ENHANCING PERFORMANCE OF CLASSIFICATION TECHNIQUES FOR EEG BASED BRAIN–COMPUTER INTERFACE

A Thesis Submitted by

Suman Dutta

Registration No: 950804011

In partial fulfillment of the requirements

For the Award of the Degree of

Doctor of Philosophy

Under the Supervision of

Dr. Mandeep Singh

Professor

Department of Electrical & Instrumentation Engineering

Dr. Amod Kumar

Professor, Electronics Engineering, NITTTR, Chandigarh

THAPAR INSTITUTE OF ENGINEERING & TECHNOLOGY

(Deemed to be University)

PUNJAB (INDIA)-147004

May 2018

DECLARATION

I hereby declare that the research work presented in this thesis titled **"Enhancing Performance of Classification Techniques for EEG based Brain-Computer Interface"** submitted for the award of the degree of Doctor of Philosophy in the Department of Electrical & Instrumentation Engineering is a bona-fide record of my own research work carried out under the supervision of **Dr. Mandeep Singh**, Professor, Electrical & Instrumentation Engineering and **Dr. Amod Kumar**, Professor, Electronics Engineering, NITTTR, Chandigarh.

The matter presented in this thesis has not been previously submitted in part or full to any other university or institution for the award of any degree in India or abroad.

S. Dutta

Date: 13.09.2019 **Suman Dutta**

We certify that the above statement made by the student is correct to the best of our knowledge and belief.

Date: 13.09.2019

13/9/19.

Dr. Mandeep Singh
Professor, EIED
Thapar Institute of Engineering & Technology

Kumar. 13/09/2019

Dr. Amod Kumar

Professor, Electronics Engineering, NITTTR, Chandigarh.

ACKNOWLEDGEMENT

First of all, I would like to bow down in front of the almighty God for His grace and bountiful blessings which enabled me to complete this research work successfully.

I am very thankful to **Prof Dr. Prakash Gopalan**, Director, Thapar Institute of Engineering and Technology, Patiala, and **Prof Dr. O.P. Pandey**, Dean of Research &Sponsored Projects, Thapar Institute of Engineering and Technology, Patiala, for providing me with the facilities to complete this research work. I would like to express my heartfelt gratitude towards **Prof Dr. R.S. Kaler**, Head, Department of Electrical and Instrumentation Engineering, Thapar Institute of Engineering and Technology, Patiala.

I would like to express my heartfelt gratitude, indebtedness, and sincere appreciation to both of my esteemed supervisors **Dr Mandeep Singh** , Professor , Electrical and Instrumentation Engineering, Thapar Institute of Engineering and Technology, Patiala and Dr **Amod Kumar**, Professor, Electronics Enginnering,NITTTR, Chandigarh for their continuous inspiration, support, guidance , encouragement , advice, patience and individual feedback throughout the course of my Ph.D study. They were easily approachable and motivated me to achieve my research goals. I feel very grateful and blessed to have worked under their supervision. I would not imagine carrying out this research without their constant motivation, support and guidance. Most of the novel ideas and solutions found in this thesis are the result of our numerous discussions .Their editorial comments were very meaningful and significant for writing this thesis.

I would like to thank the esteemed members of my Doctoral Committee: **Dr. Sanjay Kumar Jain**, Ph.D coordinator and Associate Professor, Electrical and Instrumentation Engineering, **Dr. Ravinder Agarwal**, Professor, Electrical and Instrumentation Engineering, **Dr. Vinay Kumar**, Assistant Professor, Electronics & Telecommunication Engineering for their patience, encouragement, and constructive criticism.

I am extremely grateful to my father, **Shri Mahadeb Dutta**, my mother, **Shrimati Sumita Dutta** and all other family members for their boundless love, affections and blessings.

Last and most importantly, I am extremely thankful to my wife **Mrs. Sarbani** for making this entire journey easy and comfortable through providing unconditional support and moral encouragement. At the end, special thanks to my daughter **Miss Mahashweta, whose** respect, belief in my ability and smiling face always made me enthusiastic in achieving my goal.

PUBLICATIONS BASED ON THIS DISSERTATION

Papers in SCI Journals

1. Suman Dutta, Mandeep Singh, Amod Kumar, "Classification of non-motor cognitive task in EEG based brain-computer interface using phase space features in multivariate empirical mode decomposition domain" Biomedical Signal Processing and Control, vol. 39, pp 378-389, January 2018, IF: 2.21.
2. Suman Dutta, Mandeep Singh, Amod Kumar, "Automated classification of non-motor mental task in electroencephalogram based brain-computer interface using multivariate autoregressive model in the intrinsic mode function domain" Biomedical Signal Processing and Control, vol. 43, pp 174-182, May 2018, IF: 2.21.
3. Suman Dutta, Mandeep Singh, Amod Kumar, " A novel application of multi-scale multivariate entropy as features for classifying non-motor mental tasks in EEG based BCI". (This paper has been accepted for publication in IETE journal of Research, but currently under production checklist).

ABSTRACT

Mental task (MT) Electroencephalography (EEG) is EEG recorded during performance of non-motor general mental tasks. MT based brain-computer interface (BCI) paradigm using non-motor mental tasks can be viewed as generalization of motor imagery (MI) based BCI paradigm. MT based BCI paradigm shows better potential than MI based BCI paradigm in enhancing the quality of life of the physically disabled persons. Consequently, the focus of this current research lies in MT based BCI systems. Classification of such non-motor cognitive Electroencephalography (EEG) signals produced during performance of mental task is the central challenge in developing such type of EEG based non-invasive BCI systems. The human brain shows extremely complex nonlinear and non-stationary spatiotemporal patterns of EEG signals that vary over multiple temporal scales. We believe that accurate representation of such type of complex and subtle patterns contained in the signal dynamics holds the key in enhancing the real time performance of a BCI system. But the neurological control signals driving the MT based BCI paradigm contain a number of different types of sophisticated spatiotemporal patterns which have not been identified yet. Current feature extraction algorithms relying on prior assumptions about the patterns may discard meaningful information contained in the data. Due to this, their ability to accurately identify new type of patterns is limited. The prime motivation of our work stems from this need for discovering new patterns through nonlinear signal processing or from the geometrical and topological properties of the RPS. In this dissertation, we aim at enhancing the classification performance in mental task (MT) based BCI paradigm.

We investigated the following feature extraction approaches:

1. Introducing the largest singular value of the phase space matrix in the multivariate empirical mode decomposition(MEMD) domain as feature for classifying mental task in MT based brain-computer interface.
2. Introducing Eigen values of the covariance matrix of the coefficient matrix of the multivariate autoregressive(MVAR) model in the MEMD domain.
3. Proposing multivariate multi scale entropy values as EEG features for classifying non-motor mental task EEG.
4. Singular values in the phase space of original signal as features.

In the first approach, we employed singular value decomposition (SVD) based phase space analysis of the multivariate intrinsic mode functions (IMFs) and extracted largest singular

values from the phase space matrices of the sensitive IMFs for constructing the feature vectors. With these new feature vectors, we achieved highest classification accuracy of 83.33% for binary classification between mental arithmetic and mental letter composing.

Our second approach is based on deriving multivariate autoregressive (MVAR) models of the set of relevant multivariate intrinsic mode functions (IMFs) generated from the Multivariate empirical mode decomposition (MEMD) of the multi-channel EEG signals. In this approach, the set of statistically significant Eigen values computed from the derived multivariate AR models of the set of relevant IMFs were used for constructing the feature vectors. Finally, we classified the constructed feature vectors by employing LS-SVM classifier with three different kernel functions. We achieved highest average classification accuracy of 94.3% for binary and 77.7% for three class classification.

In the third approach, we proposed multivariate multi scale entropy based complexity measures as EEG features for classifying EEG signals in MT based BCI paradigm. These entropy values computed over selected scales have been employed for constructing the feature vectors. We achieved highest classification accuracy of 100% for binary classification of the two pairs mental tasks.

We tested all our approaches on a bench mark EEG data set and evaluated the results. The accuracy, speed and consistency of the test results show efficacy of the proposed features. In this way, this thesis presents several novel results in the broad area of brain signal classification using EEG recordings which further leads to better understanding of cognitive brain dynamics and improved performance of next generation of noninvasive BCI systems.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

EEG	:	Electroencephalogram
ECOG	:	Electrocorticogram
EMG	:	Electromyogram
MEG	:	Magnetoencephalogram
EOG	:	Electrooculogram
MRI	:	Magnetic Resonance Imaging
fMRI	:	Functional Magnetic Resonance Imaging
PET	:	Positron Emission Tomography
SPECT	:	Single Photon Emission Computed Tomography
NRIS	:	Near Infrared Spectroscopy
BCI	:	Brain-Computer Interface
ERP	:	Event Related Potential
EP	:	Evoked Potential
SCP	:	Slow Cortical Potential
VEP	:	Visual Evoked Potential
SSVEP	:	Steady State Visual Evoked Potential
MI	:	Motor Imagery
SMR	:	Sensory Motor Rhythm
ERD	:	Event Related De-synchronization
ERS	:	Event Related Synchronization
EMD	:	Empirical Mode Decomposition
MEMD	:	Multivariate Empirical Mode Decomposition
IMF	:	Intrinsic Mode Function
PSR	:	Phase Space Reconstruction
RPS	:	Reconstructed Phase Space
EVD	:	Eigen Value Decomposition
SVD	:	Singular Value Decomposition
PCA	:	Principal Component Analysis
ICA	:	Independent Component Analysis
MMSE	:	Multivariate Multi Scale Sample Entropy
MMPE	:	Multivariate Multi Scale Permutation Entropy
MMFE	:	Multivariate Multi Scale Fuzzy Entropy

LS-SVM	:	Least Square Support Vector Machine
LDA	:	Linear Discriminant Analysis
KNN	:	K Nearest Neighbor
BPN	:	Back Propagation Network
ANN	:	Artificial Neural Network
PNN	:	Probabilistic Neural Network
SVM	:	Support Vector Machine
FFT	:	Fast Fourier Transform
AR	:	Auto Regressive
AAR	:	Adaptive Auto Regressive
MVAR	:	Multivariate Auto Regressive
PSD	:	Power Spectral Density
DWT	:	Discrete Wavelet Transform
WT	:	Wavelet Transform
WPT	:	Wavelet Packet Transform
STFT	:	Short Term Fourier Transform
MP	:	Matching Pursuit
CWT	:	Continuous Wavelet Transform
HMM	:	Hidden Markov Model
GMM	:	Gaussian Mixture Model
PS	:	Phase Synchronization
PLV	:	Phase Locking Value

CHAPTER 1

INTRODUCTION

1.1 Historical background of EEG based BCI system

In 1929, Hans Berger, a Professor of Psychiatry from Germany, first discovered that electrical signals produced by the human brain could be recorded from the scalp [1].He developed the method for measuring electrical signals of brain and coined the term Electroencephalography (EEG).Besides EEG, there are various other techniques, such as Electrocorticography (ECoG), Magneto encephalography (MEG) for monitoring the activities of human brain. All these techniques enable us to have better knowledge of the anatomical structure and insight of the brain function during performance of specific type of mental task. Due to very high temporal resolution, non-invasiveness, usability and low cost, EEG has become the most common technique for clinical diagnosis and brain research [2, 3].Besides clinical diagnosis, EEG has been used for developing Brain-computer interface (BCI) system also. In 1970, Jacques Vidal first used the term brain-computer interface (BCI)to describe it as “any computer- based system that produced detailed information on brain function”[4].A brain-computer interface is a “communication system that does not depend on the brain’s normal output pathways of peripheral nerves and muscles” [5-8]. These systems rely solely on the mental activity of the user to control applications such as communication, transportation or any other need of the user. The commands are generated and sent through BCI through decoding cognitive EEG signals. Thus, it provides a direct interface between human brain and a computer [9]. Although BCI is an emerging technology with great potential in diversified application areas, the main purpose is to rehabilitate the people with motor impairments through providing an alternative channel of communication and control. For people with locked in syndrome (ALS), BCI system may be their only means of communication with the outside world [10-14]. For the operation of every BCI, a neurological control signal is required. Depending upon this control signal, different BCI systems have been developed [15-22]. These systems are designed using either non-invasive EEG signals or invasive electrocorticography signals recorded from implanted electrodes. The classification of these systems is provided in chapter2.

1.1.1 Non motor mental task (MT) based BCI paradigm

In a seminal work, Keirn and Aunon [22, 75, and 76] first explored the use of general non motor Mental Task (MT) based BCI paradigm. They classified a set of five non-motor mental tasks using asymmetry of EEG power in certain frequency bands across the two hemispheres. Their work was motivated by a number of previous works that showed that there are marked differences between activities over the two hemispheres during various mental tasks [23-26]. More recent works also have demonstrated that EEG power during various mental tasks changes over a wide variety of brain regions and frequency bands [27-32]. These changes in EEG signal power is referred to as event -related desynchronization/Synchronization. This MT based BCI generalizes Motor Imagery (MI) based BCI to include other non-motor cognitive tasks, like mental arithmetic, mental letter composing, geometric figure rotation etc. A Motor imagery (MI) task is seen as mental rehearsal of any motor act without any overt motor activity[33,34].The operation and control of this MI based BCI paradigm is based on a characteristic pattern in the EEG signals known as Event-Related (De) Synchronization (ERD/ERS). During motor movements, whether real or imagined, ERD/ERS generally occurs as a decrease or increase in amplitude between the frequency ranges of 8-14Hz or 16-32 Hz, known as mu and beta rhythms respectively. These changes typically occur over the sensory motor areas of the brain [31]. Though MI based BCI systems are designed to rehabilitate people with motor impairments, it suffers from the limitation that it has very few degrees of freedom. Due to low degrees of freedom, it does not allow the user to issue more than two instructions. Apart from very few degrees of freedom, there are reports saying that not all people with motor impairments find it easy to perform imagined motor movements [35]. Keirn and Aunon overcame these limitations of MI based BCI through developing mental task (MT) based BCI which is driven by more number of non-motor mental tasks , namely, mental arithmetic task, mental letter composing etc. The MT based BCIs having more degrees of freedom yield more flexible BCI communication protocols than MI based BCIs. Furthermore, MT based BCIs provide improved performance and good user experience since the users have the option to select mental tasks as per their own comfort level. Finally, MT based BCI paradigm allows self-paced and stimulus-free control i.e. their operation also is spontaneous, independent and asynchronous i.e. Not time locked with any external stimulus. Due to these advantages, MT based BCI paradigm show enormous potential for rehabilitating the physically challenged people [36]. But the true potential of this MT based BCI paradigms have not been explored so far. Following the work of Keirn and Aunon, a number of research groups have shown improved performance of these MT based BCI paradigms. Millan, et al., have progressively improved such BCI paradigm using spontaneous mental tasks and PSD features [37-42]. The literature survey on this MT based BCI research is provided in Chapter3.

1.1.2 Motivation for the study

In spite of enormous potential of MT based BCI paradigm, such type of BCI paradigms have not been explored so far. This new BCI paradigm is still unable to meet accepted performance level of real time BCIs due to poor classification of these MT EEG signals. Enhancing the classification performance for these tasks is a big challenge. The reason for low classification performance is nonlinear, non-stationary and high dimensional nature of the signals. These MT EEG signals show complex spatiotemporal patterns that vary over multiple scales. Identifying exact type of pattern in these MT EEG signals for their accurate and reliable classification is a significant and challenging problem. Considering the significance of the problem, we kept the focus of our research on addressing the issues and challenges related to this MT based BCI paradigm.

1.2 Research goal and objectives

In this research, the main goal is focused on enhancing the performance of classification techniques for fast, reliable and accurate classification of non-motor mental tasks(MT) in general MT based BCI paradigm . Our contribution is intended to advance the current state of the art and achieve this goal by developing new feature methods through exploiting the non-linear, non-stationary and multi-dimensional nature of the EEG signals.

Through this dissertation, we described the work that we carried out for achieving our research goal and set of objectives as mentioned below

- To identify application specific mental task to be classified for a particular BCI application
- To identify the location and no of relevant EEG channels for the identified tasks.
- To extract meaningful and more discriminatory features from the underlying source components
- To develop a suitable classifier for the discrimination of the identified tasks.

1.3 Significant contributions of the dissertation

This dissertation presents research work on developing new approaches for EEG feature extraction for classifying different non-motor EEG signals in MT based BCI paradigm. We investigated four novel feature extraction approaches for improved classification of mental tasks in MT based BCI paradigm.

The major contribution of this research comes from developing the following novel feature extraction approaches:

- For the first time, we have introduced the largest singular values of the phase space trajectory matrices in the MEMD domain as feature for classifying non-motor mental task in EEG based Brain-computer interface. With these new extracted feature values, we achieved maximum classification accuracy of 83.33 % in binary classification using LS-SVM classifier with RBF kernel.
- For the first time, we have introduced statistically significant eigen values as features from the multivariate autoregressive (MVAR) model of the multivariate IMFs in the MEMD domain. With these new extracted feature values, we achieved significantly high classification accuracy of 94.43% in the binary classification of mental arithmetic task and base line task using LS-SVM classifier with Polynomial kernel function.
- For the first time, we have proposed multivariate and multi scale versions of three entropy based structural complexity measures as EEG features for classifying non-motor mental task EEG.We achieved highest classification accuracy of 100% in the binary classification of two pairs of mental task using multi-scale multivariate fuzzy entropy (MMFE).
- For the first time , we have investigated a new feature extraction approach based on combined application of singular value decomposition (SVD) and phase space reconstruction for classifying multichannel EEG signals in non-motor mental task EEG based brain-computer interface(BCI). We achieved highest classification accuracy of 96.66% for binary classification using LS-SVM classifier.

1.4 Organizational architecture of the Dissertation

This dissertation consists of the following eight chapters

- Chapter2 provides the background knowledge about the anatomy and physiology of human brain, and thereafter presents an overview of EEG signals and EEG based brain computer interfaces (BCIs).
- Chapter3 provides the comprehensive survey of the contemporary literatures for the identification of the research gaps and formulating the research goals and objectives subsequently.
- Chapter4 introduces our first novel feature extraction approach based on the combination MEMD algorithm and phase space analysis.
- Chapter5 introduces our second novel feature extraction approach based on the combination of the multivariate phase space reconstruction and multivariate Auto Regressive model(MVAR) model.
- Chapter6 introduces our third feature extraction approach based on the structural complexity analysis of the multi-channel signals using different entropy measures in a multivariate and multi scale framework.
- Chapter7 presents singular values in the RPS of the original signals as EEG features in MT based BCI.
- Chapter8 presents the conclusion and scope of future work.

CHAPTER 2

OVERVIEW OF EEG SIGNALS AND EEG BASED BRAIN-COMPUTER INTERFACES

This chapter provides an overview of the background knowledge related to EEG signals and EEG based Brain-Computer Interfaces (BCIs). For this purpose, it first introduces EEG signals and later presents an overview of EEG based BCI systems.

2.1 Overview of EEG signals

EEG signals represent electrical activities of human brain under different types of brain functions. The vast population of post synaptic cortical neurons when fired simultaneously gives rise to these signals. These signals are measured from the scalp surface by placing surface electrodes. Due to its high temporal resolution and other advantages, these signals are widely used in different applications of EEG analysis.Therefore, a brief introduction of the anatomical structure and functions of our brain along with the electrophysiology based production mechanism are given below.

2.1.1 Human Brain

It is the central organ of human nervous system having recurrent architecture of billions of neurons interacting with each other through sophisticated electrochemical mechanism.Being the main control centre, it controls the functional activities of various organs in the body. It is responsible for all activities of our body like perception, cognition, attention, emotion, memory, coordination and action [43, 45].] The brain is protected by the skull bones and suspended in cerebrospinal fluid.

This section presents a short overview of the anatomical structures of the brain, their functions and its neurophysiology for the generation of local current flows within the brain which is measured as EEG.

2.1.1.1 Functional anatomy of Human Brain

Anatomically, the entire human brain consists of “three major parts; cerebrum, cerebellum and brainstem” [45] as illustrated in Figure 2.1. The brief introduction of the three major parts are given below:

(1) **Cerebrum:** It is largest in size and plays most important role within the human brain. Out of 100 billion neurons or nerve cells, it contains 70% of the nerve cells. It is situated above the other brain structure. It controls all the neuro-physiological processes underlying various brain functions like thoughts, movements, emotions and motor functions. Cerebral cortex which is made up of neural tissues, represents the outermost layer of the cerebrum covering the outer layer of grey matter and core of white matter. The cerebrum is divided into two cerebral hemispheres such as right and left hemispheres. The two hemispheres are connected by corpus callosum. Each hemisphere is divided into four lobes - the frontal, parietal, occipital and temporal [44]. These four cortical lobes represent brain areas involved in various perceptual, cognitive and motor related brain functions.

a. **Frontal lobe** is associated with functions including self-control, reasoning, thinking, emotions, problem solving, motor development and movement.
b. **Parietal lobe** is involved in the processing of sensory perceptions related to pain sensation, tactile information, motor functions related to orientation and movement.
c. **Occipital lobe** represents the brain cortical region involved in the processing of visual information.
d. **Temporal lobe** is involved in the processing of information related to auditory stimuli, speech, memory etc.

(2) **Cerebellum**: It is known as little brain representing a major feature of hindbrain. It is smaller than human cerebrum and has two hemispheres. The number of neurons it contains is related to the number of neurons present in the neocortex.It occupies only up to10% of the total brain volume. Though it does not initiate movement, it plays dominant role in motor i.e.movement related functions of ourbody through co-ordination, precision and timing. It integrates the sensory information received from the spinal cord for fine tuning motor activity like fine motor skills, motor learning for posture and body balance control.

(3) **Brainstem**: It is the posterior part connecting the cerebrum to the spinal cord.Though small in size, it is extremely important part of our brain as it provides various tracts for connecting the nerves from motor and sensory systems to the rest of the body.It plays significant role in the control of cardiac and respiratory functions such as heart rate, breathing, sleeping, eating.

