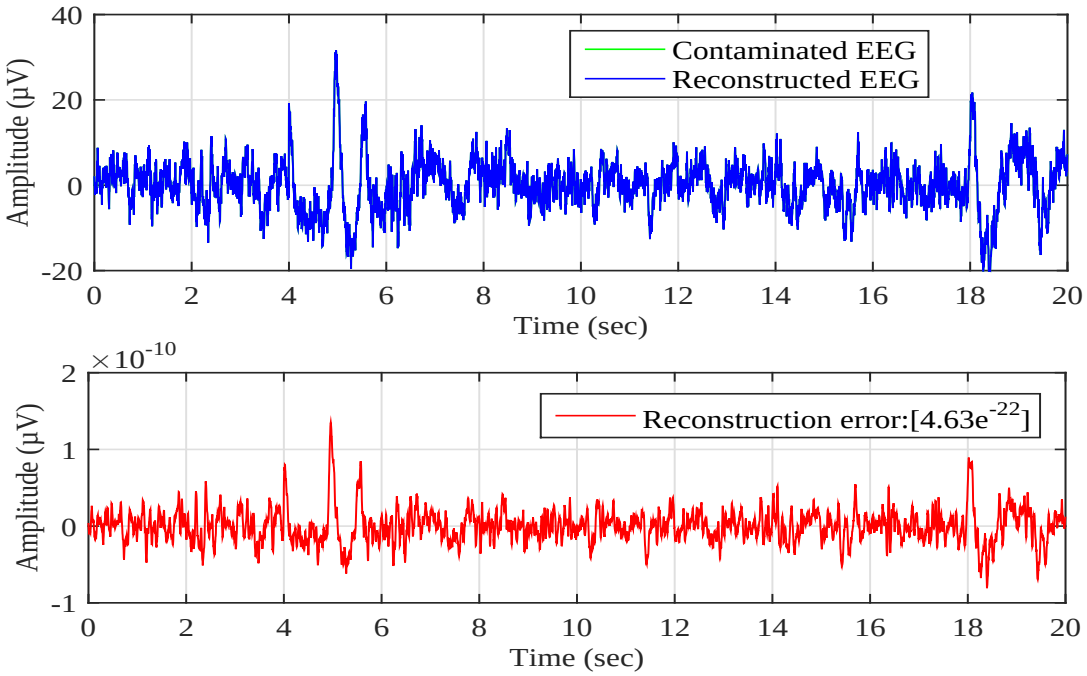

Figure 3.15: Hybrid wavelet transform technique. The original EEG and reconstructed EEG signal by simply adding subbands (top) and reconstruction error (bottom).

Chapter 4

Artifact Suppression Methods

The widely accepted research community in the brain computer interface is BCI research community. It is broadly recognized where neurological approaches are the only source of controlling any BCI system. Artifacts are unwanted signals that can impede with neurological phenomena. The artifacts could change the features of neurological phenomena or even be wrongly used as the source of control in BCI systems. The most important sources of physiological artifacts are electrooculography (EOG) and electromyography (EMG) artifacts that are careful among in BCI systems. This paper evaluates EOG and EMG artifacts related with BCI systems and the present methods for dealing with them. Based on the type of neurological approach used and the methods engaged for handling EOG and EMG artifacts, it is assumed that nearly 300 refereed journal and conference papers are reviewed and categorized. Depending on the handling EMG and EOG artifacts this study exposes the weaknesses in BCI studies associated with reporting the methods. It is not reported in the maximum BCI paper, whether or not they have measured the existence of EMG and EOG artifacts in the brain signals. Only a small percentage of BCI papers reports some automated methods for rejection or elimination of artifacts in their schemes. It is required to develop automatic methods to handle artifacts or to design BCI systems whose performance is robust to the presence of artifacts. It must be remembered that removal of artifacts involves rejecting or correcting the artifacts without distorting the signal of interest. This is primarily done in two ways: either by filtering and regression or by separating/decomposing the EEG data in to other domains. In this section, we briefly describe the artifact handling methods. Our attention throughout this chapter will be on EOG and EMG artifacts.

4.1 Terminology in Artifact Suppression Methods

Fractional Gaussian Noise (fGn): Fractional Gaussian noise (fGn) [78] is a generalization of ordinary white noise. It is a versatile model for homogeneously spreading broadband noise without any dominant frequency band, is an intrinsically discrete-time process, and may be describe as the incriment process of fractional Brownian motion (fBm) since fBm is only self-similar Gaussian process with stationary increments. Consequently, the statical properties of fGn are entirely determined by its second-order structure, which depends solely upon one single scalar parameter, H, its Hurst exponent. More precisely, $\{x_H|n|, n = ..., -1, 0, 1, ...\}$ is a fGn of index H (with $0 < H < 1$) if and only if it is a zero-mean Gaussian stationary process whose autocorrelation sequence $r_H[k] := E\{x_H[n]x_H[n+k]\}$ is

$$r_H[k] = \frac{\sigma^2}{2}\left(|k-1|^{2H} - 2|k|^{2H} + |k+1|^{2H}\right) \tag{4.1}$$

It is well known that the special case $H = \frac{1}{2}$ reduces to (discrete-time, uncorrelated) white noise, whereas other values induce non-zero correlations, either negative when $0 < H < \frac{1}{2}$ or positive when $\frac{1}{2} < H < 1$ (long-range dependence). Taking the discrete Fourier transform of Eq.(4.1), we readily obtain the power spectrum density of fGn, or

$$S_H(f) = C\sigma^2|e^{i2\pi f} - 1|^2 \sum_{k=-\infty}^{\infty} \frac{1}{|f+k|^{2H+1}} \tag{4.2}$$

with $|f| \leq \frac{1}{2}$. If $H \neq \frac{1}{2}$, we have $S_H(f) \sim C\sigma^2|f|^{1-2H}$ when $f \to 0$. It is therefore follows that fGn is a convenient model for power-law spectra at low frequencies. From its spectral properties, the particular value $H = \frac{1}{2}$, delineates two domains with contrasting behaviors. In the regime $0 < H < \frac{1}{2}$, we have $S_H(0) = 0$, and the spectrum is high-pass (sometimes referred to as an "ultraviolet" situation). On the other hand, within the range $\frac{1}{2} < H < 1$, we have $S_H(0) = \infty$, with a "$\frac{1}{f}$" "1/f"-type spectral divergence ("infrared" catastrophe). In both situations, the power-law form of the spectrum, although not exactly verified, is well approximated over most of the Nyquist frequency band. In other words, we have a quasi-linear

relation in log-log coordinates,

$$logS_H(f) \approx (1 - 2_H)log|f| + C, \tag{4.3}$$

for most frequencies $-\frac{1}{2} \leq f \leq \frac{1}{2}$.

From the paper [157], an fGn is a random field with E-dimensional random vectors defined in a d-dimensional space. It corresponds to the spin dimension E and space dimension d in the theory of critical phenomena in statistical physics. It has been shown that the fGn is invariant under a semigroup (Kadanoff block) transformation with critical exponent H, known as the Hurst coefficient in the fields of engineering and stochastic fractal processes.

In theory of probability, fractional Brownian motion, nothing like standard Brownian motion, the increments of fBm need not be independent. The fBm is a continuous-time Gaussian process BH(t) on [0, T], which begins at zero, has anticipation zero for all t in [0, T], and has the following covariance function:

$$E[B_H(t)B_H(s)] = \frac{1}{2}(|t|^{2H} + |s|^{2H} + |t - s|^{2H}) \tag{4.4}$$

where H means real number in $(0, 1)$, named the Hurst parameter or Hurst index related with the fractional Brownian motion. The unevenness of the resultant motion, with a higher value leading to a smoother motion is indicated by Hurst exponent.

The value of H determines what kind of process the fBm is:

- if $H = 1/2$ then the process is in fact a Brownian motion or Wiener process;
- if $H > 1/2$ then the increments of the process are positively correlated
- if $H < 1/2$ then the increments of the process are negatively correlated.

The increment process, $X(t) = B_H(t+1) - B_H(t)$, is known as fractional Gaussian noise.

In signal analysis, white noise is a randomly generated signal having different intensity at different frequencies, giving it a fixed power spectral density [16]. Statistical model for signals and signal sources are referred by white noise, rather than to any trarget signal. Although light that appears white generally does not

have a flat power spectral density over the visible band but white noise attractions its name from white light. The only theoretical construction noise signal is an infinite-bandwidth white noise. In practice, the bandwidth of white noise is restricted by the transmission medium, by the process of noise generation, and by finite measurements capabilities. In this way, arbitrary generated signals are measured as "white noise" if they are observed to have a flat spectrum over the range of frequencies that are relevant to the perspective signal.

Hurst Exponent (H): Hurst exponent is a measure of self similarity, predictability and the degree of long-range dependence in a time-series. It is also a measure of the smoothness of a fractal time-series based on asymptotic behavior of the rescaled range of the process. According to the Hursts generalized equation of time series, Hurst exponent H is defined as

$$H = \frac{log(R/S)}{log(T)} \tag{4.5}$$

where T is the duration of the sample of data and R/S is the corresponding value of rescaled range. R is the difference between the maximum and minimum deviation from the mean while S represents the standard deviation. Hurst exponent is estimated by plotting (R/S) versus T in loglog axes. The slope of the regression line approximates the Hurst exponent [1].

Confidence Interval (CI): A confidence interval measure the probability that a population parameter will be fall between two sets values. It can take any number of probabilities; the most common numbers may be 95% or 99%. It is the probability that a value will fall between an upper and lower bound of a probability distributions. The range of values used to estimate a population parameter of a confidence interval and is related with a specific confidence level. Constructing confidence interval around a sample means using these equations [46] that is given below:

$$CI = \bar{y} \pm z\frac{s}{\sqrt{n}} = \bar{y} \pm z\sigma_{\bar{y}} \tag{4.6}$$

where $\bar{y}$ is the sample mean, and z stands for the z-score, which is the number of standard deviations based on the confidence level. The alphabet s is the standard deviation, n is the number of samples or data length, and $\sigma_{\bar{y}}$ is the standard error of the mean.

How frequently the observed interval contains the parameter is determined by the confidence level or confidence coefficient. Whereas two-sided confidence limits form a confidence interval, their one-sided counterparts are referred to as lower or upper confidence bounds (or limits). The upper confidence limit,

$$UCL = \bar{y} + z\frac{s}{\sqrt{n}} \tag{4.7}$$

The lower confidence limit,

$$LCL = \bar{y} - z\frac{s}{\sqrt{n}} \tag{4.8}$$

The confidence coefficient indicated by 90%, 95%, 98% and 99%. In applied practice, confidence intervals are typically stated at the 95% confidence level. However, when presented graphically, confidence intervals can be shown at several confidence levels, for example 90%, 95%, and 99%. To be more confident that our interval includes $\bar{y}$, we need a wider interval: a 99% CI is wider than a 95% CI based on the same data, and a 90% CI is narrower.

Confidence intervals are frequently calculated around the estimates from statistical hypothesis tests. They may be calculated for the t test, chi-square test, analysis of variance, regression, and most other tests of inference. A 95% CI is a range of values within which 95% of the results of repeated samples from the overall population would lie; this is the most frequently reported CI level. The confidence limits are related to the P value. If one calculates the 95% CI of the difference in means between two samples, and zero is within the range of the 95% CI, then the P value will not be significant at the level less than 0.05 [154].

The major advantages of CIs [31] are that they give point and interval estimates in measurement units that should be readily comprehensible in the research situation; CIs help combine evidence over experiments. They support meta-analysis and meta-analytic thinking focused on estimation, and CIs give information about precision, and this may be more useful than a calculation of statistical power.

Sub-band Energy: The subband energy of a signal $x(n)$ is

$$E_x = \sum_{n=-\infty}^{\infty} |x(n)|^2 \tag{4.9}$$

A discrete-time energy signal is defined as one for which $0 < E < \infty$. It can reduce the dimension of feature space, but not provide the information in time-domain.

For EEG signal processing, the sub-band log energy is used in this paper and the sub-band log energy is calculated using following equation:

$$Log_2Energy = log_2E_x \tag{4.10}$$

where, E_x= subband energy.

Thresholding: In signal processing, thresholding is another technique used for signal denoising. The concept is called thresholding, when an empirical wavelet coefficient is greater than the noise level a natural estimator for a wavelet coefficient itself. Thresholding smooths time series data and filter the noise. To eliminate or suppress small value wavelet coefficients which mainly represent the noise content, the thresholding procedure is used. This process bears two basic rules: namely the keep-or-kill hard thresholding and shrink-or-kill soft thresholding familiarized by Donoho and Johnstone [45, 44]. Coefficients with absolute values lower than the threshold is set to zero are used for hard thresholding, while soft thresholding in addition shrinks the remaining nonzero coefficients toward zero. The magnitude of the threshold is changed by hard and soft thresholding. Recently, adaptive thresholding may be with mean weighted average and with gaussian weighted average. And it is proved that adaptive thresholding is better as compared to fixed thresholding.

Threshold Estimation: The next step is to calculate a threshold value to detect the artifacts in the wavelet domain. The choice of threshold value will decide both the amount of artifact suppression and the amount of distortion to the neural signal at the same time. Although the presented thresholding scheme is simple and effective, selection of a rational threshold value is a crucial task as it directly affects the denoising results. For example, choosing a very large threshold will shrink almost all the coefficients to zero and may result in over smoothing of the measured signals. On the other hand, a small value of threshold will retain the sharp edges and details but may fail to suppress the noise artifacts. Among the existing hard thresholding methods, the most popular one is the universal threshold (UT) [111, 7]. The universal threshold is defined as:

$$K = \sigma\sqrt{2LogN} \tag{4.11}$$

where, N is the signal length and σ denotes the standard deviation of the noise.

The later is estimated from the median of the detail coefficients at the first level of signal decomposition:

$$\sigma = \frac{|median(C_a)|}{0.6745} \tag{4.12}$$

where, C_a represent the wavelet coefficients a^{th} level of decomposition, k denotes the threshold, $|.|$ denotes the absolute value of elements in C_a and 0.6745 is constant value for Gaussian noise. The universal threshold may be unwarrantedly large since its dependence on the number of samples. It will yield an overly smoothed estimate and in some cases, the pseudo Gibbs phenomena may appear. In paper [97], statistical threshold (ST) is implemented which based on the statistics of the signal and mathematically, the ST is formulated as:

$$T = 1.5 * std(H_k) \tag{4.13}$$

where, T is the estimation of neuronal wide band signal magnitude using ST, $std(H_k)$ employs standard deviation of wavelet coefficients at k^{th} level. However, this particular threshold is fixed for each C_a and not optimal most of the time. By extensive testings, it has been found that this threshold is not suitable for our application as it may produce serious distortion to EEG signals. Particularly for wavelet coefficients those contain components from spikes or when the neural recording has severe large artifacts such that the data distribution violates from the typical Gaussian fitting. Hence it is proposed to the next chapter a statistical threshold method for artifact suppression in this paper.

4.2 Existing Artifact Suppression Methods

In this section, we give a comprehensive overview of techniques that can be used for the removal of artifacts from an EEG. For each method, we cite publications on its early use in EEG processing; we also explain the reasoning behind its use for the removal of artifacts from an EEG and highlight some of its advantages and deficiencies. Additionally, we mention extensions of the algorithms, if they exist.

Simple low pass, band pass or high pass filtering represents one of the first classical attempts at removing artifacts from a measured EEG. However, this is only effective when the frequency bands of the signal and interference do not

overlap [193]. With spectral overlap, which is commonplace for typical artifacts recorded along with the EEG, alternative techniques are needed such as adaptive filtering, Wiener filtering and Bayes filtering [193], as well as regression [63], EOG correction [29], blind source separation [90], and more modern attempts like the wavelet transform (WT) method [133], empirical mode decomposition (EMD) [76] and nonlinear mode decomposition (NMD) [81].

The BSS techniques are also known as component based methods since they find principal or independent components equivalent to the input EEG channels and perform processing in the transformed domain. The process is reversed by applying the inverse transformation to the corrected components, with the consequence that all EEG channels are processed and estimated simultaneously. In contrast, filtering, regression, EOG correction, the WT method, EMD and NMD estimate each artifact-corrected channel independently, in the time, frequency or time-frequency domain.

4.2.1 Linear Regression

Regression algorithms are arguably the most frequently used EEG artifact correction techniques up to the mid 90s, especially for ocular interferences, thanks to their simplicity and reduced computational demands. When one or more reference channels are available and with the premise that they properly represent all interference waveforms, then artifacts may be corrected by subtracting a regressed portion of each reference channel from the contaminated EEG. Regression may be done either in the time or frequency domains by estimating the influence of the reference waveforms on the signal of interest.

Linear regression assumes that each EEG channel is the sum of the non-noisy source signal and a fraction of the source artifact that is available through a reference channel. Then, the goal of regression is to estimate the optimal value for the factor that represents such propagation fraction. In multiple linear regression the measured signal at each electrode is influenced by more than one reference waveforms, for example vertical, horizontal and radial ocular artifacts.

Regression methods have been replaced by more sophisticated algorithms primarily because the former need one or more reference channels. The disadvantage

of the method is that limits their applicability to removing mainly EOG [29] and ECG [215] artifacts. Since other potentially more efficient algorithms emerged, like PCA and ICA that have become commonplace in most recent publications ([166, 213]). Nevertheless, and despite its drawbacks, regression is still used as the gold-standard technique to which the performance of other algorithms may be compared, see for instance [206] for ocular contamination.

4.2.2 Independent Component Analysis (ICA)

The ICA was introduced to EEG study simultaneously by different groups in [93] and has mostly replaced other approaches for the removal of artifacts from an EEG. In contrast to the possibly incorrect assumptions of PCA [82], it is the case that artifacts and brain activity are usually sufficiently independent, which explains the success of ICA for artifact removal [82]. In practice, nowadays only a few ICA algorithms such as SOBI [25], (extended) InfoMax [120] and fastICA [80] are used to process biomedical signals. Recently, the adaptive mixture of independent component analyzers (AMICA) [79] has emerged as an interesting alternative to the former methods [151]. Many more methods exist, but we do not deal with these algorithms in detail in this paper since the interested reader can find in depth explanations elsewhere (see for instance [54] and references therein).

The effectiveness of ICA is based on statistical independence of the sources and the full column rank of the mixing matrix. Even if the sources are not exactly independent, ICA based algorithms have been reported to be successful at removing artifacts from the EEG signal of interest. Due to ICA being based on statistical features, results will not be reliable if the amount of data given to the algorithm is insufficient [93]. It would be best to use all the available data, provided the artifacts and cerebral activity were spatially stationary through time; however this may not be the case. The goal then becomes using the maximum amount of data when the sources are reasonably stationary.

Constrained ICA (cICA)

- Temporally constrained ICA: The goal in this case is to obtain an output that is statistically independent of other sources but closest to a reference signal [90]. This is done by solving a constrained optimization problem

through an augmented Lagrangian function. Moreover, prior knowledge can be introduced into the model by means of reference channels that make the original matrix become augmented by a number of rows equal to the references [90, 166]. Other attempts at solving the temporally constrained ICA problem are [118, 89] where prior information of some source waveforms is used.

- Spatially constrained ICA: The idea here is to define a set of spatial constraints on the mixing matrix to represent prior knowledge or assumptions of the spatial topography of some source sensor projections [90]. The method described in [74] has been reported to have excellent performance for denoising certain types of EEG signals [2]. Precisely, the algorithm incorporates reference or constraint topographies, such that some source sensor projections are approximately known, hence limiting the degree to which certain columns of the mixing matrix may deviate from the known projections. The idea is not new, in fact in [82] the authors use spatially constrained ICA (SCICA) [83], but this latter method involves a more computationally expensive optimization.

We end by noting that using any form of constraints to solve the BSS problem involves the relaxation of strict assumptions regarding the statistical, temporal or spectral properties of the associated waveforms, in order to satisfy the temporal or spatial constraints [74].

Canonical Correlation Analysis

Canonical Correlation Analysis or CCA [37] measures the linear relation between two multi-dimensional random variables and can be applied to solve the BSS problem by taking the source vector as the first multi-dimensional random variable and a temporally delayed version of the source vector as the second multi-dimensional random variable [168, 56].

4.2.3 Filtering Methods

Simple filtering is normally not an option to remove artifacts from EEG recordings, except for narrow band artifacts like environmental line noise (50/60Hz in-

terference can be removed with a notch filter). Thus, numerous artifact removal techniques described in this section try to adapt the filter parameters to minimize the mean square error between the estimated EEG and the desired original signal. To overcome the limitation of the artifact-free signal being unknown, each method implements strategies following certain optimization criteria. In what follows be briefly describe some of the main filtering techniques employed in the removal of artifacts from the EEG.

Adaptive Filtering An adaptive filter is a system with a linear filter that has a transfer function controlled by variable parameters and a means to adjust those parameters according to an optimization algorithm [216]. The filter weights can adapt based on the feedback from output of the system and it requires a reference input to compare the desired output with the observed output. Adaptive filters iteratively adjust a vector of weights according to an algorithm of optimization. These weights model the contamination of the artifact on the EEG activity [165, 207]. Adaptive filtering represents an improvement over linear regression since propagation factors do not need to be constant or frequency independent [185, 165]. The most prevalent family of algorithms is based on least mean-squares, which is linear in complexity and convergence. Another well known family is based on recursive least-squares (RLS), which is quadratic in complexity and convergence. According to [128] RLS are superior in accuracy at removing ECG artifacts from EMG recordings. In [98], a method for removing the EOG artifacts contained in EEG signal based on adaptive filtering is used. The method uses separately recorded noisy EEG and clean EEG as two reference inputs. The noisy EEG signals with three types of EOG artifacts-horizontal eye movement, vertical eye movement and eye blinks have been recorded for five subjects. The adaptive filter, based on a least mean square (LMS) algorithm, adapts its coefficients to produce an output which matches the reference input.

Wiener Filtering Wiener filtering is another type of parametric technique, based on a statistical approach, which produces a linear time-invariant filter that minimizes the mean square error between the desired signal and its estimate [193]. The minimization is done using an estimation of the power spectral densities of the signal and artifact, hence it does not need a reference waveform. The disad-

vantages are that calibration is needed prior to usage and that it cannot run in real time. On the other hand, when properly calibrated, it can achieve a better SNR for corrected data compared to the adaptive filter [193].

Bayes Filtering

- Kalman Filtering: Kalman filter is another kind of Bayes filter which also doesn't require an external reference signal and is capable of operating in real-time. This method requires the filter models to be created prior to implementation of the algorithm and the model has to be linear. Besides, it also assumes that a-priori estimation is Gaussian and can work only with unimodal distribution. Now, the artifacts are of different waveform shapes and it is likely to be of multi-modal distribution. The a-priori estimation may also not be valid due to non-stationary, nonlinear properties of EEG signals. So, Kalman filter is not supposed to work robustly for such applications.

- Particle Filtering: Particle filter is a kind of filter based on Bayesian approach which overcomes the limitation of Kalman filter as it does not require the model to be linear or the distribution to be unimodal. But it still needs a-priori user input which may not be available always in EEG-based applications. And there has very little work been done so far to use particle filter to remove artifacts in EEG signals. Hence it is not guaranteed to be a successful choice, but one can definitely try to observe the outcome of such filter implementation in removing artifacts.

Filtering in Practice Using different forms of filtering for artifact removal from the EEG dates at least as far back as 1976, when Wright [220] obtained the best linear filter to remove EMG from EEG recordings: a least-squares Kalman filter that exploits a priori knowledge. Even if other types of filters exist and have been used in the EEG artifact removal literature, adaptive filtering is most common. It is still in use, for instance in [27] where three least-mean squares adaptive filters are employed in cascade to eliminate line interference, ECG artifacts and EOG spikes respectively. More generally, they are often used for comparison with other artifact removal methods, for instance in [100, 165]. Filtering approaches such

as adaptive, Wiener or Bayes filtering have the advantage that they can be automatized [191], however they need a measured or reliably estimated reference to operate. Some of these methods can operate on single channels, characteristic that makes them attractive for the personal health care environment [191].

Spatial Filtering The principal component analysis (PCA) is a type of spatial filter that transforms the time domain datasets into a different space by rotating axes in an N-dimensional space (where N is the number of variables or EEG channels) such that each dimension in the new space has minimum variance and the axes are orthogonal to each other [23]. The PCA uses an orthogonal transformation to convert the observations of possibly correlated variables into values of linearly uncorrelated variables called principal components, less than or equal in number to the original variables. The transformation is defined for the principal components to have the largest possible variances while being orthogonal to each other. The PCA is also known as the discrete Karhunen-Love transform.

The PCA was introduced into EEG analysis in [8], where the authors use it to empirically determine the spatial distribution of eye activity, and has since been used extensively for artifact removal [51, 107, 82]. Berg and Scherg [8] report PCA as being more effective at removing ocular artifacts than other non-BSS methodologies including regression, and for source localization [51, 8]. The greatest problem with PCA is that the assumption of orthogonality between neural activity and typical physiological artifacts does not generally hold. In fact, it has been demonstrated that PCA is unable to separate some artifactual components from brain signals, especially when they have similar amplitudes [51, 107].

4.2.4 Wavelet Transform

Using wavelet transform the EEG signal denoising is done according to following three steps: decompose the signal into a number of levels, second threshold the detail coefficients and third reconstruct the signal from the filtered representation. Artifact removal based on the WT relies on the sources of interest being decomposable on a wavelet basis, whereas artifacts cannot (depending on the type of signal, it may be the artifacts that have a better defined wavelet decomposition, for instance consider background EEG and blinks). This implies that only a few

wavelet coefficients with high absolute value should represent the signal and that wavelets coefficients with low absolute value correspond to the artifacts. Evidently, good separation of signal and noise depends on the wavelet basis and its similarity to the source signals to be preserved. Thus, the mother wavelet, the shrinkage rule and the noise level rescaling are important to the design of the noise removal method [168].

The DWT is often accompanied by threshold selecting criteria, such as Stein's unbiased risk estimate (SURE) [44] implemented in [168] or other forms of thresholding [30], such that only large enough coefficients are kept. Even though the DWT remains an interesting tool for EEG processing on its own [47, 57, 87, 104], nowadays it is more often found combined with other denoising techniques such as ICA [2], one reason for this being that the DWT is in fact unable to remove completely artifacts that overlap in the spectral domain like ECG on an EMG signal [193].

Wavelet transform has been chosen to assist labeling artifacts. The reason is that it is suitable for non-stationary signal analysis (e.g. neural signals) and is a powerful tool to detect abrupt changes or localized events mostly due to artifacts [123]. Among different wavelet transforms, the DWT is the simplest one in terms of computational complexity. However, the problem of DWT is that it is not translation invariant. Therefore small shifts in a signal can cause large changes in the wavelet coefficients and large variations in the distribution of energy in the different wavelet scales [138, 24].

Hence, denoising with DWT often introduces artifacts in the signal near discontinuities during signal reconstruction. One solution is to use stationary wavelet transform (SWT) which is translation invariant, as there is no down sampling of data involved in the algorithm [138, 24]. The SWT overcomes this translation-invariance drawback of DWT, but has redundant information and is relatively slow [37]. The design difference between DWT and SWT is the filter at each stage [38]. The approximate and detail sequences at each level of decomposition are of the same length as the original sequence. After obtaining the coefficients at j^{th} level, the algorithm up samples the filter coefficients by a factor of $2j - 1$.

4.2.5 Stationary Wavelet Transform (SWT)

The SWT decomposition process [122] is described by Eqs. (4.14) and (4.15). A signal, f, is projected onto a dyadically-spaced set of scales (spaced using a base of 2, i.e., scale $= 2^j$), or levels (level $j = log_2(scale 2^j)$), using a set of level dependent quadrature mirror decomposition filters, h_j and g_j, that have respective band-pass and low-pass properties specific to each wavelet basis [124]. The broad scale, or approximation, coefficients, A_j, are convolved separately with g_j and h_j. This process splits the A_j frequency information roughly in half, partitioning it into a set of fine scale, or detail coefficients, D_{j+1}, and a coarser set of approximation coefficients, A_{j+1}. During the next level of processing, a zero is placed in between each consecutive value found in the g_j and h_j filters (i.e., up-sampling by two) to achieve the g_{j+1} and h_{j+1} filters. This procedure can be iteratively continued until the desired level of decomposition, $j = J$, is obtained. Note that the algorithm is initiated by setting $a_0 = f$.

$$A_{j+1}(k) = \sum_n h_j(n-k)A_j(k) \tag{4.14}$$

$$D_{j+1}(k) = \sum_n g_j(n-k)A_j(k) \tag{4.15}$$

$$A_j(k) \sum_n h_j(k-n)A_{j+1}(n) + \sum_n g_j(k-n)D_{j+1}(n) \tag{4.16}$$

The A_j coefficients can be reconstructed from A_{j+1} and A_{j+1} by convolving each with the respective reconstruction filter, $h_j(n)$ or $g_j(n)$, and summing (Eq.(4.16)). Note that each reconstruction filter is also level dependent and includes $2^j - 1$ zeros between each filter coefficients. This process can be iteratively continued until the original signal, f, is recovered.

Thus, two types of coefficients are generated: approximate and detail coefficients that contain low and high frequency information respectively. The generated wavelet coefficients at different levels denote the correlation coefficients between artifactual signal and the wavelet function. The artifactual events will have larger coefficient values if they have higher correlation with the wavelet function while smaller coefficients will be generated corresponding to the actual neural activities.