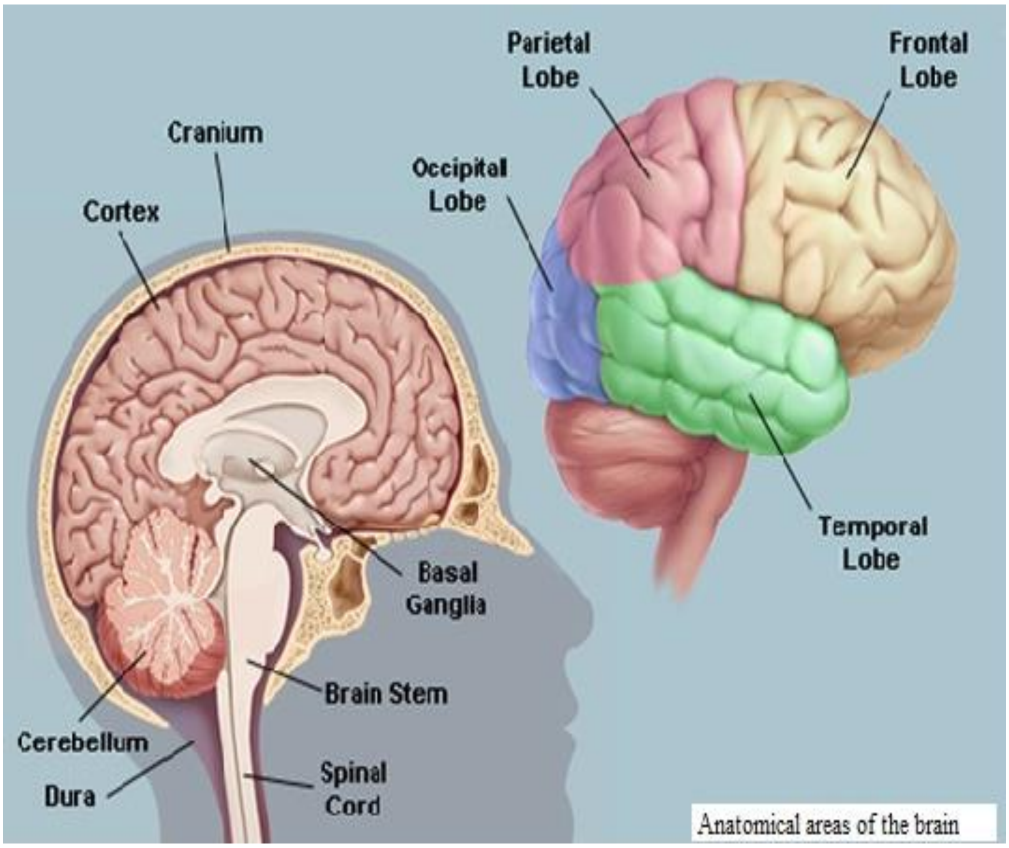


Figure 2.1: Anatomical structure of human brain

2.1.1.2. Neurophysiology of Human brain

The two types cells within human brain include nerve cells and supportive glial cells. There are about 100 billion neurons in the brain.The anatomy of nerve cell is similar to other cells. Neurons are densely interconnected with each other through synaptic junctions. The interconnections of neurons and release of neurotransmitters in response to nerve impulses gives rise to brain electrical activity.

Due to electrochemical effect, they pass messages to each other over long distances. Nerve cells consists of"three basic parts, cell body (soma), axon and dendrites" [43, 44] as shown in Figure2.2.

The nucleus of nerve cell is the main part which provides instructions to the cell for its activity. The longest and slender part of the neuron isknown as axon.It connects the nucleus of its own neuron to the shortest branch of another neuron known as dendrite.Dendrites containing many receptor sites are located on one or both ends of a nerve cell.Neurons can communicate with each other through the axon dendrite link by generating action potential.Due to the action of ionic pumps,the ionic concentration of the axon changes rapidly which is known as action potential. This action potential"allows an electrical signal to travel quickly through the axon to the next dendrite" [46]. This causes the neurons to emit and send a chemical known as neurotransmitter through the synaptic gap between the neurons for triggering an activity.

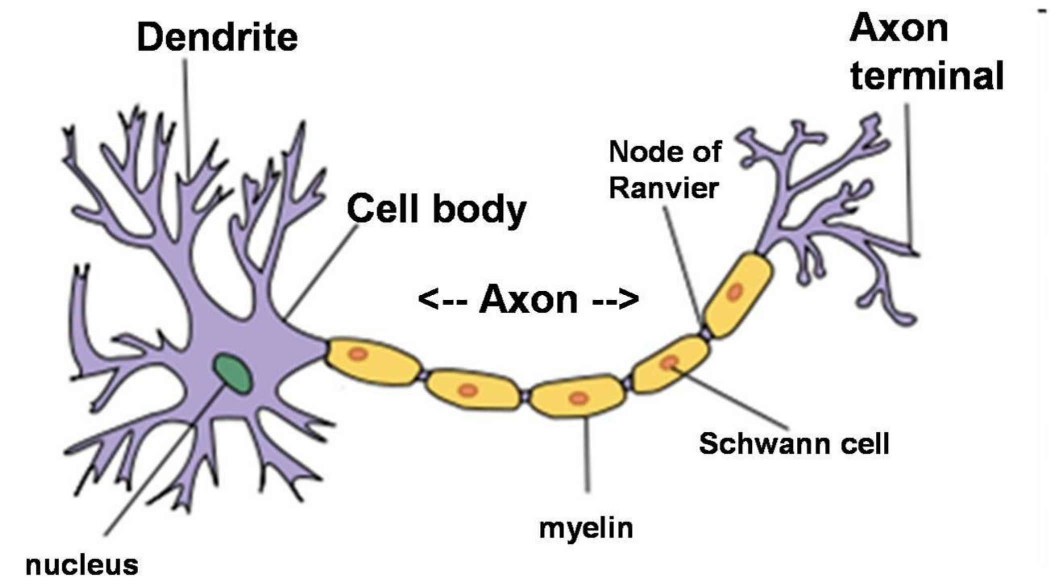


Figure 2.2: Anatomical structure of neuron.

Under the action of an electrochemical concentration gradient, neurons get activated causing flow of local currents. The neuroelectric potential generated by a nerve cell has two components, action potentials (AP) and postsynaptic potentials (PSP). AP is generated due to firing of postsynaptic neuron when the PSP reaches its threshold conduction level[46]. The APs of neurons have not enough contribution to the generation of scalp EEG or ECoG recordings due to its smaller distribution and shorter duration compared to PSPs.The summation of the PSPs of a large population of active neurons in the cortical surface is recorded as the EEG[47]. The very small voltage produced by a single nerve cell is difficult to measure accurately with latest technology.The electrical activities of neurons on the cortical surface have significant contributions in the EEG recordings.

2.1.2 Electroencephalography (EEG)

The electroencephalography is a non-invasive method for recording the neuro electric brain activity of cortical nerve cells of the brain [48]. It measures changes in electrical potentials at the surface of the scalp caused by the synchronized firing of action potentials in neurons near the cortical surface of the brain [49]. Hans Berger first introduced the EEG machine to the world in 1929[50] and later established electroencephalography (EEG) as a basic tool for clinical diagnosis of neurological disorders or diseases like epilepsy, Parkinson's disease, Alzheimer, sleep disorder, brain tumour, head injury, dementia etc. as well as in the development of BCI applications for neuro rehabilitation purposes[51-53]. He noticed that brain waves vary with the individual's state of consciousness. This revolutionary idea of Berger helped to create new branch of medical science called neurophysiology. There are two types of EEG, scalp EEG and intracranial. Scalp EEG signals are recorded from the scalp surface by placing surface electrodes with good mechanical or electrical contact. It is a non-invasive and hence painless method. But, electrocorticography (ECoG) is an invasive method for recording intracranial EEG by implanting special subdural electrodes on the cortical surface through surgery.The amplitude of brain electrical signal suffers severe attenuation from V range to μV range while propagating from cortex level to scalp surface.The bandwidth of spontaneous EEG is from below 1Hz to about 50Hz. EEG measurements are affected by the location of the recording electrodes.

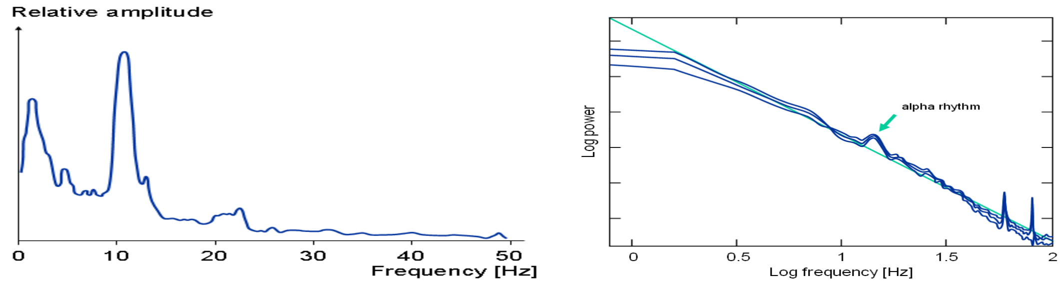


Figure 2.3: (a) Frequency spectrum of normal EEG (b) Frequency spectrum of human ECoG.

2.1.2.1 EEG Lead Systems

For EEG recording, surface electrodes are placed as per 10-20 electrode placement system, illustratedin [54, 55].

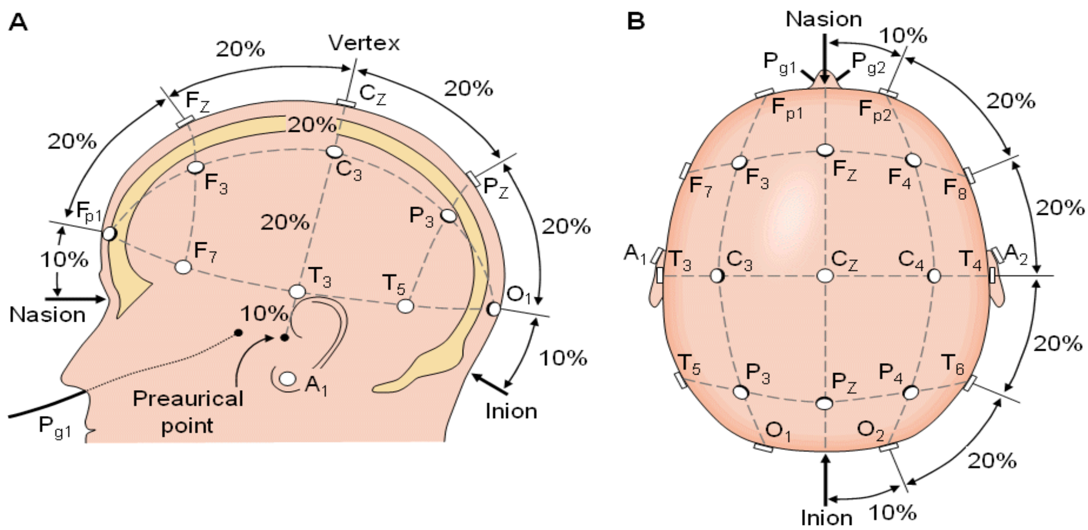


Figure 2.4:10-20 international standards for EEG electrode placement. (Jasper, 1958b).

Figure 2.4 shows location of 21 electrodes on thescalp.From these points, the skull perimeters are measured in the transverse and median planes. Electrode locations are determined by dividing these perimeters into 10% and 20% intervals. Three other electrodes are placed on each side equidistant from the neighboring points, as shown in Figure. In addition to these 21 electrodes, intermediate 10% electrode positions are also used. The locations and nomenclature of the electrodes are standardized by the American Electroencephalographic Society [56].

The placements of the electrodes are referred to as montage. Following montages are used for EEG recordings.

Bipolar montage: In this, each EEG channel represents the signal wave form that measures the potential difference between two adjacent electrodes [48, 57].

Referential montage: In this, each EEGchannel measures the potential difference between a specific electrode and a reference electrode whose position is not standardized [48, 57].

Average reference montage: In this, the common reference for each channel is derived by taking average value of the signals from the outputs of all the amplifiers[57].

Laplacian montage: In this, each channel measures the potential difference between an electrode and the weighted average of the neighborhood electrodes [57].

2.1.2.2 Frequency rhythms of the EEG signal

The human EEG signals are generated due to synchronous firing of billions of neurons. These signals contain wealth of information in its dynamical patterns. These dynamical patterns which contain meaningful information about brain activities can be decoded by examining frequency bands associated with different mental activities. The power spectrum of EEG signal contains five frequency sub bands“such as 0.5-4Hz(delta), 4-8Hz(theta), 8-13Hz(alpha), 13-30Hz(beta), and >30Hz(gamma)” [48, 57, 58].

Delta rhythm lies in the frequency sub band of 0.5 to 4 Hz characterized by highest in amplitude and the lowest in frequency.This rhythm indicates deep sleep, serious brain disorder etc.

Theta rhythm represents the frequency sub band of 4Hzto 8Hz with amplitude usually greater than 20 micro volts. This wave can be found under emotional stress like frustration or disappointment, and deep meditation.

Alpha rhythm has frequency range8Hz to 13 Hz, amplitude range 30-50 mV.

This rhythm is prominent in the posterior region (occipitallobe) of an adult with his eyes closed or is in relaxed state.It is considered as a normal waveform in most of the cases.This alpha rhythm normally disappears during attentive mental state, stressor eyes open. This frequency sub band when recorded from the sensorymotor areas is called mu rhythm.

Beta wave This EEG frequency rhythm belongs to the subfrequency band of 13Hz-30Hz. It has low amplitude and varying frequencies originated in the frontal brain area. It is generated when the brain is actively engaged active thinking, active attentions etc.

Gamma rhythm contains the frequency sub band from 30 Hz and above.This rhythm corresponds to performance ofcomplex cognitive task such as mental arithmetic task, music perception by musicians ,drivers' or pilot'scognitive load estimation and motor functions.

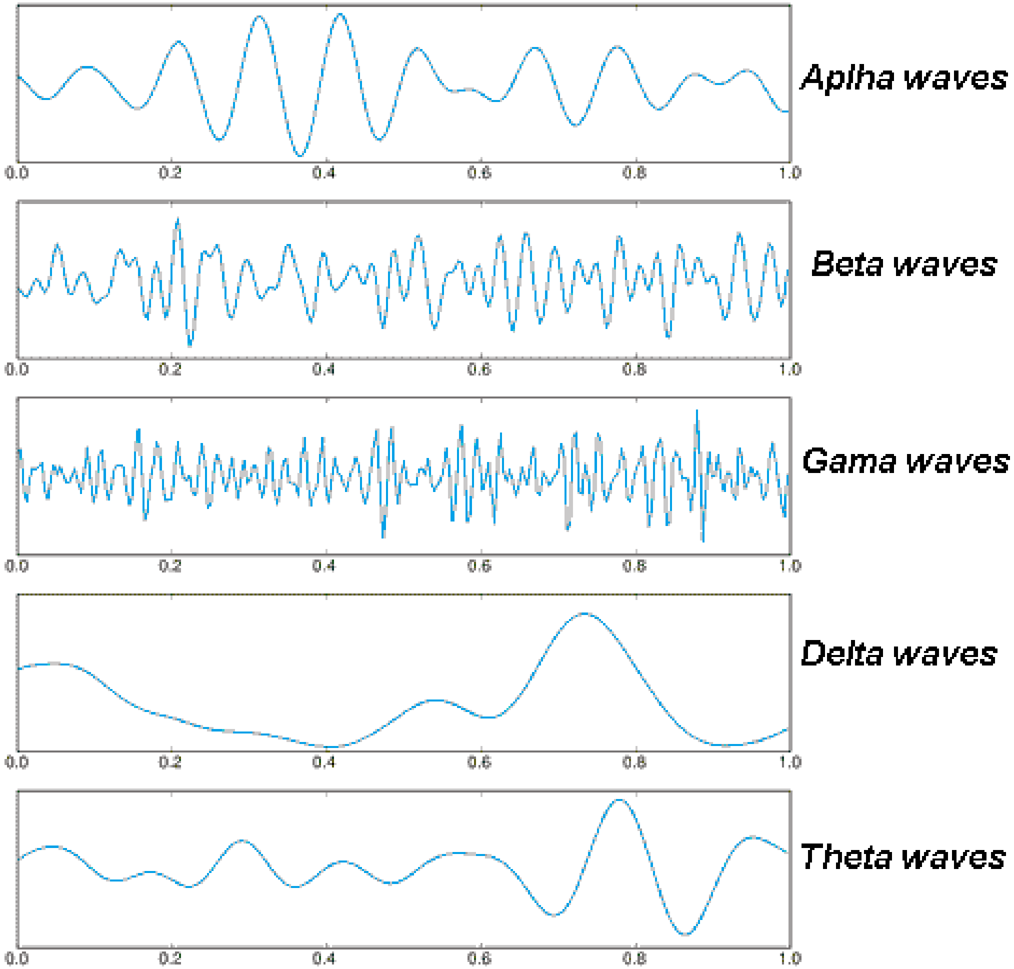


Figure2.5: Various types of normal EEG rhythm(Lotte, 2009).

2.2 Overview of EEG based BCI systems

2.2.1 Definition, Classification of BCI systems.

BCI is defined as a “communication and control channel that does not depend on the brain’s normal output channels of peripheral nerves and muscles” [9]. A BCI system relies solely on metal activity to control certain application for communication, transportation or any other need of the user. The messages and commands sent through a BCI are generated by decoding the EEG signals encoding the users intention. Thus, “it acts as a direct interface between human brain and a computer, robotor some other output device.”[59]. It is a pattern recognition system which works by recognizing the subtle dynamical patterns of brain’s electrical activity generated by external stimulus or induced by a specific cognitive task , namely mental arithmetic ,motor imagery. etc. For its operation, every BCI requires one specific control signal like P300, SCP, VEP, SSVEP, MI and non-motor imagery.

Depending upon the type of EEG control signal employed, current EEG based BCI systems are classified as ‘spontaneous’ EEG based and ‘evoked’ EEG based BCI systems. Spontaneous EEG are the measurement of continuous brain waves which are recorded without any specific stimuli such asmental task EEG, namely, motor imagery EEG, mental arithmetic EEG, mental letter composing EEG etc , while evoked EEG signal represents neuro-electric potential generated by specific stimuli, such as visual, auditory, somatosensory etc. Spontaneous EEG based BCI systems are called Asynchronous or self-paced BCI or independent BCI whereas evoked EEG based BCI systems are called Synchronous BCI systems or dependent BCI system as these are driven by some specific external stimuli.

Dependent and Independent BCI systems

A dependent BCI does not carry the messages or command through the brain’s normal outputs, but activity in these pathways is necessary to produce the brain activity that does carry the information. Good examples of this kind of BCI are most VEP, SSVEP based BCIs, etc. On the other hand, independent BCI’s do not need peripheral nerves and muscles neither to produce nor carry the messages that will control the application. Most BCI’s are considered to be independent, for example the system based on sensorymotor rhythms and Non-motor Imagery.

Synchronous and Asynchronous BCI systems

Synchronous BCI systems are cue based and time locked. They do not allow the subject to control it at any desired moment. Examples of synchronous systems are P300, SCP, and VEP. An asynchronous BCI system is self-paced i.e. spontaneous or not cue based. Examples of asynchronous BCIsystems are motor imagery(MI) or non-motor mental task(MT), audio imagery, spatial navigation imagery etc.

Motor imagery (MI) or Sensorymotor Rhythm Control (SRC) based BCI paradigm

BCI systems based on motor imagery (MI) or SRC make use of the decrease / increase in Mu (8-12Hz) and Beta rhythms (18-30Hz) in the EEG signals due to motor activity.

These are self-induced BCI systems whose operation is based on detection of changes in the brain’s mu rhythm or sensorymotor rhythm (SMR) (8-12Hz) in preparation for motor tasks (muscle movements). The act of actual physical movement or motor imagery i.e. mental rehearsal of body part movements causes change in mu rhythms. Pfurtscheller*et al*., (1993) from Graz, Austria, first explored this paradigm of BCI actively. They introduced the notion of ERDand ERS which revealed the non-phase-locked events in the frequency domain. These responses are typically detected by the change in power spectral density (PSD) which is used as feature for classifying these motor imagery tasks.

These ERD/ERS events helped to develop a five –class motor imagery BCI system (Obermeier*et al*., 2001). Pfurtscheller’s group demonstrated successfully that the hand grasp function can be restored in a tetraplegic patient using motor imagery.

For this, the tetraplegic subject is trained to learn to control their Mu and Beta rhythms in real time while performing imagined foot movement to generate a command for the paralyzed hand for grasping a cylinder.

Non-motor mental task (MT) based BCI paradigm

Several non-motor mental tasks, such as mental arithmetic of multiplication, mental letter composing, geometric figure rotation , visual counting invoke hemispheric asymmetry in the brain wave which is reflected in the EEG power spectrum or Autoregressive Coefficients of the EEG signals. Keirn and Aunon (1990) designed this BCI paradigm by selecting five non-motor cognitive tasks which can invoke brain wave symmetry. They demonstrated that it is possible to classify these non-motor cognitive tasks using frequency band power and a Bayes quadratic classifier. Milan et al developed a BCI system based on imagination of cube rotation, arithmetic tasks and word concatenating, together with motor imagery. The Oxford//London group has also developed a BCI system that uses motor imagery together with non-motor cognitive tasks. Curran et al, also from the same group investigated different cognitive tasks, such as motor imagery, auditory imageryetcto drive BCI systems. Canbrera and Nielsen investigated auditory imagery and spatial navigation imagery with higher spatial resolution than Curran et al to drive BCI systems.

Slow cortical potentials (SCP) based BCI

Slow cortical potentials (SCPs) are neurological phenomena which are reflected as EEG oscillations lasting several hundred milliseconds or several seconds. SCPs are used to drive this type of BCI paradigm. Using this neurological driver,Birbaumer*et al*., (1999, 2003) at Tubingen demonstrated this paradigm of BCI,known asThought-Translation Device (TTD). Using this BCI paradigm, ALSpatient can controla spelling device.Negative SCPs are typically associated with higher cortical activations, while positive SCPs are associated with reduced cortical activation.

P300 potentials based BCI

This BCI paradigm is driven by detecting an ERP (event related potential) well-known as P300 brain wave. Donchin and Farwell first designed a BCI through detecting a P300 wave. They designed a grid containing alphabetical characters flashing randomly. When the desired letter is flashed, P300 wave is produced to indicate selection. This BCI paradigm is highly reliable, but it has the drawback that it requires increased level of attention causing mental fatigueat a rapid rate than in other BCI paradigms.

Steady-State Visual Evoked Potentials (SSVEP)

SSVEP are brain responses elicited by a number of visual stimuli modulated at a certain frequency; these stimuli produce an oscillatory response in the EEG activity characterized by oscillations at the stimulation frequency and sometimes at harmonics or sub-harmonics of it. SSVEP responses are measured within narrow frequency bands around the visual stimulation frequency. The strongest SSVEP responses are easily recognized by analyzing the change in PSD in the EEG recorded over the primary visual (striate) cortex. The visual evoked potentials (VEPs) can be divided into three regions, namely, low (8-13Hz), medium (15-30Hz) and high frequency region(30-60Hz). This is based on their amplitudes. Low frequency region (8-13Hz) has highest amplitude while high frequency region (30-60Hz) has lowest amplitude. Sutter (1984) made the first attempt to design a BCI system based on SS-VEP. The training time required to achieve acceptable control of this BCI paradigm is little compared to MI or SCP based BCI paradigm.

2.2.2. Architecture of BCI systems

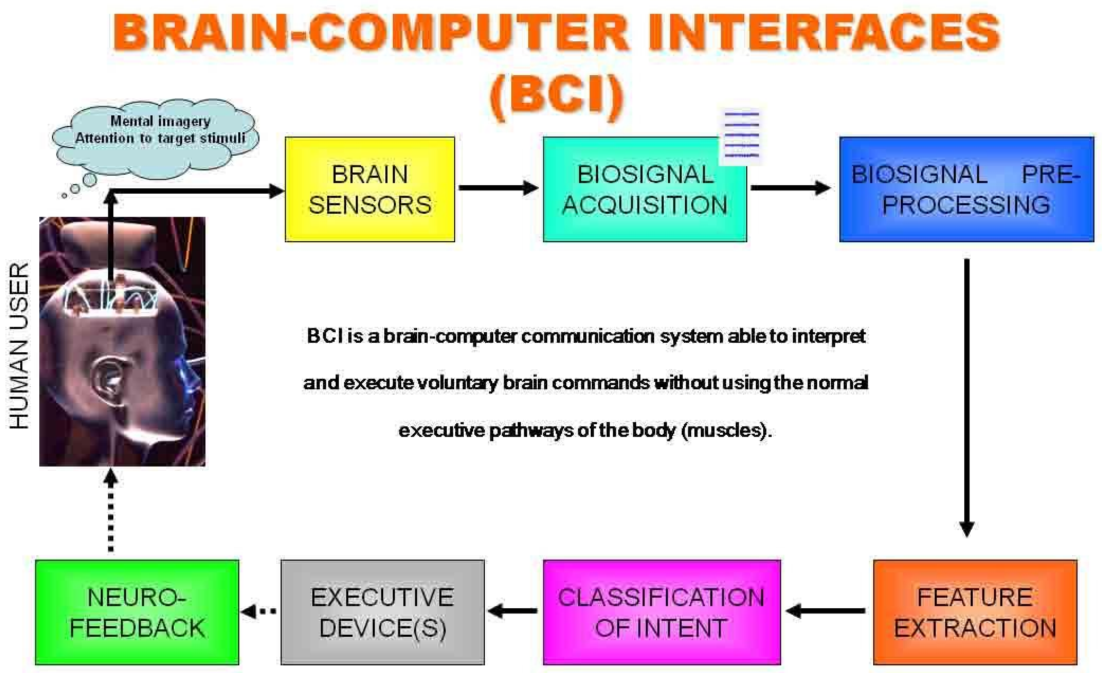


Figure 2.6 : Generalized block diagram representationof a BCIsystem

BCI as pattern recognition system consists of the following blocks as shown in Figure 2.7.

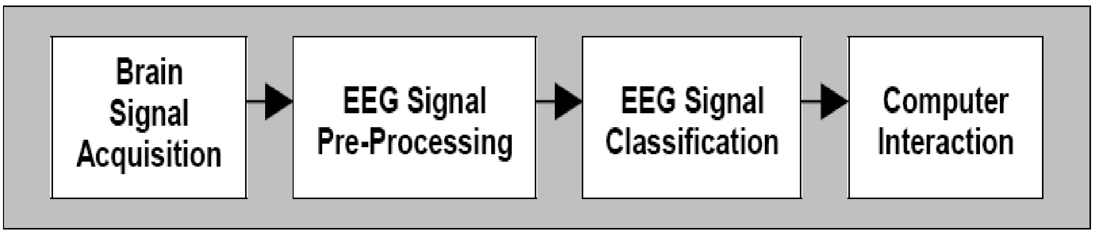


Figure2.7: Main Components of BCI

From Figure2.7, it is clear that three main aspects can be considered; namely, pre-processing i.e. spatial filtering, feature extraction and feature selection, and classification algorithm. Spatial filtering can be considered as preprocessing stage for increasing the poor spatial resolution of the signals. In the second stage, feature vector is formed by extracting the most discriminatory features from the most discriminatory EEG channels. The dimensionality of the feature vector should be reduced for faster and reliable classification. Lastly, an appropriate machine learning classifier is chosen for classifying the feature vectors into a class label. Classification algorithms aim to estimate the class of a feature vector. The most popular classification algorithms for BCI applications are LDA, SVM, Neural Network Classifier, Nonlinear Bayesian Classifier etc.

BCI systems are controlled directly by measuring brain electrical activity. Depending upon the technique of measurement, the BCIcan be either invasive or non-invasive. The non-invasive BCI approaches are based on EEG, MEG, PET, fMRI and optical imaging (NRIS). Non-invasive BCI systems are more popular than ECoGbased invasive BCIdue to health hazards involved in surgery. EEG is the brain's electrical activity recorded using electrodes attached to the scalp. It is the sum total of the potentials generated from a large population of cortical neurons activated during mental processes. It is in the micro volts range due to high attenuation by skull and scalp.

Current EEG based BCI systems use different types of neurological phenomena, such as VEP, SCP, P300 potential, Mu and/or beta rhythms with ERD/ERS, movement related potential (MRP), cognitive task related EEGs etc. Subjects induce brain activity pattern as per experimental protocol used in a particular BCI approach. The protocol to be followed by the subjects can be imaginary movements, focusing on flashing characters on a screen and so forth. The brain activity is picked up by placing surface electrodes on the scalp surface as per 10-20 electrode placement system. The recorded signals havingrange of microvoltsare normally referenced to the left and/or right mastoid.

Pre-Processing

The digital EEG data normally contains a lot of noise and artifacts such as 50/60 Hz power line interference, baseline drift, Electro-Oculogram (EOG), Electro-Myogram (EMG), Electrocardiogram (ECG) etc. Simple frequency specific filtering is not sufficient to remove these noise signals due to their overlapping spectral characteristics. Besides, the spatial resolution of EEG is very poor. Due to this, sophisticated spatial filtering methods such CSP, principal PCA and ICA are popular in the pre-processing stage to reduce these noise signals and improve its spatial resolution. This stage aims at cleaning and denoising the recorded the digital data for enhancing the relevant information embedded in the signals [60].

Feature Extraction

The neuro physiological phenomena driving an EEG based BCI are represented by certain set offeatures. So, feature extraction describesthe control or driver signal by few numerical values called "features" [60]. Feature vectors are constructed from the feature values[61].

The feature vector represents the signal pattern in reduced dimension.Classification rate depends on the optimality of the EEG data representation in the feature space and appropriate choice of classifier. It is essential to extract the task specific complex spatio-temporal oscillatory patterns accurately from the weak and highly noisy multi-channel EEG channel data recorded from the scalp. The EEG features depend on the BCIparadigm.Extracting optimal features reduce the processing time and improves classification accuracy. The choice of both pre-processing as well as feature extraction depends on the type of EEG signals and the number of electrodes.Most of EEG feature extraction approaches estimate the signal's energy distributed in the frequency domain, time domain, and time-frequency domain.

Classification

EEG based BCIs are viewed as pattern recognition which involves feature extraction and classification. Classification is the process ofassigninga class label to a feature vector extracted from a signal. For EEG based BCI, a class corresponds to a specific mental task. This step can also be termed as "feature translation". Classification of humanEEG signals is a crucial task since the signals have extremely complex characteristic, highly noisy and contaminated with artifacts [62, 63].

2.2.3 Applications of BCI systems

BCI is an emerging technology with tremendous potential to be used for rehabilitation purpose. The possible practical applications of BCI are:

1. BCI controlled wheelchairs can provide mobility assistance to the elderly and disabledpersons.
2. BCI systems can be employed for restoring function of limbs.
3. BCI systems can be employed for writinge-mails or documents.
4. BCI systems can be employed for creating Virtual reality(VR) applications .
5. BCI systems can be employed for remote monitoring of pilot's or driver's attention.
6. BCI systems can be employed for neurofeedback therapy.

2.3.4 BCI Features and its properties

Efficient BCIdesign must consider few critical properties of its features as mentioned below

- **Noise and outliers**: Poor signal to noise ratio of EEG signal causes BCI features to be noisy and full of outliers.
- **High dimensionality**: In BCI systems, the extracted feature vectors often have dimensionality due to use of severalEEGchannels.This may cause over fitting of classifier.
- **Time information**: BCI features should encode the temporal dynamics of the brain activity patterns as reflected in the EEG.

2.3.5 Survey of EEG features in MT based BCI paradigm

EEG basedBCI systems have been designed using a great variety of features.These features can be classified as linearor nonlinear, time domain or frequency domain or time-frequency domain. FFT based estimation of Band Power, Power Spectral Density frequency domain features while time domain methods include linear time series model like autoregressive (AR) model, adaptive or multivariateAR models. These conventional methods of EEG feature extraction assume linear and stationary nature of the brain activity and hence do not have the capacity to capture the nonlinear and non-stationary nature of our brain dynamics. As a result, a plethora of non-linear feature extraction methods based on Wavelet Transform (WT), Wavelet Packet Transform (WPT), Empirical Mode Decomposition (EMD), Hilbert-Huang Transform (HHT), Fractal and Chaos theory etc. have been developed by the BCI research community.

Frequency-Domain Representations

In the frequency domain representations, EEG signals are described in terms of oscillatory components and their time evolutions. It has become a standard practice in neuroscience to describe brain activity in the standard frequency bands: δ (0-4Hz), θ (4-8Hz), α (8-13Hz), β (13-30Hz), γ ($\geq$ 30Hz).

The computation of amplitude, power, energy or phase in this domain has yielded powerful insights. Consequently, it has been widely used as benchmark method for BCI feature extraction.

Though this is the most common method for feature extraction, it has the limitation that it assumes the signal to be generated from a linear and stationary process. Due to this assumption, it can capture only frequency domain information and ignore meaningful and relevant patterns that changes in time and space.

Power Spectral Density (PSD): Power Spectral Densities (PSDs) are popular features used in MT-based BCI systems. This is a frequency domain feature which estimatesthe signal power density across a spectrum of frequencies. The PSD describes the distribution of signal's power over frequency. This is estimated as the squared modulus of the FFT of asignal, scaled by a proper constant term.The plot of the average PSD spectrum shows difference at frequency sub bands.PSD values have been utilized as useful features for classifying different motor imagery tasks. It has been well established that frequency band power in the mu band (8-12Hz) from the sensorymotor areas decreases under the performance of left and right hand motor imagery tasks. PSD values have been used as main features in SSVEP based BCI systems. In general, fast Fourier transforms (FFT), wavelet transforms (WT), and autoregressive (AR) model co-efficients have been used to estimate the PSD features.

Auto Regressive (AR) Model based feature: Parametric feature extraction based on Autoregressive (AR) model is very popular in the feature extraction stage of any BCI system.An autoregressive model (AR) is a linear time series model that can be viewed as an IIR filter. In an AR model, the current term ofa timeseries is estimated by a linear weighted sum of itsprevious terms. The weights are the auto regression coefficients.A number of techniques exist for computing the AR coefficients like least square error (LSE), Burg's method, the geometric lattice method etc. In this approach, the EEG time series is fitted with a fixed orderAR modeland the coefficients are usedfor subsequentclassification.

Multivariate Autoregressive (MVAR) model based feature: EEG signals are inherently multidimensional in nature. The scalar AR model is applicable to one channel EEG time series only. Due to this, scalar AR model coefficients are weak in representing the cross channel dependency of the large scale networks in the cerebral cortex.MVARmodel overcomes thisweaknessof scalarAR modelbycapturing thefunctional couplingofthelocal field potentials from different brain areas.

Hjorth parameters: These are three time domain parameters, namely,activity, mobility and complexityfor characterizing the temporal dynamics of EEG signal in terms of amplitude, time scale and complexity.Activity is a measure of the mean power i.e.varianceof the signal.Complexity measurethedeparturefrom the sine waveand is expressed as the number of standard slopes actually seen in the signal during the average time required for one standard amplitude, as given by the mobility.

Time -Frequency Transforms: Oscillatory EEG patterns contain the meaningfuldynamic brain activity pattern indicating different information processing states of the brain. Traditional spectral analysis methods based on Fourier Transform are not adequate to characterizethese subtle oscillatory patterns buried in the noisy EEG, since the underlying neural processes generating them are inherently nonlinear and dynamic. The transient changes in the power or peak frequency of EEG waves can provide information of primary interest. The non-stationary nature of the EEG signals makes it necessary to use methods which are able to quantify their spectral content as a function of time. The investigation of features in the EEG signal requires a detailed time frequency analysis of the signal. Time-frequency methods decompose the EEG signals into a series of frequency bands and the instantaneous power is represented by the envelope of oscillatory activity, which forms the spatial patterns for a given electrode montage at a time-frequency grid.In this transform, a mother wavelet and its time shifted and scaled versions are employed as basic functions to expand the signal.Though, both wavelet transform and wavelet packet transform based methods have been used by previous researchers, these are good for linear and stationary signals. Considering the nonlinear and non-stationary nature of the EEG signals,EMD or HHT is most suitable for time frequency analysis of the signals.EMD decomposes a signal adaptively into finite number of IMF functions. Hilbert-Huang Transform (HHT) combines EMD with Hilbert Transform to produce Hilbert spectra of a nonlinear and non-stationary signal.

Common Spatial Filter (CSP): Common Spatial Pattern(CSP) has been widely used as spatial filter for classifying motor imagery EEG in BCI system. These filters obtain a new time series having better variance and discriminatory power than the original EEG signal. Employing simultaneous diagonalization of two covariance matrices, this method focus on activities with particular spatial distributions through combining data from multiple channels.

Information theory based feature: Different information theoretic measures have been proposed for quantifying complexity of EEG between the different regions of the brain. Entropy, one such measure derived from information theory, can discriminate different cognitive states. Through such complexity measures, the dynamics of functional integration in the brain behind each cognitive state can be assessed without assumptions about the underlying system. Wang et al have analyzed EEG data records related to five different cognitive mental tasks by using approximate entropy as a feature. Manoilov et al estimated the power spectral entropy (PSE) of EEG signals and classified them for BCI using NN as classifier.

Phase Synchronization based feature: Elly andPatrick employedphase synchronization measure as novel discriminatory feature for classifying EEG signals in mental task (MT) based BCI paradigm. Coupled dynamics of these functionally connected brain regions estimated through phase synchronization are characterized by distinct spatiotemporal patterns The statistical features derived from this spatiotemporal patterns have been used for accurate classification of the single-trial imagined movements.

Morphological /Fractal feature: The fractal description of EEG signals can be a useful tool for feature extraction. Fractals are objects which possess a form of self-scaling. Parts of the whole can be made to fit the whole in some way by shifting and stretching. Fractal features represent the morphology of the signals. These morphological differences can be picked up and used for several applications including feature extraction in BCI.

Chaos theory based feature: In this approach, the EEG signals are assumed to show chaotic behavior as they are generated by the chaotic dynamics our brain. EEG signals are classified by statistical and probabilistic classifier models, such as HMM based on chaotic parameters such as entropy, Lyapunov'sexponent, fractal dimension , correlation dimension,etc.

Independent Component Analysis (ICA): This is apromising statistical technique employed to linearly transform the observed multivariate random data into statistically independent from each other, not merely uncorrelated. The output of ICA is a set of maximally independent components. Multi-channel EEG signals recorded at different scalp locations are assumed to be linear mixtures of various artifacts and source components. The decomposition of such signals using this method is very useful since it extracts the significant source components related to a particular mental task. ICA based analysis, can be used to extract new features from the selected independent components that carry information about the task. Thus, it has the potential for improving theclassification score of the EEG patterns associated with specific mental task.

Source Reconstruction Methods: These methodscapture the information about the electrical activities of the cortical sourcesby solving the inverse problem. EEG inverse solutionsreconstruct the underlying neuroelectric activityby using scalp recorded EEG measurements. These methodscompensatethe distortion and smearing effect due tothe lowconductivityand volume conduction effectandimprove the spatial resolution of the EEG signals. Thetwo approaches available to solve the EEG inverse problemareparametric and imaging approach. Recently, feature extraction algorithms based on EEG inverse solutions have become very promising in enhancingthe performance of BCI systems.

Kalman Filter basedfeature: It is a tool for estimating the state of a dynamical system recursively using the input and output values of the system. This new feature extraction algorithm aims to estimate the parameters of Kalman Filter from the measured EEG data. The linear parameters of the Kalman filter are used as new feature vectorshowingbetter classification results than commonly used features such as AR coefficients,frequency bands (BP) and wavelet coefficients.

Higher Order Statistical/Polyspectral feature: These provide a way of describing biosignals such asEEG revealing the non-Gaussian nature of the signals. ShenMinfen et al. estimated the parametric bispectra under different functional states of brain and applied higher-order statistics (HOS) to investigate the non-Gaussianity aspect of the EEG signals for extracting information beyond second order or power spectra.

2.3Classification Algorithms

BCI systems can be viewed asmachine learning enabled pattern recognition systems [64, 65]. Such systems translate electrical activityof braininto a command for controllinga computer or some output device like a wheel chair or prosthetic arm. Feature extraction and classification are twoimportantalgorithms fordesigningof any real time BCI system. Classification algorithm aims at assigning a feature vector into one of a given number ofcategories [65, 66]. The feature vectoris termed as an instance and the categories are termed as classes. The feature vector contains feature values describing one aspect of the total characteristics of the instance. Classification algorithms assign class labels to the feature vectors extracted from the specific experimental data set. The performance of a BCI depends on the discriminatory power of the feature vector and the type of classifier employed.

A largenumber of classification algorithms have beenemployed in BCI research. These algorithms are used to classifypatterns in theEEG [67]. These classifiers are categorized as“linear classifiers, neural network classifiers, nonlinear Bayesian classifiers, nearest neighbor classifiers and support vector machine classifiers.”F.Lotte et al,presented a detailed review ofthe classifiersused inBCI research [61]. While performingpattern classification, classifiers face following two main problems:

The curse -of - dimensionality:

The amount of input data points needed to properly describe a class rise exponentially with the dimensionality of the feature vectors"[64, 68]. This is a major problem in BCI design which arises when the number of training data is small compared to the dimension of the feature vectors. Under this condition, the classifier gives poor results. So, it is recommended that the number of training data samples per class should be at least five to ten times of the dimensionality of the feature vectors [69, 70]. Due to inherent multidimensional nature of the EEG signals, the dimensionality of the feature vectors is high and the input training data samples is low.

The Bias-Variance tradeoff:

It has been seen that classification error depends upon three possible sources:

Noise: It represents the unwanted part within the signal which does not carry any information. This component of error is not reducible by the classifier.

Bias: It is caused due to the divergence of the estimated mapping value from the best mapping value.

Variance: It reflects the sensitivity of the classifier model to the training set used. So, high classifier performance means high classification accuracy which further needs low classification error. For obtaining lowest classification error, both the sources of bias and variance must be low. But unfortunately, there is a natural tradeoff between bias and variance. This is due to the fact that stable classifiersproduce high bias and low variance, whereas unstable classifiers producelow bias and high variance. There are several stabilization techniques forreducingvariance. Two such important techniques are (a) regularization (b) combination of classifiers [71].

Types of Classifiers:

Generative–discriminative: Generative classifier are informative classifier which derives the class model from the training dataset and predict the class of an unknown feature vector through computing the likelihood of each class. The class which is most likely is chosen. Example of such classifier is Bayes quadratic classifier. Discriminative classifier such as Support Vector Machines (SVMs) learns the classification model from the training dataset for classifying a feature vector directly [72,73].

Static-dynamic: In static classifier, e.g. Multilayer Perceptrons (MLPs), the learned input-output mapping function is static i.e. time invariant. Due to non-recurrent architecture, these classifiers cannot take into account temporal information during classification. On the other hand, dynamic classifiers, e.g. Hidden Markov Model, learn the input-output mapping function which is dynamic i.e. time varying.Due to its feedback i.e.recurrent architecture, these classifiers classify a sequence of feature vectors and thus able to catch the temporal dynamics [74].

Stable-unstable: Stable classifiers, e.g., Linear Discriminant Analysis (LDA), have a low complexity or capacity [74]. These are called stable since their performance is not affected by small variations in the trainingdata set. On the contrary, unstable classifiers, e.g., Multi-Layer Perceptron (MLP), have high degree of complexity. Due to their high complexity, the performance of such type of classifier is sensitive to small variations in the training set [71].

Regularized-unregularized: Regularization technique is employed for controlling the complexity of a classifier and thus help in avoiding overtraining of the classifier. A regularized classifier offers better generalization performance compared to a non regularized classifier and is more robust with respect to outliers [64, 65].

CHAPTER 3

LITERATURE SURVEY

3.1 Survey of EEG feature extraction methods for BCI applications

The prime motivation of our research work stems from the need of searching better EEG feature for improving the classification performance in generalMT based BCI system. Feature which can identify unknown but meaningful patterns in the EEG dynamics has better discriminatory power. Aiming at this goal, a comprehensive survey of the contemporary literature on the feature extraction methods was carried out before the actual start of the research.