4.2.6 Machine Learning

Few existing methods adopted the idea of machine learning (mostly supervised learning) for artifact separation from useful EEG signal by training a classifier with (supervised) or without (unsupervised) labeled training datasets. Using a machine learning algorithm when artifactual epochs are identified then the identified epochs are either emphasized as artifact annotator to the clinicians for helping making a decision or can be rejected before examination from clinician or before sending to automated signal processing system. Machine learning techniques are

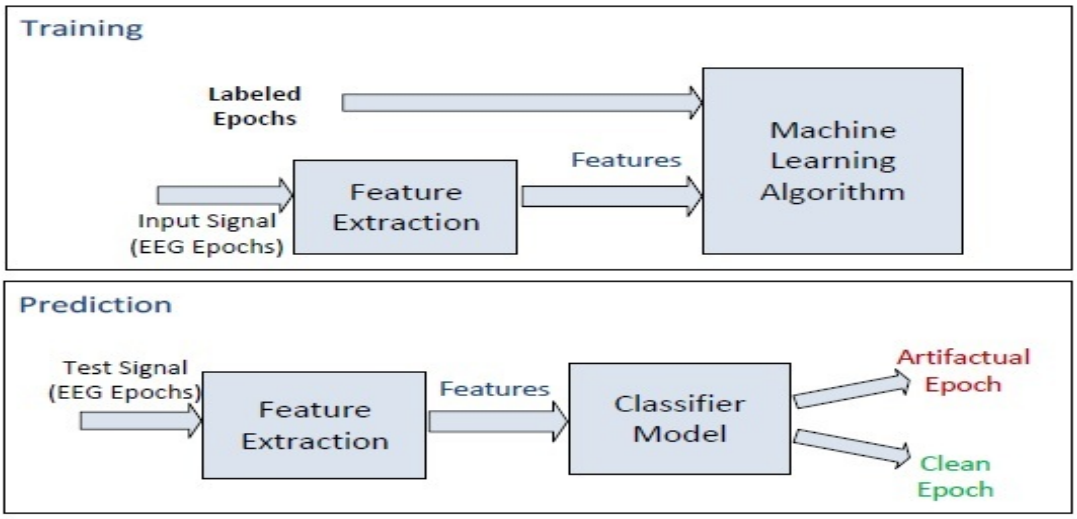

Figure 4.1: Machine learning classification for identifying artifactual epoch from clean EEG epoch [85].

mainly two types: supervised and unsupervised learning. Among supervised learning algorithms, two most popular methods used for classification between artifact and brain signals are ANN and SVM. Among unsupervised learning, k-means clustering and outliers detection are most common in this particular area of research. A basic approach to classify artifact from EEG by using the machine learning classifier is shown in Figure 4.1.

4.2.7 Nonlinear Mode Decomposition

A novel adaptive decomposition tool for time-domain signal analysis based on the synchrosqueezed wavelet transform (SqWT) [36] is nonlinear mode decomposition [81]. The fully oscillatory components and their harmonics are found using the

method that decomposes a signal into its so-called nonlinear modes. In [81], the writers state that its (NMD) qualitative and quantitative supremacy over the EMD and EEMD methods and prove that it has strong mathematical background (EMD and EEMD are empirical) and since it is more robust to noise. Also the authors further demonstrate the application of their recommended algorithm to artifact removal from a human EEG recording.

In the paper [81], the NMD procedure contains of four parts: adaptive curve extraction (first harmonic) from the SqWT, identification of possible harmonics, reliable identification of the true harmonics and reconstruction of the nonlinear modes from the SqWT. It is observe that the time-frequency representation with much better frequency resolution and the same time resolution occurred in this method than that of the WT [81, 36] due to the construction of the SqWT. Adaptive curve extraction is equivalent to finding the region of the SqWT that contains the required component (of the form $A(t)cos\phi(t)$). By inspecting regions of peaks at harmonic frequencies and extracted candidate harmonics to find the corresponding supports. In summary, the SqWT is reconstructed by summing all nonlinear modes of the main curve and of its entire true harmonics (values not belonging to the supports of the extracted curves are set to zero).

For EEG artifact suppression, it is the only reference paper on the applicability of NMD and, in general, to any form of denoising, is in the original work [81]. It appears [47, 57, 104], that a potentially motivating new tool for improving the results of artifact suppression is achieved by using the simpler WT. A post processing technique applied to the continuous wavelet transform in order to generate localized time-frequency representation of non-stationary signals is the synchrosqueezing transform (SST). The continuous wavelet transform is a projection based algorithmth at identifies oscillatory components of interest through a series of time-frequency filters knownas wavelets. The SST was originally introduced in the context of audio signal analysis and is shown to be an alternative to EMD. The SST aims to decompose a signal into constituent components with time-varying harmonic behavior. These signals are assumed to be the addition of

individual time-varying harmonic components yielding

$$s(t) = \sum_{k=1}^{K} A_k(t) cos(\theta_k(t)) + \eta(t) \tag{4.17}$$

where, $A_k(t)$ is the instantaneous amplitude, $\eta(t)$ represents additive noise, k stands for the maximum number of components in one signal, and $\theta_k(t)$ is the instantaneous phase of the k^{th} component. The instantaneous frequency of the k^{th} component is estimated from the instantaneous phase as

$$f_k(t) = \frac{1}{2\pi} \frac{d}{dt} \theta_k(t) \tag{4.18}$$

In seismic signals, the number k of harmonics or components in the signal is infinite. They can appear at different time slots, with different amplitudes, instantaneous frequencies, and they may be separated by their spectral bandwidths $\Delta f_k(t)$.

The spectral bandwidth defines the spreading around the central frequency, which in our case is the instantaneous frequency, for a completed disentangling of concepts. This magnitude is a constraint for traditional time frequency representation methods. The STFT and the CWT tend to smear the energy of the superimposed instantaneous frequencies around their center frequencies [35]. The smearing equals the standard deviation around the central frequency, which is the spectral bandwidth. The SST is able to decompose signals into constituent components with time-varying oscillatory characteristics. Thus, by using SST we can recover the amplitude $A_k(t)$ and the instantaneous frequency $f_k(t)$ for each component.

From CWT to SST: The CWT of a signal is

$$w_s(a,b) = \frac{1}{\sqrt{a}} \int s(t) \psi^*(\frac{t-b}{a}) dt \tag{4.19}$$

where, ψ^* is the complex conjugate of the mother wavelet and b is the time shift applied to the mother wavelet, which is also scaled by a. The CWT is the cross correlation of the signal $s(t)$ with several wavelets that are scaled and translated versions of the original mother wavelet. The symbols $w_s(a,b)$ are the coefficients representing a concentrated time-frequency picture, which is used to extract the instantaneous frequencies. It is observe that there is a limit to reduce the smearing

effect in the time-frequency representation using the CWT. This smearing mainly occurs in the scale dimension a, for constant time offset b show that if smearing along the time axis can be neglected, then the instantaneous frequency $w_s(a,b)$ can be computed as the derivative of the WT at any point (a,b) with respect to b, for all $w_s(a,b) \neq 0$:

$$w_s(a,b) = \frac{-j}{2\pi W_s(a,b)} \frac{\partial W_s(a,b)}{\partial b} \tag{4.20}$$

The final step in the new time-frequency representation is to map the information from the time-scale plane to the time-frequency plane. Every point (b,a) is converted to $(b, w_s(a,b))$, and this operation is called synchrosqueezing. Because a and b are discrete values, we can have a scaling step $\Delta a_k = a_{k-1} - a_k$ for any a_k where $w_s(a,b)$ is computed. Likewise, when mapping from the time-scale plane to the time-frequency plan $(b,a) \rightarrow (b, w_{inst}(a,b))$ the SST $T_s(w,b))$ is determined only at the centers w_l of the frequency range $[w_l - \Delta w/2, w_l + \Delta w/2]$ with $\Delta w = w_l - w_{l-1}$:

$$T_s(w_l,b)) = \frac{1}{\Delta w} \sum_{a_k : |w(a_k,b) - w_l| \leq \Delta w/2} W_s(a_k,b) a^{-3/2} \Delta a_k \tag{4.21}$$

The above equation shows that the time-frequency representation of the signal $s(t)$ is synchrosqueezed [73] along the frequency (or scale) axis only. The synchrosqueezing transform reallocates the coefficients of the continuous wavelet transform to get a concentrated image over the time frequency plane, from which the instantaneous frequencies are then extracted; this is an ultimate goal in EEG signal analysis. The identified frequencies are used to describe their source EEG and eventually gain a better understanding of the artifact detection.

4.2.8 Stationary Subspace Analysis (SSA)

Recently, used one of the popular blind source separation algorithm is stationary subspace analysis (SSA) that factorizes a multivariate time series into two components termed as stationary signal and non-stationary signal. In EEG analysis, the electrodes on the scalp record the activity of a large number of sources located inside the brain. These sources can be stationary or non-stationary, but they are not discernible in the electrode signals, which are a mixture of these sources. The

SSA allows the separation of the stationary from the non-stationary sources in an observed time series.

According to the SSA model [229], the observed multivariate time series $s(t)$ is assumed to be generated as a linear superposition of stationary sources $s^s(t)$ non-stationary sources $s^n(t)$ as follows:

$$s(t) = Mn(t) = \begin{bmatrix} M_s & M_n \end{bmatrix} \begin{bmatrix} s^s(t) \\ s^n(t) \end{bmatrix}$$

where, M is an unknown but time-constant mixing matrix; $n(t)$ is the mixture of stationary sources $s^s(t)$ and non-stationary sources $s^n(t)$; M_s and M_n are the basis of the stationary and non-stationary subspace respectively. Given samples from the time series $s(t)$ the aim of SSA is to estimate the inverse mixing matrix M^{-1} separating the stationary from non-stationary sources in the mixture.

Using the estimated mixing matrix, $\hat{M} = \left[\hat{M}_s \hat{M}_n\right]$ given by SSA, these artifactual components $X_{art}(t)$ are projected back to EEG channel, and artifacts in EEG data $s_{art}(t)$ are calculated as:

$$s_{art}(t) = \hat{M}_n X_{art}(t) \tag{4.22}$$

Finally, this ocular activity is removed from the EEG recording to yield the clean EEG data by the following formula:

$$\hat{s}(t) = s(t) - s_{art}(t) \tag{4.23}$$

where, $\hat{s}(t) \approx S_{EEG}(t)$. Thus the pure EEG of any channel is reconstructed using Eq.(4.23). The EOG suppression results for a single channel (multichannel is used for SSA experiment) of recorded electroencephalography are illustrated in Figure 4.2, in which the separated electro-oculogram and purified EEG signals are shown in the second and third rows respectively.

4.2.9 Discrete Wavelet Transform (DWT)

In discrete wavelet transform (DWT), only the lower frequency band is decomposed, giving a right recursive binary tree structure whose right lobe represents the lower frequency band and its left lobe represents the higher frequency band. The frequency band $[\frac{f_m}{2} : f_m]$ of each detail scale of the DWT is directly related to the sampling rate of the original signal, which is given by $[f_m = \frac{f_s}{2^J}]$, where

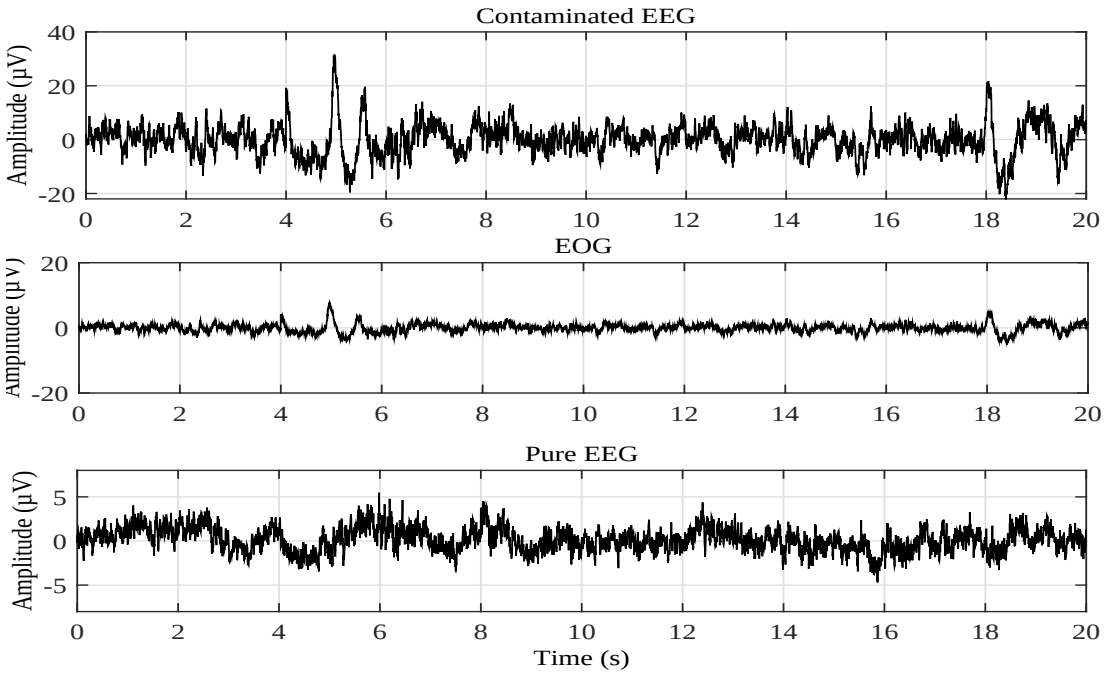

Figure 4.2: Example of the separation of clean EEG from the contaminated data using SSA. The artificially generated EEG signal (top), the segregated EOG signal (middle), and pure EEG signal (bottom).

f_s is the sampling frequency, and J is the level of decomposition. The highest frequency that the signal could contain, from Nyquist theorem, would be $\frac{f_s}{2}$. Frequency bands corresponding to seven decomposition levels with Daubechies 4 (db4) mother wavelet which is choosed for this filter. The wavelet coefficients $F_{j,k}$ succeeding to the signal $s(t)$ can be obtain as

$$F_{j,k} = \int_{-\infty}^{\infty} s(t) W_{j,k}(t) dt \tag{4.24}$$

where, j and k are the scaling and shifting factors, respectively. The mother wavelet $W_{j,k}$ is used for reconstruction as

$$S(t) = \sum_{k=-\infty}^{\infty} F_{j,k} W_{j,k}(t) dt \tag{4.25}$$

The wavelet transform (WT) is the decomposition consists of observing the signal at different resolution levels and different translations in time by bandpass filtering [224]. The strength of WT based signal decomposition lies in using short high frequency basis functions and long low frequency ones to isolate different characteristics of the signal. In such decomposition the signal is represented as finite

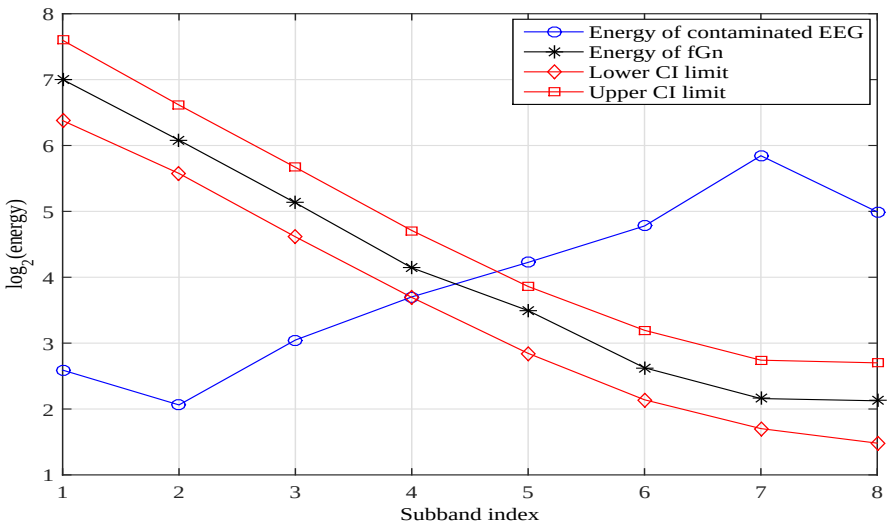

Figure 4.3: The selection of threshold from subband energy using DWT. The selection of threshold subband index of contaminated EEG channel based on the subband energy of fGn. The 5^{th} subband exceeds the upper boundary of CI and hence the 5^{th} one of EEG is selected as the highest order subband index to represent the pure EEG signal.

set band-passed signals of different frequency bands. Subband signals are reconstructed from the detail and approximate coefficients denoted as $C_1, C_2, ..., C_j$ and R_j respectively.

The analyzed signal of the channel is represented as [139]:

$$\tilde{s}(t) = \sum_{j=1}^{J} q_j(t) + q_{J+1}(t) \tag{4.26}$$

where, q_j is the j^{th} subband corresponding to the detail coefficient C_j at the j^{th} level and is the $(J+1)^{th}$ subband reconstructed from the approximate coefficient R_j of the channel.

A noise assisted DWT based approach is implemented here to reduce the low frequency noise from single channel EEG. At the end of the decomposition, the signal $s(t)$ is represented as

$$\tilde{s}(t) = \sum_{j=1}^{J+1} q_j(t) \tag{4.27}$$

where, $\tilde{s}(t) \approx s(t)$. The pure EEG signal of the single channel can be estimated by summing up the lower order subbands as:

$$\hat{s}(t) = \sum_{j=1}^{D} q_j(t) \tag{4.28}$$

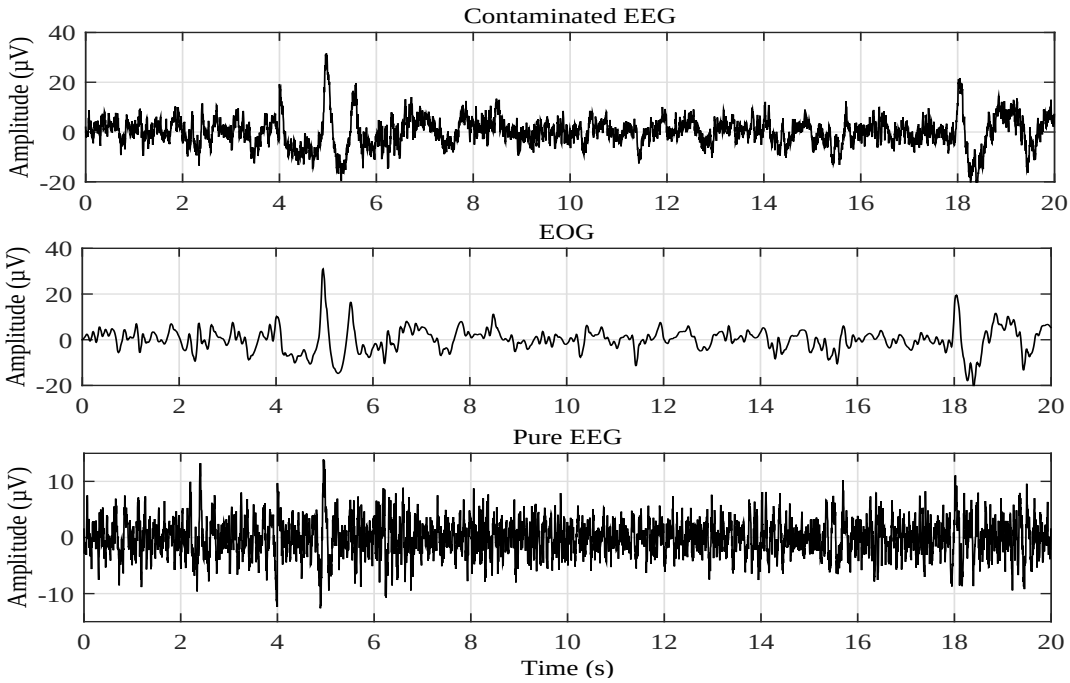

Figure 4.4: The separation of clean EEG from the contaminated data using DWT. Separation of EOG artifact from real (recorded) EEG signal using wavelet denoising method. The recorded (contaminated with EOG) EEG signal (top), the separated EOG artifact (middle) and the pure EEG signal (bottom).

where, $q_j(t)$ is the j^{th} subband of the channel. Here, the subject is to find the critical (threshold) subband with index D such that the subbands of indices $1, 2, 3, ..., D$ are responsible for relatively lower frequency pure EEG component.

The low frequency noise suppression from EEG is achieved by using subband energy based data adaptive thresholding [139]. The fGn is also used as the reference signal for adaptive threshold detection. The subband energy of fGn is decreased with increasing its center frequency when the subband decomposition is performed with dyadic filter bank. The wavelet transform is itself a dyadic decomposition and hence it produced the similar nature of subband energy distribution of fGn as obtained by using EMD. The determination of threshold subband based on the subband energy is illustrated in Figure 4.3. The pure EEG results for a single channel of contaminated EEG are shown in Figure 4.4.

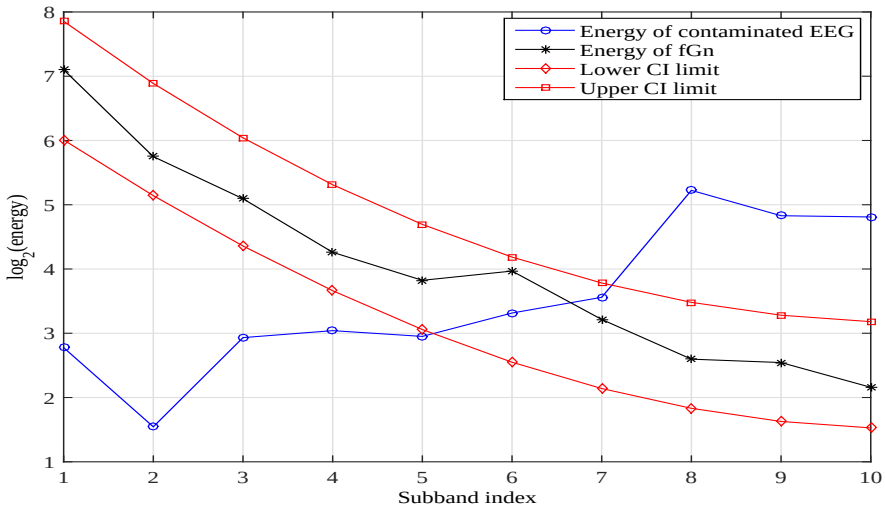

Figure 4.5: Selection of IMF of the artificially contaminated EEG signal using EMD. The selection of starting IMF (lowest order) to extract the low frequency component of mixed signal. Here, the 8^{th} IMF is selected. Its energy exceeds the upper limit of 95% confidence interval of the IMFs energies of fGn.

4.2.10 Empirical Mode Decomposition (EMD)

Based on the principle of the empirical mode decomposition (EMD) technique [163], the signal $s(t)$ is represented as

$$s(t) = \sum_{g=1}^{G} c_g(t) + r_G(t) \tag{4.29}$$

where, $c_1(t), c_2(t), ..., c_G(t)$ are all of the intrinsic mode functions included in the signals and $r_G(t)$ is a negligible residue. Here, G is the total number of intrinsic mode functions. The completeness of the decomposition is given by the Eq.(4.29).

The analyzing EEG signal $s(t)$ consists of a slowly varying trend (EOG) superimposed to a high frequency fluctuating process $h(t)$, and the trend is expected to be captured by intrinsic mode functions of large indices. A process of detrending $s(t)$ which corresponds to estimating relates to the computation of the partial, fine-to-coarse, reconstruction.

$$\hat{s}(t) = \sum_{d=1}^{D} c_d(t) \tag{4.30}$$

where, D is the larger IMF index prior contamination by the trend. Each of the IMF $c_d(t); d = 1, 2, ..., D$ represent relatively higher frequency oscillations i.e. $\hat{s}(t) \approx h(t)$. The optimized $d = D$ is chosen when the energy at index d departs

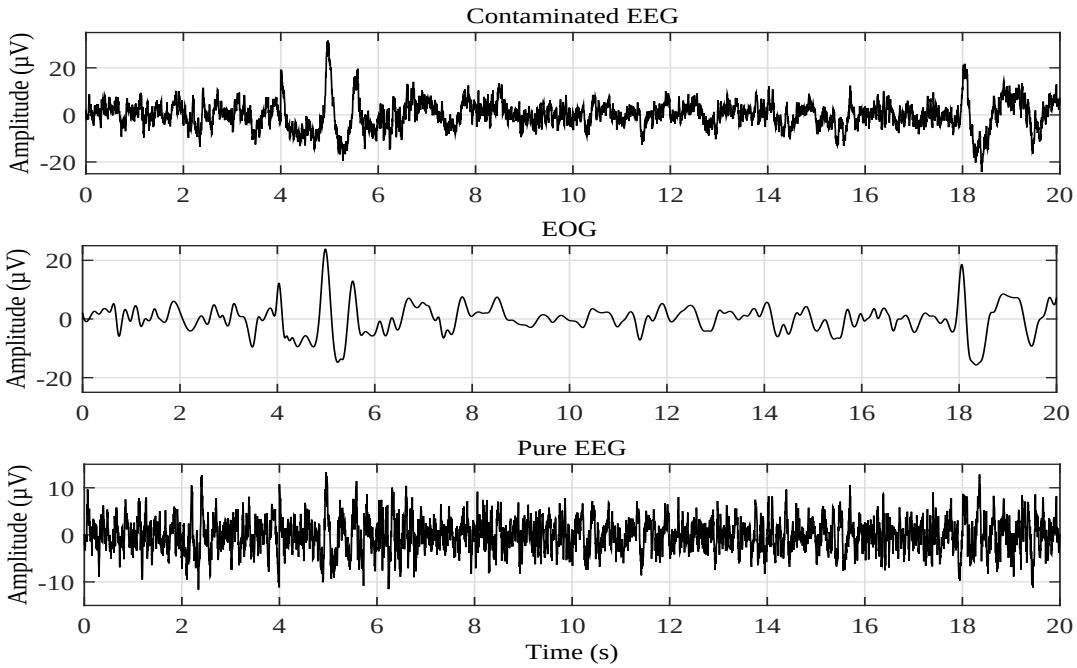

Figure 4.6: The empirical mode decomposition based data adaptive filtering technique is used to separate EOG from the contaminated EEG. The top panel shows the artificially contaminated EEG signal and the other panels illustrate EOG (middle) and purified EEG signal (bottom).

significantly from the energy of the reference signals. According to Eq.(4.30), the pure EEG is calculated.

The energy of the IMFs with different local scales is distributed in different frequency bands. The fractional Gaussian noise is used here as the reference signal. The energies of the intrinsic mode functions of fractional Gaussian noise are computed and then its upper and lower limits of 95% confidence interval (CI) are derived. The upper limit of CI of subband energy of fractional Gaussian noise (fGn) is considered as the threshold. The logarithmic energy (with base 2) of individual intrinsic mode function of EEG is compared with the threshold level. The energy of the n^{th} IMF of EEG signal exceeds the threshold and hence it is selected to remove the artifacts [163].

It is observed in Figure 4.5 that the 7^{th} IMF is touched the upper limit of confidence interval. The 7^{th} one is the starting point of lower frequency components. The electro-oculogram is separated by summing the IMFs 7 to 10 as well as the residue. By subtracting EOG from raw EEG, we get the purified EEG that reflects the clean EEG in Figure 4.6.

The EMD has been criticized due to its little robustness to noise [81] (its

performance is greatly influenced by white noise), but has gained a lot of attention in the last few years in many fields [77] and in particular for processing biomedical signals. The robustness of the original algorithm has been improved with the ensemble empirical mode decomposition (EEMD) [222].

Enhanced/Ensemble Empirical Mode Decomposition: Knowing the fact that EMD approach is very sensitive to noise which often leads to mode mixing difficulty, enhanced empirical mode decomposition (EEMD) is introduced. Therefore, EEMD is proposed which uses an average of number of ensembles (IMFs) from EMD as the optimal IMFs thus it provides a noise-assisted data analysis method [192]. The EEMD has likewise been used in the context of EEG (in this case in conjunction with CCA. To select the components of EEG signal, EMD also requires thresholding, such as the methodology proposed in [168, 103]. At the end, it is concluded that EMD has been modified to the multivariate location for example with the multivariate EMD (MEMD) [126, 205] and the turning tangent EMD (2T-EMD) [54], however these methods are rarely used for EEG due to their computational complexity.

4.3 Proposed Artifact Suppression Method

The role of signal processing is crucial in the development of a real-time brain computer interface. The principal attention of EEG pre-processing is to remove the dominant artifacts successfully before its further processing. The performance of the application EEG potentially depends on its proper cleaning. The state of the art online BCI application paradigms suffers from lower accuracies due to excessive artifacts caused by muscular (eye-movements, muscle tremors, etc.) and other electromagnetic interferences. The contemporary EEG cleaning methods exhibit several problems in the real time implementation of BCI. Although EMD based approach is more effective in artifact suppression, it takes longer processing time. Another potential method is the ICA based BSS which damages the intelligibility of the original EEG and degrades the accuracy of BCI classification. The existing artifact suppression techniques are not fully usable as they are not suitable to be implemented in real time in pre-processing of the practical BCI.

The target of this research paper is the development of the novel online methods to preprocess and clean EEG multichannel signals. The model based approach multiband decomposition (hybrid wavelet transform) can be employed to achieve the goal of this research.

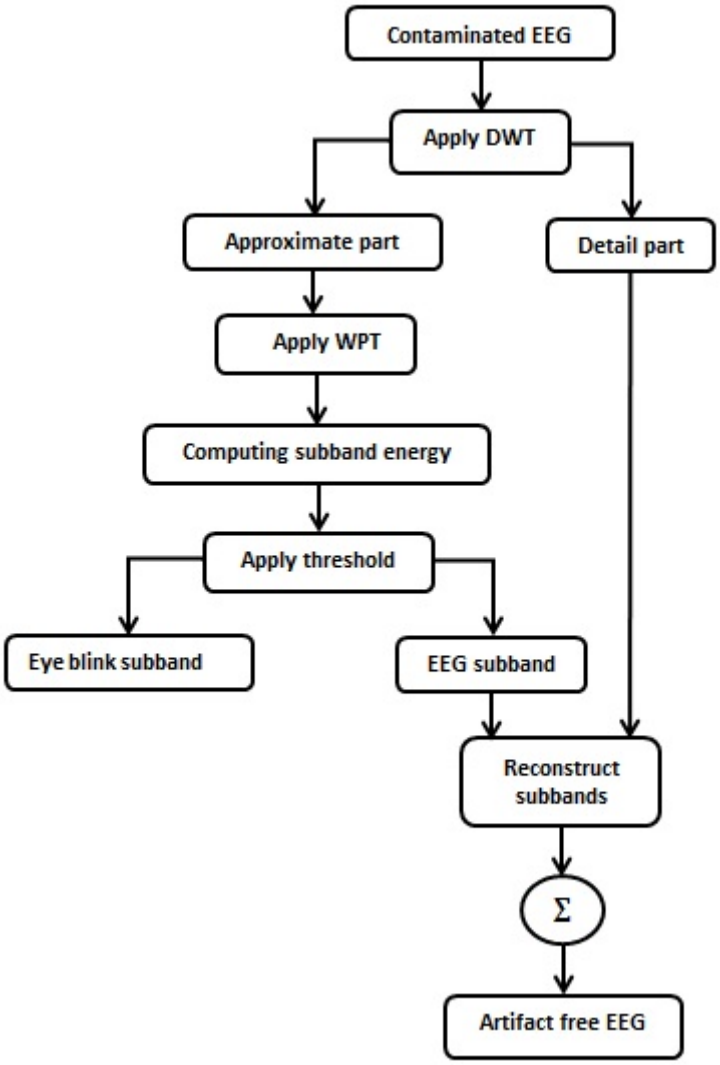

Figure 4.7: This flowchart shows the different components of the proposed system for artifact removal in hybrid wavelet transform (HWT) method. In this block diagram, discrete wavelet transform, wavelet packet transform and threshold methods are applied serially for EOG separation.