In a seminal work, **Keirn and Aunon** (1990)[22, 75,76] first explored the use of non-motor mental tasks (MT) in MT based BCI paradigm. They designed the experimental protocol and acquired the EEG signals related to five non-motor imagery mental tasks such as mental arithmetic task, mental letter composing task etc. These tasks were chosen as these could invoke hemispheric brainwave asymmetry. In their experimental study, six electrodes were placed over the left and right central, parietal and occipital areas of the cortex. They classified different pairs of these non-motor mental tasks using power spectral asymmetry ratio and Burg AR model coefficients with Bayesian Quadratic (BQ) classifier. They demonstrated the necessity to use (δ, θ, α and β) bands and AR features to enhance the classification results. First, they divided the EEG signals into ¼ s segments and computed 36 AR features for each segment. BQ classifiers were employed to classify different combinations of task pairs for five subjects. They achieved highest average classification accuracy of84.6% across all task pairs and over all the five subjects. Their study showed that the AR method was superior and could differentiate between two mental tasks for each subject.

C.W.Anderson *et al* (1998)[78-81] employed MVAR model of EEG for classifying a set of five non-motor mental tasks using neural network (NN) classifier. They continued the work of Keirn and Aunon using the same EEG data set and derived both scalar AR model as well as multivariate AR model from the raw EEG signals without employing any decomposition technique. They used scalar and multivariate AR coefficients in addition to Karhunen-Loeve (KL) transform and correlation matrix eigenvalues as input features.Using 80% of the data set to train the feed forward back propagation neural network (BPNN) , they were able to achieve average classification accuracy ranging from 86.1% to 91.4% for each subject. They achieved highest average classification accuracy of 91.4% for binary classification using multivariate AR coefficients and feed forward neural network trained.

R. Palaniapan *et al* (2002)[84] (2005)[85] proposed a new mental task (MT) based brain-computer interface(BCI) paradigm using fuzzy ARTMAP(FA) neural network classifier. based on the five non-motor EEG data set acquired by Keirn and Aunon. They employed two different spectral analysis methods to obtain the frequency content(i.e. PSD) of the EEG signals from 0 to 50 Hz. They computed power spectral density (PSD) values using two methods: Wiener –Khinchine(WK) theorem and Burg's sixth order autoregressive(AR) method. In the WK method, they applied two different lag windows ,Turkey and Parzen.Each EEG signal from a particular mental task is segmented with a 0.5s window, i.e. for a length of 125 data points giving 20 patterns for each mental task per session. . A total of 200 patterns are obtained for the five mental tasks performed by each subject across two sessions

Feature vectors were constructed from the concatenation of 300 PSD values in the range of 0 to 50 Hz computed from the six channels. They employed FA classifier instead of any parametric classifier or types of NN architecture due to its less training time and ability for incremental learning. FA classifier was trained to classify three mental tasks where each mental task is represented by the 300 PSD values(from 6 channels of EEG).PSD values from half of available patterns chosen randomly were used for training and PSD values from the remaining half of the patterns were used for testing the FA.

They obtained average classification accuracy of 94.43% using WK-Parzan method, 93.67% using AR, 91.08% using WK-Tukey method for four subjects.

D. Garrett *et al*(2003)[86] employedAR model coefficients as features with one linear (linear discriminant analysis,LDA)and two nonlinear classifiers(neural networks and support vector machines) for classifying spontaneous EEG during the performance of five non-motor mental tasks. They used the most popular and benchmark EEG data set acquired by Keirn and Aunon. They computed sixth order autoregressive (AR) models for each channel independently for the data within each segment after discarding the segments containing eye blinks. Therefore, data in each segment were reduced from 750 dimensions (125 samples x6 channels) to 36 dimensions(6 AR coefficients x6 channels).The classification algorithms were trained using full set of five trials and tested using remaining five trials. They obtained classification accuracies of approximately 45%, 53%, and 52% using LDA, neural networks and SVM respectively .They adopted genetic algorithm (GA) based feature selection and showed that nonlinear classifiers produced slightly better classification results than linear classifiers.

N.J.Huan *et al* (2004)[87]investigated the performance of different mental task combinations for a two-state brain-computer interface(BCI)design for individual subjects using Keirn and Aunon's data. They employed six different feature extraction methods for extracting features from five non-motor mental task EEG signals for four subjects. In all the methods, they derived AR mode of order six from 125 data points. They employed multilayer perceptron neural network and linear discriminant analysisas classifiers and obtained classification accuracy of 93.10% and 97.00% respectively.

N.J. Huan *et al* (2005)[88]classified non-motorEEG signalsusingfeaturesfromfixed AR and adaptive ARmodels. Five different mental tasks from 4 subjects were used in the experimental study and combinations of 2 different mental tasks are studied for each subject. Four different feature extraction methods were used . Multilayer perceptron (MLP) neural network (NN) trained by the back propagation (BP) algorithmwasused to classify these features into the different categories representing the mental tasks. The best results for FAR was 92.70% while for AAR was only 81.80%.

N.Y. Liang *et al* (2006)[89] subdivided the EEG signals into ½ s segments with an overlap of ¼ s between adjacent segments. In the next step, they computed a sixth order AR model for fitting into half-second EEG segments and constructed 36 -D feature vector (6 channels X6 coefficients) from the coefficients of the model for each segment estimated through the Burg method. Using the constructed feature vectors , they classified five non-motor mental tasks using Keirn and Aunon's benchmark EEG database[22] employing three classifiers , namely BPNN, SVMs and Extreme Learning Machine (ELM). They obtained highest testing accuracy(%) of 55.76 , 60.86, 63.96 using BP , ELM and 1-against all SVM classifiers respectively without smoothing their outputs and 78.24 , 86.70 and 85.75 respectively with smoothing their outputs. They showed that ELM needs about 1 to 2 orders of magnitude less training time compared with SVMs and BPNN. The classification performance of ELM was compared with a BPNN classifier and Support Vector Machine (SVM) classifier. ELM gave similar classification accuracy,but its training time was less compared with both SVM and BPNN.

C.W. Anderson *et al.* (2006)[90] employed generalized SVD to separate multichannel electroencephalogram (EEG) into components for filtering out artifacts. Short-time principal component analysis of time-delay embedded EEG was used to represent windowed EEG data for classifying five mental tasks using committee of decision trees.

R. Palaniapan (2006)[91] improved the performance of MT based BCIby computingspectral power and asymmetry ratio at the gamma band (24-37Hz) in addition to the lower frequency bands.Using the extracted featurevectors, he trained Elman neural network (ENN) with resilient back propagation algorithm to classifydifferent combinations of two mental tasks.They segmented EEG signal for each mental task into 20 segments with length 0.5 s, so each EEG segment was 125 samples in length. The EEG segments were filtered twice using Elliptic filters into five frequency bands such as 0-3 Hz(delta), 4-7Hz(theta), 8-13 Hz(alpha), 14-20Hz(beta) and 24-37Hz(gamma). They computed band power (BP) and asymmetry ratio features from each of these five spectral bands. Each spectral band gave 15 features (six band powers and nine asymmetry ratios), thereby giving a total of 75 features. They compared the performance of 60 features from four bands excluding gamma band with 75 features from five frequency bands including gamma band.

The classification results showed that with the inclusion of additional gamma band features , less iterations were required which further resulted in reduced training time. Though there would be some additional computation required for the gamma band features, it was negligible (for both the methods, the computation of features for each EEG segment took 0.046 s). The differences in ENN testing times for Methods A and B were also negligible (both

C.W. Anderson *et al* (2007)[92]embeddedEEG signals in time and applied short time principal component analysis (STPCA) for capturing the EEG variations in space and time. They extracted features from this representation which were classified via a standard linear discriminant analysis (LDA). They tested their approach on two EEG datasets – five task dataset i.e. non-motor mental task (MT) EEG dataset acquired by Keirn and Aunon.and three task dataset i.e. BCI competition III, motor imagery Dataset V. The training trials were randomly partitioned into 80% for constructing the classifier and 20% for evaluating it. With 5- fold cross-validation , they obtained classification accuracy of 80% with one subject from the five task dataset and 55% accuracy from the three task dataset.

L. Zhiwei and Minfen (2007)[93] proposed a method for EEG classification based on wavelet packet entropy (WPT) and support vector machine(SVM).They extractedwavelet packet entropy (WPT)featuresforclassifyingfive non-motor mental task EEG signals usingSupport Vector Machine (SVM) with three kernel functions i.e. radial basis function (RBF) , linear and polynomial . They divided the EEG signals into 1 s segments that overlapped by 0.8s and decomposed each EEG segment of each channel using seven level wavelet packet decompositionwith db4 and extracted four spectrum bands (delta, thetas, alpha, and beta). Feature vectors were constructed through computing the entropy values from these bands and the resulting entropy feature vectors were used for training and testing a SVM classifier. They tested their approach on the benchmark five task EEG dataset acquired by Keirn and Aunon . SVM classifier was trained to discriminate between 2-class , 3-class , 4-class and 5-class. With their proposed approach, they achieved an average classification accuracy of 76.3% for five class for subject 1 and 68.5% for subject2. The best accuracy was obtained using linear kernel where averaged classification accuracy is 93.0% for subject1 and 87.5% for subject2 for 2-class classification. For 3 class classification, they obtained average classification accuracy of 91.4% for subject 1 using SVM with RBF kernel.

P.F. Diez ***et al*** (2009)[94] proposed a feature extraction approach based on empirical mode decomposition (EMD) for classifying five mental tasks from the benchmark mental task EEG database acquired by Keirn and Aunon. They adopted univariate approach and extracted six different time domain features such as Root Mean Square (RMS), Variance, Shannon Entropy , Lempel-Ziv Complexity Value , and Central and Maximum Frequencies from the IMFs of all the channels for constructing high dimensional feature vectors of 180 dimensions. They employed Wilks' lambda parameter for reducing the feature dimension through selecting the most important variables. Finally, they employed two classifiers such as Linear Discriminate Analysis(LDA) and Neural Network(NN) classifier for classification and achieved average classification accuracyof91 ±5% and 87 ±5%respectively. The high classification accuracy concludes the effectiveness of EMD to process nonstationary and nonlinear signals such as EEG . Their test results offer better performances than other traditional methods ,like power spectral analysis.

M.F.Kaleem ***et al*** (2010)[95]proposed a novel feature extraction approachfor classifyingEEG signals using EMDand Teager energy operator.They decomposed the EEG signals (6 EEG signals per task per subject, for a total of 5 tasks over multiple trials) into their constituent oscillatory modes, called intrinsic mode functions, and separated out the trend from the signal. Teager energy operator was used to compute the average energy of both the detrended signal and the trend. The average energy was used to construct separate feature vectors with a small number of parameters for the detrended signal and the trend.Based on these parameters, one-versus-one classification of mental tasks was performed using Linear Discriminant Analysis.They tested their approach on the benchmark mental task (MT) EEG database acquired by Keirn and Aunon. Using both feature vectors, they achieved an average correct classification rate of more than 85% which demonstrated the effectiveness of their method.

Rifai Chai *et al* (2012) [96] classified three non-motor imagery mental tasks for BCI control of wheelchair using EEG signals. They employed genetic algorithm (GA) based optimization of an artificial neural network (ANN) for classifying any combination of three tasks out of six non-motor imagery mental tasks. They obtained the best triplet combination out of six non-motor imagery mental tasks for generating wheelchair steering commands effectively.For the feature extraction process, they computed spectral band power (BP) from the four EEG rhythms using FFT of each one second segment of each channel. From the computed band power of each rhythms, they computed asymmetry ratios for all the channels and constructed the feature vectors . They utilized a 3-layer feed forward neural network with one hidden layer and log-sigmoid function as activation function. For avoiding the problems of convergence to local minima and sensitivity of conventional gradient- descent (GD) technique, they utilized genetic algorithm (GA) as global search method for optimizing the training parameters of the back-propagation algorithm based feed-forward neural network . They selected the number of hidden neurons corresponding to highest classification accuracy using GA method. . They obtained average classification accuracy ranging from 76% to 85% and information transfer rate (ITR) varying from 0.5 to 0.8 bits per trial.

A.Gupta andR.K.Agarwal (2012) [97] classified mental tasks in EEG based BCI system by extracting the task relevant features from EMDbased decomposition of the non-linear and non-stationary EEG time series. The relevant and more discriminatory features which provide maximum classification accuracy were selected using ratio of scatter matrices, Chernoff distance measure and linear regression. These features decrease the training time of a classifier and provide better generalization capacity. The performance of different mental task using different measures used for feature selection is compared and evaluated in terms of classification accuracy. Experimental results show that there is significant improvement in classification accuracy with features selected using all feature selection methods and in particular with ratio of scatter matrices. Linear regression performs better in comparison to other two methods and Support Vector Machine requires lesser number of features to build the model.

D.Vidaurre *et al* (2013)[98] proposeda new feature extraction method based on sparse autoregressive features for multiple signal classification. In this method, previous estimation of the order of the AR model is not required, the order is selected in a data driven way.They used a maximum AR order of 15 . They tested their approach on the benchmark non-motor mental task EEG data set collected Keirn and Aunon.They divided each trial into ten one-second sub periods of 250 sample points to get ten instances per trial. They considered each group of five trials as a different data set andtrained individual models for each subject. They got 100 instances for training and testing each experiment. They employed different classification algorithms such as LDA, LR and SVM for ten pairwise i.e. binary classification problems . With SVM classifier, they obtained highest classification accuracy of 99% for the best pair of activities between mental letter composing and geometric figure rotation task.Their sparse autoregressive features outperformed the performance of the state -of -the art AR features. With regularization, their approach can be applied to data sets with few instances.

M.Tolic and F.Jovic (2013)[99] presented a method based on discrete wavelet transform (DWT) and Backpropagation Neural network (BPNN) for classifying five differentnon motor EEG signals.They decomposed each EEG signal into ½ s segments and subsequently decomposed each half second segment by applying DWT . For the decomposition , they used transform level 5 and 'db4' (Daubechies4).Discrete wavelet energy was computed at various decomposition levels generated from this decomposition.Feature vectors were constructed from these computed wavelet energies and subsequently fed to a neural network (NN) classifier for its training and testing.. They tested their feature extraction approach on Keirn and Aunon's benchmark five task EEG data base. They achieved mean 90.75%classification accuracy for all five tasks and 99.87%mean classification accuracy for binary classification between any task and baseline. Mean classification accuracy for the recognition of all five tasks was 90.75% and mean classification accuracy for the recognition of two tasks (baseline and any other mental task) was 99.87%. The same procedure was also used on the motor imagery dataset. A mean classification accuracy of 68.21% suggests alternative methods of feature extraction for motor imagery tasks.

M.Hariharan *et al* (2014)[100]employedStockwell transform (ST) for investigating the time-frequency characteristics of EEG signals of different mental tasks. From this time-frequency representation, they extracted featurefor classifyingEEG dynamics underlying different mental tasks. They tested their approach on Keirn and Aunon's benchmark EEG database. They employed three classifiers such as k-means nearest neighbourhood (kNN), linear discriminant analysis (LDA) and support vector machine (SVM) for testing the strength of the proposed features. They achieved average classification accuracy ranging between 84.72% and 98.95% for multiclass (five mental tasks) problems. They employed ten-fold cross validation method for getting consistency of the classification result.

Xiaoou Li *et al* (2014)[101] proposed classification of mental task EEG signals using wavelet packet entropy and Granger causality as features. They investigated employing multiple kernel learning support vector machine(MKL-SVM) with gradient descent optimization for classifying different mental and cognitive tasks in EEG- based brain computer interface (BCI) systems. For the SVM,they defined thekernel function as linear combination of polynomial kernel and radial basis function.Their experimental results showed better classification performance compared with the SVM based on a single kernel. For mental tasks, they obtained average accuracies of 99.20%, 81.25%, 76.76%, and 75.25% for 2-class, 3-class, 4-class, and 5-class classification respectively. Their results indicates the effectiveness of their approach for mental task classification in EEG based human-computer interaction (HCI).

A.Gupta and R.K. Agrawal(2015)[102] obtained a minimal set of relevant and non-redundant features for enhancing performance of mental task classification through investigating six popular multivariate filter methods based on different criteria : distance measure, causal effect and mutual information. Their experimental results demonstrated improved classification accuracy as well as considerably reduced computation time with using each of these six multivariate feature selection methods. Among the investigated combinations of feature extraction and selection methods, the combination of wavelet transform and linear regression performs the best. They validated their experimental results through performing ranking analysis and statistical tests.

Z.M.Lwin(2015)[103] proposed a novel time-frequency based EEG feature extraction approach by employing Matching Pursuit(MP) algorithm. They tested their approach on the benchmark five task EEG data set acquired by Keirn and Aunon. After removing the artifacts caused by eye blink through ICA method, they segmented each 10 second trial EEG data into ten one second segments. MP algorithm was applied to one second EEG segments of each channel and 45 atoms were extracted per channel. So a total of 270 atoms were extractedfor the six channel EEG segments of one second duration. Finally, feature vectors of dimension 36 were constructed from the six statistical coefficients i.e. mean, variance, minimum ,maximum and sum of the six parameter values like modulus, amplitude, scale, frequency, position etc of each atom.

Least square support vector machine (LS-SVM) multiclass classifier with Gaussian RBF kernel was employed for the classification of EEG signals. They classified five mental tasks usingtwo multi-class support vector machine (SVM) classifiers with 'One-Vs-One' and 'One-Vs-All' scheme. Using 17 features (including mean, var, std, max, min, sum) , they achieved highest classification accuracy of 89.73% with 'One-Vs-One' multiclass classifier. They concluded that combination of mean, variance and minimum is very suitable for classification of MP based features with less execution time and small number of features.

M. Hendel *et al* (2016)[104]proposed a new two stage hybrid structure with supervised and unsupervised learning for classification of EEG signals. In the preprocessing step, after removal of artifacts , the EEG signals were divided into 46.1s EEG segments .DWT was applied to extract the five main sub-bands (δ, θ, α, β, and ϒ bands) from each segment and each channel using 5 stages of decomposition and Daubechies 4 (db4). Feature vectors having 150 components were constructed using five descriptors from all the five bands and all the six channels.The first stage consists of a Self Organizing Map (SOM) that reduces the original variable set (150 components obtained from DWT decomposition) into a smaller feature set, enabling to cluster the redundant irrelevant features and select the best descriptors to enhance the performance of the final classifier. The second stage uses a Quadratic Loss Multi-Class Support Vector Machine (M-SVM). It receives as input the obtained feature set from the first stage and outputs class posterior probability estimates. They validated their approach on the benchmark EEG database of Keirn and Aunon that consists on EEG recordings of 7 subjects performing five different mental tasks namely: Baseline, Multiplication, Letter, Rotation, and Counting.Using five distinct trials and five-fold cross-validation , they achieved average classification accuracy ranging from 81.73% to 91.90%.

Out of five trials, first three trials were used for training, fourth trial was used for validation and the fifth trial was used for testing.

A.Gupta and D.Kumar(2017)[105] proposed employment of FCM followed by empirical wavelet transform(EWT) for better representation of EEG signal which further enhances performance of mental task classification in BCI system. They decomposed the signal into desired number of support (segment) through applying EWT.FCM clustering algorithm was employed for avoiding overlapping segments obtained from the EWT decomposition. For compact representation of each segment, they calculated eight statistical parameters : root mean square, Lempel-Ziv complexity measure ,Shannon entropy, central frequency , maximum frequency, variance, skewness, and kurtosis .Feature vectors of dimension 144 (3 EWT segments x 8 parameters x 6 channels) were constructed from the concatenation of these eight parameters obtained from the six channels. For reducing the high feature dimension, they utilized three well-known multivariate feature selection methods viz. Bhattacharyya distance (BD), ratio of scatter matrices (SR), and linear regression (LR).They tested their approach for binary classification of ten pairs of five mental tasks using the benchmark EEG dataset (Keirn and Aunon,1990). Their experimental findings employing these methods endorsed that feature selection enhances the performance of mental task classification considerably . Ranking mechanism and Friedman’s statistical test were performed for ranking and comparing different combinations of feature extraction and feature selection methods. They obtained highest gain in classification accuracy using the combination of LR and FWET with respect to the best combination of EWT with or without feature selection. They obtained 100% classification accuracy in some cases binary combination of mental tasks for subject-2 and subject-7.

S.Dutta *et al* (2018a)[106] presented a new framework for EEG feature extraction based on the combination of multivariate empirical mode decomposition (MEMD) and phase space reconstruction (PSR) for classifying a small set of non-motor cognitive task EEG signals in mental task based multi-task brain computer interface(BCI) system. They employed phase space analysis of the intrinsic mode functions (IMFs) and used the largest singular values of the phase space trajectory matrices corresponding to a subset of sensitive IMFs for forming the feature vectors. The set of ten feature vectors corresponding to ten different trials of a specific mental task were partitioned into training set and testing set. The first seven feature vectors i.e. 70% of the data set was used for training and remaining three feature vectors i.e. 30% were used for testing the classifier.They obtained highest classification accuracy of 83.33% for binary i.e. pair wise classification of three pairs of mental task using a LS-SVMclassifier with RBF kernel function and tenfold cross validation.

S.Dutta *et al* (2018b)[107] proposed a new feature extraction approach based on the multivariate autoregressive model (MVAR) models of the task sensitive IMFgroups in the multivariate EMDdomain.They constructed the feature vectors from the sub set of Eigen values from the coefficient matrices of the MVAR model in the MEMD domain. . The set of ten feature vectors corresponding to ten different trials of a specific mental task were partitioned into training set and testing set. The first seven feature vectors i.e. 70% of the data set was used for training and remaining three feature vectors i.e. 30% were used for testing the classifier. On the testing data set, they achieved highest value of average classification accuracy of 94.43% for binary classification of the first pair of mental task i.e. baseline and mental arithmetic task using polynomial kernel and 91.65% for the second pair i.e. mental arithmetic and mental letter composing task using radial basis function (RBF) with ten-fold cross validation. They achieved highest value of average classification accuracy of 77.77% on the test data set for three class classification employing OneVsOne scheme of multiclass SVM classifier.