4.3.1 Hybrid Wavelet Transform (HWT)

The HWT method is based on a novel combination of discrete wavelet decomposition, and wavelet packet tree. In DWT, at every recursive decomposition level,

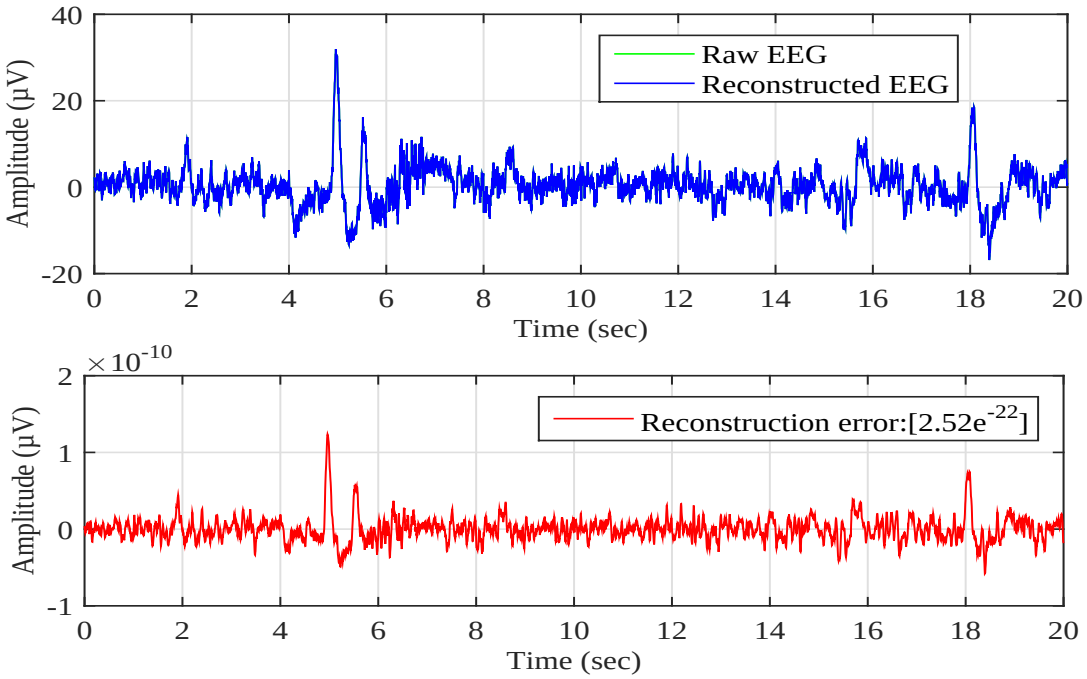

Figure 4.8: The results of hybrid wavelet transform technique for raw EEG data. This illustrated the original EEG and reconstructed EEG signal by simply adding subbands (top) and reconstruction error (bottom).

signal is down-sampled by a factor of 2 which results in number of output coefficients at every level as half of the input signal to that level. Therefore, number of samples to be convolved at every stage reduces to half compared to the input signal length. Due to the discrepancy of down-sampling at every decomposition level in DWT, WPT is also preferred. In the proposed method, as shown in Figure.4.7, firstly the DWT is applied on the contaminated EEG data. After three levels decomposition of EEG signal using DWT (using Eq.(4.27).), then WPT is applied on the approximate coefficient signal to detect ocular artifact zone to decompose it up to three levels using Daubechies 4 (db4) mother wavelet as a basis function. The WPT evenly divides a frequency band into 2^L parts where L is the number of decomposition level. With the example $L = 3$, the approximate signal is decomposed by three layers wavelet packet, the original signal is equivalent to Eq.(3.5).

4.3.2 Proposed Algorithm for Artifact Suppression

The block diagram of the proposed method is illustrated in Figure 4.7. Here the energy based subband thresholding with HWT decomposition technique is

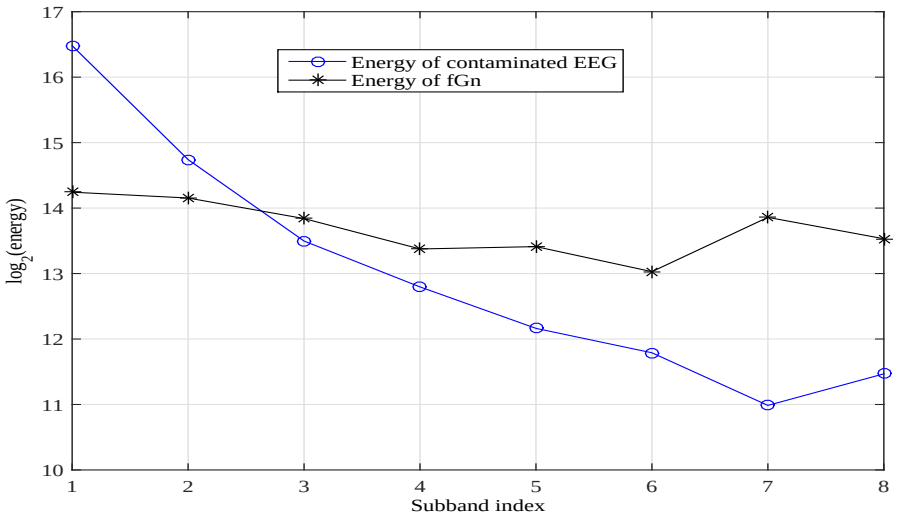

Figure 4.9: The selection of threshold from subband energy of contaminated EEG channels (by applying proposed HWT) based on the subband energy of fGn. In this observation, selection of the index of subbands of EEG signal from which the low frequency components can be extracted. The subbands energy of fGn is considered as the threshold. The logarithmic energy (with base 2) of individual subband of EEG is compared with the threshold level. The energy of the 2^{nd} subband index of EEG signal exceeds the threshold and hence it is selected to remove the artifacts.

proposed. In this implementation, fractional Gaussian noise (fGn) is used to determine the threshold level derived from the analysis data. The reference signal (fGn) magnitude has used here the following statistical threshold formula before hybrid wavelet decomposes.

$$fGn(t) = \lambda * m(t) * std(s(t)) \tag{4.31}$$

where, λ is a multiplication factor and it ranges $3 < \lambda < 4$, $m(t)$ is the reference signal for adaptive thresholding using in HWT, std stands for standard deviation and $S(t)$ is the EEG signal with the frequency $0 \sim fs/2$.

According to the proposed (HWT) method, the algorithm for separating electro-oculogram artifacts from EEG signal is given bellow:

(i) Perform DWT on the contaminated EEG signal together with the fGn up to 3 levels yielding a finite set of subbands.

(ii) Decompose the resultant approximate coefficient of DWT using WPT to up to L (=3) levels to select the artifactual coefficient. The corresponding subbands are reconstructed. The subbands of each channel are arranged using Eq.(3.6).

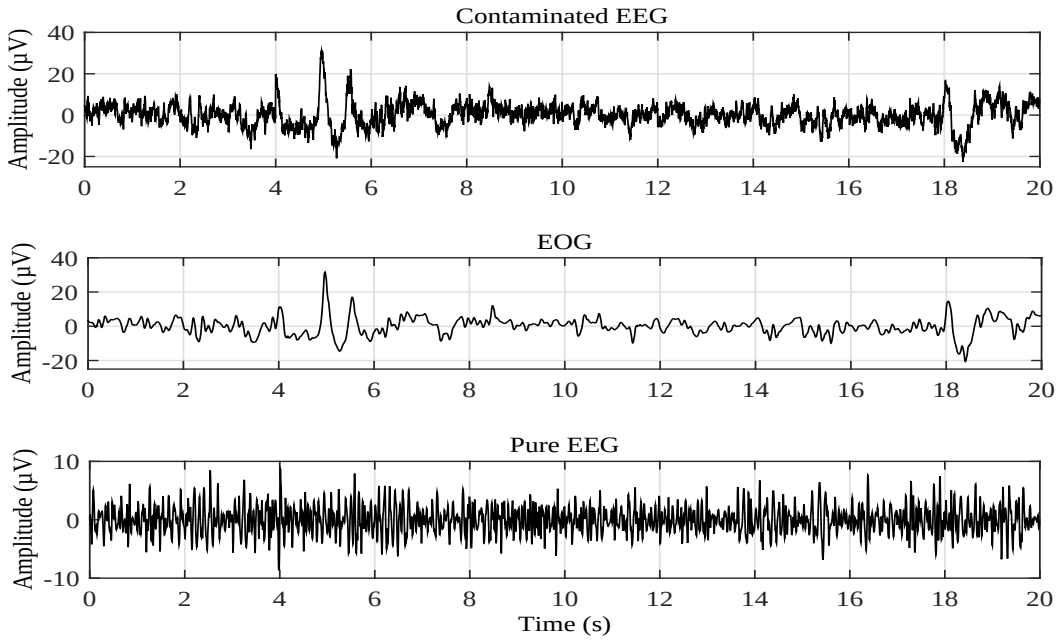

Figure 4.10: Mixture of EOG like contaminated EEG signal and pure EEG. (Top): contaminated EEG signal. (Middle): EOG signal. (Bottom): the pure EEG signal using HWT. According to the experimental figure on the contaminated EEG, the proposed HWT is removed the most EOG artifacts from the mixed data and at the same time preserved the most EEG signal as shown in Figure. As a result, the purified EEG signal is found as completely artifact free in HWT method.

(iii) Compute the log2 energies of all the subbands of raw EEG as well as fGn obtained by WPT.

(iv) Compute the log_2 energies of all the subbands of raw EEG as well as fGn obtained by WPT.

(v) Find the lowest order subband with energy exceeding the energy of fGn.

(vi) Thus selected b^{th} (in Figure 4.9, b=2) subband is the starting index of constructing electro-oculogram (EOG) signal.

(vii) The rest of the subbands are used to reconstruct the pure EEG signals.

The EOG suppression results for a single channel of recorded EEG is illustrated in Figure 4.10 in which the separated EOG and purified EEG signals are shown in the second and third rows respectively.

After this phenomenon; the EEG signal artifact is removed according to the proposed method and that is illustrated in Figure 4.7. In this Figure, subband

thresholding based HWT decomposition technique is proposed. Besides this the artifact free EEG subband and the detail coefficients of DWT are reconstructed by upsampled by a fctor 2 and convolution. The original signal, reconstructed signal using proposed HWT method and the reconstruction error are illustrated in Figure 4.8. It is noted that the signal reconstruction error of EMD is smaller than DWT and WPT. But EMD is not suitable for real time implementation because of computational complexity and proposed HWT is more suitable to clean EEG for implementation of online BCI.

Finally, the HWT is applied to a contaminated EEG. In such hybrid wavelet decomposition the signal is represented as finite set band-passed signals of different frequency bands in WPT part. The log_2 energies of the orderly sequence subband coefficients are computed. Using Eq.(4.31), the magnitude of reference signal (fGn) is calculated before it is decomposed. The WPT technique is applied on statistically thresholded fGn for calculating its log_2 energy. After calculating energy of the every subband of measured EEG and fGn, the energy based subband is selected as artifact which exceeding the energy of fGn is illustrated in Figure 4.9. In Figure 4.9, the subband index are plotted in order to analogy with other methods. The artifact free EEG signals are reformed by summing up the reconstructed subband. The separation results of purified EEG are shown in Figure 4.10 and the purified EEG signal is found as completely artifact free.

Chapter 5

EEG Classification for BCI Implementation

The neural signal such as Electroencephalogram (EEG) signal classification for motor imagery movements has been a great challenge in the design and improvement of Brain Computer Interfaces (BCIs). There are mainly two challenges. The first one is the variability in the recorded EEG data, which manifests across trials as well as across individuals. Consequently, features that are more discriminative need to be identified before any pattern recognition technique can be applied. The second challenge is in the pattern recognition domain. The number of data samples in a class of interest, e.g. a specific action, is a small fraction of the total data, which is composed of samples corresponding to all actions of all users. Building a robust classifier when learning from a highly unbalanced dataset is very difficult; minimizing the classification error typically causes the larger class to overwhelm the smaller one [183].

5.1 Existing Classifiers Used in BCI

This section reviews the classification algorithms used to design BCI systems. They are several types of classifiers, mainly divided into five different categories: linear classifiers, neural networks, nonlinear bayesian classifiers, nearest neighbor classifiers and combinations of classifiers. The most important characteristics of different classifiers is the essential properties for BCI applications which are highlighted in the following:

5.1.1 Linear Classifiers

Linear classifiers are discriminant algorithms that use linear functions to distinguish classes. They are probably the most popular algorithms for BCI applications. Two main kinds of linear classifiers have been used for BCI design, namely, Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM).

1. Linear Discriminant Analysis: The aim of LDA (also known as Fisher's LDA) is to use hyperplanes to separate the data representing the different classes. For a two-class problem, the class of a feature vector depends on which side of the hyperplane the vector is (see Figure 5.1).

   LDA assumes normal distribution of the data, with equal covariance matrix for both classes. The separating hyperplane is obtained by seeking the projection that maximize the distance between the two classes means and minimize the inter class variance. To solve an N-class problem ($N > 2$) several hyperplanes are used. The strategy generally used for multiclass BCI is the "One Versus the Rest" (OVR) strategy which consists in separating each class from all the others. This technique has a very low computational

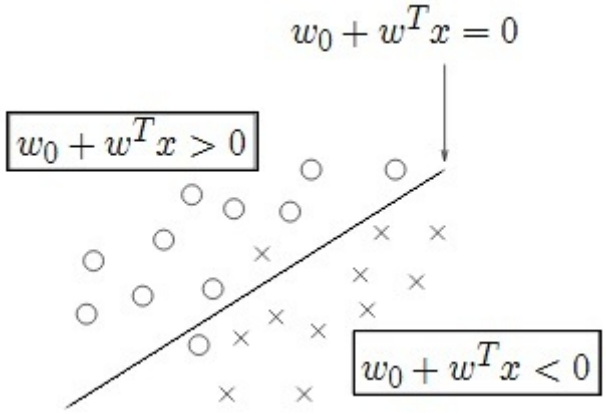

Figure 5.1: A hyperplane which separates two classes: the "circles" and the "crosses" [116].

   requirement which makes it suitable for online BCI system. Moreover this classifier is simple to use and generally provides good results. Consequently, LDA has been used with success in a great number of BCI systems such as motor imagery based BCI, P300 speller, multiclass or asynchronous BCI. The main drawback of LDA is its linearity that can provide poor results on complex nonlinear EEG data [58].

A Regularized Fisher's LDA (RFLDA) has also been used in the field of BCI. This classifier introduces a regularization parameter that can allow or penalize classification errors on the training set. The resulting classifier can accomodate outliers and obtain better generalization capabilities. As outliers are common in EEG data, this regularized version of LDA may give better results for BCI than the non-regularized version [11]. Surprisingly, RFLDA is much less used than LDA for BCI applications [116].

2. Support Vector Machine: The SVM is a widely used classifier utilizing the method of supervised machine learning. The goal of SVM is to construct an optimal hyperplane using training data that separate two or more datasets for classification in the test data. The optimal hyperplane is constructed in order to obtain the maximal margin from the nearest samples of different datasets, known as the support vectors (see Figure 5.2). In cases where the datasets are not linearly separable in the original finite dimensional space, the data can be re-mapped into a sufficiently higher dimensional space. This mapping is conducted by using a defined kernel function, which presumably ensures an easier separation of the datasets. The hyperplane defined in the higher dimensional space can be viewed as a non-linear separating hyperplane in the original finite dimensional space. Thus, this approach is also known as a non-linear SVM [169].

   Support vector machine has been used widely for classification of electroencephalogram (EEG) signals for the diagnosis of neurological disorders such as epilepsy and sleep disorders. SVM shows good generalization performance for high dimensional data due to its convex optimization problem. The incorporation of prior knowledge about the data leads to a better optimized classifier. The basic support vector classifier is very similar to the perceptron. Both are linear classifiers, assuming separable data. In perceptron learning, the iterative procedure is stopped when all samples in the training set are classified correctly. For linearly separable data, this means that the found perceptron is one solution arbitrarily selected from an (in principle) infinite set of solutions. In contrast, the support vector classifier chooses one particular solution: the classifier which separates the classes with maximal

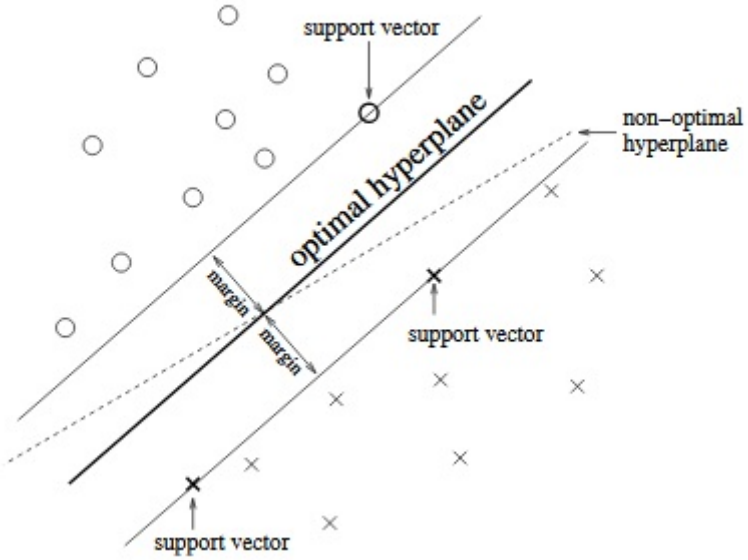

Figure 5.2: SVM find the optimal hyperplane for generalization [116].

margin. The margin is defined as the width of the largest tube not containing samples that can be drawn around the decision boundary. It can be proven that this particular solution has the highest generalization ability.

The support vector classifier has many advantages. A unique global optimum for its parameters can be found using standard optimization software. Nonlinear boundaries can be used without much extra computational effort. Moreover, its performance is very competitive with other methods. A drawback is that the problem complexity is not of the order of the dimension of the samples, but of the order of the number of samples. For large sample sizes $N_S > 1000$ general quadratic programming software will often fail and special-purpose optimizers using problem-specific speedups have to be used to solve the optimization [187] .

5.1.2 Neural Networks

Artificial neural networks are computational systems originally inspired by the human brain. They consist of many computational units, called neurons, which perform a basic operation and pass the information of that operation to further neurons. The operation is generally a summation of the information received by the neuron followed by the application of a simple, non-linear function. In most

neural networks, these neurons are then organized into units called layers. The processing of neurons in one layer usually feeds into the calculations of the next, though certain types of networks will allow for information to pass within layers or even to previous layers. The final layer of a neural network outputs a result, which is interpreted for classification or regression. Neural Networks (NN)[116] are, together with linear classifiers, the category of classifiers mostly used in BCI research. Let us recall that a NN is an assembly of several artificial neurons which enables to produce nonlinear decision boundaries. This section first describes the most widely used NN for BCI, which is the MultiLayer Perceptron (MLP). Then, it brifly presents other architectures of neural network used for BCI applications.

1. MultiLayer Perceptron: An MLP is composed of several layers of neurons: an input layer, possibly one or several hidden layers, and an output layer. Each neuron's input is connected with the output of the previous layer's neurons whereas the neurons of the output layer determine the class of the input feature vector. Neural networks and thus MLP, are universal approximators, i.e., when composed of enough neurons and layers, they can approximate any continuous function. Added to the fact that they can classify any number of classes, this makes NN very flexible classifiers that can adapt to a great variety of problems. Consequently, MLP, which are the most popular NN used in classification, have been applied to almost all BCI problems such as binary or multiclass, synchronous or asynchronous BCI. However, the fact that MLP are universal approximators makes these classifiers sensitive to overtraining, especially with such noisy and non-stationary data as EEG, e.g., [5]. Therefore, careful architecture selection and regularization is required [88]. A MultiLayer Perceptron without hidden layers is known as a perceptron. Interestingly enough, a perceptron is equivalent to LDA and, as such, has been sometimes used for BCI applications [26].

2. Other Neural Network Architectures: Other types of NN architecture are used in the field of BCI. Among them, one deserves a specific attention as it has been specifically created for BCI: the Gaussian classifier [135]. Each unit of this NN is a Gaussian discriminant function representing a class prototype. According to its authors, this NN outperforms MLP on BCI data and can

perform efficient rejection of uncertain samples. As a consequence, this classifier has been applied with success to motor imagery [182] and mental task classification, particularly during asynchronous experiments. Besides the Gaussian classifier, several other NN have been applied to BCI purposes, in a more marginal way. They are given below: [169]

- Learning Vector Quantization (LVQ) Neural Network [155]
- Fuzzy ARTMAP Neural Network;
- Dynamic Neural Networks such as the Finite Impulse Response Neural Network (FIRNN), Time-Delay Neural Network (TDNN) or Gamma dynamic Neural Network (GDNN);
- RBF Neural Network;
- Bayesian Logistic Regression Neural Network (BLRNN);
- Adaptive Logic Network (ALN);
- Probability estimating Guarded Neural Classifier (PeGNC) [50].

5.1.3 Nonlinear Bayesian Classifiers

This section introduces [169] two Bayesian classifiers used for BCI: Bayes quadratic and Hidden Markov Model (HMM). Although Bayesian Graphical Network (BGN) has been employed for BCI, it is not described here as it is not common and, currently, not fast enough for real-time BCI [164]. All these classifiers produce nonlinear decision boundaries. Furthermore, they are generative, which enables them to perform more efficient rejection of uncertain samples than discriminative classifiers. However, these classifiers are not as widespread as linear classifiers or neural networks in BCI applications.

1. Bayes quadratic: Bayesian classification aims at assigning to a feature vector the class it belongs to with the highest probability. The Bayes rule is used to compute the so-called a posteriori probability that a feature vector has of belonging to a given class. Using the MAP (Maximum A Posteriori) rule and these probabilities, the class of this feature vector can be estimated.
   Bayes quadratic consists in assuming a different normal distribution of data.

This leads to quadratic decision boundaries, which explains the name of the classifier. Even though this classifier is not widely used for BCI, it has been applied with success to motor imagery and mental task classification.

2. Hidden Markov Model (HMM): The HMMs are popular dynamic classifiers in the field of speech recognition. An HMM is a kind of probabilistic automaton that can provide the probability of observing a given sequence of feature vectors. Each state of the automaton can modelize the probability of observing a given feature vector. For BCI, these probabilities usually are Gaussian Mixture Models (GMM). HMM are perfectly suitable algorithms for the classification of time series. As EEG components used to drive BCI have specific time courses, HMM have been applied to the classification of temporal sequences of BCI features and even to the classification of raw EEG. HMM are not much widespread within the BCI community but these studies revealed that they were promising classifiers for BCI systems. Another kind of HMM which has been used to design BCI is the Input-Output HMM (IOHMM). The IOHMM is not a generative classifier but a discriminative one. The main advantage of this classifier is that one IOHMM can discriminate several classes, whereas one HMM per class is needed to achieve the same operation.

5.1.4 Nearest Neighbor Classifiers

The classifiers presented in this section [169] are relatively simple. They consist in assigning a feature vector to a class according to its nearest neighbor(s). This neighbor can be a feature vector from the training set as in the case of k Nearest Neighbors (kNN), or a class prototype as in Mahalanobis distance. They are discriminative nonlinear classifiers.

1. k Nearest Neighbors: The aim of this technique is to assign to an unseen point the dominant class among its k nearest neighbors within the training set. For BCI, these nearest neighbors are usually obtained using a metric distance. With a sufficiently high value of k and enough training samples, kNN can approximate any function which enables it to produce nonlinear

decision boundaries. KNN algorithms are not very popular in the BCI compmunity, probably because they are known to be very sensitive to the curse-of-dimensionality, which made them fail in several BCI experiments. However, when used in BCI systems with low-dimensional feature vectors, kNN may prove to be efficient.

2. Mahalanobis distance: Mahalanobis distance based classifiers assume a Gaussian distribution for each prototype of the class. Then, a feature vector is assigned to the class that corresponds to the nearest prototype, according to the so-called Mahalanobis distance. This leads to a simple yet robust classifier, which even proved to be suitable for multiclass or asynchronous BCI systems. Despite its good performances, it is still scarcely used in the BCI literature.

5.1.5 Combinations of Classifiers

In most papers related to BCI, the classification is achieved using a single classifier. A recent trend, however, is to use several classifiers, combined in different ways. The classifier combination strategies used in BCI applications are the following [169]:

1. Voting: The voting approach was used to determine the class designation. While using voting, several classifiers are being used, each of them assigning the input feature vector to a class. The final class will be that of the majority. Voting is the most popular way of combining classifiers in BCI research, probably because it is simple and efficient. For instance, voting with LVQ NN, MLP or SVM have been attempted.

2. Boosting: Boosting consists in using several classifiers in cascade, each classifier focusing on the errors committed by the previous ones. It can build up a powerful classifier out of several weak ones, and it is unlikely to overtrain. Unfortunalty, it is sensible to mislabels which may explain why it was not succesfull in one BCI study. To date, in the field of BCI, boosting has been experimented with MLP and Ordinary Least Square (OLS).

3. Stacking: Stacking consists in using several classifiers, each of them classifying the input feature vector. These classifier are called level-0 classifiers. The output of each of these classifiers is then given as input to a so-called meta-classifier (or level-1 classifier) which makes the final decision. Stacking has been used in BCI research using HMM as level-0 classifiers, and an SVM as meta-classifier.

4. Random subspaces: The random subspace technique consists in generating new training sets from the original one and in training a different classifier for each one of these new training sets [188]. The final decision is made thanks to majority voting. When using random subspaces, the new training sets are generated by using only a subset of the features from the original training set, these features being randomly selected. The main advantage of such a method is that it enables to reduce the dimensionality while still using all the available features, through several classifiers. This method has been used for BCI with decision trees and kNN [188].

   The main advantage of such techniques is that a combination of similar classifiers is very likely to outperform one of the classifiers on its own. Actually, combining classifiers is known to reduce the variance and thus the classification error.

5.2 Terminology for BCI Classification

Brain computer interfaces combine knowledge and techniques from neuroscience, signal processing and machine learning. Besides having the hardware to record brain signals, it is also important to know which parts of the brain are responsible for certain mental processes and how the signals in the brain behave under these mental processes. Using mental processes to activate brain regions to control a device or a computer program is called paradigms. The most used paradigm is the motor imagery (MI) paradigm. Here, the subject imagines a movement with one of its limbs to activate areas in the Sensory Motor Cortex (SMC). The activation in this area produces changes in certain frequency ranges which can be captured with signal processing techniques like the discrete Fourier transform. However,

people suffering from ALS are not suitable for the MI paradigm since their motor neurons are damaged. Here, other mental paradigms need to be used, e.g., audio, mental subtraction, mental navigation, word association and mental rotation.

After using signal processing techniques to focus on certain regions and frequency ranges, features need to be extracted from the signals to classify them and translate them into different commands for the devices or computer programs. For this purpose, classifiers like LDA, SVM or logistic regression are commonly used. These classifiers are able to recognize patterns in the signals and discriminate between them. Discriminating the signals makes it possible to translate these signals into different commands.

The present section [115] is dedicated to classification methods for BCI. However, most pattern recognition/machine learning pipelines, and BCIs are no exception, not only use a classifier, but also apply feature extraction/selection techniques to represent EEG signals in a compact and relevant manner. In particular for BCI, EEG signals are typically filtered both in the time domain (band-pass filter), and spatial domain (spatial filter) before features are extracted from the resulting signals. The best subsets of features are then identified using feature selection algorithms, and these features are used to train a classifier.

Feature Extraction: As there are many ways in which EEG signals can be represented (e.g. [6, 113, 121]), the two most common types of features used to represent EEG signals are frequency band power features and time point features. Band power features represent the power (energy) of EEG signals for a given frequency band in a given channel, averaged over a given time window (typically 1 second for many BCI paradigms). Band power features can be computed in various ways [13, 72], and are extensively used for BCIs exploiting oscillatory activity, i.e. changes in EEG rhythm amplitudes. As such, band power features are the gold standard features for BCI based on motor and mental imagery for many passive BCI aiming at decoding mental states such as mental workload or emotions, or for SSVEP-based BCIs.

Time point features are a concatenation of EEG samples from all channels. Typically, such features are extracted after some pre-processing, notably band-pass or low-pass filtering and down-sampling. They are the typical features used

to classify ERP, which are temporal variations in EEG signals amplitudes time-locked to a given event/stimulus. These are the features used in most P300-based BCI.

Both types of features are benefitted from being extracted after spatial filtering. Spatial filtering consists of combining the original sensor signals, usually linearly, which can result in a signal with a higher signal-to-noise ratio than that of individual sensors. Spatial filtering can be data independent, e.g. based on physical consideration regarding how EEG signals travel through the skin and skull, leading to spatial filters such as the well-known Laplacian filter or inverse solution based spatial filtering. Spatial filters can also be obtained in a data-driven and unsupervised manner with methods such as PCA or ICA. Finally, spatial filters can be obtained in a data driven manner, with supervised learning, which is currently one of the most popular approaches. Supervised spatial filters include the well-known Common Spatial Patterns (CSPs), dedicated to band-power features and oscillatory activity BCI.