M.M. Rahman *et al* (2018) [108] proposed a novel scheme for mental task classification using reflection coefficients of EEG data as robust feature. They computed reflection coefficients directly from autocorrelation function of the EEG data and investigated their applicability as robust feature for classifying mental task in BCI system. The reflection coefficients utilized in this scheme offer some major advantages , such as noise robustness , variation of values within a certain boundary and ease of computing via recursive relations. For determining the number of reflection coefficients to be considered ,for constructing feature vectors, they performed detailed statistical analysis on the first few reflection coefficients and decided to use only the first two such coefficients for each channel. For an l channel EEG data, the feature dimension will be 2l. They tested the validity of their approach on the publicly available benchmark EEG data set collected by Keirn and Aunon. For effective classification of the feature vectors, they employed SVM with three kernel functions such as linear, quadratic and polynomial and leave-one-out cross-validation process.Here the dimension of the feature vectors is 12. With such a low feature dimension, they achieved average classification accuracies more than 89.16% from different subjects .They demonstrated that reflection coefficients offer a robust and effective feature vector which can be classified with very high level of accuracy and low computational time. They investigated the effect of channel selection and concluded that considering all channels for feature extraction provide best classification.

J.K. Nuamah ***et al*** **(2018**) [109] computed task engagement indices (TEI) from EEG recorded from six healthy participants who performed five separate cognitive tasks. They used task engagement indices as features and support vector machines (SVMs) for identification and offline classification of cognitive engagement during a specific mental task . They designed six separate multiclass SVMs to classify five cognitive tasks for the participants. They obtained average classification accuracy of 93.33±8.16% across the six participants.

CHAPTER 4

LARGEST SINGULAR VALUES IN THE PHASE SPACE ANALYSIS OF INTRINSIC MODE FUNCTIONS OF NON-MOTOR EEG SIGNALS

We present our first feature extraction approach based on the phase space analysis of the IMFs generated by MEMD based decomposition of the six channels EEG signals. Our proposed approach consists of three stages, in the first stage; the application of MEMD to multichannel EEG data gave rise to adaptive i.e. data driven decomposition of the multivariate time series data into a set of IMF groups. All the member IMFs within a group has common oscillatory frequency but different amplitude and cortical origin. In the second stage, a small subset of IMF groups was selected according to their task correlation factor and subsequently represented in the two dimensional phase space through their trajectory matrices. In the third stage, largest singular values of the trajectory matrices corresponding to these sensitive IMFs were employed for constructingthe feature vectors. Finally, the extracted feature vectors were fed to a LS-SVMclassifier for binary i.e. pair wise classification of these mental task EEG signals. With these new feature vectors, it is shown that LS-SVM with RBF kernel provides accuracy of 83.33% in classifying between mental arithmetic and mental letter composing. The performance of this classifier was evaluated on various parameters such as accuracy, specificity and sensitivity. The classification results show the potential of the proposed approach for classifying any non-linear a non- stationary signal.

4.1 Introduction

Fast, reliable and accurate classification of mental task is a crucial problem for designing real time BCI systems. Combination of high discriminatory feature with machine learning algorithm will lead to enhanced performance of BCI system. But search for high discriminatory feature is an open research problem for improved classification of these EEG signals. Keirn and Aunon [22, 75, 76] proposed that EEG signals could distinguish between various mental tasks accurately. They designed the experimental protocol and acquired EEG signals related to a small set of five non-motor imagery mental tasks.Their work was motivatedby the previous work by T.Fernández, et al(1995)[77].They used AR model coefficients and band power asymmetry ratio as features with quadratic Bayesian classifier for classifying these five non-motor mental tasks. Later, Charles W Anderson et al [78-81] employed multivariate AR (MVAR) model of EEG for the classification of same

EEG data set related to five non-motor mental tasks. They continued the work of Keirn and Aunon [22,75,76], and derived both scalar AR andMVARmodel from the original signals. But these approaches have the limitation that they assumed the underlying signal generating mechanism to be linear and stationary. Due to this assumption of linearity and stationarity, these methods are not adequate in capturing the complex nonlinear dynamics contained in these signals. Extracting subtle information from these EEG signals by analyzing their extremely complex pattern is a formidable task as they are notoriously noisy, nonlinear and non-stationary. A plethora of feature extraction methods have been employed for the time-frequency analysis of many non-stationary signals representing a wide range of natural phenomena such as seismological signals, biomedical signals etc. But the common drawback of all these techniques is that they decompose a signal based on apriori fixed bases.They give sub-optimal localization in the joint time-frequency plane which makes their performance inadequate. This has given rise to data driven method called, EMD whose performance has been established to be adequate in many cases of nonlinear and non-stationary real world time series data such as earthquake data,winds, ocean, acoustic signals, mechanical vibration signals, biomedical signals etc. In contrast to wavelet and other time-frequency based decomposition, EMD is a fully data driven algorithm which does not require any apriori basis function for the multi scale decomposition of the signal.

EMD decomposes a signal into a set of oscillatory modes known as IMFs based on local characteristic time scale of the data [110]. EMD obtains the oscillatory modes(scales) adaptively and considers the signal dynamics at the local level, making it a natural choice for generating the data scales required for multi scale analysis.Therefore, applying EMD or its multivariate extension i.e. MEMD [111] is quite reasonable for any nonlinear and non-stationary signal like EEG. Pablo F.Diez et al [94] employed univariate empirical mode decomposition (EMD)for classifying same set of fivemental tasks from thememental task EEG data set. They adoptedunivariate approach andextracted fourdifferent time domain features (RMS, Variance, Shanon Entropy, and Lempel –Ziv complexity) from the IMFs. M.Kaleem et al [95] employed EMD and Teager Energy operator for classifying five mental tasks from the benchmark data set. They also adopted univariateapproach using one EEG channel only. For catering to the need of real world multichannel data, Rehman and Mandic [111] developed multivariate extension of the standard univariate EMD algorithm known asMEMD algorithm. But most of this research work had focused on one channel EEG signal with standard EMD algorithm. Standard EMD algorithm is univariate in nature. Due to this, these research works are based on the sequential decomposition of the multichannel EEG channels instead of simultaneous decomposition.But sequential analysis of multichannel EEG data using standard EMD algorithm gives rise to the twin problems of mode mixing and mode alignment. MEMD algorithm enabled us to circumvent these twin problems of its univariate counterpart by generating equal number of IMFs for all data channels.

The review of literature has confirmed that brain activity exhibits some kind of chaotic type i.e. deterministic nonlinear behaviour. Therefore, applying nonlinear methods such as EMD or its multivariate extension i.e. MEMD coupled with phase space reconstruction is quite reasonable for EEG time series analysis. Nonlinear measures hold the capacity to capture subtle changes occurring in any non-stationary signal since the system dynamics become prominently visible in the reconstructed phasespace. So, nonlinear dynamical methods using phase space reconstruction provide the ability to go deeply into the subtle dynamics shown by the signal more accurately. This has motivated the researchers across different domains to employ phase space analysis as a promising tool for modeling any nonlinear phenomenon.It has been shown that the PSR of IMFs is like fertile ground for investigating new features for classifying EEGs. Lee, Lim, Kim et al [112] used Euclidian distance based features computed from PSR of wavelet coefficients for detecting epileptic seizure EEG signals.But these research work done using phase space reconstructions so far had focused on the detection of various pathological conditions. Besides, they employed one channel EEG signals with standard EMDalgorithm.To the best of our knowledge,no research report is available in the literature presenting combined application of MEMD on multichannel EEG signals related to non-motor mental task for BCI applications.

The main focus is to illustrate the substantial advantages of employing combination of MEMD based decomposition and phasespace reconstructions for extracting nonlinear features from a subset of sensitive IMFs. Our main contribution lies in extracting SVD based novel features from the 2D PSR of sensitiveIMFs instead of original EEG signals.Due to the amplitude and phase modulated (AM-FM)characteristics of the IMFs;their phase space trajectory matrices follow special geometry. In viewof this, we have been motivated to use singular value decomposition (SVD) of the trajectory matrices since the singular values can unravel the special geometrical structure of the IMFs.Our contribution in this research is twofold (a)combined application of MEMD algorithm with phase space reconstruction of resultingIMFs (b)extracting new features based on SVD decomposition of the trajectory matrices of the sensitive IMFs.

4.2 Data and Methodology

4.2.1 Mental Task(MT) EEG Database

Motor imagery(MI) based BCI systems are designed to rehabilitate people with motor impairments. Although this MI based BCI paradigm has been widely used in BCI research , it has two important limitations – it does not allow the user to issue more than two instructions due to few degree of freedom. Apart from its very few degree of freedom, not all people with motor impairment find it easy to perform imagined motor movements [35]. These limitations of MI based BCI have been overcome by Keirn and Aunon through developing mental task (MT) based BCI which is driven by a small set of five non-motor mental tasks , namely, mental arithmetic task, mental letter composing etc. The MT based BCIs having more degrees of freedom yield more flexible BCI communication protocols than MI based BCIs. Furthermore, MT based BCIs provide improved performance and good user experience since the users have the option to select mental tasks as per their own comfort level. Finally, MT based BCI paradigm allows self-paced and stimulus-free control i.e. their operation also is spontaneous, independent and asynchronous i.e. not time locked with any external stimulus. Due to these advantages, MT based BCI paradigms show better potential for people with severe motor disabilities to communicate with the external world just by thinking only [36]. All these advantages of non-motor imagery (MT) based BCI has motivated us to use the benchmark EEG data set collected by Keirn and Aunon, from Purdue University. This EEG data set is old , but it has been considered as the benchmark EEG data set for non-motor imagery mental task by the BCI researcher community.Using this EEG data set, it has been possible to classify a set of mental task with high accuracy. The EEG response of these non-motor mental tasks such as mental arithmetic, mental letter composing invokes brainwave asymmetry. The EEG channels or electrode locations selected under these non motor mental tasks corresponds to fronto parietal regions of the cortex. There are motor imagery (MI) EEG data publicly available , we did not use those since we were interested in non-motor imagery (MT) EEG based BCI.

They performed experiments on seven subjects, 21 to 48 years old. The subjects were seated in a sound controlled booth with dim lighting and noiseless fans (for ventilation).

An Electro-Cap elastic electrode was used to record EEG signals from six cortical locations such as C3, C4, P3, P4, O1 and O2 based on the international 10-20 electrode placement system. The impedance of all the electrodes was kept below 5K Ohms. These scalp electrodes were referenced to two electrically linked mastoids, A1 and A2.The signals were recorded for duration of 10 seconds during each trial of a task.The sampling frequency was 250 samples per second. EEG signals were recorded on 7 subjects, who performed five differentmental tasks.

This EEG data set is available online

athttp://www.cs.colostate.edu/eeg/main/data/1989_Keirn_and_Aunon

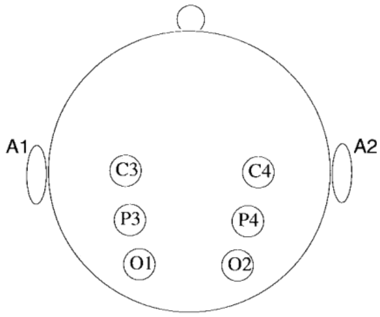
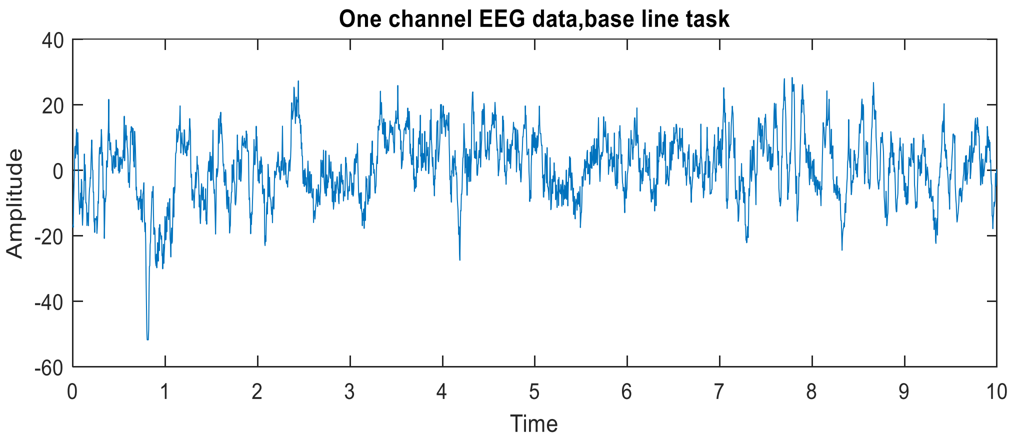
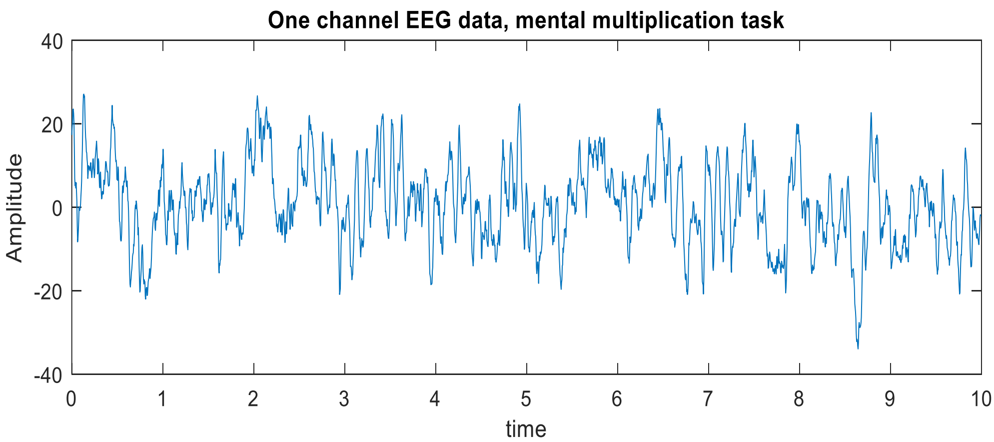


Figure 4.1: Electrodeplacementsaccording to the 10-20 system.





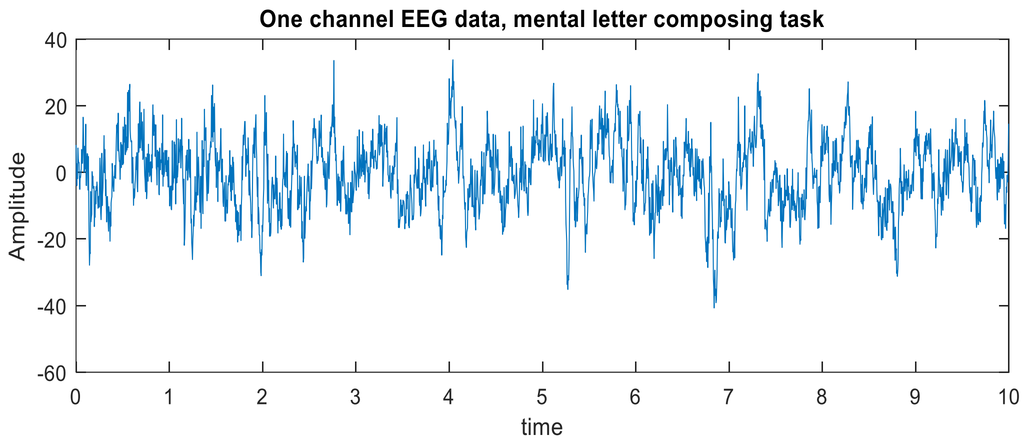


Figure 4.2: EEG signal waveforms of (a) baseline (b) mental multiplication task (c) mental letter composing task

The above Figure4.2 shows single channel EEG waveforms corresponding to single trial performance of baseline and two non-motor mental tasks, namely mental multiplication and letter composing task.

4.2.2. Methodology:

In this section, the block diagram of the proposed methodology has been described along with the short introduction of the signal processing and machine learning methods used in this work.

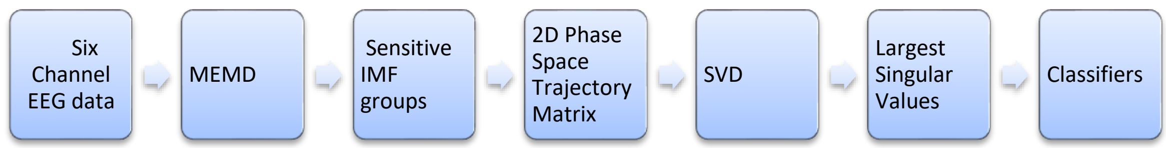


Figure4.3: Block diagram representation of the proposed methodology

4.2.2.1. Multivariate Empirical Mode Decomposition

For catering to the need of real world multichannel data, Rehman and Mandic [111] extended the standard univariate EMD algorithm to multivariate data known as MEMD algorithm.In this algorithm, the local mean of n –dimensional signals are computed by the multiple n-dimensional envelopes, which are generatedby taking signal projections along different directions in n-variatespaces.Fora uniform set ofdirection vectorstoprojectthe signal, low discrepancy Hammersley sequences [113] are used to obtain quasi-uniform pointson high dimensional spheres [114]. The IMFs are obtained iteratively through sifting process [115].

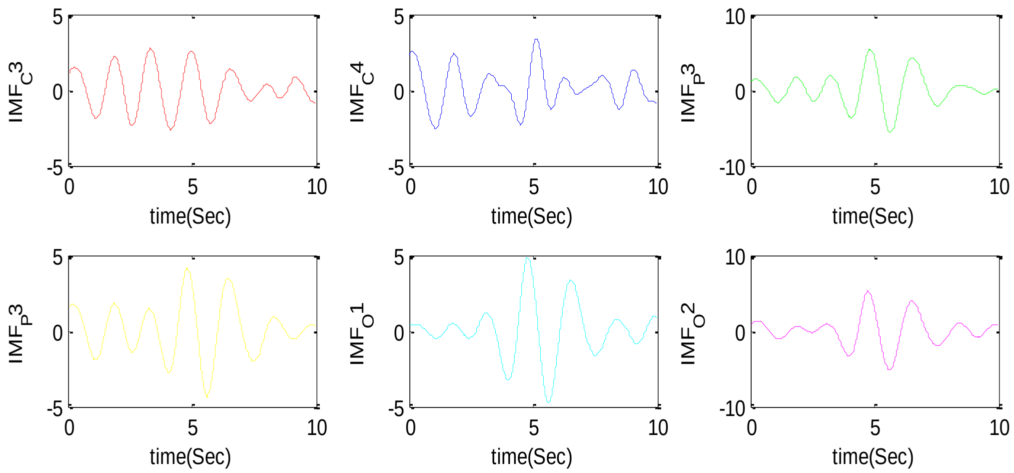
This sifting process stops when all the K projected signals fulfill the stopping criterion [116].The details of MEMD are outlined as given in [111].

4.2.2.2 MEMD based decomposition and selection of most sensitive IMF group using their power spectra

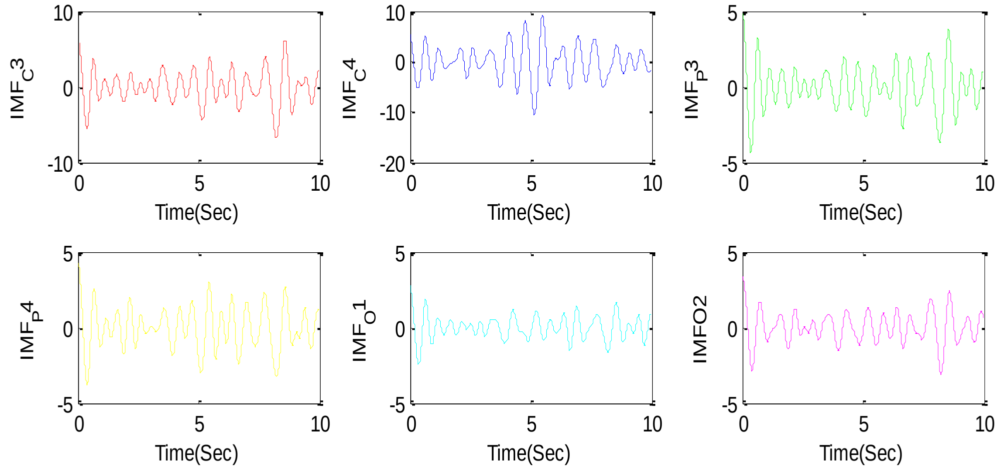
The multivariate EMD (MEMD) algorithm based adaptive decomposition of the six channel EEG data corresponding to single trial performance of each task generated twelve groups of intrinsic mode function (IMF) and one residue function depicting the trend only. Each IMF group visualized as time domain waveforms in Figure 4.4 represents one frequency level and are generated in descending order of high frequency i.e. the first IMF group represents highest frequency level and subsequent IMF groups will have decreasing frequency levels. All these time domain waveforms representing constituent IMF groups reveal discriminatory mono-component oscillatory mode common across the six EEG channels.

The temporal structure of these oscillatory waveforms reveal discriminatory feature which becomes more prominent in its reconstructed phase space .