While spatial filtering followed by either band power or time points feature extraction are by far the most common features used in current EEG-based BCIs, it should be mentioned that other feature types have been explored and used. Firstly, an increasingly used type is connectivity features. Such features measure the correlation or synchronization between signals from different sensors and/or frequency bands. This can be measured using features such as spectral coherence, phase locking values or directed transfer functions, among many others. Researchers have also explored various EEG signal complexity measures or higher order statistics as features of EEG signals. Finally, rather than using vectors of features, recent research has also explored how to represent EEG signals by covariance matrices or by tensors (i.e. arrays and multi-way arrays, with two or more dimensions), and how to classify these matrices or tensors directly.

Feature Selection: A feature selection step can be applied after the feature extraction step to select a subset of features with various potential benefits. Firstly, among the various features that one may extract from EEG signals, some may be redundant or may not be related to the mental states targeted by the BCI. Secondly, the number of parameters that the classifier has to optimize is positively

correlated with the number of features. Reducing the number of features thus leads to fewer parameters to be optimized by the classifier. It also reduces possible overtraining effects and can thus improve performance, especially if the number of training samples is small. Thirdly, from a knowledge extraction point of view, if only a few features are selected and/or ranked, it is easier to observe which features are actually related to the targeted mental states. Fourthly, a model with fewer features and consequently fewer parameters can produce faster predictions for a new sample, as it should be computationally more efficient. Fifthly, collection and storage of data will be reduced. Three feature selection approaches have been identified [106]: the filter, wrapper and embedded approaches. Many alternative methods have been proposed for each approach.

Filter methods rely on measures of relationship between each feature and the target class, independently of the classifier to be used. The coefficient of determination, which is the square of the estimation of the Pearson correlation coefficient, can be used as a feature ranking criterion. The coefficient of determination can also be used for a two-class problem, labelling classes as 1 or +1. The correlation coefficient can only detect linear dependencies between features and classes though. To exploit non-linear relationships, a simple solution is to apply non-linear preprocessing, such as taking the square or the log of the features. Ranking criteria based on information theory can also be used e.g. the mutual information between each feature and the target variable. Many filter feature selection approaches require estimations of the probability densities and the joint density of the feature and class label from the data. One solution is to discretize the features and class labels. Another solution is to approximate their densities with a non-parametric method such as Parzen windows. If the densities are estimated by a normal distribution, the result obtained by the mutual information will be similar to the one obtained by the correlation coefficient. Filter approaches have a linear complexity with respect to the number of features. However, this may lead to a selection of redundant features.

Feature selection has provided important improvements in BCI, e.g. the stepwise linear discriminant analysis (embedded method) for P300-BCI and frequency bands selection for motor imagery using maximal mutual information (filtering

methods). Other popular methods used in EEG-based BCIs notably include filter methods such as maximum relevance minimum redundancy (mRMR) feature selection [166, 180] or R2 feature selection. It should be mentioned that five feature selection methods, namely information gain ranking, correlation-based feature selection, Relief (an instance-based feature ranking method for multiclass problems), consistency based feature selection and 1R Ranking (one-rule classification) have been evaluated on the BCI competition III data sets. Amongst ten classifiers, the top three feature selection methods were correlation-based feature selection, information gain and 1R ranking, respectively.

5.3 Classification Approach with Artifact Suppression

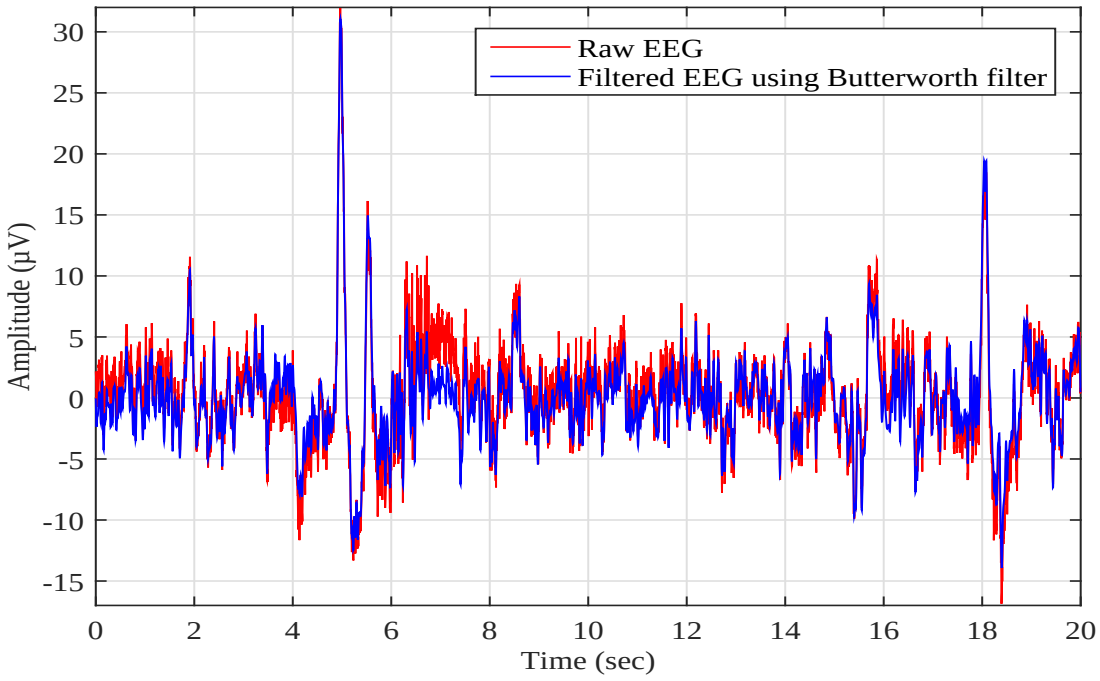

Figure 5.3: Time domain representation of raw EEG and filtered EEG signal using Butterworth band pass technique.

5.3.1 Preprocessing

Generally, the procedure of transforming raw data into an arrangement that is more suitable for further study and interpretable for the user is called preprocess-

ing. For instance of EEG data, preprocessing usually refers to eliminate noise from the data to get closer to the real neural signals. There are several causes for preprocessing of EEG data. Firstly, the signals that are selected from the scalp are not necessarily an accurate representation of the signals originating from the brain, as the spatial information gets lost. Secondly, EEG data tends to contain a lot of noise and artifacts which can produce weaker EEG signals. The well-known artifacts such as eye blinking or muscle movement can contaminate the data and distort the representation. Finally, to separate the essential neural signals from random neural activity that happens during EEG recordings.

In this experiment, a class is assigned to each trial, i.e., the discrete classification of each class is considered. For each dataset and trial, the data from time segment located from 0.5s to 2.5s after the visual cue instructing the subject to perform MI is considered. A fourth-order Butterworth band pass filter is used to filter useful components (0.5-30Hz) from each trial which is illustrated in Figure 5.3. It represented the time domain view of the raw EEG and Butterworth band pass filtered EEG.

5.3.2 Spatial Filtering and Features Extraction

The CSP is a feature extraction technique used in signal processing for separating a multivariate signal into additive sub-components. The technique used to design spatial filters such that the variance of the filtered data from one class is maximized while the variance of the filtered data from the other class is minimized. Thus, the resulting feature vectors increase the discriminability between the two classes by means of minimize the intra class variance and maximize the inter class variance [162]. This property builds CSP as one of the most effective spatial filters for EEG signal processing. The method of CSP was first introduced to EEG analysis for detection of abnormal EEG [102], and effectively applied on movement-related EEG for the classification purpose [138, 2]. The target of the CSP is to project the multichannel EEG data into low dimensional spatial subspace with a projection matrix using linear transformation.

For details explanation of the CSP algorithm, assume the original EEG data matrix E_k^i from trial i for class k. The dimension of each E_k^i is $D \times T$, where D

is the number of channels and T is the number of samples per channel. For the explanation, the EEG data of a single trial ($i = 1$) is represented $E_{k\in(L,R)}$ as where L denotes left hand and R denotes right hand t movement. The normalized spatial covariance of the EEG for left hand movement, C_L and right hand movement, C_R can be calculated as:

$$C_L = \frac{E_L E_L'}{tr(E_L E_L')}, C_R = \frac{E_R E_R'}{tr(E_R E_R')} \tag{5.1}$$

where E_L and E_R represent the original EEG matrices for left hand and right hand movement respectively, $()^{'}$ is the transpose operator and $tr(.)$ represents the sum of the diagonal elements of any given matrix. The composite spatial covariance, C is the sum of the averaged normalized spatial covariance $\bar{C}_L$ and $\bar{C}_R$. The $\bar{C}_L$ and $\bar{C}_R$ are estimated by averaging over all the trials of each class. The composite spatial covariance, C is calculated as

$$C = \bar{C}_R + \bar{C}_L = \sum \lambda \sum^{'} C = \bar{C}_R + \bar{C}_L = \sum \lambda \sum^{'} \tag{5.2}$$

where $\sum$ is the matrix of Eigen vectors and λ is the diagonal matrix of Eigen values. The averaged normalized spatial covariance $\bar{C}_L$ and $\bar{C}_R$ are transformed as

$$J_L = X\bar{C}_L X' \quad \text{and} \quad J_R = X\bar{C}_R X' \tag{5.3}$$

where $X = \frac{\sum^{'}}{\sqrt{\lambda}}$ is the whitening transformation matrix. $\bar{J}_L$ and $\bar{J}_R$ share common eigenvectors and the sum of corresponding eigenvalues for the two matrices will always be one. If $\bar{J}_L = Y\Lambda_L Y'$ and $\bar{J}_R = Y\Lambda_R Y'$ then , $\Lambda_L + \Lambda_R = I$ where I is the identity matrix. Since the sum of two corresponding eigenvalues is always one, a high eigenvalue for$\bar{J}_L$ means that a high variance for EEG in left hand movement and a low variance for the EEG in right hand movement (low eigenvalue for$\bar{J}_R$) and vice versa. The classification operation is done based on this property. The projection of whitened EEG onto the eigenvectors Y corresponding to the largest Λ_L and Λ_R will give feature vectors that significantly enhance the discrimination ability.

The goal of the CSP is to find B spatial filters to create a projection matrix W of dimension $N \times B$ (each column is a spatial filter). The projection matrix W is represented as

$$W = Y'X \tag{5.4}$$

The projection matrix W linearly transforms the original EEG into uncorrelated components according to:

$$Z = WE \tag{5.5}$$

The original EEG, E can be reconstructed by $E = W^{-1}Z$ where W^{-1} is the inverse matrix of W. The columns of W^{-1} are spatial patterns that describe the variance of the EEG. The first and last columns contain the most discriminatory spatial patterns that explain the high variance of one class and the low variance of the other.

For successful calculation of spatial filters, the CSP algorithm is greatly used algorithm. In this step, the CSP algorithm is used to perform the spatial filtering action on training setoff EEG signal. Using Eq.(5.5) the filtering is accomplished by linearly transforming the EEG measurements. Both training and testing set of EEG, CSP features are then extracted from each of the trial of the samples. The process is completed by projecting the EEG data onto the CSP filters. When the filters w are obtained, CSP feature extraction consists in filtering the EEG signals using the w and then computing the resulting signals variance. The CSP features are generated as the log variance of the projected signals.

$$f = \log(wC_{ct}w^T) \tag{5.6}$$

where, C_{ct} is the current trial covariance matrix, w is projection matrix and T is transpose of a matrix. It is common to select 3 pairs of CSP spatial filters, corresponding to the 3 largest and smallest eigen-values, hence resulting in a trial being described by 6 CSP features [114]. To extract features from the data, the CSP algorithm with CSP filter pairs m ($m = 1, 2,$ and 3) is used in this experiment.

5.3.3 Classification Method Used in the Experiment

A classifier is a technique that utilizes various independent variable values (features) as input and predicts the corresponding class to which the independent variable belongs [76]. In the EEG signal analysis, the features can be any kind of extracted information from the signal, such as energy, entropy, power etc. and the class can be the type of task or the stimulus used during the recording. A classifier has a number of parameters that need to be learned from training data.

The learned classifier is a model of the association between the features and the classes. For example, for a given feature x of a class y, the classifier is a function f that predicts the class $y = f(x)$. After the learning, the classifier is able to predict new instances that have not been used in the training data. Thus, the performance of the classifier is tested on a different set of instances.

The LDA, also known as Fishers linear discriminant analysis is a technique used to find a linear combination of features that separates two or more classes of data. It is typically used as a dimensionality reduction step before classification [95]. It reduces dimensionality but at the same time preserves as much of the class discriminatory information as possible. The goal of the LDA is to use a separating hyperplane that maximally separate the data representing the different classes. The hyperplane is found by selecting the projection, where the same classes are projected very close to each other and the distance between two classes means is as maximum as possible [114].

Let as assume that we have K classes, each containing N observations x_i. The within-class scatter, $\tilde{S}_w$ for all K classes can be calculated as:

$$\tilde{S}_w = \sum_{k=1}^{K} f_k S_w^k \tag{5.7}$$

where the within-class covariance matrix S_w^k and the fraction of data f_k are calculated according to the following formulas:

$$S_w^k = \sum_{i=1}^{N_K} (x_i^k - \mu^k)(x_i^k - \mu^k)^T \tag{5.8}$$

$$f_k = \frac{N_k}{\sum_{j=1}^{K} N_j} \tag{5.9}$$

where N_k is the number of observations of k^{th} class and μ^k indicates mean of the all observations x_i for k^{th} class. The between class scatter $\tilde{S}_b$ for all K classes is calculated as:

$$\tilde{S}_b = \sum_{k=1}^{K} f_k S_b^k \tag{5.10}$$

where the between class covariance matrix, S_b^k can be estimated as

$$S_b^k = \sum_{k=1}^{K} (\mu^k - \mu)(\mu^k - \mu)^T \tag{5.11}$$

where μ indicates the mean of the all observations x_i for all classes. The main objective of LDA is to find a projection matrix that maximizes the ratio of the determinant of $\tilde{S}_b$ to the determinant of $\tilde{S}_w$. The projections that providing the best class separation are eigenvectors with the highest eigenvalues of matrix M:

$$M = \frac{\tilde{S}_b}{\tilde{S}_w} \tag{5.12}$$

Since the matrix M is asymmetric, the calculation of eigen-vectors can be difficult. This difficulty can be minimized by using generalized eigenvalue problem [102]. Now, the aim of the LDA is to seek $(K-1)$ projections $[y_1, y_2, y_3, ..., y_{k-1}]$ by means of $(K-1)$ projection vectors. The transformed data set y is obtained as a linear combination of all input features x with weights W.

$$y = x^T W \tag{5.13}$$

where $W = [w_1, w_2, w_3, ..., w_H]$ is a matrix form with the H eigen-vectors of matrix M associated with the highest eigenvalues. The LDA reduces the original feature space dimension to H. The LDA performs well when the discriminatory information of data depends on the mean of the data. But it does not work for the variance depended discriminatory informative data. Also, the performance of the LDA is not good for nonlinear classification.

The CSP based classification is implemented with LDA to classify the EEG trials using the ranked CSP features. The LDA classifier uses a linear hyper plane to separate feature vectors from two classes. The intercept b and normal vector a of this hyperplane are computed as follow:
slope of the discriminant hyperplane:

$$a^T = C^{-1}(\mu_1 - \mu_2)^T \tag{5.14}$$

bias of the discriminant hyperplane:

$$b = -\frac{1}{2}(\mu_1 + \mu_2)^T \tag{5.15}$$

with μ_1 and μ_2 being the mean feature vectors for each class and C the covariance matrix of both classes. With LDA, for an input feature vector x, the classification output is

$$\text{LDA}_{out} = ax^T + b \tag{5.16}$$

where, x=feature vector. If this output is positive, the feature vector is assigned to the first class, otherwise it is assigned to the second [114]. During LDA training the discriminant hyperplane coefficients k_i are computed for every training set ranked feature i ($i = 1, ..., 2m$ where m is the CSP filter pair). The LDA accuracy score of each test set ranked CSP feature θ_j ($j = 1, ..., 2m$) is now measured by the computed coefficients. For every value of k_i and their corresponding test set features θ_j , LDA accuracy scores are measured as $S_{m,j} = k_j \Theta \theta_j$ where the symbol Θ represents LDA test operator. So, it classifies a set of data using a trained LDA classifier and provides the resulting score (LDA output) and classification accuracy (%).

5.3.4 Proposed Method for Classification

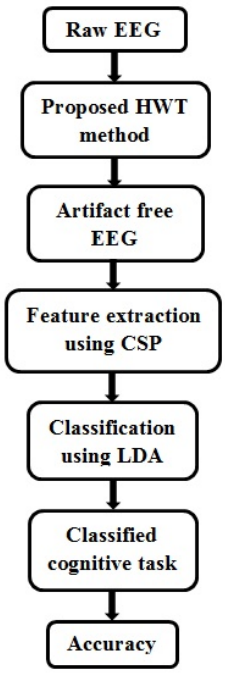

Figure 5.4: Overview of Motor Imagery (MI) EEG signal classification using proposed artifact removing method. We applied this approach to a group of 9 healthy subjects and found a significant classification performance enhancement as compared to other artifact separation approaches.

In this section, a number of channels that cover the motor cortex of the brain are selected in the preprocessing stage. The training and test sets of the raw EEG are separated and individually filtered by a band pass filter. The preprocessing stage is applied before proposed HWT artifact removing method. Figure 5.4,

shows the block diagram of the proposed method for classification accuracy which ensure the proposed HWT artifact removing method is perfectly removed the artifact from raw EEG signal. The method is subdivided into two stages for EEG signal processing and machine learning. A detail of each stage is described below for MI classification. Our BCI system is composed of four main modules (see Figure 5.4):

- An Artifact Detection Module;
- An Artifact Removal Module;
- A Feature Extraction Module; and
- A Feature Classification Module.

An Artifact Detection Module: Therefore, eight EEG segments are obtained each second. The artifact detection algorithm is first applied to each EEG segment, before that segment is processed by the artifact removal, feature extraction and feature classification modules.

An Artifact Removal Module: We propose to remove the artifacts using the HWT with an adaptive thresholding mechanism. As shown in Figure 4.7, the wavelet coefficients generated by the artifact detection module are used in our artifact removal algorithm to denoise the EEG signals. The denoised signals are obtained by performing an inverse HWT on the thresholded wavelet coefficients energy. The performance of the proposed algorithm is compared with those of other artifact removal algorithms such as SSA, DWT, EMD algorithms. The performance evaluation criteria used are provided in the next chapter.

A Feature Extraction Module: After processing the EEG signals by the artifact detection and removal modules, the feature extraction and classification modules are applied next. Common Spatial Pattern method was employed for feature extraction in this study. Details of the algorithm are described in previous section with the example of discriminating left hand vs. right hand motor imagery. The filtered signal corresponding to the desynchronization of the left hand motor cortex is characterized by a strong motor rhythm during imagination of right hand movements, and by an attenuated motor rhythm during left hand

imagination. This criterion is exactly what the CSP algorithm optimizes: maximizing variance for the class of right hand trials and at the same time minimizing variance for left hand trials.

A Feature Classification Module: In this section, the feature vector framed out of all the features of different type of signals is given as input to the classifiers. Based on various performance indices, the performance of classifiers are evaluated for this method. The stepwise LDA selects the features that best discriminate between the left and right classes. The whole process of classification method is summarized as follows:

(i) Apply Motor Imagery (MI) raw EEG to proposed HWT artifact removing technique for clean EEG.

(ii) Learn spatial filters by the artifact free EEG using CSP algorithm.

(iii) Extract CSP features by the spatial filters.

(iv) Extracted features are fed to a LDA classifier.

(v) According to the artifact suppression and classification method, finally, display the estimated classified result in percentage.

This straightforward method can be used to design a well-organized BCI system with better accuracy. Wavelet packet transform is employed to decompose raw EEG signals. Thereafter, EEG signals with effective frequency sub-bands are grouped and reconstructed. EEG feature vectors are extracted from the reconstructed EEG signals with common spatial patterns (CSP). A public dataset (BCI Competition IV-II-a) is employed to validate the proposed method. These results show that the proposed method is effective for two-class motor imagery classification.

Chapter 6

Experimental Results and Discussion

One of the most faced challenges for EEG signal processing applications and is an open research problem is artifact recognition and reduction/elimination. In EEG-based health-care applications, the electrodes are placed on the scalp and therefore, the recordings are most lying to artifacts and interferences. To make it is difficult for simple signal preprocessing technique to identify them from EEG, because of the variety of artifacts and their overlapping with signal of interest in both spectral and temporal domain, even sometimes in spatial domain. Sometimes results in poor performance both in terms of signal distortion and artifact removal when the use of simple filtering or amplitude threshold to remove artifacts. It is a great challenge to develop appropriate methods for artifact detection and reduction with the help of recent development in signal processing techniques/algorithms in the past decade and a half. Still now it is an active area of research because there is no universal complete solution yet. However, we realize that there is a gap between designed algorithm and its target application after careful reviewing most of the relevant artifact detection and removal algorithms/methods in the literature. Maximum existing methods are not application-specific and therefore these methods suffer from high computational burden. Considering this scenario, we present a comparative analysis of the existing methods/algorithms with their advantages, limitations and application-specific challenges. First of all, we present different artifact handling methods and then proposed a hybrid wavelet transform denoising method.

6.1 Data Description

The real EEG data collected from well-known publicly available brain computer interface (BCI) competition IV-2b dataset is used to evaluate the proposed method. The data set consists of EEG data from 9 subjects where all subjects are right-handed, having normal or corrected-to-normal vision. Each subject consists of five sessions that are recorded on different days. Each session comprising ten trials, two classes of the motor imagery (MI) of left hand movement (class-1) and right hand movement (class-2), six runs, 20 trials per run and 120 trials per session. The variable of the signal holds 6 channels (the first 3 are EEG signals and the last 3 are EOG signals). The sampling rate is 250 Hz and the subjects are recommended to imagine the corresponding hand movement over a period of 4 seconds results in 1000 samples per channel for every trial. The cue-based screening sessions are consisted of 20 trials per run and 120 trials per session for two classes of imagery. The trials containing artifacts as noted by authorities are marked with 0 corresponding to a clean trial and 1 corresponding to a trial holding an artifact. According to the instruction, 100 trials are obtained as clean trials and 20 trials tenured as artifacts. Among the clean trials, 51 trials for left hand and 49 trials for right hand movement MI. Along with contaminated 20 trials, the 9 trials are for left hand and 11 trials are for right hand movement.

6.2 Experimental Results of EEG Cleaning

The performance of the proposed hybrid wavelets transform (HWT) based method is first tested with simulated data and then the experiments are conducted with real EEG data. The amplitude of all the channels are normalized before processing. The scaling factors of normalization are stored to get back the original amplitude of EEG channel after separating artifact and EEG. If l is the scaling factors of normalization for any EEG channel, the separated EEG with original scale is obtained as:

$$S_{EEG}(t) = l * \hat{S}_{EEG}(t) \tag{6.1}$$

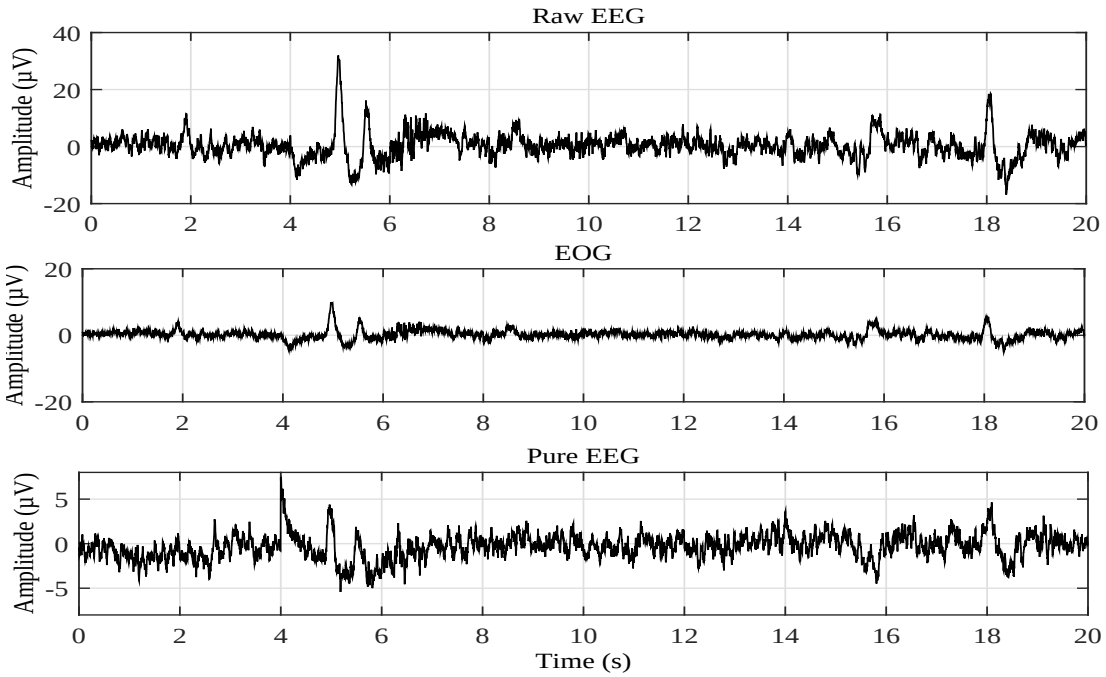

Figure 6.1: The cleaning approach of EEG from the raw data using SSA method. The raw EEG signal (top panel), the separated EOG signal (middle panel), and pure EEG signal (bottom panel). From the illustration, it is obvious that the extracted EOG carry the original information of EEG signal and pure EEG signal cut its valuable information.

where, $S_{EEG}(t)$ is the separated EEG and $\hat{S}_{EEG}(t)$ is the normalized EEG. The EEG signal is first artificially corrupted with EOG and the performance of HWT is compared with SSA, EMD and DWT. In the previous chapter-artifact suppression, the existing artifact removing methods and proposed HWT are tested with the artificially corrupted EEG signal with detail description. In this section, firstly, the HWT based method is tested with the artificially corrupted EEG signal (in summary) and after that the signal is comprised with wavelet based method, DWT and empirical mode decomposition (EMD) technique. For comparison purposes, the time series contaminated EEG data is decomposed into multiple subbands using DWT, EMD, and proposed HWT.

The potential multiresolution decomposition approach, DWT which is analyzing a non-stationary signal like EEG. In this study, db4 is used with 7 level decomposition. It is another approach for EOG artifact reduction. The separation of clean EEG from the contaminated data using DWT is shown in Figure 4.4. The pure EEG do not show any large EOG artifacts but EOG signal contains some original signal information. But it is faster than EMD.

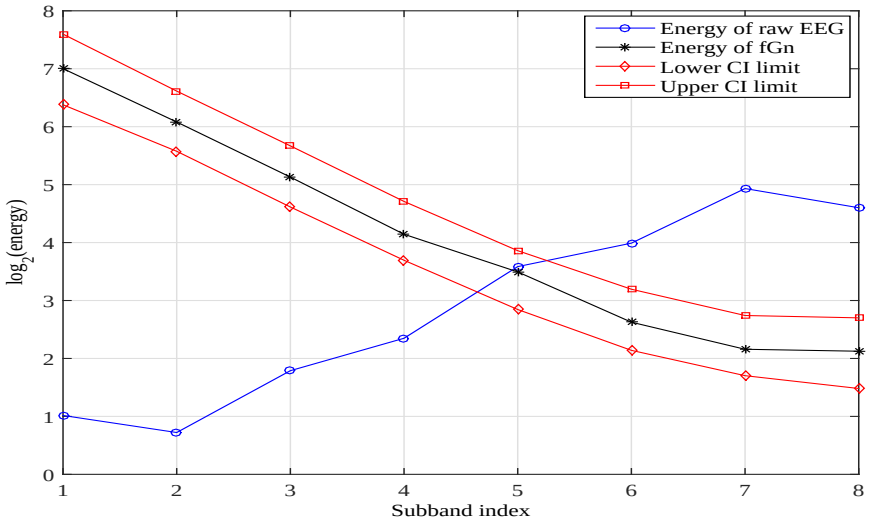

Figure 6.2: The selection of threshold subband index of raw EEG based on the subband energy of fGn. The 6^{th} subband exceeds the upper boundary of CI and hence the 6^{th} one of EEG is selected as the highest order subband index to represent the pure EEG signal.

The EMD is a popular decomposition algorithm which is used here to compare the result of HWT. The EMD method decomposes a signal based on its intrinsic time scales. The experimental result for artifact suppression is illustrated in Figure 4.5 and Figure 4.6. In this Figure 4.6, the method not only reduced EOG artifacts but also eliminated the original information.

Finally, the HWT is applied to a contaminated EEG. In such hybrid wavelet transform, the energy based subband is selected as artifact which exceeding the energy of fGn is illustrated in Figure 4.9. The separation results of purified EEG are shown in Figure 4.10 and the purified EEG signal is found as completely artifact free. Also, it is faster than EMD and DWT.

After analyzing the artificially contaminated EEG signal in different filtering approach with fGn as the reference signal, now is the time to apply all methods to real EEG signals. The separation of EEG contaminated by a known EOG signal is already illustrated at the beginning of this section. All the implemented method is repeated for the raw EEG data. The separation of EOG artifact using different existing method and proposed HWT methods are illustrated below.