(a)



(b)



(c)

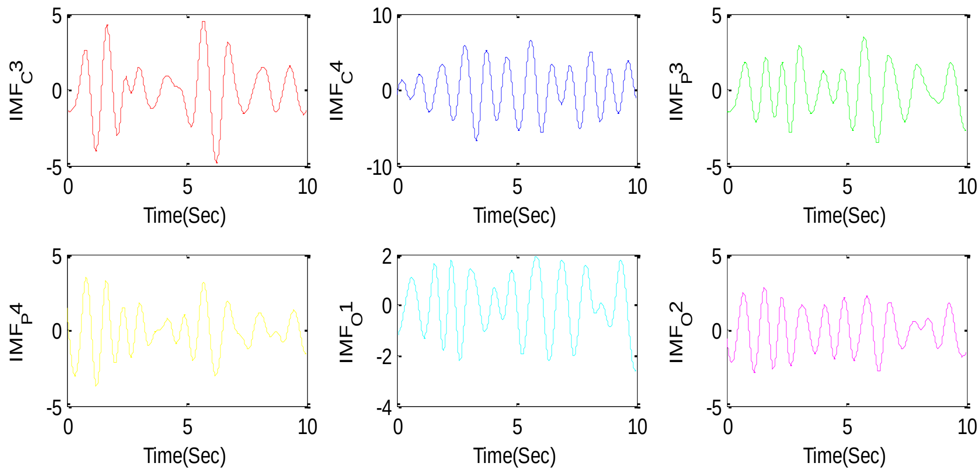


Figure 4.4: Waveformsof IMF components of the most sensitive IMF groups(a) Base Line Task (b) Mental Multiplication Task (c) Mental Letter Composing Task

The above Figure4.4 visualizes the waveforms i.e. the variation in amplitude of each component intrinsic mode functions (IMFs) within the most sensitive IMF group for three mental tasks. The IMF groups are generated from the MEMD based adaptive decomposition of the six channel EEG signals corresponding to single trial performance of three classes of non-motor mental task. The six component IMFs within an IMF group have same oscillatory frequency but have different amplitudes and cortical locations. Each of these IMF group represents mono-component or narrow band oscillatory component of brain activity dynamics at particular cortical location.

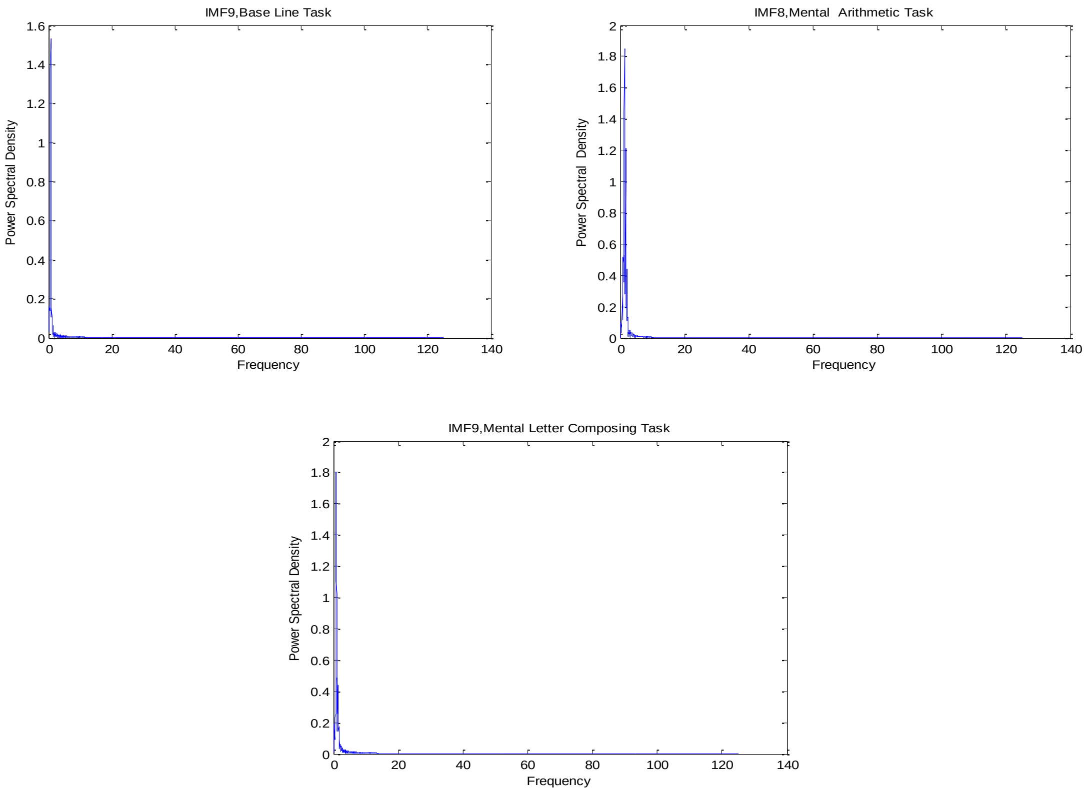


Figure 4.5: PowerSpectralDensity(PSD) of relevantIMFs (a) Baseline (b) Mentalarithmetic (c) Mental letter composing task.

4.2.3 LS-SVMClassifier

This version of SVM employ kernel function for solving a set of linear equations. They used all training points for deriving this model of LS-SVM. Due to this approach, the cost of complexity and computation time is significantly reduced in LS-SVM. The derivation of the LS-SVM is as follows

Minimize

$$J(w,b,e) = \frac{1}{2}w^{\tau}w + \frac{\gamma}{2}\sum_{i=1}^{N} e_i^2 \qquad (1)$$

subject to equality constraints

$$y_i[w^{\tau}g(x_i) + b] = 1 - e_i, i = 1,2,\dots,N \qquad (2)$$

Where e=(e_1, e_2, e_3, e_4, ...e_i)^T , $w$is thed-dimensionalweightvectorandb is a bias, g(x)is a mappingfunctionthat mapsxintod-dimensional space.TheLagrangian multiplierα_i can be defined for equation (3) as

$$L(w,b,e;\alpha) = J(w,b,e) - \sum_{i=1}^{N} \alpha_i \{y_i[w^{\tau}g(x_i) + b] - 1 + e_i\} \qquad (3)$$

Whereα_i are Lagrange multipliers and$\alpha = (\alpha_1, \alpha_2 \ldots, \alpha_N)^T$.After solving (5), the LS-SVM classifier is obtained as

$$f(x) = sign[\sum_{i=1}^{N} \alpha_i y_i K(x, x_i) + b] \quad (4)$$

where K(x, xi) =$g^T(x)\, g\,(x_i)$ is a kernel function.

Statistical parameters for performance evaluation

The performance of the LS-SVM classifier can be evaluated by computing different performance parameters such as accuracy, sensitivity, specificity, positive predictive value, negative predictive value etc.

Accuracy(ACC) : It is the ratio of number ofcorrectly classified samples to total number of samples.

$$ACC(\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100$$

Sensitivity (SEN): It is the ratio of the number of true positive samples to the total number of positive samples and given by:

$$SEN(\%) = \frac{TP}{TP + FN} \times 100$$

Specificity (SPF): It is the ratio of the number of true negative samples to the total number of negative samples and given by:

$$SPE(\%) = \frac{TN}{TN + FP} \times 100$$

WhereTP and TN represents the total number of true positive and true negative events, respectively, and FP and FN represents the total number of erroneously positively events and erroneously negative events respectively.

2DPhaseSpaceRepresentation (PSR) of Intrinsic Mode Functions and Singular Value Decomposition (SVD) of PhaseTrajectoryMatrices

Reconstructed phase space (RPS) is a fertile framework for investigating the nonlinear dynamics of a signal in the time domain. In this RPS, a signal provides a snapshot of the temporal dynamics of the system generating it more accurately.In this work, we reconstructed the two dimensional phase space (2D PSR) of the sensitiveIMFs by assuming time lag =1 according toTakens. It is evient from Figure4.6, that 2D PSR of IMFs show elliptical patterns. SingularValuedecomposition (SVD) of this 2Dphase space trajectory matrix reveals the hidden structure of a signal such as stability, invariance of proportion and rotation.

We employed SVD method to decompose the phase trajectory matrix and the singular valuesσarecalculated:

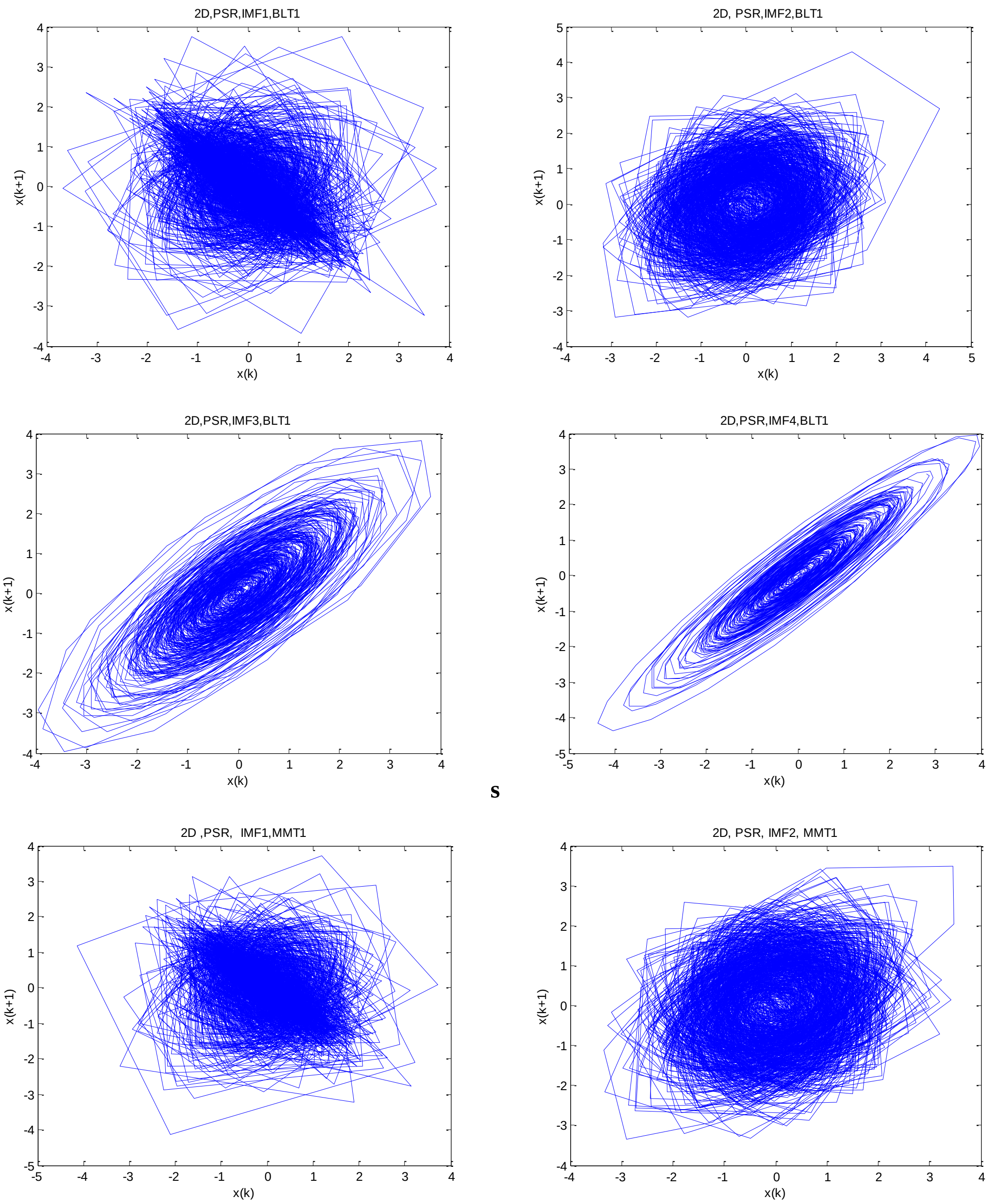
$$\sigma = [\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{44}] \tag{5}$$

Where $\sigma_{11} \gg \sigma_{22} \gg \sigma_{33} \gg \sigma_{44}$

4.3 Results

In this study, we used six channel EEGsignals corresponding to single trial performance(10 second duration)of three non-motor mental tasks.From the Keirn and Aunon's benchmark EEG data set related to a set of five non-motor mental tasks, we have chosen three tasks i.e. mental arithmetic task, mental letter composing task and baseline task. There is no specific reason for selecting only three out of five tasks, we could have used all the five tasks. Any combination of these tasks can be chosen depending upon the targeted application. Mental arithmetic task can be used for the application of building an EEG based mental calculator . Similarly, letter composing task can be used for building an EEG based mental type writer. And combinations of two, three or four of these tasks are to be used for the application of EEG based multitask BCI . In our research, we selected the combination of mental arithmetic and mental letter composing tasks keeping in view of building EEG based mental calculator as well as mental type writer.To account for their extremely complex nonlinear,non-stationary nature, we applied multivariate empirical mode decomposition (MEMD) on the multivariateEEG data to find out their inherent oscillations in the data known as IMFs.The application of MEMD gave rise to a set of IMFgroups,where each IMF within a group has common frequency but different amplitudes of oscillation. From the set of generated IMF groups, we selected a subset off four sensitive IMF groups(IMF_7,IMF_8,$IMF_9$and IMF10). These four IMF groups were selected due to their higher degree of correlation with their respective original signal. This enabled us to get rid of the unwanted noise and artifacts within the signal. The oscillatory temporal behavior of the IMFs can be best studied through reconstructing the dynamics of these sensitive IMFs in the phase space. For phasespace reconstruction, we used the value of embedding dimension, m= 2and time delay(τ=1)for forming the embedding dimension vector and time delay vectors respectively.The2D PSR corresponding to each IMF of a task is shown in Figure 4.6. The geometrical and topological property of the generated phase space trajectory matrices encodes wealth of information regarding spatio-temporal dynamics. We further applied singular value decomposition (SVD)on the covariance matrices of phasespace trajectory matrices for their structural analysis using their singular values.We considered only the largest singular value of each IMF group for forming the featurevectors since,it has dominant control over the geometrical structure of these phase space matrix. Each trial of a task is represented through its feature vector comprising the largest singular values of the four sensitiveIMFs.All the ten feature vectors for ten trials of a specific task and the total thirty feature vectors(ten per task)for thirty trials of three task is shown in the Table 4.1.

2D phase space trajectory of sensitive intrinsic mode functions



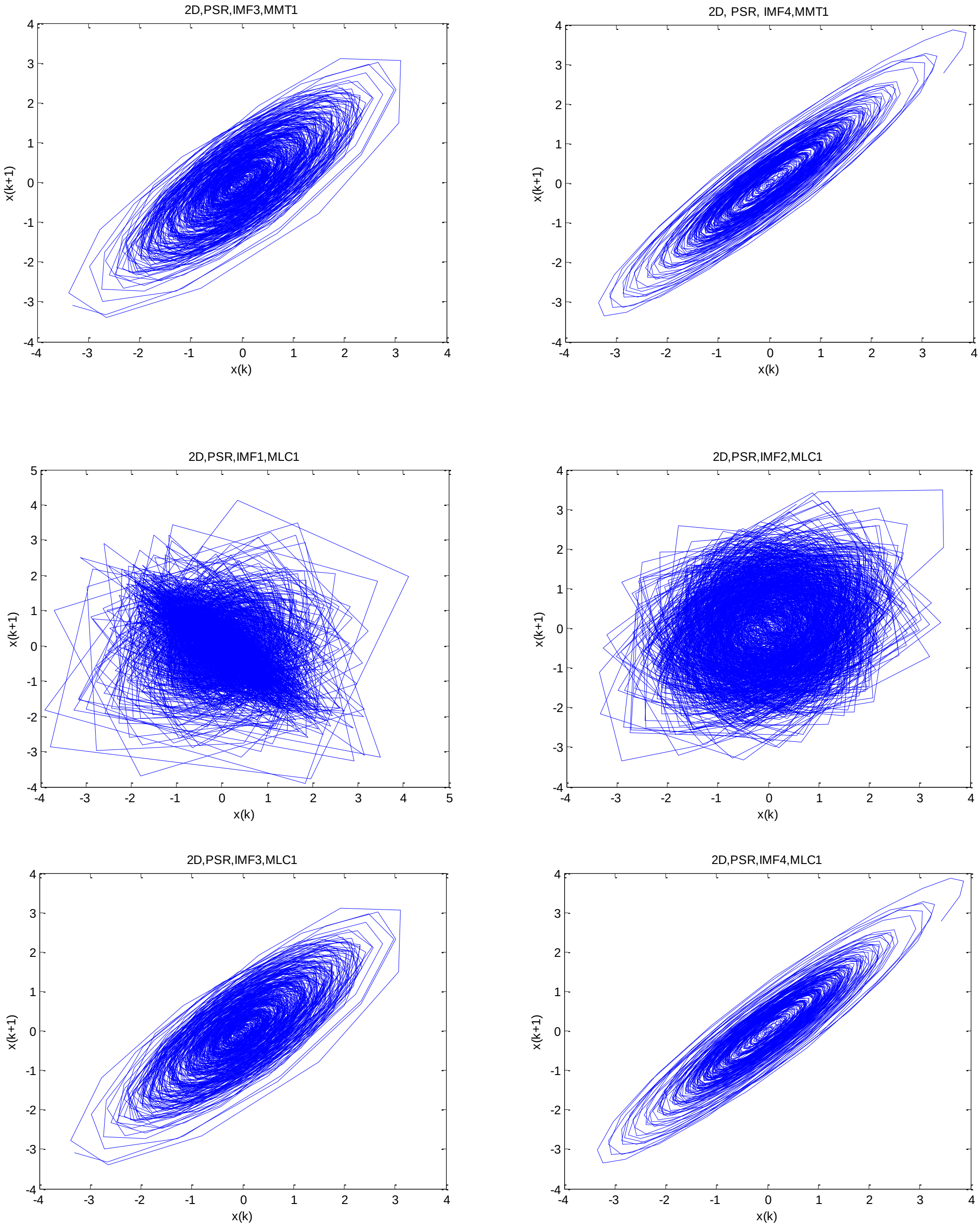


Figure4.6: 2D PSR of first four IMFs of (i) base line EEG: (a) IMF_1, (b) IMF_2, (c) IMF_3, (d) IMF_4

(ii) Mental multiplication task EEG (a) IMF_1, (b) IMF2(c) IMF_3 (d) IMF_4

(iii) Mental letter composing task EEG (a) IMF_1, (b) IMF_2, (c) IMF_3, (d) IMF4

Table4. 1: Mean and standard deviation of Largest Singular Values for ten trials of sensitive IMFs

IMF No	Baseline(mean±std)	Mental arithmetic(mean±std)	Mental letter composing(mean ± std)
7	112.2164±18.0784	158.4629±30.9921	131.3931±19.1819
8	198.6472±52.7957	214.1005±56.7523	176.5351±42.1999
9	104.6329±36.5002	102.9417±40.6245	93.2217±36.3279
10	94.1504±32.3748	97.6126±32.9566	113.6577±51.7180

Table 4.2(a): Performance of LS-SVM classifierfor binary classification of mental task

Pair of mental task	ACC	SEN	SPE	PPV	NPV
Base Line Vs Mental Arithmetic	66.66	60	100	100	33.33
Mental Arithmetic Vs Mental Letter Composing	83.33	100	66.66	75	100
Base Line Vs Mental Letter Composing	50	66.66	33.33	50	50

Table 4.2(b): Performance of k- Nearest Neighbour classifier for binary classification of mental task

Pair of mental task	ACC	SEN	SPE	PPV	NPV
Base Line Vs Mental Arithmetic	50	33.33	66.66	50	50
Mental Arithmetic Vs Mental Letter Composing	33.33	00	66.66	00	40
Base Line Vs Mental Letter Composing	66.66	66.66	66.66	66.66	66.66

Table 4.2 (c): Performance of LDAclassifier for binary classification of mental task.

Pair of mental task	ACC	SEN	SPE	PPV	NPV
Base Line Vs Mental Arithmetic	66.66	60	100	100	33.33
Mental Arithmetic VsMental Letter Composing	83.33	100	66.66	75	100
Base Line VsMental Letter Composing	50	66.66	33.33	50	50

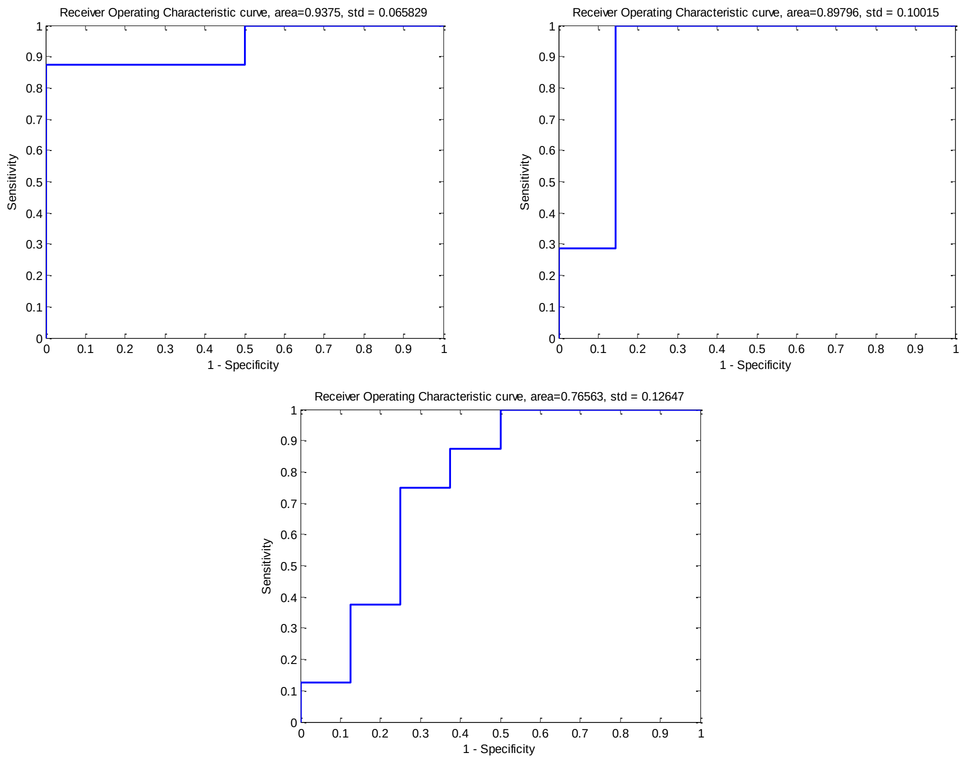


Figure 4.7: ROC of LS-SVM classifiers for three pairs of mental task.

Table 4.3: Summary of Classification Results obtained using proposed methodology and other existing methodologies

Authors	Features used	Classifier used	Classification accuracy (%)
Kern and Anon(1990)	FFT based band power Scalar AR model coefficients Asymmetry ratio	Bayesian Quadratic classifier	84.6%
Charles W Anderson, Z Latko,Sijercie (1996)	AR model co-efficients	Back Propagation Neural Network (BPN)	71%
		Probabilistic Neural Network (PNN)	38%
Charles W .Anderson , Erik A. Stolz , SanyogitaShamsunder(1998)	Multivariate AR model,Scalar AR coefficients, Eigen values of Correlation matrix	Feed forward neural network trained via error back propagation algorithm	86.1 to 91.4%
RamaswamyPalaniappan,RaveendranParamesran, Shogo Nishida, Naoki Saiwai (2002)	Power Spectral Density (PSD)	Fuzzy ARTMAP	93.67%
Deon Garret, David A .Peterson, Charles W Anderson, Michael H.Thaut, (2003)	AR Model Co-efficientPower Spectra	SVM/ANN/LDA	52%/53%/45%
Nan-Ying Liang,ParamasivamSaratchandran,Guang-Bin-HuangNarasimhanSundarajan (2006)	AR Model Co-efficient	BPNN/SVM/Extreme Learning Machine	55.76% /60.86 %/63.96%
M.F.Kaleem et al	Teager Energy based Average Energy	LDA	85%
Diez et al (2009)	RMS, Variance, Shanon Entropy, Lempel-Ziv complexity etc in EMD domain	LDA/NN	91%/87%
Martina Tolic, FranjoJovik (2013)	DWT based discrete wavelet energy features	Back Propagation Neural Network	90.75% (for multiclass) 99.87% (for binary)
Zin Mar Lwin(2015) Mie Mie Thaw	Gabor based Matching Pursuit(MP)	Multiclass SVM Classifier	(a) 89.73% (OneVsOne) (b) 52.43%(OneVs All)
MouniaHendel, AbdelkaderBenyetton, FatihaHendel (2016)	-	Hybrid Classifier using Self Organizing Map and Probabilistic Quadratic Loss Multiclass SVM	81.73 % to 91.9%
SumanDutta, Manddep Singh , Amod Kumar	Largest singular values extracted from phase space trajectory in MEMD domain	LS-SVM	83.33%(binary)

4.4 Discussion

Phase space reconstruction is a technique that allows us to have a better look into the underlying dynamics and develop new insights on the working mechanisms of our brain.The phase space trajectory matrix generated from phasespace reconstruction represents the underlying nonlinear dynamics of the system.We are motivated to extract new features from the geometrical and topological properties of the phasespace trajectory matrices represented by2D PSRs or3D PSRs. The covariance matrix of this phase space trajectory matrix has a special geometry. We intend to get benefit from this special geometry by using the singular values of this matrix as distinguishing features of the underlying dynamics of a multivariate time series. Covariance matrix of a data matrix can be viewed as a linear transformation having a geometrical structure.This geometrical structure decides the shape and orientation of the data matrix.The Eigen values or the singular values determine the geometrical structure of this covariance matrix which in turn decides the shape and orientation of the phase space trajectory matrix. In this research,our main contribution comes in not only developing a particular approach for EEG feature extraction, but in suggesting that the combined application of powerful techniques like PSR, EMD and SVDetc can provide us a fertile ground for searching novel features having better discriminatory power .

In this study, we proposed to use the largest singular values of the phase space trajectory matrices correspondingto a number of sensitive intrinsic mode functions (IMFs) as a novel feature for classifying threenon-motorcognitivetasks in EEG based BCIsystem. We used six channel EEGdata corresponding to single trial performance(10 second duration)of three non-motor mental task.The six channels EEG signals were adaptively decomposed into intrinsic mode functions through MEMD based decomposition. From the largeset of generated IMFs, we considered only four (IMF_7-IMF_{10}) sensitive IMFs as their dynamics reveals meaningful information in the reconstructed phases pace. We applied SVD on the phase space trajectory matrices corresponding to the sub set of sensitiveIMFs.We considered the largest singular value of the phase space matrices corresponding to the four sensitive IMFs. We constructed feature vectors representing each trial EEG of this task using these largest singular values. Finally,these feature vectors were used for training and testing three different classifiers-least square support vector machine (LS-SVM) , K-nearestNeighborhood classifier and linear discriminant teanalysis (LDA) for classifying these task related EEG signals.The results of binary i.e. pair wise classification for the three classifiers is shown in the Tables 4.2(a)-(c). The highest classification accuracy obtained with this approach is 83.33% for the pair - mental arithmetic and mental letter composing task usingRBF kernel. The area under ROC curve as shown in Figure 4.7 is a useful parameter to quantify the overall performance of the classifier. Beside ROC, the other performance parameters such as SEN, SPE, PPV, NPV have been computed as shown in the tables.Table 4.3 presents only the summary of the classification accuracies obtained with our proposed feature extraction methodologies along with other approaches using same EEG data set. The table is not intended to present a comparison of their performances.Through this table, we intend to present a snapshot of the classification accuracies obtained by BCI researchers in classifying same set of non-motor mental task in MT based BCI paradigm. Though the same benchmark EEG data sets had been used in testing the effectiveness of the presented approaches, classification accuracies obtained through these approaches cannot be compared due to various factors. Firstly, though the presented approaches classified mental tasks using the same benchmark EEG data set , they employed different types of classifiers trained and tested with different types of feature vectors. Secondly, in all the presented approaches, the way of partitioning the data set into training data set and testing data set was not same. Besides , the methods employed for cross-validation were also different. For comparison of classification accuracy , different classifiers trained and tested on the same feature vectors are to be used or same classifier trained and test with different feature vectors are to be used. Hence, the classification approaches presented in the table cannot be compared as they are based on different feature extraction approaches and different classifiers.

The advantages of our proposed approach are as follows:

(a) We have avoided the twin problems of mode mixing and mode alignment faced by standard EMD algorithm by using MEMD algorithms.
(b) 2D PSRs (Phase Space Representations) of the generated IMFs enabled us to dive into the subtle brain dynamics more accurately.
(c) The extracted features are suitable for real time implementation.

4.5 Conclusions

In this work, we investigated the applicability of a new framework for EEG feature extraction based on combination of two powerful methods- MEMD based decomposition and phase space reconstruction.The largest singular values computed from the phase space trajectory matrices of the sensitive IMFs enabled us to generate very useful feature space for classifying EEG signals related to different cognitivetask. The extracted features are in time domain and this makes our approach suitable for real time implementation. The extracted feature vectors had been classified by LS-SVM classifier with different kernel functions and the results were compared with LDA and KNN classifiers. The classification performance indicates that the LS-SVM with RBF kernel has great potential. Till date, phase space analysis of EEG signals in the MEMD domain is very limited. Our proposed methodology provides a fertile ground for further analysis of EEG signals in diverse application domains. Novel features computed from the 3D PSR of the sensitive IMFs and combining them with the proposed singular value features may form the basis of future work for enhancing the performance of the classifiers.

CHAPTER5

MULTIVARIATE AUTOREGRESSIVE (MVAR) MODEL OF INTRINSIC MODE FUNCTIONS FOR ANALYSIS OF ELECTROENCEPHALOGRAM

In this chapter, we have introduced a new feature extraction based on the multivariate autoregressive model (MVAR) model of the sensitive intrinsic mode function (IMF) groups in the multivariate empirical mode decomposition (MEMD) domain.We computed Eigen values from the coefficient matrix of the MVARmodel for classifying three different non-motor cognitive tasks in EEG based BCI. In the first stage, the application of MEMD to multichannel EEG data gave rise to adaptive i.e. data driven decomposition of the multivariate time series data into a large number ofIMFgroups. In the secondstage, the sensitive IMFgroups were selected according to their task correlation factor .MVAR model of order six was developed from the five sensitive IMF groups and finally the eigen values of the correlation matrix derived from the coefficient matrix were employed for forming the feature vectors. At the last stage, the extracted feature vectors were fed to a LS-SVM classifier for automatic classification of mental task EEG signals. We tested our approach on the mental task EEG data sets of three subjects.We achieved highest value of average classification accuracy of 94.43% for binary classification of the first pair of mental task i.e. baseline and mental arithmetic task using polynomial kernel and 91.65% for the second pair i.e.arithmetic and letter composing task using radial basis function (RBF) withten-fold cross validation. We achieved highest value of average classification accuracy of 77.77% for three class classification employing OneVsOne scheme of multiclass SVM classifier.The performance of the binary classifier was evaluated on various parameters such as accuracy, specificity and sensitivity. The encouraging results show the potential of the proposed approach for classifying any nonlinear and non-stationary signals.

5.1 Introduction

EEG signal represents neuro-electric activities of human brain under different mental states. Hence, these signals offer the possibility of classifying different cognitive task for the development brain computer interface (BCI). Keirn and Aunon [22, 75, 76] proposed that EEG signals could distinguish between various mental tasks accurately. They designed the experimental protocol and acquired EEG signals related to five non-motor imagery mental tasks which invoke hemispheric brainwave asymmetry. Their work was motivated by the previous work by Fernández, T.et al(1995)[77] .They used AR model coefficients and band power asymmetry ratio as features with quadratic Bayesian classifier for classifying five non-motor mental tasks EEG signals which was acquired by them. Later, Charles W Anderson et al [78-81] employed MVAR model of EEG for the classification of same set of five non-motor mental tasks [22]. They continued the work ofKeirn and Aunon and derived both scalar AR model as well as multivariate AR model from the raw EEG signals without employing any decomposition technique. Due to the assumption of linearity, these methods are not adequate in capturing the complex nonlinear dynamics contained in these signals. Extracting subtle information from these EEGsignals by analyzing their extremely complex patternis a formidable task as they are notoriously noisy, nonlinear and non-stationary. .Fast, reliable and accurateclassification of EEG signals related to performance of different cognitive task is the central challenge for designing real time BCI systems. Combination of best feature extraction algorithm with suitable machine learning algorithm will lead to enhanced performance of BCI system. But searching for features having more discriminatory power is yet an open research problem for improved classification of these EEG signals. Traditional FFT based non-parametric and parametric methodsof EEG analysis assume linearity and stationary of the signals. All such complexities of the EEG signals call for their accurate representation in the joint time-frequency plane. But the time-frequency representation of a signal is not unique.A plethora of advanced feature extraction methods have been employed from the joint time-frequency analysis of many non-stationary signals representing a wide range of natural phenomena such as seismological signals, biomedical signals etc. But the common drawback of all these techniques is that they decompose a signal based on apriori fixed bases.They give sub-optimal localization in the joint time-frequency plane which makes their performance inadequate. This has given rise to the development of data driven method called,EMD whose performance has been established to be adequate in many cases of nonlinear and non-stationary real world time series data such as earthquake data,winds,ocean acoustic signals, mechanical vibration signals, biomedical signals etc..In contrast to wavelet and other time-frequency based decomposition, EMD is a fully data driven algorithm for the multi scale decomposition of the signals.EMD decomposes a signal into a set of oscillatory modes known as intrinsic mode functions (IMFs) based on local characteristic time scale of the data [110].EMD obtains the oscillatory modes adaptively through considering the signal dynamics at the local level. Therefore, applying EMD or its multivariate extension i.e. MEMD [111] is quite reasonable for any nonlinear and non-stationarysignal likeEEG .Pablo F.Diez et al [94] employed univariate empirical mode decomposition (EMD)for classifying same set of five mental task from the mental task EEG data set[22]. They adopted univariate approach and extracted four different time domain features (RMS, Variance, Shannon Entropy, and Lempel –Ziv complexity) from the IMFs. M.Kaleem et al [95]employed EMD and TeagerEnergy operator for classifying five mental tasks from the benchmark data set [22]. They also adopted univariate approach using one EEG channel only. For catering to the need of real world multichannel data, Rehman andMandic [111] developed multivariate extension of the standard univariate EMD algorithm known asMEMD algorithm. But research work done so far employing EMDhas predominantly focused on the detection of various pathological conditions. Besides, they employed one channel EEG signals with standard EMD algorithm. Standard EMD algorithm is univariate in nature.Due to this, these research works are based on the sequential decomposition of the multichannel EEG channels instead of simultaneous decomposition.But sequential analysis of multichannel EEG data using standard EMD algorithm gives rise to the twin problems of mode mixing and mode alignment. MEMDalgorithm enabled us to circumvent these twin problems of its univariate counterpart by generating equal number of IMFs for all data channels.

Till date, to the best of our knowledge, published research work on EEG feature extraction methods combiningEMD or MEMDwith MVAR for classifying non-motor mental task in EEG basedBCI systemis not available. Due to this,there is ample scope for undertaking further research in this direction through deriving MVARmodel based feature extraction in the MEMD domain .Therefore, the main focus of our research was to extend the work of Charles W Anderson et al [78-81] through computing multivariate AR model in the MEMD domain . Our contribution in this research is two fold(a) combiningMEMD algorithm with MVARmodel(b) extracting the non-zero Eigenvalues as features from the covariance matrix of the sensitiveIMFgroups.The novelty of our approach stems from adopting a multivariate and nonlinear approach through combining MEMD based decomposition and multivariateAR model.Though the multivariate AR model is a linear model, but it offers the advantage of modeling the inter-regional dependency withinbrain electrical activity.

5.2 Data and Methodology

5.2.1. EEGDatabase

The benchmark EEG database used in this research was collected byKeirn and Aunon, from Purdue University and is available online at http://www.cs.colostate.edu/eeg/main/data/1989_Keirn_and_Aunon.

5.2.2Methodology

In this section,the block diagram of the proposed methodology has been described along with the short introduction of the associated signal processing and machine learning methods used in this work. Our proposed feature extraction approach which is based on the combination of MEMD based data driven decomposition of the six channelEEG data and multivariate autoregressivemodel(MVAR)consists of a number ofstages as shown in Figure 5.1

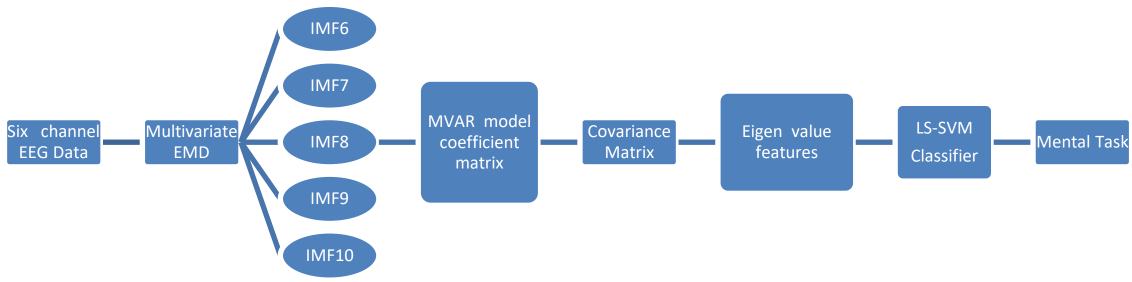


Figure5.1: Block diagram representation of the proposed methodology

We employed multi channel EEG data since single channel EEG data cannot capture the neural activity dynamics which is distributed over the entire cortical surface. In the first stage, we applied MEMD algorithm on the single trial, six channel EEG data for decomposing the same into a large number of IMF groups. MEMD enabled us to capture the nonlinear and non-stationary nature of the signals. In the second stage, we selected a small subset of task sensitive IMF groups (IMF_6- IMF_{10}) according to their task correlation factor. This selection of sensitive IMF groups enabled us to get rid of unwanted noises and artifacts present in the other IMF groups. In the third stage, multivariate AR models of order six were computed from all the five selected IMFgroups. In the fourthstage,a set of non-zero eigen values were computed from the covariance matrices corresponding to coefficient matrices of the six orderMVAR models in the IMFdomain. The eigenvalues represent the magnitude of spreads (variances) of the data.Finally, feature vectors were constructed from these computed eigen values and subsequently used for training and testing a LS-SVM classifier.

5.2.2.1Multivariate Empirical Mode Decomposition

For catering to the need of real world multichannel data, Rehman and Mandic [111] extended the standard univariate EMD algorithm to multivariate data known asMEMD algorithm. This MEMD decomposes multivariate signal into several IMF groups where each IMF group has the same length and components containing the same frequency distribution, in the same order of the group. In this algorithm, the local mean of n –dimensional signals are computedby the multiple n-dimensional envelopes, which are generatedby taking signal projections along different directions in n-variatespaces.Fora uniform set ofdirection vectorstoprojectthe signal, low discrepancy Hammersley sequences [113] are used to obtain quasi-uniform points on high dimensional spheres [114]. The IMFs are obtained iteratively through sifting process [115]. This sifting process stops when all the K projected signals fulfill the stopping criterion [116].

The details of MEMD are outlined in [111].

5.2.2.2 Selection of most sensitive IMF group using their power spectra

The multivariate EMD (MEMD) algorithm based adaptive decomposition of the six channel(C_3,C_4,P_3, P_4, O_1,O_2) EEG data corresponding to single trial performance of each task generated twelve number of intrinsic mode function (IMF) groups and one residue function depicting the trend only. Each IMF group represents one frequency level and are generated in descending order of high frequency i.e. the first IMF group represents highest frequency level and subsequent IMF groups will have decreasing frequency levels.All these constituent IMF groups represent mono-component oscillatory mode common across the six EEG channels.

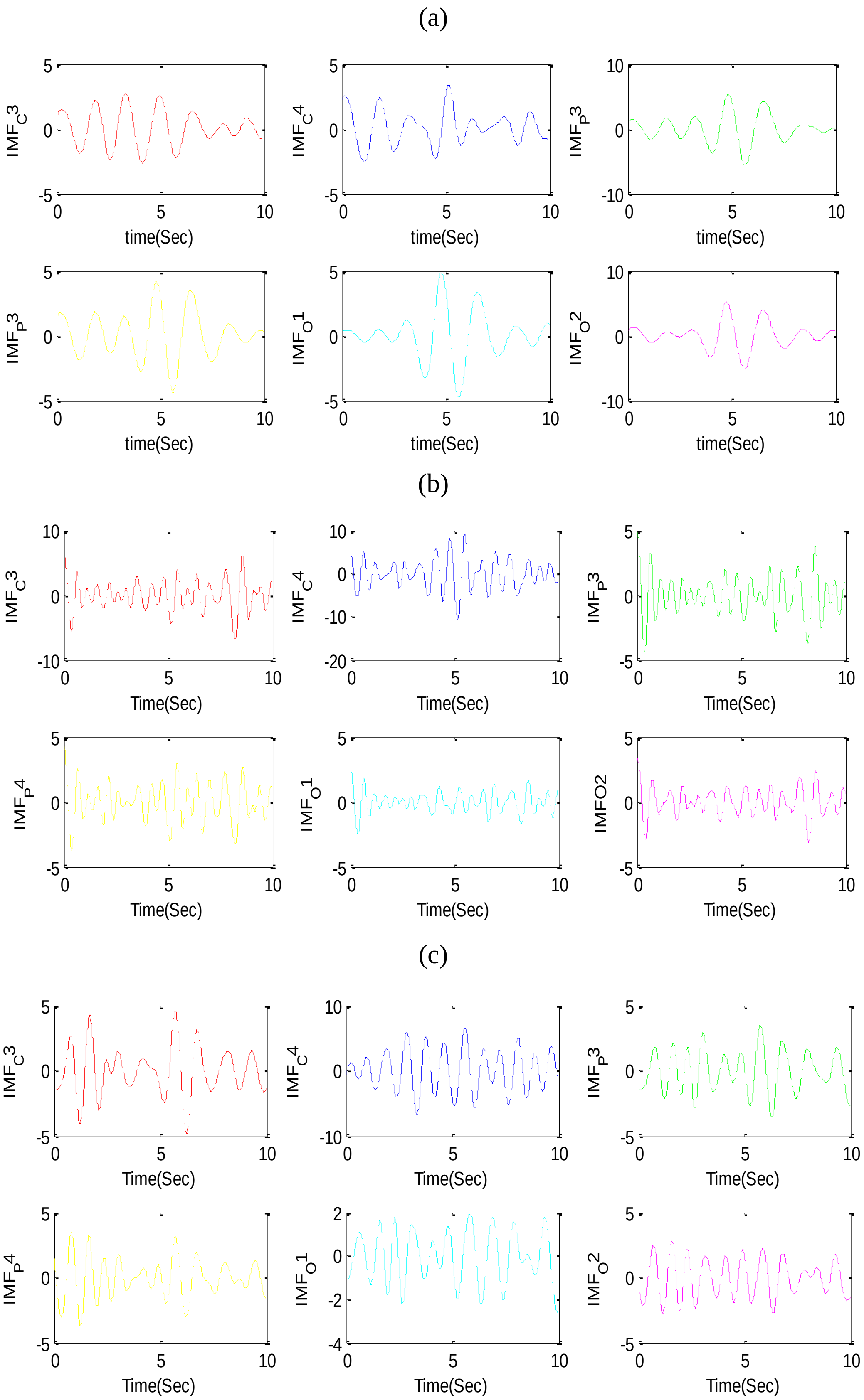


Figure5.2: Waveforms ofIMF components of the most sensitive IMF groups for (a) Base Line Task (b) Mental Multiplication Task (c) Mental Letter Composing Task

The above Figure 5.2 visualizes the waveforms i.e. the variation in amplitude of the component intrinsic mode functions (IMFs) within the most sensitive IMF group for three mental tasks. The IMF groups are generated from the MEMD based adaptive decomposition of the six channel EEG signals corresponding to single trial performance of three classes of non-motor mental task.The six component IMFs within an IMF group have same oscillatory frequency but have different amplitudes. Each of these IMF group represents mono-component or narrow band oscillatory component.