Figure 6.1 illustrates the SSA method for EOG artifact reduction using raw EEG. The separated purified EEG and EOG signals are visually represented in time domain here. The first raw, second raw and the third raw are represented

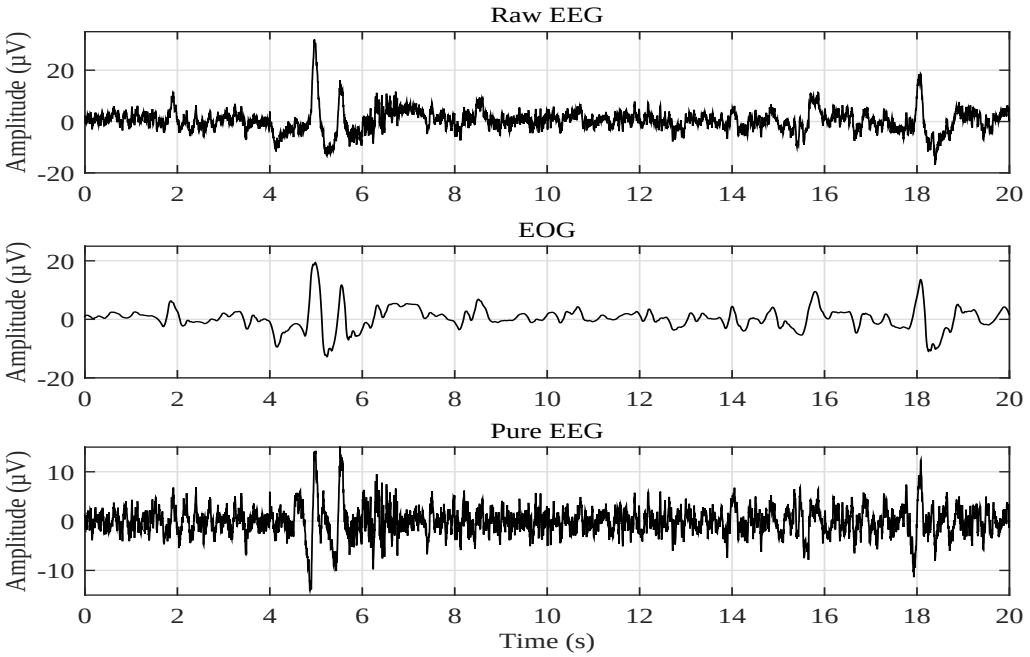

Figure 6.3: Separation of EOG artifact from real (recorded) EEG signal using wavelet denoising method. The recorded (contaminated with EOG) EEG signal (top), the separated EOG artifact (middle) and the pure EEG signal (bottom).

raw EEG, EOG and clean EEG signal respectively. The SSA is removed the less EOG artifacts from the raw data and at the same time the most eye blink artifacts are retained in pure EEG signal.

The wavelet decomposition is also a so called data adaptive technique. A thresholding based denoising method using wavelet transform is implemented to separate the EOG signal from real EEG is illustrated in Figure 6.2 and Figure 6.3. The determination of threshold subband based on the subband energy is illustrated in Figure 6.2. It is observed that the 5^{th} subband exceeds the upper limit of confidence interval. The 5^{th} one is the starting point of lower frequency components. The electrooculogram (EOG) is separated by summing the subbands 5 to 8. The separation results of purified EEG are shown in Figure 6.3.

The EMD based data adaptive filtering approach with fGn as the reference signal is applied to separate the artifacts from real EEG signals presented in Figure 6.4 and Figure 6.5. The selection of the starting IMF is illustrated in Figure 6.4 in which the 6^{th} IMF is selected as the starting point of low frequency component. The separated purified EOG and EEG signals are visually represented in time domain shown in Figure 6.5.

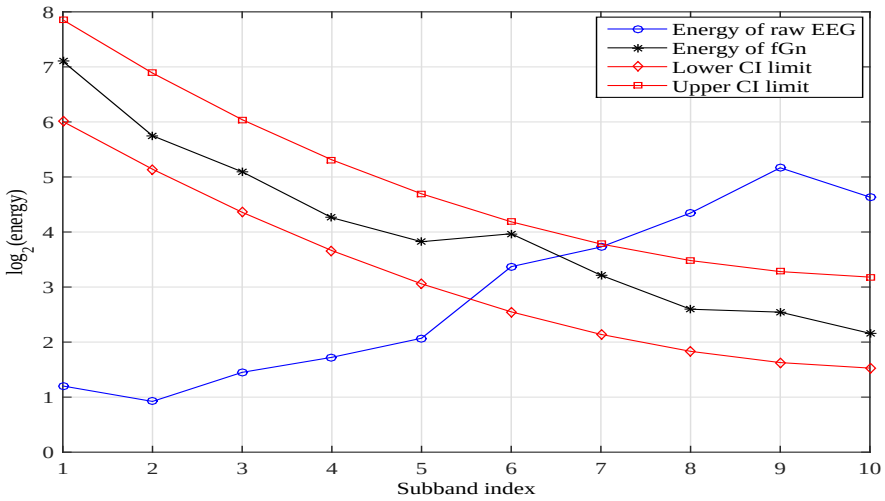

Figure 6.4: Selection of IMF of the real EEG signal using EMD. The selection of starting IMF (lowest order) to extract the low frequency component of mixed signal. Here the 7^{th} IMF is selected. Its energy exceeds the upper limit of 95% confidence interval of the IMFs energies of fGn.

Lastly, the proposed method is evaluated using EOG contaminated EEG signals (real EEG). The HWT is introduced to separate low frequency artifacts from raw EEG using suband energy based criterion and it is illustrated in Figure 6.6. A data adaptive subband filtering is employed here to separate the lower frequency noise. The threshold of filtering is determined by comparing the energy of individual subband of raw EEG with that of the reference signals fGn. The first three subbands mostly contain the noise. According to the experimental results on the raw EEG, HWT is removed the most EOG artifacts from the raw data and at the same time preserved the most EEG signal as shown in Figure 6.7. As a result, the purified EEG signal is found as completely artifact free in HWT method.

The spectrum of EEG signals separated by using SSA, EMD, DWT, and proposed HWT approach are illustrated in Figure 6.8 and Figure 6.9 for measured EEG and raw EEG respectively. Also, the time domain representation of the EEG signal of all methods is shown in Figure 6.10 and Figure 6.11 for contaminated EEG and raw EEG respectively.

In Figure 6.8, the green, black, blue, cyan, magenta and red solid line spectra represent EEG, contaminated EEG, clean EEG spectrum using HWT, EEG spectrum using DWT, EEG spectrum using EMD and EEG spectrum using SSA respectively. The artifact reduction EMD and DWT methods are omitted the

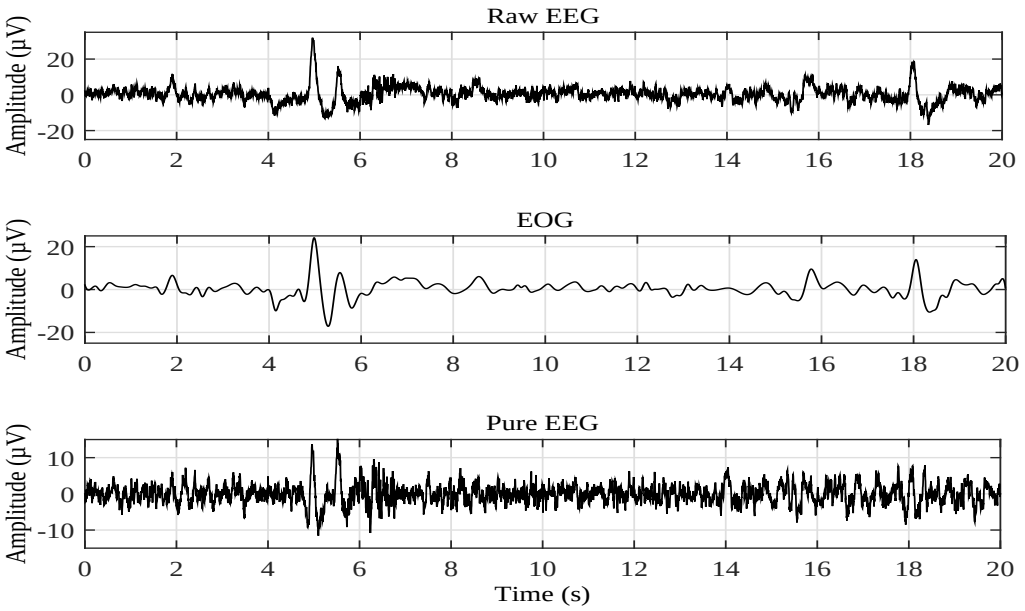

Figure 6.5: The empirical mode decomposition based data adaptive filtering technique is used to separate electro-oculogram (EOG) from the raw electroencephalography (EEG). The top panel shows the real EEG signal and the other panels illustrate EOG (middle) and purified EEG signal (bottom).

delta (<4Hz) band and part of theta (4∼8Hz) band but retained the other EEG frequency bands. The proposed HWT technique is well preserved the EEG frequency bands except part of the delta rhythm. Delta wave is found in deep sleep, so it may be avoided.

In Figure 6.10, the black, green,red, magenta, cyan, and blue solid line signal represent contaminated EEG, EEG, clean EEG using SSA, clean EEG using EMD, EEG using DWT and pure EEG using HWT respectively. The EMD method reduced the EOG but it contained some artifacts with EEG signal. The wavelet based transform DWT produced artifact free signal but computational complexity of DWT is higher than HWT (Figure 6.16). On the other hand, proposed hybrid wavelet based transform HWT is found as completely artifacts free pure EEG without losing information and it takes low computational time.

Beside the visual results, the spectrum of the separated EOG signals are also used as the factor of performance measurement of the separation algorithms. It is well known that the low frequency EOG signal is mixed up with the pure EEG during recording. The frequency range of EOG is usually from 0 to 2.8 Hz and

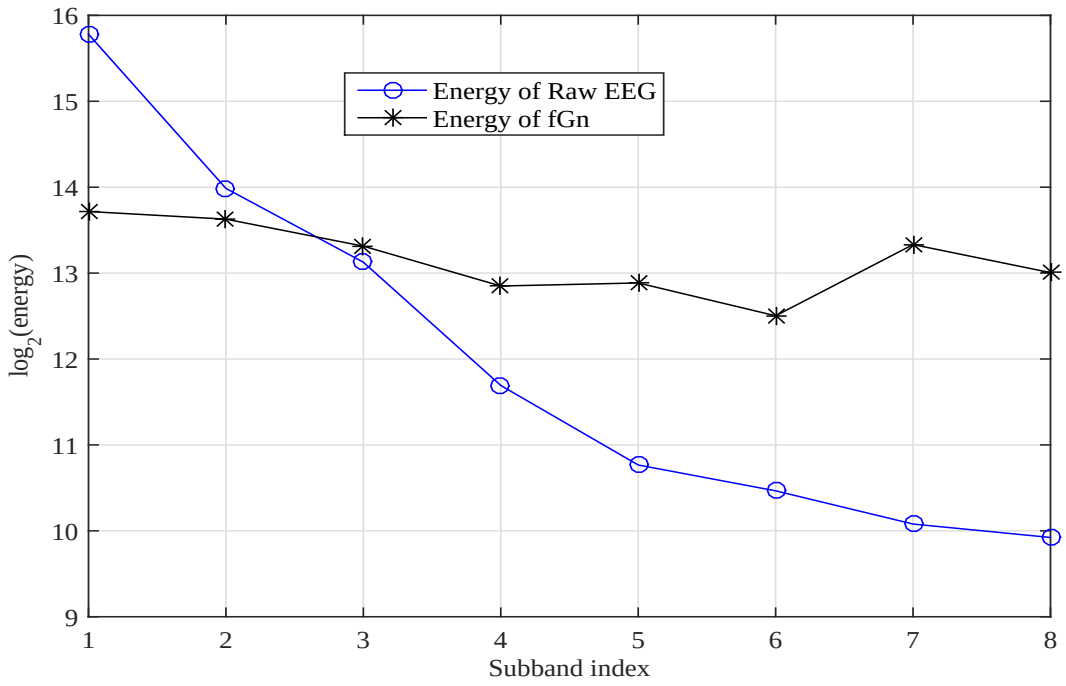

Figure 6.6: The selection of threshold subband energy of raw EEG channels (by applying proposed HWT) based on the subband energy of fGn. In this observation, selection of the index of subbands of the EEG signal from which the low frequency components can be extracted. The subbands energy of fractional Gaussian noise (fGn) is considered as the threshold. The logarithmic energy (with base 2) of individual subband of EEG is compared with the threshold level. The energy of the 2^{nd} subband index of EEG signal exceeds the threshold and hence it is selected to remove the artifacts.

the high frequency part in the spectra comes from the EEG signal. The spectrum of the measured EOG is also included along with the spectrum of extracted EOG for better comparison in Figure 6.13. The spectrum of EOG signals separated by using SSA, EMD, DWT and HWT based filtering approach are illustrated in Figure 6.13 for contaminated EEG and Figure 6.15 for real EEG. It is illustrated that most of the noise energy is concentrated in the lower frequency part of the spectra. The main strength of the EOG from 0 to 2.8 Hz range are understandable overlapping of all spectra and it is very common for three methods and after that range, whereas in HWT approach, no neural activities in the brain correlated with EOG artifacts. However, it must be noted that the relatively lower-energy high-frequency part of the signals plays an important part in conveying information existing three methods-SSA, EMD, DWT [229, 163, 139] are spread beyond the frequency range of actual EOG signals. Also the high frequency components

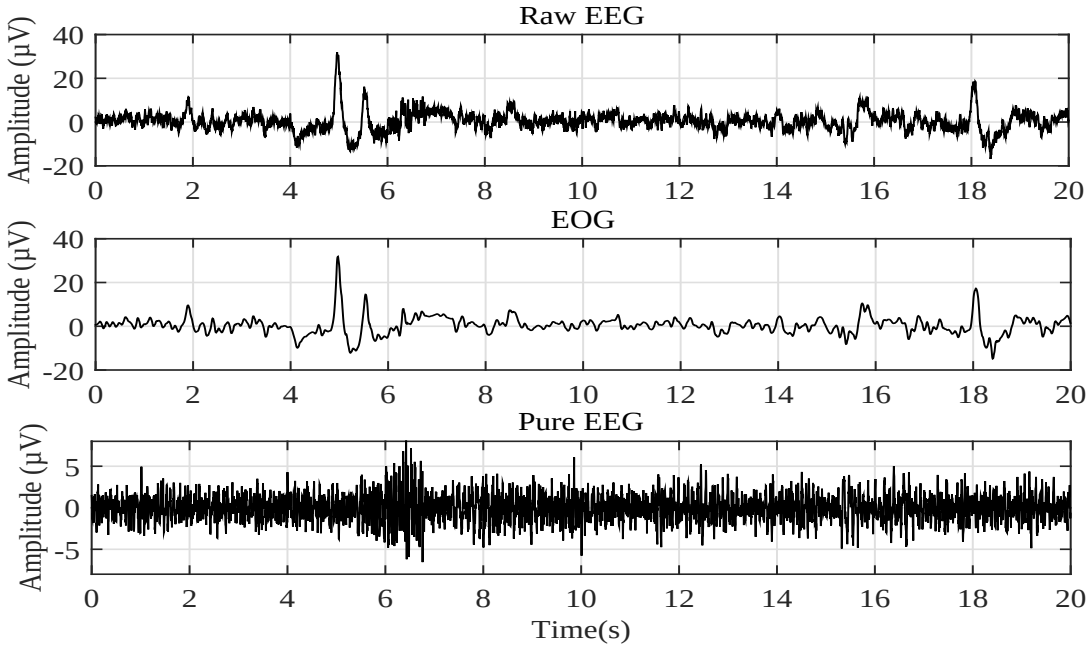

Figure 6.7: Separation of EOG artifact using proposed HWT method. The normal recording of EEG contaminated by EOG which appears as higher energy and lower frequency trend (top panel); the separation of EOG artifact is performed by proposed HWT algorithm (middle panel); purified EEG is extracted from the recorded raw EEG by subtracting the separated EOG (bottom panel) [s: second].

are also with reasonable energies. It is illustrated that the spectrum of wavelet approach (DWT) and other methods (EMD, SSA) capture some energies come from the high frequency EEG signal i.e., the target EOG is not separated properly but includes some signal components of EEG signals. It is noticed from the spectrum of the EOG that the measured EOG signal includes a much amount of EEG signals.

The time domain representation of the EOG signal of all methods is shown in Figure 6.12 and Figure 6.14 for contaminated EEG and raw EEG respectively. In Figure 6.12, the top panel is indicated for EOG signal and the remaining panel are extracted EOG from SSA, EMD and DWT method respectively. The bottom panel shows the extracted EOG by the proposed HWT technique. All the three methods extracted EOG accurately except SSA method. Besides the time domain representation of the EEG signal of all methods is shown in Figure 6.10 and Figure 6.11 for contaminated EEG and raw EEG respectively.

The HWT based filtering is obtained by combining the subbands as frequency

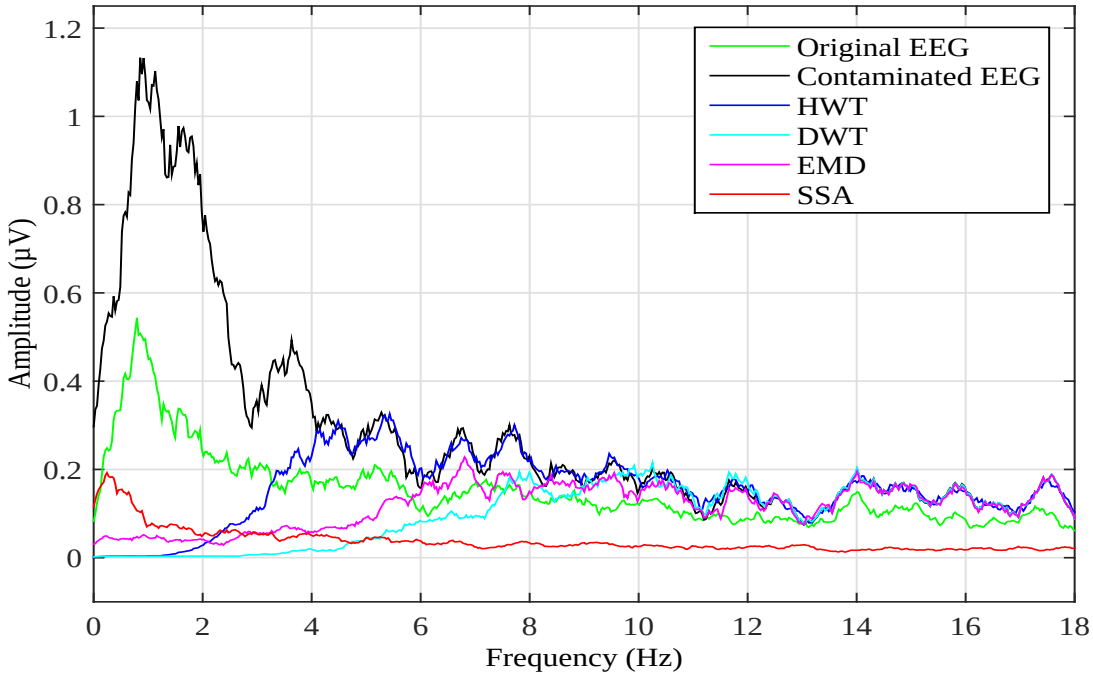

Figure 6.8: The graphical representation of the EEG spectra which contaminated by the EOG and compared it with other EEG separated by different four methods. The green, black, blue, cyan, magenta and red solid line spectra represent original EEG, contaminated EEG, clean EEG spectrum using HWT, DWT, EMD and SSA respectively. The artifact reduction EMD and DWT methods omitted the delta (<4Hz) and part of theta (4∼8Hz) bands but retained the other EEG frequency bands. The SSA method cut the original EEG frequency; it only shows the delta rhythm. The proposed HWT technique is well preserved the EEG frequency bands except part of the delta rhythm. Delta wave is found in deep sleep, so it may be avoided.

varying filtering. It is suitable to process the non-stationary signals. In the proposed method, the EOG is considered as the trend on the recorded EEG signals. If the EOG artifact is viewed as the trend (relatively low frequency with higher amplitude compared to EEG signal), the presented approach will be able to remove it.

To evaluate the performance of the proposed approaches in a real environment, the raw EEG and clean EEG for the mentioned four methods are analyzed in Figure 6.9 (in frequency domain) and Figure 6.11 (in time domain). In Figure 6.11, the first raw, second raw, third raw, fourth raw and the fifth raw are represented raw EEG, clean EEG using SSA, clean EEG using EMD, clean EEG using DWT and clean EEG using proposed method, HWT respectively. It is observed that the purified EEG signal contains more original information although the artifact has cancelled out. It is obvious that using SSA, EMD and DWT filter for artifact

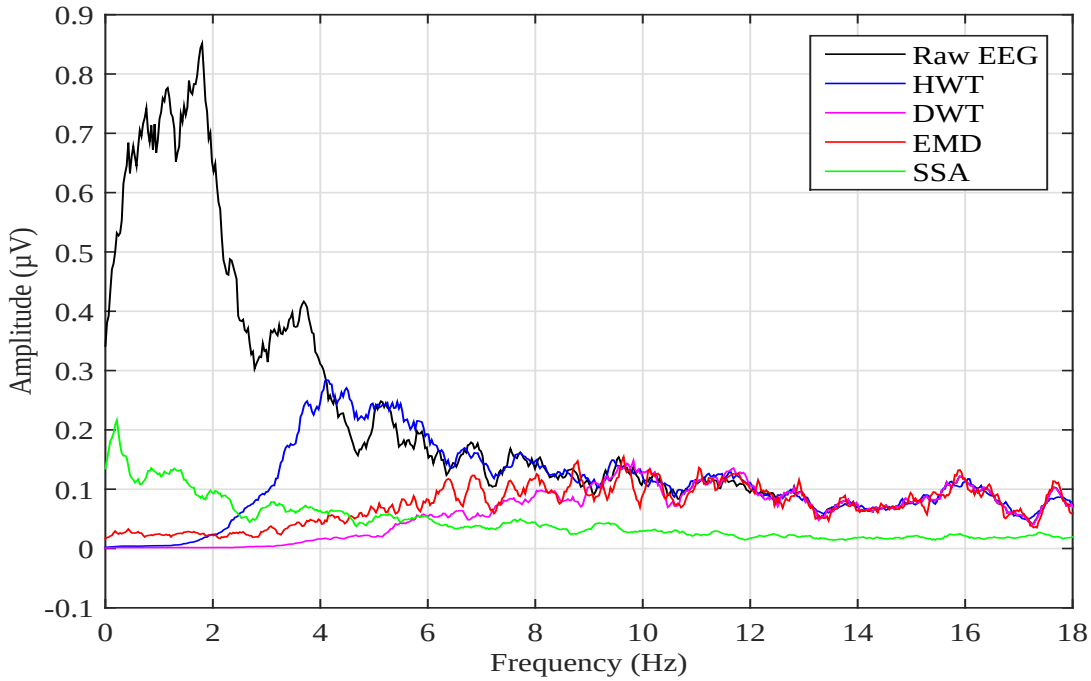

Figure 6.9: The graphical representation of raw EEG and separated EEG spectra of raw EEG contained different four methods. The black, blue, magenta, red and green solid line spectra represent raw EEG, clean EEG spectrum using HWT, DWT, EMD and SSA respectively. The artifact reduction EMD and DWT methods omitted the delta (<4Hz) and part of theta (4∼8Hz) bands but retained the other EEG frequency bands. The SSA method cut the original EEG frequency and shows the delta rhythm. The proposed HWT technique is well preserved the EEG frequency bands except part of the delta rhythm.

reduction, underlying EEG or low frequency cerebral data may be lost. In order to lessen the data loss HWT method is used. From above figure, it is experimental that the HWT based method is best for reduce the EOG from raw EEG without cutting the information and it help to get clean EEG. It is important to stress that the proposed method separates broad frequency content of neurophysiological signals from very strong EOG interference. In this aspect, the research results are presented here on the behavior of the HWT. It is observed graphically that HWT provides better result in comparison with SSA, DWT and EMD. So, it is an efficient technique for improving the quality of EEG signal in biomedical application.

The performance of the proposed artifact removing method is evaluated by classifying the imagined movement EEG data. The proposed approach for classification is applied to the publicly available datasets of BCI competition IV which

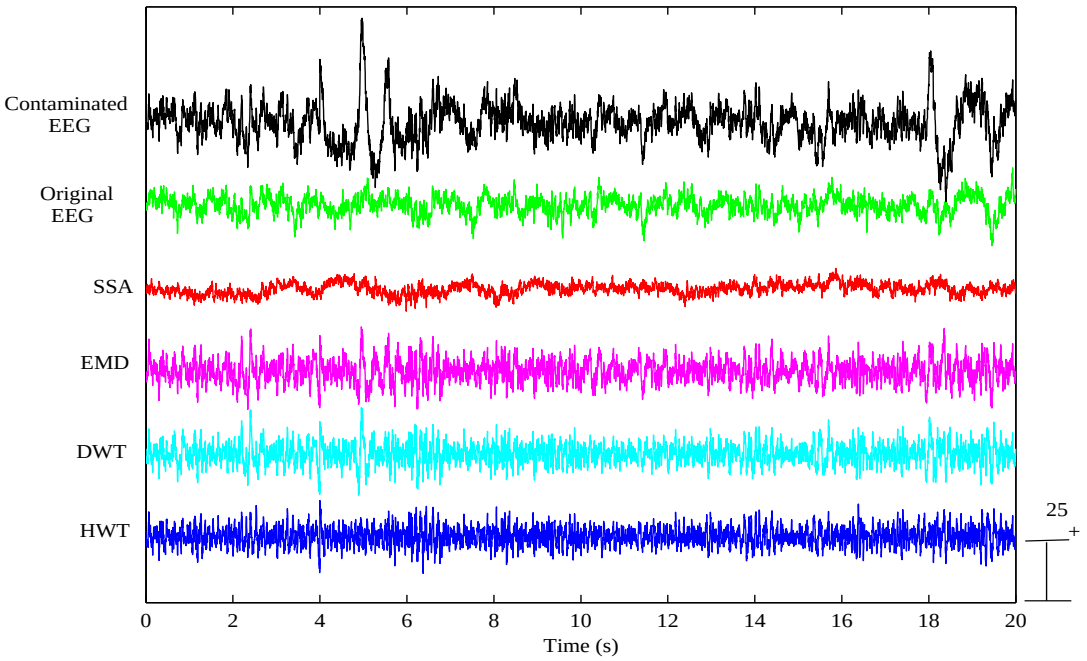

Figure 6.10: Visual comparison of the simulation results corresponding to the contaminated EEG signals (i.e. mixed-up of EEG and EOG signal), the EEG source signal, and the corrected EEG signals after applying proposed HWT and different ocular artifact removal methods.

is illustrated in Figure 5.4. These datasets contain MI EEG signals. The motor imagery classification experiment has described in previous chapter. The classification result express in percentage accuracy. The performances of artifact suppression are measured by performance metrics, computational cost and classification accuracy which will be describe in section 6.3.1. On the basis of the classification table and computational cost, it is proved that the proposed hybrid method is perfect for real time implementation.

6.2.1 Performance Analysis of Artifact Suppression

It is always challenging to evaluate the performance of artifact elimination methods found in the literature. The artifact removing can be done either by visually by specialized which is subjective but the process is not standard or by artificial or semisynthetic data is reconstructed whether perfectly accurate or not. Since there is neither any ground truth data available nor any universal or standard quantitative metric(s) used in the literature that can capture both amount of artifact removal and distortion. So, it is quite difficult to compare some artifact removal methods based on their ability to remove artifacts. Performance assess-

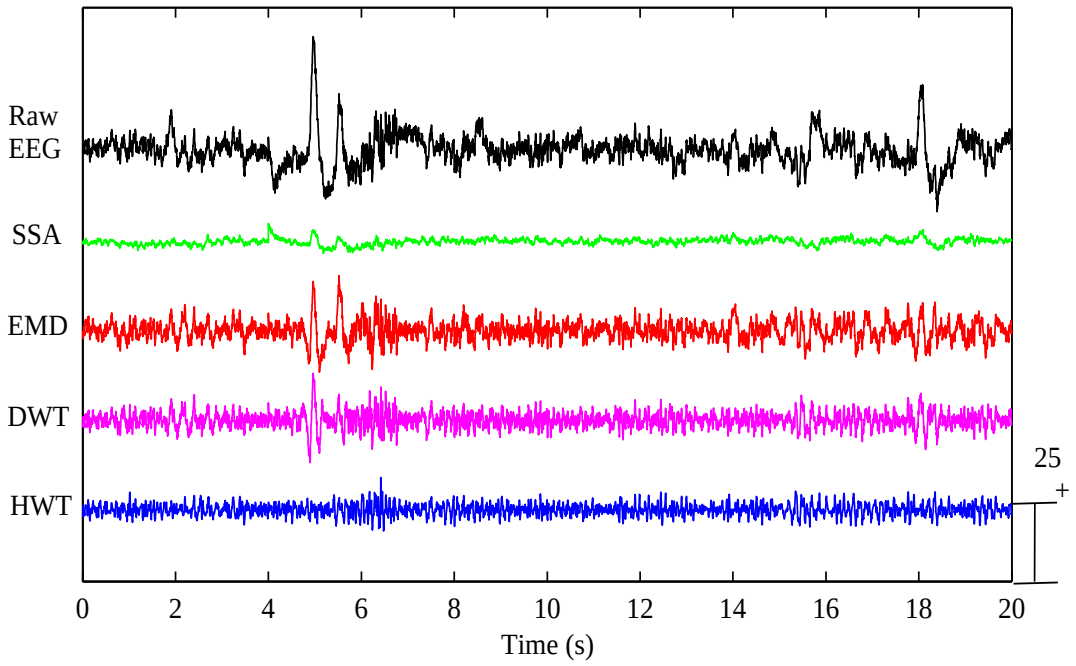

Figure 6.11: Time domain representation of raw and pure EEG signal separated by SSA, EMD, DWT and proposed HWT technique.

ment is a related part of EEG signal processing that is required before a method can be used reliably in a clinical condition. The major challenge for evaluating the performance of the artifact removing algorithm is that the noiseless signal is not known as a priori. Regardless of the fact that simulated data can be used to improve and assess artifact removal algorithms to obtain earliest indications after that the methods need to be evaluated with real data. Therefore, it is necessary to develop tools that allow researches in the field to accurately measure and compare the performance of new and present algorithms to select the optimal one for a certain scenario [206].

For evaluating the performance of a classifier and validate the performance of the proposed method an appropriate criterion is an important concern to make correct predictions. In this paper, the performance of the proposed method is assessed by signal to artifact ratio, mutual information and calculating classification accuracy and evaluating performance of the classifier in terms of training, testing and validating performances.

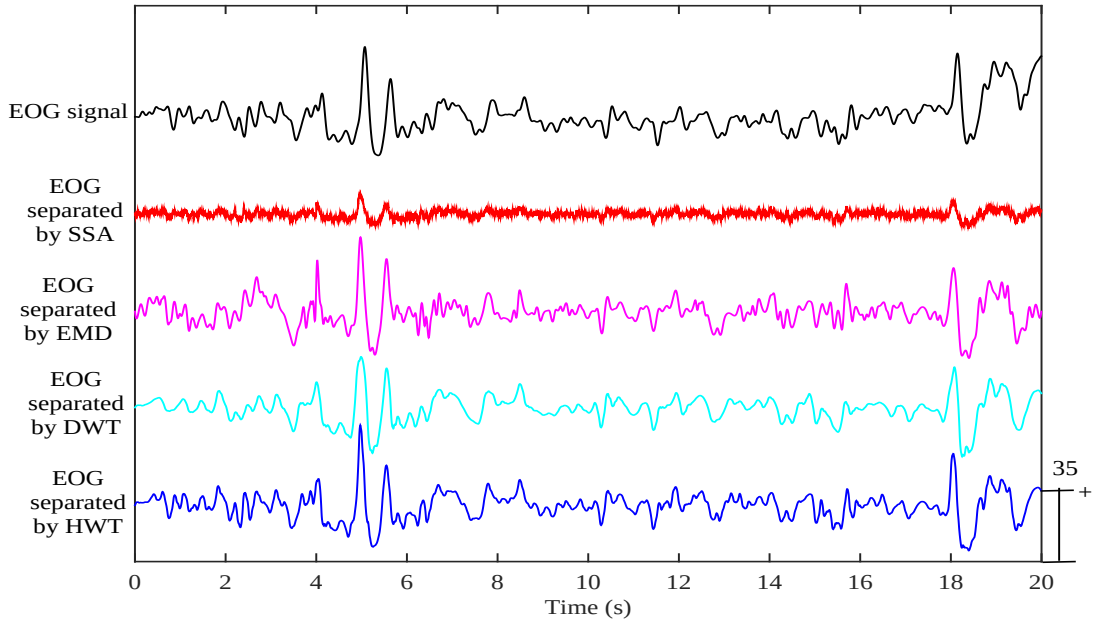

Figure 6.12: Illustration of the original EOG which contaminated the EEG and compared it with other EOG separated by proposed HWT and different three methods.

Performance Metrics for Artifact Removing

The parameters, factors or measures of quantitative assessment used for measurement, contrast or to track presentation or invention are termed as metrics. Experts use metrics to compare the performance of different methods. It is very difficult to estimate the performance of artifact deduction algorithms because a worthy estimate of the clean EEG activity is usually unavailable. As a result, several studies do not quantify the performance of their proposed approach. As an alternative, they use qualitative graphical comparison, i.e., unclean EEG signals and the corrected or denoised EEG signals are plotted and qualitatively compared [201]. Unluckily, such qualitative measures are subjective. A number of scholars therefore have attempted to quantify the performance by using criteria such as the ratio between the functions of spectral density and the corrected and the raw EEG signals [172] and expert scoring [173].

Artifact subspace reconstruction (ASR) is a metric that measure the performance of algorithm has been compared with the most popular artifact removal method. ASR is the most common and widespread online method of eliminating transient, high-amplitude-related artifacts from different sources like eye blinks,

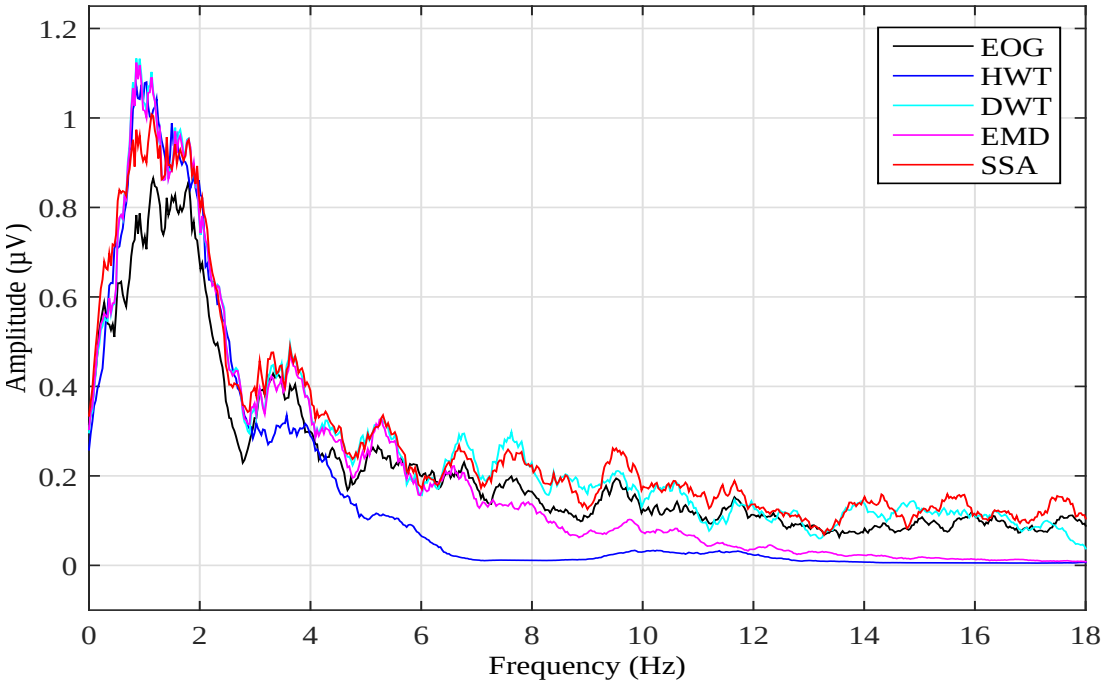

Figure 6.13: Illustration of the EOG spectra which contaminated the EEG and compared it with other EOG separated by existing three artifact suppression methods and proposed hybrid method.

bursts of muscle, and movement while retrieving essential EEG background activities that lie in the subspace spanned. The method is presented with EEGLAB, and comparison has been performed with a threshold of three standard deviations without rejecting any channel [110].

To evaluate the performance of an artifact removal algorithm another approach uses simulated EEG data. In this case, contaminated EEG signal is produced using the artifacts which are manually added to clean EEG signals and the artifact removal algorithm is then applied to the simulated signals. The clean EEG signals should be known with this process. Therefore, assessment criteria such as a correlation coefficient [28], and errors in time or frequency domains can be used to evaluate the performance. Different statistical performance metrics have been used to objectively compare various combinations of ocular artifact (OA) removal in time and frequency domain. For time domain comparison, correlation coefficient, mutual information, signal to artifact ratio and normalized mean square error have been evaluated. For frequency domain comparison, time frequency analysis has been utilized. Based on this rationale, we generated contaminated EEG signals and investigated the performance of the different artifact removing

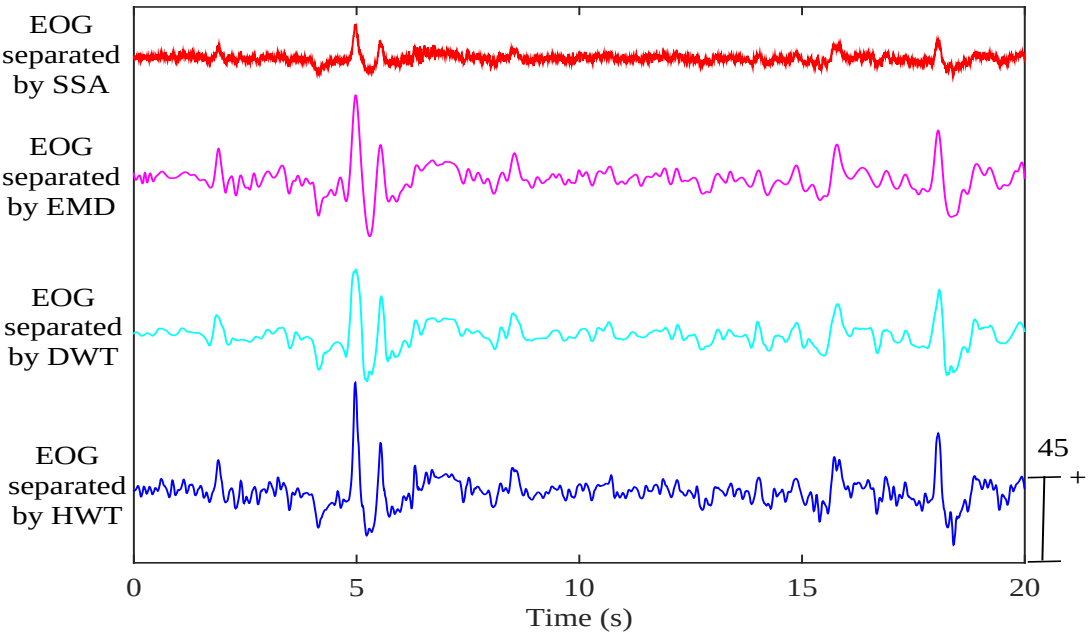

Figure 6.14: Time domain representation of EOG separated by three different methods like SSA, EMD, DWT and proposed HWT technique.

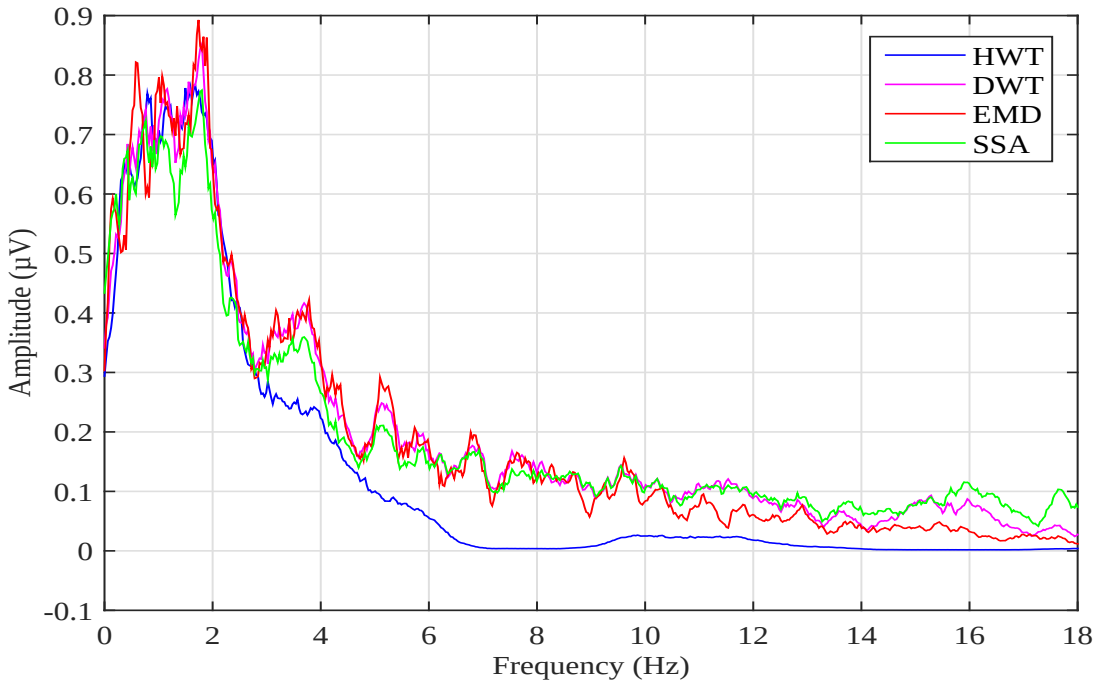

Figure 6.15: Comparison of separated EOG spectra of raw EEG for the proposed HWT and other three existing methods.

algorithms.

Mean Square Error (MSE): It is used to describe similarity between the original signal and de-noised signal [18].

$$MSE = \frac{1}{N}\sum_{i=1}^{N}[s_i(t) - \hat{s}_i(t)]^2 \tag{6.2}$$

Here, $s_i(t)$ stands for contaminated EEG signal, $\hat{s}_i(t)$ is artifact free EEG signal and N for the signal length or the number of samples. And the lower is the MSE value,the better is the method for the artifact removing.

Normalized Mean Square Error(NMSE): Normalized mean squared error (NMSE) is an information processing system of the overall deviations between expected and measured values. It is defined as [97]:

$$NMSE = \frac{1}{N}\sum_{i=1}^{N}\frac{[s_i(t) - \hat{s}_i(t)]^2}{\overline{s_i(t)}\ \overline{\hat{s}_i(t)}} \tag{6.3}$$

Root-Mean-Square Error (RMSE): For evaluation of de-noising algorithm we used Root-Mean-Square Error (RMSE), which is good statistical index for case, when the original clear signal is known. RMSE is defined as [106]:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1}[s_i(t) - \hat{s}_i(t)]^2} \tag{6.4}$$

where $s_i(t)$ stands for contaminated EEG signal, $\hat{s}_i(t)$ is artifact free EEG signal and N for the signal length or the number of samples in both signals. The values of RMSE close to 0 mean that filtering technique reduces noise efficiently with minimal changes in useful signal. RMSE represents the difference between two signals. The smaller the difference is between two signals, the smaller the RMSE value becomes.

The Relative Root Mean-Squared Error (RRMSE): The RRMSE was used as the first evaluation measure to assess the effect of muscle artifact removal, which is defined as follows [21]:

$$RRMSE = \frac{RMS[s(t) - \hat{s}(t)]}{RMS[s(t)]} \tag{6.5}$$

with $\hat{s}(t)$ representing the reconstructed EEG signal after muscle artifact removal.

Correlation Coefficient (CC): In order to provide a quantitative measure of performance for the proposed artifact removal method, the correlation coefficient

(CC) between the extracted eye blink (EB) artifact source and the original EEGs and the artifact removed EEGs are computed [144]. A quantitative measure of some types of correlation and dependence, meaning statistical relationships between two or more observed data values are illustrated by correlation coefficient. It is a measure of the linear correlation between two signals $x(t)$ and $\hat{s}(t)$, giving a value between $+1$ and -1 inclusive, where 1 is total positive correlation, 0 is no correlation, and -1 is total negative correlation. Values of $+1$ and -1 indicate perfect increasing and decreasing linear fits, respectively. It is widely used in the sciences as a measure of the degree of linear dependence between two signals. The correlation coefficient (CC) between the original EEG in each channel and its reconstructed counterpart after removing muscle artifact was calculated in [21, 60, 144]. The correlation criterion which uses second-order statistics, compares the linear relationship between the recorded EEG signals and the reconstructed signal after removing ocular artifact. The correlation coefficient computed as follows:

$$CC = \frac{cov[x(t), \hat{s}(t)]}{\sigma_{x(t)}\sigma_{\hat{s}(t)}} = \frac{\sigma_{x(t)\hat{s}(t)}}{\sqrt{\sigma_{x(t)x(t)}\sigma_{\hat{s}(t)\hat{s}(t)}}} \tag{6.6}$$

where, $\sigma_{x(t)\hat{s}(t)} = cov[x(t), \hat{s}(t)]$ is a covariance matrix of the two signals and $\sigma_{x(t)x(t)} = \sigma^2_{x(t)}$ is the standard deviation.

Time-Frequency Analysis (TFA): Time and frequency components can be analyzed simultaneously using the wavelet decomposition tool of EEGLAB toolbox (Matlab, CA, US). This allows qualitative comparison of the signals before and after artifact denoising [106].

Power Spectral Density (PSD): PSD is used to evaluate the frequency content of a signal. In this paper [14], PSD of the EEG signals are computed using a nonpara-metric method known as the Welch method. The method is based on averaging a periodogram spectrum data over overlapping data segments.

Mean Absolute Error (MAE): Mean absolute error is defined to measure the distortion across different frequency bands, delta (0.5-4Hz), theta (4-8Hz), alpha (8-12Hz), beta (12-30Hz), and gamma (30-40Hz).

$$MAE = |P_{inEEG} - P_{outEEG}| \tag{6.7}$$

where $EEGout$ is the reconstructed EEG from the proposed method and $EEGin$ is the artifact-free EEG. An effective method is supposed to remove all artifacts

so that the output EEG is as close as possible to the pure EEG; and P denotes the power spectrum density (PSD). PSD is estimated using Welch method with different parameters [127] 200 sample points as the length of window and 5 sample points overlap. The average PSD for each frequency band was calculated for all subjects.

Most of the published articles evaluated their method in terms of some qualitative plots. In addition, very few of them quantified the distortion to desired EEG signals due to the removal effect. Therefore, it's not fair to tell which performs best based on the study [86]. The performance of the proposed artifact removal algorithm has been evaluated both in terms of the quantity of artifact removal and the amount of distortion it brings into the signal of interest. Several efficiency metrics have been calculated in both time and spectral domain to quantify such evaluation. In order to have fair evaluation and clear idea we also have considered the amount and duration of artifacts present in the signals. In order to determine whether the method is successful at removing ocular artifact (OA) from EEG, the performance is assessed using three performance criteria i.e. signal to artifact ratio (SAR), mutual information (MuI) and classification accuracy. The performance metrics are described in details below:

Signal to Artifact Ratio (SAR): The metrics commonly employed to represent the energy of the signal compared to the energy of the artifact is the signal to artifact ratio (SAR) [166, 184].

$$\begin{aligned} SAR_{dB_{(\text{contaminated EEG})}} &= 20\log10\frac{\text{rms}[x(t)]}{\text{rms}[s(t)-\hat{s}(t)]} \\ SAR_{dB_{(realEEG)}} &= \frac{1}{N}\sum_{i=1}^{N}20\log10\frac{rms[s(t)]}{rms[s(t)-\hat{s}(t)]} \end{aligned} \tag{6.8}$$

Here, $x(t)$ is the clean EEG signal. $s(t)$ stands for contaminated EEG signal, $\hat{s}(t)$ is artifact free EEG signal and N for the signal length or the number of samples.

Raw EEG signals suffer from low signal to artifact ratio. To get better artifact free EEG signal, the SAR value must be higher. The SAR is used to compare the ratio after denoising to the original ratio of the artifactual EEG signal. The proposed HWT algorithm is compared with other eye blink artifact removal algorithms.

The better the de-noising, the greater the signal to artifact ratio and mutual information. Artifact getting rid of effect of different methods are shown in Table 6.1 and Table 6.2. As known from Table 6.1, the signal to artifact ratio after hybrid wavelet decomposition is maximum for artificially corrupted EEG and become better artifact removing method. A comparison of SSA, EMD, DWT and HWT analysis for artificially corrupted EEG is summarized (using$SAR_{dB_{(contaminatedEEG)}}$formula) in Table 6.1

Table 6.2 also shows the comparison results of the average SAR values (using $SAR_{dB_{(realEEG)}}$ formula)from SSA, DWT, EMD and HWT methods, respectively for raw EEG. It is seen that the HWT based technique yields the higher SAR (for a single channel) result than SSA, DWT and EMD. The results demonstrate that the performance of wavelet based hybrid denoising method is very effective in removing ocular artifact.

Mutual Information (MuI): It is used statistically to measure how much information one random variable contains about the other random variable. Here, mutual information [229], is calculated (a nonparametric measure of relevance between the two random variables) to find how much information are artifact free EEG signal. In this paper, MuI index is adopted for the quantification of the mutual dependence of the pure EEG signals and the artifact-rejected EEG datasets using the next formula:

$$MuI(\text{inEEG,outEEG}) = \sum_{a \varepsilon inEEG} \sum_{b \varepsilon outEEG} j(a,b) * \log \frac{j(a,b)}{j_1(a) j_2(b)} \tag{6.9}$$

where, $j(a,b)$ is the joint probability distribution function (pdf), and j_1 (a) and j_2 (b) are the marginal probability distribution functions of inEEG and outEEG respectively. In this paper, $s(t)$ stands for contaminated EEG signal≈ inEEG and $\hat{s}(t)$ is artifact free EEG signal ≈ out EEG.

If the mutual information between artifact-free EEG and output EEG from proposed method is large, it means they are closely related and the method is better for artifact removing. Moreover, the mean and standard deviation of MuI (contaminated EEG) for four methods are reported in Table 6.1. To evaluate the similarity between the raw EEG signal and corrected signal quantitatively within all the artifact free intervals of the recording, the SAR and MuI are calculated

between them. Here, the criterion to select an artifact free segment is that no samples of the EOG exceeded 12μV. The mean and standard deviation of SAR and MuI (raw EEG) for four methods are listed in Table 6.2.

Using the above equation, the experimental results point out that the SAR of both measured and real EEG signals could be obviously improved in different conditions and the mutual information (MuI) on averaged of synthetic and real world. By comparing with the other existing artifact removal method it is demonstrated that the proposed HWT method improved the SAR and MuI both significantly better than the SSA, DWT and EMD-based method.

Table 6.1: Comparison of various artifact removing methods and proposed HWT method for artificially contaminated EEG data.

No.	Methods	SAR in dB	MuI
1.	SSA [229]	-6.0228	0.1513 ± 0.0846
2.	DWT [138]	-5.3210	0.5367 ± 0.2435
3.	EMD [138]	-5.1680	0.6722 ± 0.3536
4.	HWT	-4.6864	0.6996 ± 0.3542

Table 6.2: A comparative summary of stationary subspace analysis, discrete wavelet, empirical mode and proposed hybrid wavelet transform for real EEG data.

No.	Methods	SAR in dB	MuI
1.	SSA [229]	0.4101 ± 0.8738	0.2438 ± 0.2576
2.	DWT [138]	1.6983 ± 0.6190	0.7303 ± 0.2784
3.	EMD [138]	2.2601 ± 1.1498	0.7967 ± 0.2010
4.	HWT	3.0889 ± 1.7041	0.8562 ± 0.2273

For real EEG, the average respective values of SAR and MuI for all the methods are enumerated in Table 6.2, and demonstrate the improved performance of the proposed method over conventional methods (SSA, DWT, EMD). The table shows that the HWT based technique yields the best SAR result and MuI value compared to other methods. Based on the performance metrics values, it is observed that the hybrid wavelet transform filter with db4 mother wavelet is able to filter out more artifact compare to SSA, DWT and EMD.

6.3 Experimental Results for Motor Imagery (MI) Classification

Motor imagery is the interpretation of the person motor intention into control signals through motor imagery conditions. For example, EEG signals accompany with in the μ and β rhythms, ($\mu : 8-12$ Hz) and ($\beta : 18-26$ Hz) decrease in specific motor cortex area when the left hand movement may generate this signal. Different applications such as controlling a mouse or playing a computer game could be used according to the motor imagery rhythms. The subject might not need any training with new artificial intelligence techniques. However, in this technique it is always better to have some training before using the motor imagery systems [160]. For the performance measurement of the proposed artifact removing method, the Motor Imagery (MI) classification method is used.

6.3.1 Performance Metrics for EEG Signal Classification

To evaluate BCI performance, one must bear in mind that different components of the BCI loop are at stake [200]. Regarding the classifier alone, the most basic performance measure is the classification accuracy. This is valid only if the classes are balanced [49], i.e. with the same number of samples per class and if the classifier is unbiased, i.e. it has the same performance for each class [174]. If these conditions are not met, the Kappa metric or the confusion matrix are more informative performance measures [49]. The sensitivity-specificity pair, or precision, can be computed from the confusion matrix. When the classification depends on a continuous parameter (e.g. a threshold), the receiver operating characteristic (ROC) curve, and the area under the curve (AUC) are often used.

Classifier performance is generally computed offline on pre-recorded data, using a hold-out strategy: some datasets are set aside to be used for the evaluation, and are not part of the training dataset. However, some authors also report cross validation measures estimated on training data, which may over-rate the performance.

The contribution of classifier performance to overall BCI performance strongly

depends on the orchestration of the BCI subcomponents. This orchestration is highly variable given the variety of BCI systems (co-adaptive, hybrid, passive, self- or system- paced). For evaluating the online performance of the BCI system, there are several existing performance metrics for EEG signal classification. Classifier performance was measured for accuracy, sensitivity, specificity, precision, and the Kappa statistic [4], each defined as follows:

$$Kappa(k) = \frac{(P_o - P_e^c)}{(1 - P_e^c)} \tag{6.10}$$

where P_o characterizes the possibility of overall agreement between label assignments, classifier and true process; and P_e^c indicates the chance agreement for total labels; i.e., the sum of the proportion of cases allotted to class multiplies in proportion to true labels of that specific class in the data set. The kappa measure ranges between 1 and -1, where 1 indicates the best correct classification and -1 indicates the worst.

In the paper [43], considering receiver operating characteristic (ROC) parameters such as true positive (TP), true negative (TN), false positive (FP), and false negative (FN), the performance is compared. True positive (TP) means that EEG segment from autistic subject is correctly analyzed as autistic class. ROC graph indicates the trustworthiness of the classifier. The classification performance is evaluated in terms of sensitivity, specificity, and overall accuracy as in the next equations:

$$\text{Accuracy} = \frac{TN + TP}{TN + FP + TP + FN} \times 100\% \tag{6.11}$$

Sensitivity and specificity [187] were used as a performance measure for two classes. In order to analyze the output data obtained from the application, sensitivity (true positive ratio) and specificity (true negative ratio) are calculated by using confusion matrix. The sensitivity value (true positive, same positive result as the diagnosis of expert neurologists) was calculated by dividing the total of diagnosis numbers to total diagnosis numbers that are stated by the expert neurologists. Sensitivity, also called the true positive ratio, is calculated by the formula:

$$Sensitivity, TPR = \frac{TP}{TP + FN} \times 100\% \tag{6.12}$$

On the other hand, specificity value (true negative, same diagnosis as the

expert neurologists) is calculated by dividing the total of diagnosis numbers to total diagnosis numbers that are stated by the expert neurologists. Specificity, also called the true negative ratio, is calculated by the formula:

$$Specificity, TNR = \frac{TN}{TN + FP} \times 100\% \tag{6.13}$$

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \times 100\% \tag{6.14}$$

A common metric that can be used to compare different tests is the area under the ROC curve (AUC). It is used to measure the test accuracy. ROC curve explains two dimensional visualization of ROC curve set for classifiers performance. The easiest probability is to compute the area under the ROC curve which is part of the area of the unit square. Subsequently the value of AUC will always satisfy the following differences:

$$0 \leq AUC \leq 1 \tag{6.15}$$

It is clear that if the AUC is close to 1 (area of unit square) AUC indicates very good test.

The Information transfer rate (ITR) is an important factor to measure the performance of a BCI. The classification accuracy, BCI performance was also evaluated by ITR [219]. The ITR is a standard measure of communication systems, which indicates the amount of information communicated per unit of time. The ITR can be expressed in the form below:

$$ITR = T\left\{log_2 k + Alog_2 A + (1-A)log_2\left(\frac{1-A}{K-1}\right)\right\} \tag{6.16}$$

where, K is the total number of commands, A is the classification accuracy, and T (seconds/selection) is the average time for a selection. Depending on individual target, classification performances were calculated with different T [19]. Some of the studies showed a high ITR by increasing the number of targets and improving the accuracy of target selection [36].

Time Complexity: An important attribute of an algorithm is its computational complexity, defined as the number of floating point operations required to execute the algorithm (flops). In practice, knowing whether an algorithm is computationally efficient is as important as knowing its performance for certain

(online) applications. Typically, there exists a trade-off in between speed and accuracy, thus if two approaches offer similar results in terms of the quality of the denoised signal then the faster method should be preferred.

The execution speed of an algorithm is directly related to its implementation,

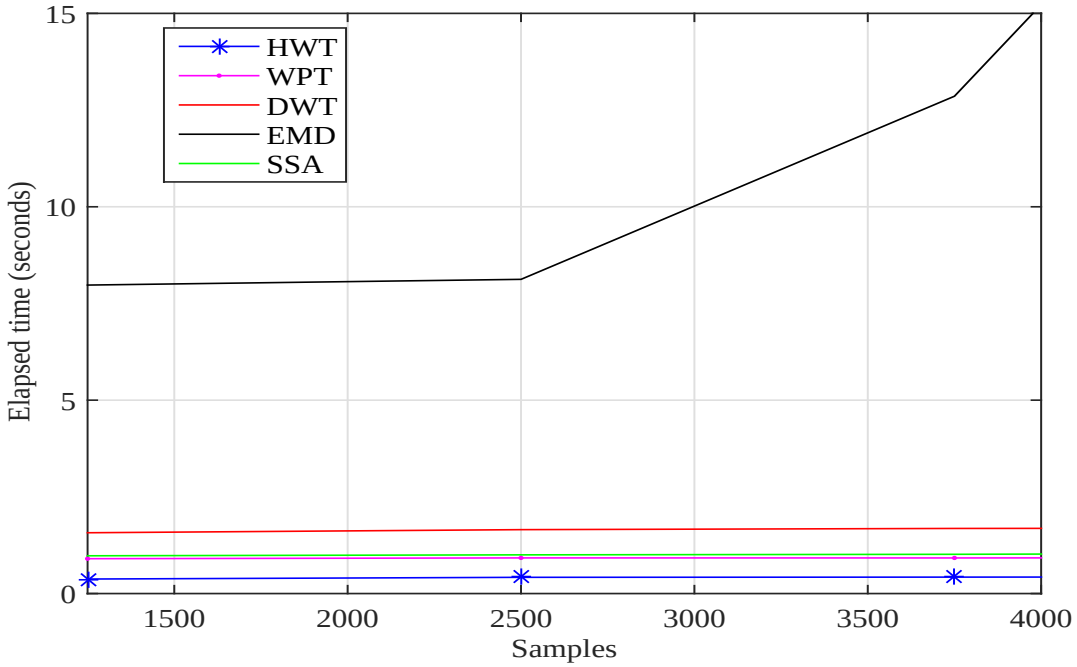

Figure 6.16: Comparison of computational time for the five methods. Time required by proposed HWT, WPT, DWT, EMD and SSA to remove the artifact. The required time is increased with increasing the number of samples for EMD. But the required time of proposed HWT technique is comparatively low and remaining unchanged with increasing the number of samples than other four methods like SSA, DWT, WPT and EMD.

not only it depends on the theoretical computational cost. The problem is that most authors do not report the computational cost of their algorithms, neither theoretically nor in practice. What is more, speed is of prior concern for online applications (for instance patient monitoring in real time), whilst we want to determine what algorithm provides denoised EEG of best quality. So, the computational time is a very important issue in an online BCI system. It is also a metric to measure the computational effectiveness of the method [139]. To quantify the computational complexity of the methods discussed in this study, the processing time under different methods is given in Figure 6.16. The relative artifact removing time required by proposed HWT and other methods as a function of the number of samples is illustrated. The required time is remaining unchanged with

increasing the number of samples for four methods like SSA, DWT, WPT and proposed HWT. Also, it clearly shows that the computational complexity of EMD is very much higher than that of other methods. Because it depends on number of samples. Hence, EMD is not comparable with HWT for online implementation of EOG artifact cleaning. The HWT is required less computational time than DWT and WPT. Although DWT and WPT can be used for online implementation to clean EEG. But it is expected that the proposed approach is quite faster compared with the other wavelet approaches and reduced computational cost discussed in this study. All algorithms are implemented in MATLAB (release R2013a) and executed on the same computer with Windows 7 Ultimate (Intel(R) Core(TM) $i5-4590$ CPU, 3.30 GHz processor, 8 GB RAM).

Classification Accuracy: A problem of todays brain computer interface (BCI) systems is that the performance in controlling a BCI can decreased rapidly over time. This is due to the non-stationarity of recorded EEG signals. Furthermore, the motivation of the subject can drop if the subject does not experience any success in controlling a BCI. For improving the classification performance of motor imagery (MI) signals, technologically advanced numerous techniques are used in Brain Computer Interfacing (BCI). Still there is scope for improvement of performance using various techniques. The classifiers performances were computed using the most commonly used parameters, is the accuracy. The accuracy of a classifier is the percentage of the test set which is classified by the classifier. This parameter is defined as follows [4].

$$\text{Accuracy} = \frac{\text{Total no. of correctly classified instances}}{\text{Total no.of instances}} \times 100$$

The classification accuracy is higher for better motor imagery(MI) classification. Comparison of classification accuracy in different two (motor imagery left hand and right hand movement) tasks for nine subjects is presented in Table 6.3. In this Table, the performance of proposed method and other artifact suppression methods are comparing for real EEG data using LDA classifier.

Comparison with Existing Methods: Comparison has been made first using simulated signals and then with real experimentally acquired EEG data corrupted by artifacts. For the real EEG data the classification accuracy is compared before cleaning the artifacts and after cleaning the artifacts using different artifact

Table 6.3: Comparison of classification accuracies (in %) of the proposed method with other methods using real EEG data.

Subjects	Raw EEG	After bandpass filtering (0.5 ∼ 30Hz)	After cleaning with		
			EMD	**DWT**	**HWT**
1.	77.08	65.97	55.55	83.33	89.58
2.	52.08	53.47	54.16	51.38	55.55
3.	89.58	89.58	75.69	95.13	96.52
4.	56.25	52.77	49.30	61.11	62.5
5.	50	51.38	41.66	49.30	64.58
6.	56.94	56.25	45.13	65.97	65.97
7.	59.72	69.44	48.61	72.91	79.16
8.	95.13	95.83	90.27	93.75	96.53
9.	93.05	92.36	88.19	93.05	93.05
Average	**69.9**	69.67	60.95	73.99	**78.16**

suppression methods like EMD, DWT and Butterworth bandpass filter. Table 6.3, gives a remarkable description of average classification accuracy in different tasks for each method. In this table, the first column shows the number of subjects participating in this experiment. The second column shows nine subjects average classification accuracy without artifact removing method, only used raw EEG using CSP and LDA methods. The third column of the table shows the MI classification accuracy using Butterworth band pass filter as an artifact removing method. The fourth column of the table shows the MI classification accuracy using other existing artifact suppression method and the proposed HWT method. From this table it is obvious that the classification accuracy of the proposed hybrid wavelet transform method is higher than other artifact removing methods. So, from above performance parameters it is easily understood that proposed technique is better than other artifact removing methods.

The results of improved classification responses with and without artifact removing experiment are summarized using CSP feature extraction method and LDA classifier. The LDA based classification is more noticeable in HWT artifact removing technique due to the proposed EEG enhancement. It must be noted

that subjects 1, 5 and 7 showed the remarkable improvement by more than 12%. The results of this study shall lead directly to a new improved MI classification in BCI paradigms after artifact removing. The EMD based MI cleaning approach and the DWT are also implemented to compare its classification performance with HWT based method. It is noted that the classification accuracy of EMD is lower than that of HWT based algorithm as illustrated in Table 6.3. We have demonstrated that the use of HWT based artifact cleaning approach enhances the LDA based classification accuracy results. Classification accuracy varies with respect to subject. The highest classification performance was found in the proposed HWT. From these results, it seems that the accuracy of the HWT method is found to be above 78%. Whereas, in the paper [117], the classification accuracy of three dataset is 75.5%. It is obvious that the classification accuracy of the proposed hybrid wavelet transform method is higher than other artifact removing methods.

The optimal performance of a BCI depends on the effective removal of ocular artifacts from EEG recordings. In this paper, we proposed a novel method for automatic identification and reduction of ocular artifacts from EEG signals by combined DWT, WPT and thresholding. The performance and effectiveness of the proposed frame work was illustrated using simulated and real EEG datasets. Results show that the proposed method can effectively remove ocular artifacts as well as it can preserve the neuronal activity related EEG signals. The results demonstrate, additionally, that the proposed method outperforms the conventional techniques.

Chapter 7

Conclusions and Future Works

This chapter summarizes the overall contributions of this thesis and highlights their significances in neural information processing system. In addition, we also present the future prospects of this work.

7.1 Contributions

A novel method based on wavelet transform and wavelet packet is introduced as a data adaptive time domain filtering approach to separate the EOG artifact from the recorded EEG signals. The performance of the algorithm is tested using both contaminated and real signals. Pure EEG signal is recently using a much for developing brain computer interfacing (BCI), whereas the contaminated EOG signal would create some problems in such applications. The proposed data adaptive hybrid wavelet transform (HWT) method efficiently separates the EOG artifact without changing the amplitude and other necessary properties of the EEG signals. The HWT is presented to reduce the subband space compared to WPT. In blind source separation algorithm like SSA is shown to be effective to remove ocular artifact, but it rely on multiple channel data and can also result in the elimination of neural activities. A thresholding based denoising method using wavelet transform is implemented to separate the EOG signal. The empirical mode decomposition is also a so called data adaptive technique. But it takes high computational time. The computational time of the DWT and HWT algorithm is also very low in comparison to the EMD algorithm. Hence, EMD is not comparable with DWT and HWT for online implementation of EEG cleaning. Based on SAR and MuI performance metrics, it is practical that the SSA, EMD and DWT method cut the original information with EOG which is absent in HWT.

The proposed HWT method separates electro-oculogram while keeping the scale of electroencephalography amplitude undistorted. The method reduced EOG artifacts but did not eliminate the original information. So, hybrid wavelet transform is better in removing the EOG artifact compare to other three methods. Thus this work is expected to be useful for neural signal processing and analysis community in the long run and will provide platform to further improve the understanding of our brain for different applications, both in clinical and non-clinical applications.

In summary, we have demonstrated that the proposed artifact removal method hybrid wavelet transform with an adaptive thresholding mechanism and MI EEG classification technique improves the classification accuracy. The proposed method outperforms other artifact handling methods and provides the following advantages:

- It allows real-time processing;
- It does not require additional EOG/EMG channels to detect and remove artifacts;
- It allows adaption to the characteristics of a given signal, resulting in minimal distortion in EEG signals even in the case of false artifact detection;
- It is fully automated.

7.2 Future Works

In this work, a new data-driven HWT method for artifact removal in single-channel EEG is presented. However, the proposed method shows improved performance for removing ocular activities from EEG signals as compared to the conventional techniques. For the purposes of demonstration, we have so far considered only low frequency eye blink artifacts, as this is the most common and troublesome artifactual component encountered in EEG signals. The most salient features of eye blink artifacts include the variance, kurtosis, Shannons entropy and range of amplitude. These features were identified and selected based on their magnitudes being significantly higher in conjunction with an eye blink than in its absence. However,

the list of artifactual components other than the eye blink artifacts and their corresponding descriptive features is not considered in this study. We anticipate that future research should serve to identify the most descriptive features of other artifactual components, and simultaneously remove the artifacts without substantial loss of cerebral signals of interest. Our future research directions include hardware implementation and optimization of effective OA removal technique for a single channel EEG system, real-time OA removal, feature extraction, and cognitive load classification to monitor brain engagement in natural environment within a wearable embedded system. Therefore, our future research will focus on additional methods that automatically select the optimal wavelet function for the proposed algorithm. It is also of interest to extend the proposed algorithm to a multivariate version and find out if and how it can improve the effectiveness of the algorithm in denoising EEG signals. The HWT would be applied for possible removal of multiple artifacts present in any number of multi channel EEG signals. Besides, in the present study only spontaneous EEG is considered rather than event related potentials (ERPs). To this extent, it is our future plan to implement the proposed method for ERP based BCI. Besides, in the present study the proposed method only deals with ocular activities and removal/reduction of other types of artifactual activities is yet to be examined in forthcoming work. One possible extension of the proposed method is to include features that can be used to identify muscle artifacts (e.g., eye movement, chewing, swallowing, clenching, head movement, body movement and so on.) rather than EOG and removing from EEG signal. Another essential extension is to use features based on the correlation with ECG artifact. Extended group of extracted features (such as: frequency band power features and time domain features) can be used in LDA to identify and classify MI activities. Other classifiers such as SVM, NN, MLP, K-NN, binary decision tree and so on can be considered to compare the classification performance with LDA. Future work may involve a comparison with other existing methods in terms of classification accuracy as well as information transfer rates, sensitivity, selectivity, AUC and kappa using large dataset.

Bibliography

[1] U Rajendra Acharya, S Vinitha Sree, G Swapna, Roshan Joy Martis, and Jasjit S Suri. Automated eeg analysis of epilepsy: a review. *Knowledge-Based Systems*, 45:147–165, 2013.

[2] Muhammad Tahir Akhtar, Wataru Mitsuhashi, and Christopher J James. Employing spatially constrained ica and wavelet denoising, for automatic removal of artifacts from multichannel eeg data. *Signal Processing*, 92(2):401–416, 2012.

[3] Tarik Al-ani and Dalila Trad. Signal processing and classification approaches for brain-computer interface. In *Intelligent and Biosensors*. InTech, 2010.

[4] Hafeez Ullah Amin, Aamir Saeed Malik, Rana Fayyaz Ahmad, Nasreen Badruddin, Nidal Kamel, Muhammad Hussain, and Weng-Tink Chooi. Feature extraction and classification for eeg signals using wavelet transform and machine learning techniques. *Australasian Physical & Engineering Sciences in Medicine*, 38(1):139–149, 2015.

[5] Divya Balakrishnan and Sadasivan Puthusserypady. Multilayer perceptrons for the classification of brain computer interface data. In *Bioengineering Conference, 2005. Proceedings of the IEEE 31st Annual Northeast*, pages 118–119. IEEE, 2005.

[6] Ali Bashashati, Mehrdad Fatourechi, Rabab K Ward, and Gary E Birch. A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals. *Journal of Neural engineering*, 4(2):R32, 2007.

[7] Alexandre Baussard, Fred Nicolier, and Frédéric Truchetet. Rational multiresolution analysis and fast wavelet transform: application to wavelet shrinkage denoising. *Signal Processing*, 84(10):1735–1747, 2004.

[8] Patrick Berg and Michael Scherg. Dipole models of eye movements and

blinks. *Electroencephalography and clinical Neurophysiology*, 79(1):36–44, 1991.

[9] Patrick Berg and Michael Scherg. A multiple source approach to the correction of eye artifacts. *Electroencephalography and clinical neurophysiology*, 90(3):229–241, 1994.

[10] Alexander Bertrand, Vojkan Mihajlovic, Bernard Grundlehner, Chris Van Hoof, and Marc Moonen. Motion artifact reduction in eeg recordings using multi-channel contact impedance measurements. In *Biomedical Circuits and Systems Conference (BioCAS), 2013 IEEE*, pages 258–261. IEEE, 2013.

[11] Benjamin Blankertz, Gabriel Curio, and Klaus-Robert Müller. Classifying single trial eeg: Towards brain computer interfacing. In *Advances in neural information processing systems*, pages 157–164, 2002.

[12] Valentina Bono, Wasifa Jamal, Saptarshi Das, and Koushik Maharatna. Artifact reduction in multichannel pervasive eeg using hybrid wpt-ica and wpt-emd signal decomposition techniques. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, pages 5864–5868. IEEE, 2014.

[13] Nicolas Brodu, Fabien Lotte, and Anatole Lécuyer. Comparative study of band-power extraction techniques for motor imagery classification. In *Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), 2011 IEEE Symposium on*, pages 1–6. IEEE, 2011.

[14] Christiaan Burger and David Jacobus van den Heever. Removal of eog artefacts by combining wavelet neural network and independent component analysis. *Biomedical Signal Processing and Control*, 15:67–79, 2015.

[15] Salvatore Calcagno, Fabio La Foresta, and Mario Versaci. Independent component analysis and discrete wavelet transform for artifact removal in biomedical signal processing. *American Journal of Applied Sciences*, 11(1):57, 2014.

[16] B Carter. Op amps for everyone, texas instruments. *ISBN*, 80949487:10–11, 2009.

[17] Nazareth P Castellanos and Valeri A Makarov. Recovering eeg brain signals:

artifact suppression with wavelet enhanced independent component analysis. *Journal of neuroscience methods*, 158(2):300–312, 2006.

[18] Wu Chao, Wang Juncheng, Zheng Yi, Wang Zhongqiu, and Ni Xiuhui. Application and simulation of wavelet packet transform in underwater acoustic signal de-noising.

[19] Xiaogang Chen, Zhikai Chen, Shangkai Gao, and Xiaorong Gao. A high-itr ssvep-based bci speller. *Brain-Computer Interfaces*, 1(3-4):181–191, 2014.

[20] Xun Chen, Chen He, and Hu Peng. Removal of muscle artifacts from single-channel eeg based on ensemble empirical mode decomposition and multiset canonical correlation analysis. *Journal of Applied Mathematics*, 2014, 2014.

[21] Xun Chen, Aiping Liu, Joyce Chiang, Z Wang, Martin McKeown, and Rabab Ward. Removing muscle artifacts from eeg data: Multichannel or single-channel techniques. 16:1–1, 01 2015.

[22] Xun Chen, Aiping Liu, Hu Peng, and Rabab K Ward. A preliminary study of muscular artifact cancellation in single-channel eeg. *Sensors*, 14(10):18370–18389, 2014.

[23] Mike X Cohen. *Analyzing neural time series data: theory and practice*. MIT Press, 2014.

[24] Ronald R Coifman and David L Donoho. Translation-invariant de-noising. In *Wavelets and statistics*, pages 125–150. Springer, 1995.

[25] Pierre Comon and Christian Jutten. *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press, 2010.

[26] Marco Congedo, Fabien Lotte, and Anatole Lécuyer. Classification of movement intention by spatially filtered electromagnetic inverse solutions. *Physics in Medicine & Biology*, 51(8):1971, 2006.

[27] A Garcés Correa, E Laciar, HD Patino, and ME Valentinuzzi. Artifact removal from eeg signals using adaptive filters in cascade. In *Journal of Physics: Conference Series*, volume 90, page 012081. IOP Publishing, 2007.

[28] Maite Crespo-Garcia, Mercedes Atienza, and Jose L Cantero. Muscle artifact removal from human sleep eeg by using independent component analysis. *Annals of biomedical engineering*, 36(3):467–475, 2008.

[29] Rodney J Croft and Robert J Barry. Removal of ocular artifact from the

eeg: a review. *Neurophysiologie Clinique/Clinical Neurophysiology*, 30(1):5–19, 2000.

[30] Rodney J Croft, Jody S Chandler, Robert J Barry, Nicholas R Cooper, and Adam R Clarke. Eog correction: a comparison of four methods. *Psychophysiology*, 42(1):16–24, 2005.

[31] Geoff Cumming and Sue Finch. Inference by eye: confidence intervals and how to read pictures of data. *American Psychologist*, 60(2):170, 2005.

[32] Ian Daly, Nicoletta Nicolaou, Slawomir Jaroslaw Nasuto, and Kevin Warwick. Automated artifact removal from the electroencephalogram: a comparative study. *Clinical EEG and neuroscience*, 44(4):291–306, 2013.

[33] Ian Daly, Floriana Pichiorri, Josef Faller, Vera Kaiser, Alex Kreilinger, Reinhold Scherer, and Gernot Müller-Putz. What does clean eeg look like? In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 3963–3966. IEEE, 2012.

[34] Ian Daly, Reinhold Scherer, Martin Billinger, and Gernot Müller-Putz. Force: Fully online and automated artifact removal for brain-computer interfacing. *IEEE transactions on neural systems and rehabilitation engineering*, 23(5):725–736, 2015.

[35] Ingrid Daubechies. A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models. *Wavelets in medicine and biology*, pages 527–546, 1996.

[36] Ingrid Daubechies, Jianfeng Lu, and Hau-Tieng Wu. Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. *Applied and computational harmonic analysis*, 30(2):243–261, 2011.

[37] Wim De Clercq, Anneleen Vergult, Bart Vanrumste, Wim Van Paesschen, and Sabine Van Huffel. Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram. *IEEE transactions on Biomedical Engineering*, 53(12):2583–2587, 2006.

[38] Maarten De Vos, Wouter Deburchgraeve, PJ Cherian, Vladimir Matic, RM Swarte, Paul Govaert, Gerhard Henk Visser, and Sabine Van Huffel. Automated artifact removal as preprocessing refines neonatal seizure detection. *Clinical Neurophysiology*, 122(12):2345–2354, 2011.

[39] Maarten De Vos, Katharina Gandras, and Stefan Debener. Towards a truly mobile auditory brain–computer interface: exploring the p300 to take away. *International journal of psychophysiology*, 91(1):46–53, 2014.

[40] Arnaud Delorme and Scott Makeig. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1):9–21, 2004.

[41] Arnaud Delorme, Jason Palmer, Julie Onton, Robert Oostenveld, and Scott Makeig. Independent eeg sources are dipolar. *PloS one*, 7(2):e30135, 2012.

[42] Arnaud Delorme, Terrence Sejnowski, and Scott Makeig. Enhanced detection of artifacts in eeg data using higher-order statistics and independent component analysis. *Neuroimage*, 34(4):1443–1449, 2007.

[43] Ridha Djemal, Khalil AlSharabi, Sutrisno Ibrahim, and Abdullah Alsuwailem. Eeg-based computer aided diagnosis of autism spectrum disorder using wavelet, entropy, and ann. *BioMed research international*, 2017, 2017.

[44] David L Donoho and Iain M Johnstone. Adapting to unknown smoothness via wavelet shrinkage. *Journal of the american statistical association*, 90(432):1200–1224, 1995.

[45] David L Donoho and Jain M Johnstone. Ideal spatial adaptation by wavelet shrinkage. *biometrika*, 81(3):425–455, 1994.

[46] Valerie J Easton and John H McColl. Statistics glossary v1. 1. 1997.

[47] Edson Estrada, Homer Nazeran, Gustavo Sierra, Farideh Ebrahimi, and S Kamaledin Setarehdan. Wavelet-based eeg denoising for automatic sleep stage classification. In *Electrical Communications and Computers (CONIELECOMP), 2011 21st International Conference on*, pages 295–298. IEEE, 2011.

[48] Mehrdad Fatourechi, Ali Bashashati, Rabab K Ward, and Gary E Birch. Emg and eog artifacts in brain computer interface systems: A survey. *Clinical neurophysiology*, 118(3):480–494, 2007.

[49] Mehrdad Fatourechi, Rabab K Ward, Steven G Mason, Jane Huggins, Alois Schlögl, and Gary E Birch. Comparison of evaluation metrics in classification applications with imbalanced datasets. In *Machine Learning and Applications, 2008. ICMLA'08. Seventh International Conference on*, pages

777–782. IEEE, 2008.

[50] Torsten Felzer and B Freisieben. Analyzing eeg signals using the probability estimating guarded neural classifier. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(4):361–371, 2003.

[51] S P Fitzgibbon, D MW Powers, K J Pope, and C R Clark. Removal of eeg noise and artifact using blind source separation. *Journal of Clinical Neurophysiology*, 24(3):232–243, 2007.

[52] SP Fitzgibbon, D DeLosAngeles, TW Lewis, DMW Powers, EM Whitham, JO Willoughby, and KJ Pope. Surface laplacian of scalp electrical signals and independent component analysis resolve emg contamination of electroencephalogram. *International Journal of Psychophysiology*, 97(3):277–284, 2015.

[53] Patrick Flandrin, Gabriel Rilling, and Paulo Goncalves. Empirical mode decomposition as a filter bank. *IEEE signal processing letters*, 11(2):112–114, 2004.

[54] Julien Fleureau, Jean-Claude Nunes, Amar Kachenoura, Laurent Albera, and Lotfi Senhadji. Turning tangent empirical mode decomposition: a framework for mono-and multivariate signals. *IEEE Transactions on signal Processing*, 59(3):1309–1316, 2011.

[55] Arthur Flexer, Herbert Bauer, Jürgen Pripfl, and Georg Dorffner. Using ica for removal of ocular artifacts in eeg recorded from blind subjects. *Neural Networks*, 18(7):998–1005, 2005.

[56] Ola Friman, Magnus Borga, Peter Lundberg, and Hans Knutsson. Exploratory fmri analysis by autocorrelation maximization. *NeuroImage*, 16(2):454–464, 2002.

[57] Jianbo Gao, Hussain Sultan, Jing Hu, and Wen-Wen Tung. Denoising nonlinear time series by adaptive filtering and wavelet shrinkage: a comparison. *IEEE signal processing letters*, 17(3):237–240, 2010.

[58] Gary N Garcia, Touradj Ebrahimi, and J-M Vesin. Support vector eeg classification in the fourier and time-frequency correlation domains. In *Neural Engineering, 2003. Conference Proceedings. First International IEEE EMBS Conference on*, pages 591–594. IEEE, 2003.

[59] Pando Georgiev, Fabian Theis, Andrzej Cichocki, and Hovagim Bakardjian. Sparse component analysis: a new tool for data mining. *Data mining in biomedicine*, pages 91–116, 2007.

[60] Hosna Ghandeharion and Abbas Erfanian. A fully automatic ocular artifact suppression from eeg data using higher order statistics: Improved performance by wavelet analysis. *Medical Engineering and Physics*, 32(7):720–729, 2010.

[61] Germán Gómez-Herrero. Automatic artifact removal (aar) toolbox v1. 3 (release 09.12. 2007) for matlab. *Tampere University of Technology*, 2007.

[62] Irina I Goncharova, Dennis J McFarland, Theresa M Vaughan, and Jonathan R Wolpaw. Emg contamination of eeg: spectral and topographical characteristics. *Clinical neurophysiology*, 114(9):1580–1593, 2003.

[63] Gabriele Gratton, Michael GH Coles, and Emanuel Donchin. A new method for off-line removal of ocular artifact. *Electroencephalography and clinical neurophysiology*, 55(4):468–484, 1983.

[64] Rémi Gribonval and Sylvain Lesage. A survey of sparse component analysis for blind source separation: principles, perspectives, and new challenges. In *ESANN'06 proceedings-14th European Symposium on Artificial Neural Networks*, pages 323–330. d-side publi., 2006.

[65] Carlos Guerrero-Mosquera and Angel Navia-Vázquez. Automatic removal of ocular artefacts using adaptive filtering and independent component analysis for electroencephalogram data. *IET signal processing*, 6(2):99–106, 2012.

[66] Joseph T Gwin, Klaus Gramann, Scott Makeig, and Daniel P Ferris. Removal of movement artifact from high-density eeg recorded during walking and running. *Journal of neurophysiology*, 103(6):3526–3534, 2010.

[67] Sebastian Halder, Michael Bensch, Jürgen Mellinger, Martin Bogdan, Andrea Kübler, Niels Birbaumer, and Wolfgang Rosenstiel. Online artifact removal for brain-computer interfaces using support vector machines and blind source separation. *Computational intelligence and neuroscience*, 2007, 2007.

[68] Hans Hallez, Bart Vanrumste, Roberta Grech, Joseph Muscat, Wim De Clercq, Anneleen Vergult, Yves D'Asseler, Kenneth P Camilleri, Simon G

Fabri, Sabine Van Huffel, et al. Review on solving the forward problem in eeg source analysis. *Journal of neuroengineering and rehabilitation*, 4(1):46, 2007.

[69] Mehdi Bagheri Hamaneh, Numthip Chitravas, Kitti Kaiboriboon, Samden D Lhatoo, and Kenneth A Loparo. Automated removal of ekg artifact from eeg data using independent component analysis and continuous wavelet transformation. *IEEE Transactions on Biomedical Engineering*, 61(6):1634–1641, 2014.

[70] Satoshi Hara, Yoshinobu Kawahara, Takashi Washio, Paul Von BüNau, Terumasa Tokunaga, and Kiyohumi Yumoto. Separation of stationary and non-stationary sources with a generalized eigenvalue problem. *Neural networks*, 33:7–20, 2012.

[71] MM Hartmann, K Schindler, TA Gebbink, G Gritsch, and T Kluge. Pureeeg: automatic eeg artifact removal for epilepsy monitoring. *Neurophysiologie Clinique/Clinical Neurophysiology*, 44(5):479–490, 2014.

[72] Pawel Herman, Girijesh Prasad, Thomas Martin McGinnity, and Damien Coyle. Comparative analysis of spectral approaches to feature extraction for eeg-based motor imagery classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 16(4):317–326, 2008.

[73] Roberto H Herrera, Jiajun Han, and Mirko van der Baan. Applications of the synchrosqueezing transform in seismic time-frequency analysis. *Geophysics*, 79(3):V55–V64, 2014.

[74] Christian W Hesse and Christopher J James. On semi-blind source separation using spatial constraints with applications in eeg analysis. *IEEE Transactions on Biomedical Engineering*, 53(12):2525–2534, 2006.

[75] Jing Hu, Chun-sheng Wang, Min Wu, Yu-xiao Du, Yong He, and Jinhua She. Removal of eog and emg artifacts from eeg using combination of functional link neural network and adaptive neural fuzzy inference system. *Neurocomputing*, 151:278–287, 2015.

[76] Norden E Huang, Zheng Shen, Steven R Long, Manli C Wu, Hsing H Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, and Henry H Liu. The empirical mode decomposition and the hilbert spectrum for nonlinear and

non-stationary time series analysis. In *Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences*, volume 454, pages 903–995. The Royal Society, 1998.

[77] Norden E Huang and Zhaohua Wu. A review on hilbert-huang transform: Method and its applications to geophysical studies. *Reviews of geophysics*, 46(2), 2008.

[78] Norden Eh Huang. *Hilbert-Huang transform and its applications*, volume 16. World Scientific, 2014.

[79] Aapo Hyvarinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE transactions on Neural Networks*, 10(3):626–634, 1999.

[80] Aapo Hyvärinen and Erkki Oja. One-unit learning rules for independent component analysis. In *Advances in neural information processing systems*, pages 480–486, 1997.

[81] Dmytro Iatsenko, Peter VE McClintock, and Aneta Stefanovska. Nonlinear mode decomposition: a noise-robust, adaptive decomposition method. *Physical Review E*, 92(3):032916, 2015.

[82] Nicole Ille, Patrick Berg, and Michael Scherg. Artifact correction of the ongoing eeg using spatial filters based on artifact and brain signal topographies. *Journal of clinical neurophysiology*, 19(2):113–124, 2002.

[83] Nicole Ille, Roland Beucker, and Michael Scherg. Spatially constrained independent component analysis for artifact correction in eeg and meg. *Neuroimage*, 13(6):159, 2001.

[84] Jorge Iriarte, Elena Urrestarazu, Miguel Valencia, Manuel Alegre, Armando Malanda, César Viteri, and Julio Artieda. Independent component analysis as a tool to eliminate artifacts in eeg: a quantitative study. *Journal of clinical neurophysiology*, 20(4):249–257, 2003.

[85] Md Kafiul Islam. *Artifact Characterization, Detection and Removal from Neural Signals*. PhD thesis, 2015.

[86] Md Kafiul Islam, Amir Rastegarnia, and Zhi Yang. Methods for artifact detection and removal from scalp eeg: a review. *Neurophysiologie Clinique/Clinical Neurophysiology*, 46(4):287–305, 2016.

[87] Darshan Iyer and George Zouridakis. Single-trial evoked potential estimation: comparison between independent component analysis and wavelet denoising. *Clinical Neurophysiology*, 118(3):495–504, 2007.

[88] Anil K Jain, Robert P. W. Duin, and Jianchang Mao. Statistical pattern recognition: A review. *IEEE Transactions on pattern analysis and machine intelligence*, 22(1):4–37, 2000.

[89] Christopher J James and Oliver J Gibson. Temporally constrained ica: an application to artifact rejection in electromagnetic brain signal analysis. *IEEE Transactions on Biomedical Engineering*, 50(9):1108–1116, 2003.

[90] Christopher J James and Christian W Hesse. Independent component analysis for biomedical signals. *Physiological measurement*, 26(1):R15, 2004.

[91] Kevin Jones, Henri Begleiter, Bernice Porjesz, Kongming Wang, and David Chorlian. Complexity measures of event related potential surface laplacian data calculated using the wavelet packet transform. *Brain topography*, 14(4):333–344, 2002.

[92] Carrie A Joyce, Irina F Gorodnitsky, and Marta Kutas. Automatic removal of eye movement and blink artifacts from eeg data using blind component separation. *Psychophysiology*, 41(2):313–325, 2004.

[93] Tzyy-Ping Jung, Scott Makeig, Colin Humphries, Te-Won Lee, Martin J Mckeown, Vicente Iragui, and Terrence J Sejnowski. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37(2):163–178, 2000.

[94] Tzyy-Ping Jung, Scott Makeig, Marissa Westerfield, Jeanne Townsend, Eric Courchesne, and Terrence J Sejnowski. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clinical Neurophysiology*, 111(10):1745–1758, 2000.

[95] Mehmed Kantardzic. *Data mining: concepts, models, methods, and algorithms*. John Wiley & Sons, 2011.

[96] Alon S Keren, Shlomit Yuval-Greenberg, and Leon Y Deouell. Saccadic spike potentials in gamma-band eeg: characterization, detection and suppression. *Neuroimage*, 49(3):2248–2263, 2010.

[97] Saleha Khatun, Ruhi Mahajan, and Bashir I Morshed. Comparative study

of wavelet-based unsupervised ocular artifact removal techniques for single-channel eeg data. *IEEE journal of translational engineering in health and medicine*, 4:1–8, 2016.

[98] Rahul Kher and Riddhish Gandhi. Adaptive filtering based artifact removal from electroencephalogram (eeg) signals. In *Communication and Signal Processing (ICCSP), 2016 International Conference on*, pages 0561–0564. IEEE, 2016.

[99] Joep JM Kierkels, Geert JM van Boxtel, and Leo LM Vogten. A model-based objective evaluation of eye movement correction in eeg recordings. *IEEE Transactions on biomedical engineering*, 53(2):246–253, 2006.

[100] Murielle Kirkove, Clémentine François, and Jacques Verly. Comparative evaluation of existing and new methods for correcting ocular artifacts in electroencephalographic recordings. *Signal Processing*, 98:102–120, 2014.

[101] Manousos A Klados, Christos Papadelis, Christoph Braun, and Panagiotis D Bamidis. Reg-ica: a hybrid methodology combining blind source separation and regression techniques for the rejection of ocular artifacts. *Biomedical Signal Processing and Control*, 6(3):291–300, 2011.

[102] Zoltan Joseph Koles. The quantitative extraction and topographic mapping of the abnormal components in the clinical eeg. *Electroencephalography and clinical Neurophysiology*, 79(6):440–447, 1991.

[103] Yannis Kopsinis and Stephen McLaughlin. Development of emd-based denoising methods inspired by wavelet thresholding. *IEEE Transactions on signal Processing*, 57(4):1351–1362, 2009.

[104] 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*, 3(4):338, 2006.

[105] V Krishnaveni, S Jayaraman, S Aravind, V Hariharasudhan, and K Ramadoss. Automatic identification and removal of ocular artifacts from eeg using wavelet transform. *Measurement science review*, 6(4):45–57, 2006.

[106] Jakub Kužílek. Independent component analysis: Applications in ecg signal processing. 2013.

[107] Terrence D Lagerlund, Frank W Sharbrough, and Neil E Busacker. Spatial

filtering of multichannel electroencephalographic recordings through principal component analysis by singular value decomposition. *Journal of clinical neurophysiology*, 14(1):73–82, 1997.

[108] Yandong Li, Zhongwei Ma, Wenkai Lu, and Yanda Li. Automatic removal of the eye blink artifact from eeg using an ica-based template matching approach. *Physiological measurement*, 27(4):425, 2006.

[109] Yuanqing Li, Andrzej Cichocki, and S-I Amari. Blind estimation of channel parameters and source components for eeg signals: a sparse factorization approach. *IEEE Transactions on Neural Networks*, 17(2):419–431, 2006.

[110] Chin-Teng Lin, Chih-Sheng Huang, Wen-Yu Yang, Avinash Kumar Singh, Chun-Hsiang Chuang, and Yu-Kai Wang. Real-time eeg signal enhancement using canonical correlation analysis and gaussian mixture clustering. *Journal of Healthcare Engineering*, 2018, 2018.

[111] Chan-Cheng Liu, Tsung-Ying Sun, Shang-Jeng Tsai, Yu-Hsiang Yu, and Sheng-Ta Hsieh. Heuristic wavelet shrinkage for denoising. *Applied Soft Computing*, 11(1):256–264, 2011.

[112] Javier Lopez-Calderon and Steven J Luck. Erplab: an open-source toolbox for the analysis of event-related potentials. *Frontiers in human neuroscience*, 8, 2014.

[113] Fabien Lotte. A tutorial on eeg signal-processing techniques for mental-state recognition in brain–computer interfaces. In *Guide to brain-computer music interfacing*, pages 133–161. Springer, 2014.

[114] Fabien Lotte. Signal processing approaches to minimize or suppress calibration time in oscillatory activity-based brain–computer interfaces. *Proceedings of the IEEE*, 103(6):871–890, 2015.

[115] Fabien Lotte, Laurent Bougrain, Andrzej Cichocki, Maureen Clerc, Marco Congedo, Alain Rakotomamonjy, and Florian Yger. A review of classification algorithms for eeg-based brain-computer interfaces: A 10-year update. *Journal of neural engineering*, 2018.

[116] Fabien Lotte, Marco Congedo, Anatole Lécuyer, Fabrice Lamarche, and Bruno Arnaldi. A review of classification algorithms for eeg-based brain–computer interfaces. *Journal of neural engineering*, 4(2):R1, 2007.

[117] Fabien Lotte and Cuntai Guan. Regularizing common spatial patterns to improve bci designs: unified theory and new algorithms. *IEEE Transactions on biomedical Engineering*, 58(2):355–362, 2011.

[118] Wei Lu and Jagath C Rajapakse. Approach and applications of constrained ica. *IEEE transactions on neural networks*, 16(1):203–212, 2005.

[119] Ruhi Mahajan and Bashir I Morshed. Unsupervised eye blink artifact denoising of eeg data with modified multiscale sample entropy, kurtosis, and wavelet-ica. *IEEE journal of Biomedical and Health Informatics*, 19(1):158–165, 2015.

[120] Scott Makeig, Anthony J Bell, Tzyy-Ping Jung, and Terrence J Sejnowski. Independent component analysis of electroencephalographic data. In *Advances in neural information processing systems*, pages 145–151, 1996.

[121] Scott Makeig, Christian Kothe, Tim Mullen, Nima Bigdely-Shamlo, Zhilin Zhang, and Kenneth Kreutz-Delgado. Evolving signal processing for brain–computer interfaces. *Proceedings of the IEEE*, 100(Special Centennial Issue):1567–1584, 2012.

[122] Stephane Mallat. Zero-crossings of a wavelet transform. *IEEE Transactions on Information theory*, 37(4):1019–1033, 1991.

[123] Stephane Mallat. *A wavelet tour of signal processing: the sparse way*. Academic press, 2008.

[124] Stephane G Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE transactions on pattern analysis and machine intelligence*, 11(7):674–693, 1989.

[125] Nadia Mammone and Francesco C Morabito. Enhanced automatic wavelet independent component analysis for electroencephalographic artifact removal. *Entropy*, 16(12):6553–6572, 2014.

[126] Danilo P Mandic, Naveed ur Rehman, Zhaohua Wu, and Norden E Huang. Empirical mode decomposition-based time-frequency analysis of multivariate signals: The power of adaptive data analysis. *IEEE signal processing magazine*, 30(6):74–86, 2013.

[127] Malik M Naeem Mannan, Myung Y Jeong, and Muhammad A Kamran. Hybrid icaregression: automatic identification and removal of ocular arti-

facts from electroencephalographic signals. *Frontiers in human neuroscience*, 10:193, 2016.

[128] C Marque, C Bisch, R Dantas, S Elayoubi, V Brosse, and C Perot. Adaptive filtering for ecg rejection from surface emg recordings. *Journal of electromyography and kinesiology*, 15(3):310–315, 2005.

[129] Joseph W Matiko, Stephen Beeby, and John Tudor. Real time eye blink noise removal from eeg signals using morphological component analysis. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 13–16. IEEE, 2013.

[130] Brenton W McMenamin, Alexander J Shackman, Lawrence L Greischar, and Richard J Davidson. Electromyogenic artifacts and electroencephalographic inferences revisited. *Neuroimage*, 54(1):4–9, 2011.

[131] Brenton W McMenamin, Alexander J Shackman, Jeffrey S Maxwell, David RW Bachhuber, Adam M Koppenhaver, Lawrence L Greischar, and Richard J Davidson. Validation of ica-based myogenic artifact correction for scalp and source-localized eeg. *Neuroimage*, 49(3):2416–2432, 2010.

[132] BRENTON W McMENAMIN, Alexander J Shackman, Jeffrey S Maxwell, Lawrence L Greischar, and Richard J Davidson. Validation of regression-based myogenic correction techniques for scalp and source-localized eeg. *Psychophysiology*, 46(3):578–592, 2009.

[133] Unser Michael. A review of wavelets in biomedical applications. *Invited paper, Proc. IEEE*, 84(4), 1996.

[134] Bogdan Mijovic, Maarten De Vos, Ivan Gligorijevic, Joachim Taelman, and Sabine Van Huffel. Source separation from single-channel recordings by combining empirical-mode decomposition and independent component analysis. *IEEE transactions on biomedical engineering*, 57(9):2188–2196, 2010.

[135] Jd R Millán, Frederic Renkens, Josep Mourino, and Wulfram Gerstner. Non-invasive brain-actuated control of a mobile robot by human eeg. *IEEE Transactions on biomedical Engineering*, 51(6):1026–1033, 2004.

[136] Andrea Mognon, Jorge Jovicich, Lorenzo Bruzzone, and Marco Buiatti. Adjust: An automatic eeg artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*, 48(2):229–240, 2011.

[137] Md Khademul Islam Molla, ATM Jahangir Alam, Munmun Akter, AR Shoyeb Ahmed Siddique, and M Sayedur Rahman. Analysis of inter-annual climate variability using discrete wavelet transform. In *Computational Intelligence Techniques in Earth and Environmental Sciences*, pages 155–171. Springer, 2014.

[138] Md Khademul Islam Molla, Md Rabiul Islam, Toshihisa Tanaka, and Tomasz M Rutkowski. Artifact suppression from eeg signals using data adaptive time domain filtering. *Neurocomputing*, 97:297–308, 2012.

[139] Md Khademul Islam Molla, Toshihisa Tanaka, Tatsuhiko Osa, and Md Rabiul Islam. Eeg signal enhancement using multivariate wavelet transform application to single-trial classification of event-related potentials. In *Digital Signal Processing (DSP), 2015 IEEE International Conference on*, pages 804–808. IEEE, 2015.

[140] Fabio Morbidi, Andrea Garulli, Domenico Prattichizzo, Cristiano Rizzo, and Simone Rossi. Application of kalman filter to remove tms-induced artifacts from eeg recordings. *IEEE Transactions on Control Systems Technology*, 16(6):1360–1366, 2008.

[141] Th Mulder. Motor imagery and action observation: cognitive tools for rehabilitation. *Journal of neural transmission*, 114(10):1265–1278, 2007.

[142] Suresh D Muthukumaraswamy. High-frequency brain activity and muscle artifacts in meg/eeg: a review and recommendations. *Frontiers in human neuroscience*, 7, 2013.

[143] Xavier Navarro, Fabienne Porée, and Guy Carrault. Ecg removal in preterm eeg combining empirical mode decomposition and adaptive filtering. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 661–664. IEEE, 2012.

[144] Kianoush Nazarpour, Yodchanan Wongsawat, Saeid Sanei, Jonathon A Chambers, Soontorn Oraintara, et al. Removal of the eye-blink artifacts from eegs via stf-ts modeling and robust minimum variance beamforming. *IEEE Transactions on Biomedical Engineering*, 55(9):2221–2231, 2008.

[145] Hoang-Anh T Nguyen, John Musson, Feng Li, Wei Wang, Guangfan Zhang, Roger Xu, Carl Richey, Tom Schnell, Frederic D McKenzie, and Jiang

Li. Eog artifact removal using a wavelet neural network. *Neurocomputing*, 97:374–389, 2012.

[146] N Nicolaou and Slawomir J Nasuto. Automatic artefact removal from event-related potentials via clustering. *The Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology*, 48(1-2):173–183, 2007.

[147] Luis Fernando Nicolas-Alonso and Jaime Gomez-Gil. Brain computer interfaces, a review. *Sensors*, 12(2):1211–1279, 2012.

[148] Hugh Nolan, Robert Whelan, and RB Reilly. Faster: fully automated statistical thresholding for eeg artifact rejection. *Journal of neuroscience methods*, 192(1):152–162, 2010.

[149] Borna Noureddin, Peter D Lawrence, and Gary E Birch. Time-frequency analysis of eye blinks and saccades in eog for eeg artifact removal. In *Neural Engineering, 2007. CNE'07. 3rd International IEEE/EMBS Conference on*, pages 564–567. IEEE, 2007.

[150] Sebastian Olbrich, Johannes Jödicke, Christian Sander, Hubertus Himmerich, and Ulrich Hegerl. Ica-based muscle artefact correction of eeg data: What is muscle and what is brain?: Comment on mcmenamin et al. *Neuroimage*, 54(1):1–3, 2011.

[151] Jason A Palmer, Ken Kreutz-Delgado, and Scott Makeig. Amica: An adaptive mixture of independent component analyzers with shared components. *Swartz Center for Computatonal Neursoscience, University of California San Diego, Tech. Rep*, 2012.

[152] Danny Panknin, Paul Von Bünau, Motoaki Kawanabe, Frank C Meinecke, and Klaus-Robert Müller. Higher order stationary subspace analysis. In *Journal of Physics: Conference Series*, volume 699, page 012021. IOP Publishing, 2016.

[153] Hong Peng, Bin Hu, Qiuxia Shi, Martyn Ratcliffe, Qinglin Zhao, Yanbing Qi, and Guoping Gao. Removal of ocular artifacts in eegan improved approach combining dwt and anc for portable applications. *IEEE journal of biomedical and health informatics*, 17(3):600–607, 2013.

[154] Ross E Petty, Ronald M Laxer, Carol B Lindsley, and Lucy Wedderburn. *Textbook of Pediatric Rheumatology E-Book*. Elsevier Health Sciences, 2015.

[155] Gert Pfurtscheller, Doris Flotzinger, and Joachim Kalcher. Brain-computer interfacea new communication device for handicapped persons. *Journal of Microcomputer Applications*, 16(3):293–299, 1993.

[156] Trieu TH Pham, Rodney J Croft, Peter J Cadusch, and Robert J Barry. A test of four eog correction methods using an improved validation technique. *International Journal of Psychophysiology*, 79(2):203–210, 2011.

[157] Hong Qian. Fractional brownian motion and fractional gaussian noise. In *Processes with Long-Range Correlations*, pages 22–33. Springer, 2003.

[158] T Radüntz, J Scouten, O Hochmuth, and B Meffert. Eeg artifact elimination by extraction of ica-component features using image processing algorithms. *Journal of neuroscience methods*, 243:84–93, 2015.

[159] BS Raghavendra and D Narayana Dutt. Wavelet enhanced cca for minimization of ocular and muscle artifacts in eeg. *World Academy of Science, Engineering and Technology*, 57(6):1027–32, 2011.

[160] Rabie A Ramadan and Athanasios V Vasilakos. Brain computer interface: control signals review. *Neurocomputing*, 223:26–44, 2017.

[161] Luke Rankine, Nathan Stevenson, Mostefa Mesbah, and Boualem Boashash. A nonstationary model of newborn eeg. *IEEE Transactions on Biomedical Engineering*, 54(1):19–28, 2007.

[162] Rajesh PN Rao. *Brain-computer interfacing: an introduction*. Cambridge University Press, 2013.

[163] Md Rashed-Al-Mahfuz, Md Rabiul Islam, Keikichi Hirose, and Md Khademul Islam Molla. Artifact suppression and analysis of brain activities with electroencephalography signals. *Neural regeneration research*, 8(16):1500, 2013.

[164] Siamak Rezaei, Kouhyar Tavakolian, Ali Moti Nasrabadi, and S Kamaledin Setarehdan. Different classification techniques considering brain computer interface applications. *Journal of Neural Engineering*, 3(2):139, 2006.

[165] S Romero, MA Mañanas, and Manel J Barbanoj. Ocular reduction in eeg signals based on adaptive filtering, regression and blind source separation. *Annals of biomedical engineering*, 37(1):176–191, 2009.

[166] Sergio Romero, Miguel A Mañanas, and Manel J Barbanoj. A comparative

study of automatic techniques for ocular artifact reduction in spontaneous eeg signals based on clinical target variables: a simulation case. *Computers in biology and medicine*, 38(3):348–360, 2008.

[167] Vandana Roy and Shailja Shukla. Automatic removal of artifacts from eeg signal based on spatially constrained ica using daubechies wavelet. *International Journal of Modern Education and Computer Science*, 6(7):31, 2014.

[168] Doha Safieddine, Amar Kachenoura, Laurent Albera, Gwénaël Birot, Ahmad Karfoul, Anca Pasnicu, Arnaud Biraben, Fabrice Wendling, Lotfi Senhadji, and Isabelle Merlet. Removal of muscle artifact from eeg data: comparison between stochastic (ica and cca) and deterministic (emd and wavelet-based) approaches. *EURASIP Journal on Advances in Signal Processing*, 2012(1):127, 2012.

[169] Chong Yeh Sai, Norrima Mokhtar, Hamzah Arof, Paul Cumming, and Masahiro Iwahashi. Automated classification and removal of eeg artifacts with svm and wavelet-ica. *IEEE journal of biomedical and health informatics*, 2017.

[170] R Sameni. The open-source electrophysiological toolbox (oset), version 3.1, 2014. *URL http://www. oset. ir*.

[171] Reza Sameni, Mohammad B Shamsollahi, Christian Jutten, and Gari D Clifford. A nonlinear bayesian filtering framework for ecg denoising. *IEEE Transactions on Biomedical Engineering*, 54(12):2172–2185, 2007.

[172] Alois Schloegl, Andreas Ziehe, and Klaus-Robert Müller. Automated ocular artifact removal: comparing regression and component-based methods. 2009.

[173] Alois Schlögl, Claudia Keinrath, Doris Zimmermann, Reinhold Scherer, Robert Leeb, and Gert Pfurtscheller. A fully automated correction method of eog artifacts in eeg recordings. *Clinical neurophysiology*, 118(1):98–104, 2007.

[174] Alois Schlogl, Julien Kronegg, J Huggins, and S Mason. 19 evaluation criteria for bci research. *Toward brain-computer interfacing*, 2007.

[175] Bill Scott. Developments in eeg analysis, protocol selection, and feedback delivery. *Journal of Neurotherapy*, 15(3):262–267, 2011.

[176] Alexander J Shackman, Brenton W McMenamin, Heleen A Slagter, Jeffrey S

Maxwell, Lawrence L Greischar, and Richard J Davidson. Electromyogenic artifacts and electroencephalographic inferences. *Brain topography*, 22(1):7–12, 2009.

[177] Supriya Shete and Revathi Shriram. Comparison of sub-band decomposition and reconstruction of eeg signal by daubechies9 and symlet9 wavelet. In *Communication Systems and Network Technologies (CSNT), 2014 Fourth International Conference on*, pages 856–861. IEEE, 2014.

[178] Leor Shoker, Saeid Sanei, and Jonathon Chambers. Artifact removal from electroencephalograms using a hybrid bss-svm algorithm. *IEEE Signal Processing Letters*, 12(10):721–724, 2005.

[179] Liu Shucong, Cheng Lina, and Li Lixin. Research on seismic signals denoising method based on multi-threshold wavelet packet. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 9(2):297–306, 2016.

[180] Balbir Singh and Hiroaki Wagatsuma. A removal of eye movement and blink artifacts from eeg data using morphological component analysis. *Computational and mathematical methods in medicine*, 2017, 2017.

[181] Ana M Skupch, Peter Dollfuss, Franz Furbass, Gerhard Gritsch, Manfred M Hartmann, Hannes Perko, Ekaterina Pataraia, Gerald Lindinger, and Tilmann Kluge. Spatial correlation based artifact detection for automatic seizure detection in eeg. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 1972–1975. IEEE, 2013.

[182] S Solhjoo and MH Moradi. Mental task recognition: A comparison between some of classification methods. In *BIOSIGNAL 2004 International EURASIP Conference*, pages 24–26, 2004.

[183] Sumit Soman et al. High performance eeg signal classification using classifiability and the twin svm. *Applied Soft Computing*, 30:305–318, 2015.

[184] Mumtaz Hussain Soomro, Nasreen Badruddin, Mohd Zuki Yusoff, and Aamir Saeed Malik. A method for automatic removal of eye blink artifacts from eeg based on emd-ica. In *Signal Processing and its Applications (CSPA), 2013 IEEE 9th International Colloquium on*, pages 129–134. IEEE,

2013.

[185] Leif Sörnmo and Pablo Laguna. *Bioelectrical signal processing in cardiac and neurological applications*, volume 8. Academic Press, 2005.

[186] N Stevenson, L Rankine, M Mesbah, and B Boashash. Newborn eeg seizure simulation using time–frequency signal synthesis. In *Proc. APRS Workshop on Digital Image Computing*, pages 145–151, 2005.

[187] Abdulhamit Subasi and M Ismail Gursoy. Eeg signal classification using pca, ica, lda and support vector machines. *Expert Systems with Applications*, 37(12):8659–8666, 2010.

[188] Shiliang Sun. Ensemble learning methods for classifying eeg signals. In *International Workshop on Multiple Classifier Systems*, pages 113–120. Springer, 2007.

[189] Li Surui, Shi Lijuan, Li Zhenxin, Dong Bingchao, et al. Motor imagery eeg discrimination using hilbert-huang entropy. *Biomedical Research*, 28(2), 2017.

[190] Kevin Sweeney. *Motion Artifact Processing Techniques for Physiological Signals*. PhD thesis, National University of Ireland Maynooth, 2013.

[191] Kevin T Sweeney, Hasan Ayaz, Tomás E Ward, Meltem Izzetoglu, Seán F McLoone, and Banu Onaral. A methodology for validating artifact removal techniques for physiological signals. *IEEE transactions on information technology in biomedicine*, 16(5):918–926, 2012.

[192] Kevin T Sweeney, Seán F McLoone, and Tomas E Ward. The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique. *IEEE transactions on biomedical engineering*, 60(1):97–105, 2013.

[193] Kevin T Sweeney, Tomás E Ward, and Seán F McLoone. Artifact removal in physiological signalspractices and possibilities. *IEEE transactions on information technology in biomedicine*, 16(3):488–500, 2012.

[194] Desney Tan and Anton Nijholt. Brain-computer interfaces and human-computer interaction. In *Brain-Computer Interfaces*, pages 3–19. Springer, 2010.

[195] Akaysha C Tang, Jing-Yu Liu, and Matthew T Sutherland. Recovery of

correlated neuronal sources from eeg: the good and bad ways of using sobi. *Neuroimage*, 28(2):507–519, 2005.

[196] Akaysha C Tang, Barak A Pearlmutter, Natalie A Malaszenko, and Dan B Phung. Independent components of magnetoencephalography: single-trial response onset times. *Neuroimage*, 17(4):1773–1789, 2002.

[197] Akaysha C Tang, Barak A Pearlmutter, Michael Zibulevsky, and Scott A Carter. Blind source separation of multichannel neuromagnetic responses. *Neurocomputing*, 32:1115–1120, 2000.

[198] Akaysha C Tang, Matthew T Sutherland, and Christopher J McKinney. Validation of sobi components from high-density eeg. *NeuroImage*, 25(2):539–553, 2005.

[199] AR Teixeira, AM Tomé, EW Lang, P Gruber, and A Martins da Silva. On the use of clustering and local singular spectrum analysis to remove ocular artifacts from electroencephalograms. In *Neural Networks, 2005. IJCNN'05. Proceedings. 2005 IEEE International Joint Conference on*, volume 4, pages 2514–2519. IEEE, 2005.

[200] Eoin Thomas, Matthew Dyson, and Maureen Clerc. An analysis of performance evaluation for motor-imagery based bci. *Journal of neural engineering*, 10(3):031001, 2013.

[201] KH Ting, PCW Fung, CQ Chang, and FHY Chan. Automatic correction of artifact from single-trial event-related potentials by blind source separation using second order statistics only. *Medical engineering & physics*, 28(8):780–794, 2006.

[202] Arjon Turnip. Automatic artifacts removal of eeg signals using robust principal component analysis. In *Technology, Informatics, Management, Engineering, and Environment (TIME-E), 2014 2nd International Conference on*, pages 331–334. IEEE, 2014.

[203] Arjon Turnip and Edy Junaidi. Removal artifacts from eeg signal using independent component analysis and principal component analysis. In *Technology, Informatics, Management, Engineering, and Environment (TIME-E), 2014 2nd International Conference on*, pages 296–302. IEEE, 2014.

[204] Michael Unser and Akram Aldroubi. A review of wavelets in biomedical

applications. *Proceedings of the IEEE*, 84(4):626–638, 1996.

[205] Naveed ur Rehman, Cheolsoo Park, Norden E Huang, and Danilo P Mandic. Emd via memd: multivariate noise-aided computation of standard emd. *Advances in Adaptive Data Analysis*, 5(02):1350007, 2013.

[206] Jose Antonio Urigüen and Begoña Garcia-Zapirain. Eeg artifact removalstate-of-the-art and guidelines. *Journal of neural engineering*, 12(3):031001, 2015.

[207] Saeed V Vaseghi. *Advanced digital signal processing and noise reduction*. John Wiley & Sons, 2008.

[208] Anneleen Vergult, Wim De Clerq, Katrien Vanderperren, Bart Vanrumste, Wim Van Paesschen, and Sabine Van Huffel. Canonical correlation analysis to remove muscle artifacts from the eeg. In *Abstracts of the 1st COST Neuromath Workgroups Meeting*, pages 40–40, 2007.

[209] Ricardo Nuno Vigário. Extraction of ocular artefacts from eeg using independent component analysis. *Electroencephalography and clinical neurophysiology*, 103(3):395–404, 1997.

[210] Paul Von Bünau, Frank C Meinecke, Franz C Király, and Klaus-Robert Müller. Finding stationary subspaces in multivariate time series. *Physical review letters*, 103(21):214101, 2009.

[211] Paul Von Bünau, Frank C Meinecke, Simon Scholler, and Klaus-Robert Müller. Finding stationary brain sources in eeg data. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 2810–2813. IEEE, 2010.

[212] De Maarten Vos, Stephanie Riès, Katrien Vanderperren, Bart Vanrumste, Francois-Xavier Alario, Van Sabine Huffel, and Boris Burle. Removal of muscle artifacts from eeg recordings of spoken language production. *Neuroinformatics*, 8(2):135–150, 2010.

[213] Garrick L Wallstrom, Robert E Kass, Anita Miller, Jeffrey F Cohn, and Nathan A Fox. Automatic correction of ocular artifacts in the eeg: a comparison of regression-based and component-based methods. *International journal of psychophysiology*, 53(2):105–119, 2004.

[214] ZhenYu Wang, Peng Xu, TieJun Liu, Yin Tian, Xu Lei, and DeZhong Yao.

Robust removal of ocular artifacts by combining independent component analysis and system identification. *Biomedical Signal Processing and Control*, 10:250–259, 2014.

[215] Markus Waser and Heinrich Garn. Removing cardiac interference from the electroencephalogram using a modified pan-tompkins algorithm and linear regression. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 2028–2031. IEEE, 2013.

[216] Bernard Widrow and Samuel D Stearns. *Adaptive signal processing*, volume 15. Prentice-hall Englewood Cliffs, NJ, 1985.

[217] Irene Winkler, Stephanie Brandl, Franziska Horn, Eric Waldburger, Carsten Allefeld, and Michael Tangermann. Robust artifactual independent component classification for bci practitioners. *Journal of neural engineering*, 11(3):035013, 2014.

[218] Irene Winkler, Stefan Haufe, and Michael Tangermann. Automatic classification of artifactual ica-components for artifact removal in eeg signals. *Behavioral and Brain Functions*, 7(1):30, 2011.

[219] Jonathan R Wolpaw, Niels Birbaumer, Dennis J McFarland, Gert Pfurtscheller, and Theresa M Vaughan. Brain–computer interfaces for communication and control. *Clinical neurophysiology*, 113(6):767–791, 2002.

[220] Stuart C Wright. *Filtering of Muscle Artifact from the Electroencephalogram*. PhD thesis, Electronic Systems Laboratory, Department of Electrical Engineering, Massachusetts Inst. of Technology, 1976.

[221] Dongrui Wu, Jung-Tai King, Chun-Hsiang Chuang, Chin-Teng Lin, and Tzyy-Ping Jung. Spatial filtering for eeg-based regression problems in brain-computer interface (bci). *IEEE Transactions on Fuzzy Systems*, 2017.

[222] Zhaohua Wu and Norden E Huang. Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in adaptive data analysis*, 1(01):1–41, 2009.

[223] Yun-jie Xu and Shu-dong Xiu. A new and effective method of bearing fault diagnosis using wavelet packet transform combined with support vector machine. *JCP*, 6(11):2502–2509, 2011.

[224] Ronggen Yang and Mingwu Ren. Wavelet denoising using principal component analysis. *Expert systems with Applications*, 38(1):1073–1076, 2011.

[225] Siyang Yin, Yuelu Liu, and Mingzhou Ding. Amplitude of sensorimotor mu rhythm is correlated with bold from multiple brain regions: a simultaneous eeg-fmri study. *Frontiers in human neuroscience*, 10, 2016.

[226] Xinyi Yong, Mehrdad Fatourechi, Rabab K Ward, and Gary E Birch. Automatic artefact removal in a self-paced hybrid brain-computer interface system. *Journal of neuroengineering and rehabilitation*, 9(1):50, 2012.

[227] Xinyi Yong, Rabab K Ward, and Gary E Birch. Artifact removal in eeg using morphological component analysis. In *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pages 345–348. IEEE, 2009.

[228] Xinyi Yong, Rabab K Ward, and Gary E Birch. Generalized morphological component analysis for eeg source separation and artifact removal. In *Neural Engineering, 2009. NER'09. 4th International IEEE/EMBS Conference on*, pages 343–346. IEEE, 2009.

[229] Hong Zeng and Aiguo Song. Removal of eog artifacts from eeg recordings using stationary subspace analysis. *The Scientific World Journal*, 2014, 2014.

[230] Hong Zeng, Aiguo Song, Ruqiang Yan, and Hongyun Qin. Eog artifact correction from eeg recording using stationary subspace analysis and empirical mode decomposition. *Sensors*, 13(11):14839–14859, 2013.

[231] Chunyu Zhao and Tianshuang Qiu. An automatic ocular artifacts removal method based on wavelet-enhanced canonical correlation analysis. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 4191–4194. IEEE, 2011.

[232] Qinglin Zhao, Bin Hu, Yujun Shi, Yang Li, Philip Moore, Minghou Sun, and Hong Peng. Automatic identification and removal of ocular artifacts in eegimproved adaptive predictor filtering for portable applications. *IEEE transactions on nanobioscience*, 13(2):109–117, 2014.