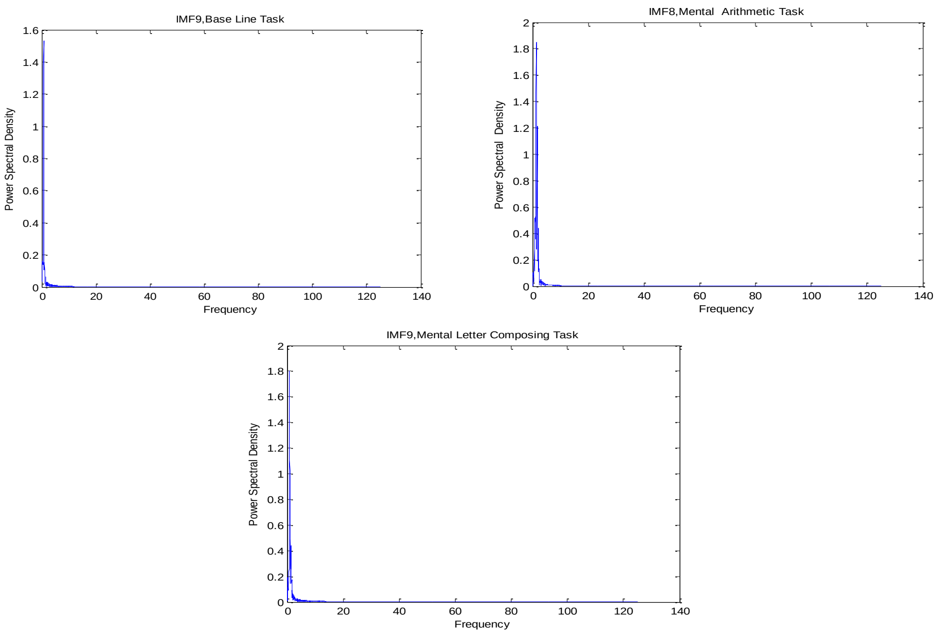


Figure5.3: Powerspectraof the most sensitive IMF groupsforthe three tasks.

All the twelve number of generated IMF groups were analyzed through their power spectral density (PSD) obtained using Welch method. From Figure5.3, it is observed that for each trial of a mental task, a particularIMF group is having highest value ofPSD i.e. the signal power is concentrated in that oscillatory component represented by the same IMF. This particular IMF group may be considered to be the most relevant for that task.For each and every trial of mental arithmetic task, the most sensitive IMF group is found to be IMF8 whereas for both mental letter composing and baseline task, the most sensitive IMF group is found to be IMF9.

5.2.2.3 Multivariate Autoregressive (MVAR) model

Auto regressive (AR) model is a linear time series model used for analysis of scalar i.e. univariate time series data. In this model, the present value of the variable is expressed as the weighted linear sum of its previous values.

An autoregressive model of order pis defined by

$$x_t = -\sum_{k=1}^{p} a_k . x_{t-k} + e_t \qquad (6)$$

Where x_t is the data at sampled point n, $a_k$are co-efficientof the ARmodel and e_t is Gaussian white noise with mean zero.

The model order, p can be selected as order with minimum Akaike Information Criterion (AIC) [81], and AR coefficients are estimated byBurg's method [117]. Attempt is made to model the current value of the variable as weighted linear sum of its previous values. The order of the model is the number of preceding observations used and the weights characterize the time series. Multivariate AR model is multivariate extension of scalar AR model i.e.it is a linear model of multivariate time series. In this, the vector of current values of all variables is modeled as a weighted linear combination of previous observations plus a random, uncorrelated input

$$\mathrm{x(k)} = -\mathrm{A(1)x(k-1)} - \mathrm{A(2)x(k-2)} - \mathrm{A(3)x(k-3)} + \ldots\ldots\ldots\ldots\ldots - \mathrm{A(p)x(k-p)} + \mathrm{e(k)} \qquad (7)$$

The multivariate AR model coefficient matrices from all six channels are calculated using a multivariate extension of the Levinson algorithm [118]. Although AIC yielded an order of two[119], we used a six order AR model as a previous study[22] with similar data achieved good performance using a scalar AR model of order six . MVAR models can capture the cross- channel information which may be relevant to classification of mental task. MultivariateAR models have been employed for applications in control, communications and sensory array processing [119-121].

5.2.2.4 Covariance matrix and eigen vector decomposition

Covariance matrix is always a symmetric matrix having diagonal elements as variances and off-diagonal elements as covariances of the data. It defines both the spread (variance) and orientation (covariance) of a data matrix i.e. it defines the shape of the data matrix. Eigen values and eigen vectors define a covariance matrix uniquely i.e. covariance matrix can be represented by a vector whose direction is given by the direction of largest spread of the data. The largest eigen vector points into the direction of the largest variance of the data where the largest Eigenvaluegives the magnitude of the spread (variance) in this direction.The Eigen values represent the magnitude of variance in the direction of largest spread of the data i.e. in the Eigen vector direction.

5.3 Results

The Eigen values were computed from the EVD of the correlation matrix corresponding to the sixth order MVAR model of the sensitive IMF groups. This eigen vector decomposition of the correlation matrix generates feature space for the classification of EEG signals related to different mental task. Only the non-zero Eigen values were used for constructing the feature vectors corresponding to single trial (10 sec) performance of a specific mental task. Table 5.1(a)-5.1(c) shows the mean and standard deviation of the non-zeroEigen value features for ten trials of a mental task. The extracted features are in time domain and this makes our approach suitable for real time implementation. The extracted feature vectors were classified by LS-SVM classifier with various kernel functions.

Table 5. 1(a): Mean and standard deviation of Eigen values of sensitive IMF groups for ten trials of a task for subject#1

Sensitive IMF groups	Baseline (mean±std)	Mental arithmetic (mean±std)	Mental letter composing (mean ± std)
6	0.8760±0.2459	2.4998±0.1013	1.0556±0.5665
7	1.1937±0.3617	2.7127±0.2759	1.3265±0.7199
8	1.7243±0.5369	3.6647±0.8143	1.7071±0.6530
9	2.2134±0.3519	4.5365±1.1398	2.0256±0.6532
10	2.5721±0.2641	7.7793±2.9735	2.5846±0.4594

Table5. 1(b): Mean and standard deviation of Eigenvalues of sensitive IMF groups for ten trials of a task for subject#3

Sensitive IMF groups	Baseline (mean±std)	Mental arithmetic (mean±std)	Mental letter composing (mean ± std)
6	1.0735±0.2588	2.4485±0.0797	0.9305±0.2777
7	1.5992±0.3994	2.6160±0.0739	1.4682±0.4193
8	2.0214±0.3408	2.7962±0.1677	1.9753±0.3070
9	2.3306±0.1502	3.9327±1.1426	2.2690±0.2255
10	2.5097±0.0746	6.2179±1.8355	2.5057±0.1134

Table 5.1(c): Mean and standard deviation of Eigenvalues of sensitive IMFs for ten trials of a task for subject#5

Sensitive IMF groups	Baseline (mean±std)	Mental arithmetic (mean±std)	Mental letter composing (mean ± std)
6	0.8071±0.2260	2.2884±0.1679	0.8345±0.2076
7	1.1293±0.4993	2.5383±0.1125	1.5233±0.3933
8	1.6170±0.4431	2.6710±0.1522	1.8719±0.3766
9	2.0524±0.3909	3.5478±0.8024	2.3044±0.1883
10	2.4752±0.1521	5.0118±1.5628	2.5260±0.0598

Table5.2(a): Average classification accuracy of LS-SVM classifier for three subjects for classifying a pair of mental task

Pair of Mental Task	Linear Kernel Function	Polynomial Kernel Function	Radial Basis Kernel Function
Baseline Vs Mental Arithmetic	91.65%	94.43%	91.65%
Mental Arithmetic Vs Mental Letter Composing	88.88%	88.88%	91.65%
Baseline Vs Mental Letter Composing	66.66%	65.16%	54.16%

Table 5.2(b): Average classification accuracy of multi-class LS-SVM classifier for three subjectsforthreeclassclassificationof mental task

Linear Kernel	Polynomial Kernel	Radial Basis Function
77.77%	74.07%	74.07%

5.4 Discussion

In this research, we extended the work of C.W. Anderson et al [81] through computing sixth order multivariate AR (MVAR) models from the IMF groups of a six channel, non-motor EEGsignals in theMEMD domain. The application ofMEMD algorithm on the six channel EEG data gave rise to finite number of IMF groups where each IMF group contains six component IMFs. All the six component IMFs of a particular IMF group have common frequency of oscillation but they originate from different cortical locations. Due to simultaneous decomposition of the six channels EEG data by the MEMD algorithm, equal number of mode aligned IMFs were generated per channel. This enabled us to circumvent the twin problems of mode mixing and mode alignment faced by standard univariate EMD algorithm when applied to multichannel data for sequential i.e. channel by channel analysis of the same. From the set of generated IMF groups, a small subset of task sensitive groups was selected as per their degree of correlation with a specific mental task. In the next stage, we computed six order MVAR model from these sensitive IMF groups. ThoughMVAR is a linear model, it can capture the inter regional dependency within the multivariate data.

In this work, we performed binary i.e. pair wise classification as well as the three class classification of three non-motor cognitive tasks. For binary classification, only one SVM is required, but for multi-class classification, a number of SVMs are combined according to schemes like OneVsOne, OneVsAll, and ECOC etc.

One LS-SVM with three different kernel functions was employed for binary classification task because of its high performance in pattern recognition problems. We employed tenfold cross-validation for assessing the pair wise classification of mental task using LS-SVM classifier. The results of binary i.e. pair wise classification is shown in Table 5.2(a). The highest average classification accuracy obtained with this approach is 94.43% for the first pair i.e. baseline- mental arithmetic using polynomial kernel function.The results of binary classification for the mental task pair between mental letter composing and baseline as shown in the third row of Table 5.2(a) is poor. For all the three kernels, achieved classification accuracy is low.

From this poor classification accuracy, we may generate the hypothesis that the dynamical patterns of brain activity under mental letter composing task is similar to that of under baseline i.e brain activity patterns under mental letter composing task does not change significantly with respect to baseline i.e relaxed state.

The highest average classification accuracy for three class classification is77.77% using linear kernel function and OneVsOne scheme of multiclass SVM.

The benefits of our proposed multivariate approach are highlighted as follows:

(a) MEMD algorithm enabled us to circumvent the twin problems of mode mixing and mode alignment faced by standard EMD algorithm while analyzing the EEG channels sequentially.

(b) The filter bank property of MEMD enabled us to select a subset of meaningful i.e. task sensitive IMF groups and get rid of the unwanted. Besides this, the use of the covariance matrix and its Eigen values as features are in time domain which make them suitable for real time implementation.

(c) Due to use of multichannel EEG data and our multivariate approach, we could capture the dynamical changes at multiple spatio-temporal scales.

(d) MVAR models extracted from the IMFs groups provide better representation of the coupled dynamics relevant to our classification.

(e) The Eigen values of the covariance matrices corresponding to the co-efficient matrices of the MVAR model capture the geometric property of the signals.

(f) We employed LS-SVM as binary classifier as well as three class classifier for its high performance in pattern recognition problems.

Table 5.3: Comparison of the proposed methodology

Authors	Features used	Classifier used	Classification accuracy (%)
Kern and Anon(1990)	FFT based band power Scalar AR model coefficients Asymmetry ratio	Quadratic Bayesian classifier	81% -87.5%
Charles W Anderson, Z Latko,Sijercie (1996)	AR model co-efficient	Back Propagation Neural Network (BPN)	71%
		Probabilistic Neural Network (PNN)	38%
Charles W .Anderson , Erik A. Stolz , SanyogitaShamsunder (1998)	Multivariate AR model, Scalar AR coefficients, Eigen values of Correlation matrix	Feed forward neural network trained via error back propagation algorithm	86.1 to 91.4%
RamaswamyPalaniappan, RaveendranParamesran, Shogo Nishida, Naoki Saiwai (2002)	Power Spectral Density (PSD)	Fuzzy ARTMAP	83.5%
Deon Garret, David A .Peterson, Charles W Anderson, Michael H.Thaut, (2003)	AR Model Co-efficient Power Spectra	SVM/ANN/LDA	76%(highest)
CK HO, M Sasaki(2005)	Wavelet transform	Neural Network	75%
Nan-Ying Liang, ParamasivamSaratchandran, Guang-Bin-Huang NarasimhanSundarajan (2006)	AR Model Co-efficient	Extreme learning Machine	50%
Diez et al (2009)	RMS, Variance, Shanon Entropy, Lempel-Ziv complexity etc in EMD domain	Linear Discriminant Analysis(LDA)	91.17%
Martina Tolic, FranjoJovik(2013)	DWT based features	Back Propagation Neural Network	90.75% (for multiclass) 99.87% (for binary)
Zin Mar Lwin(2015) Mie Mie Thaw	Gabor based Matching Pursuit(MP)	Multiclass SVM Classifier	(c) 89.73% (OneVsOne) (d) 52.43%(OneVsAll)
S.K. Agarwal(2015) Saatvik Shah Rajesh Kumar	Power Spectral Density by Welch Periodogram	Backtracking search optimization based neural classifier (BSANN)	83.32%(highest)
MouniaHendel, AbdelkaderBenyetton, FatihaHendel (2016)	-	Hybrid Classifier using Self Organizing Map and Probabilistic Quadratic Loss Multiclass SVM	81.73 % to 91.9%
Our current work	Eigen values extracted from the covariance matrices in MEMD domain	LS-SVM	94.3% (binary) 77.77% (three class)

5.5 Conclusions

In this study, we explored the applicability of eigenvalues of the covariance matrix corresponding to co-efficient matrix of the multivariate AR model as novel features for classifying three non-motor cognitive task in EEG based BCI system. The coefficient matrix of the sixth order multivariate AR model corresponding to the most sensitive IMF group generates very useful feature space for classifying mental task EEG signals. The extracted features are in time domain and this makes our approach suitable for real time implementation.The extracted feature vectors were classified byLS-SVM classifier.The highest value of average classification accuracy of 94.3%obtained for binary classification exceeds 91.4% ofthe same in previous works .Besides;we obtained average classification accuracy of77.7%for three class classification.Our proposed approach for analysis of EEG signals using MVAR model in the MEMDdomain is novel and quite encouraging. It opens up a number of directions for further analysis of EEG signals in diverse application domains. Future research may involve extracting eigen value features from the 2Dor 3DPSRof the most sensitive IMF or a number of sensitive IMFs and the results should be compared with the current research approach.

CHAPTER6

MULTIVARIATE MULTISCALE ENTROPY BASED COMPLEXITY MEASURES AS FEATURES FOR EEG CLASSIFICATION IN MT BASED BCI

In this chapter, we propose multivariate multi scale version of three traditional entropy based complexity measures, namely, sample entropy, permutation entropy and fuzzy entropy as EEG features for classifying a small set of non-motor mental task in mental task (MT) EEG BCI. These entropy measures hold the capacity to quantify the structural complexity manifested in the long range cross-correlation structure present across the data channels of any multivariate data set. These entropy values were computed from a small subset of five sensitive intrinsic mode functions (IMFs) generated through multivariate empirical mode decomposition (MEMD) of the six channel EEG data corresponding to single trial performance of a mental task. These computed entropy values from each trial of a task have been employed separately to form the overall feature matrix for all the thirty trials (ten trials per task of a single subject) of three mental tasks. Finally, 70% data from this feature matrix have been employed for training and remaining 30% for testing a LS-SVM classifier. We achieved highest classification accuracy of 100% for the two pairs (baseline-mental arithmetic, mental arithmetic- mental letter composing) using multi-scale multivariate fuzzy entropy (MMFE) values and tenfold cross validation. The classification results for the other two features i.e. multi-scale multivariate sample entropy (MMSE) and multi-scale multivariate permutation entropy (MMPE) were also promising and support in favor of their potentials to be used as EEG feature in BCI system and other application of EEG analysis.

6.1 Introduction

A BCI is a system that “measures central nervous system (CNS) activity and converts it into artificial output that replaces, restores, enhances, supplements, or improves natural CNS output and thereby changes the ongoing interactions between the CNS and its external or internal environment” [6-8]. It is a pattern recognition system which works by recognizing the subtle dynamical patterns of brain’s electrical activity induced by a specific mental task, namely mental arithmetic or mental letter composing [21]. The development of such type of real time system requires fast, reliable and accurate classification of EEG signals related to different cognitive task. But due to its complex characteristic and extremely noisy background, classification of these signals remains the central challenge in developing efficient and effective BCI system. Combination of better EEG feature with optimum machine learning algorithm may lead to enhanced performance of BCI systems. This has made search for better EEG feature an open research problem for real time BCI system development. Though it is not universally accepted till date, "with the development of nonlinear dynamics in recent years, more and more evidences indicate that human brain is a nonlinear dynamic system and EEG signal can be regarded as its output (D.Duke and W.Pritchard, 1991; W.S. Pritchard and D.W. Duke, 1995; P.Faure and H.Kom, 2001; H.Kom and P.Faure, 2003)" [122-125]. A chaotic system is a semi- deterministic nonlinear dynamical system. To obtain better classification result, many researchers are trying to use chaotic measures (entropy, fractal dimension, largest Lyapunov exponent etc) as better EEG features. In 1985, Babloyantz et al [126] first put forward that "II and IV stage EEG signals of human sleep cycle are chaotic" i.e. a kind of nonlinear dynamics. Therefore, people started analyzing EEG signals through a number of nonlinear chaotic measures like entropy, fractal dimension, Lyapunov exponent etc. A number of traditional entropy based complexity measures such as approximate entropy(Pincus, S.M, 1991)[127], Kolmogorov entropy(Grassberger,P et al, 1993)[128], sample entropy (Richman, J.S et al, 2000)[129], permutation entropy (Bandt,C et al , 2002)[130] ,fuzzy entropy(Chen.W et al, 2007)[131]etc were developed to quantify the subtle and hidden complexity existing in the EEG signal. Various versions of these entropy measures were employed to characterize EEG signals in different pathological states as well as during complex cognitive task like mental arithmetic, mental letter composing etc.

Modern neuroscience has established that human brain activity patterns show structural complexity which is manifested in the long range spatio-temporal correlation structure present across the data channels. Meaningful information regarding brain states is encoded in such type of structural complexity measures. Quantifying such type of structural complexity is the key for understanding brain functions under different mental states. This has necessitated the development of complexity measures which can quantify structural complexity over multiple spatio-temporal scales. This has further necessitated the development of multivariate multi-scale entropy (MMSE) by extending standard MSE algorithm to multivariate time series (Ahmed et al., 2011, Ahmed et al., 2012)[132]. But till date, no research report is available in the literature exploring the applicability of these structural complexity measures for classification of EEG data related to non-motor mental task. Therefore, the main motivation and novelty of our research work stems from exploiting the structural richness of multichannel EEG data through computing multivariate multi-scale version of different entropy based complexity measures ,namely ,MMSE , MMPE and MMFE for analysis and classification of EEG signals related to non-motor cognitive task The main contribution of our research comes in demonstrating the applicability of these multivariate multi-scale entropy values as features for classifying non-motor mental task in EEG based BCI system.

6.2. Data and Proposed Methodology

6.2.1. EEG Database

In this research, we have used benchmark EEG data set acquired by Keirn and Aunon (1990) [22] from Purdue University. They chose a small set of five non-motor mental task including baseline and designed the experimental protocol for the acquisition of the same.These tasks can invoke brain wave asymmetry in the activity patterns and hence offers the possibility to be classified for BCI system. This EEG dataset is available online at http://www.cs.colostate.edu/eeg/main/data/1989_Keirn_and_Aunon.

6.2.2. Proposed Methodology

Our proposed feature extraction approaches based on structural complexity measures for classifying six channel non-motorEEG data is shown through the block diagram as shown in Figure6.1.

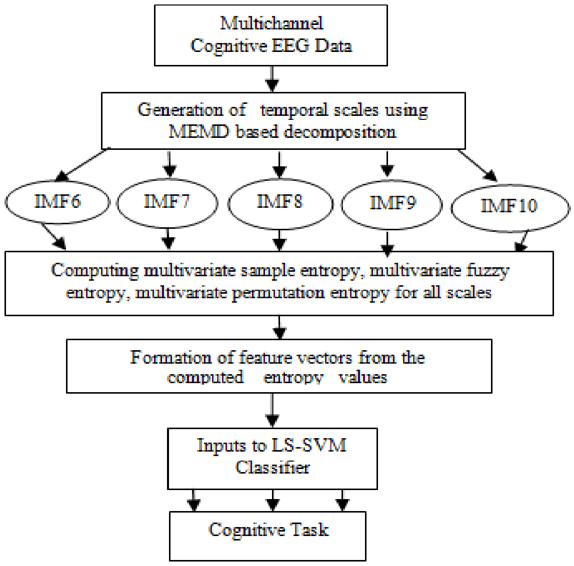


Figure 6.1: Proposed methodology

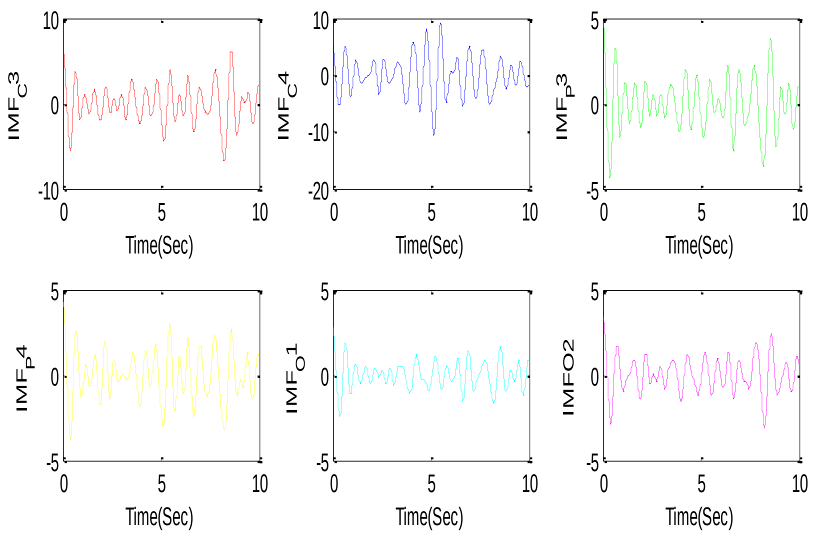
From the above Figure 6.1,it is apparent that in the first step,temporal scales are generated by applying MEMD algorithm or by coarse graining procedure. Once the temporal scales are generated, the three different multivariate versions of entropy values are computed over these temporal scales.Feature vectors are constructed from these computed entropy values. Finally, the extracted feature vectors are employed for training and testing a LS-SVM classifier.

6.2.2.1. Multivariate Empirical Mode Decomposition

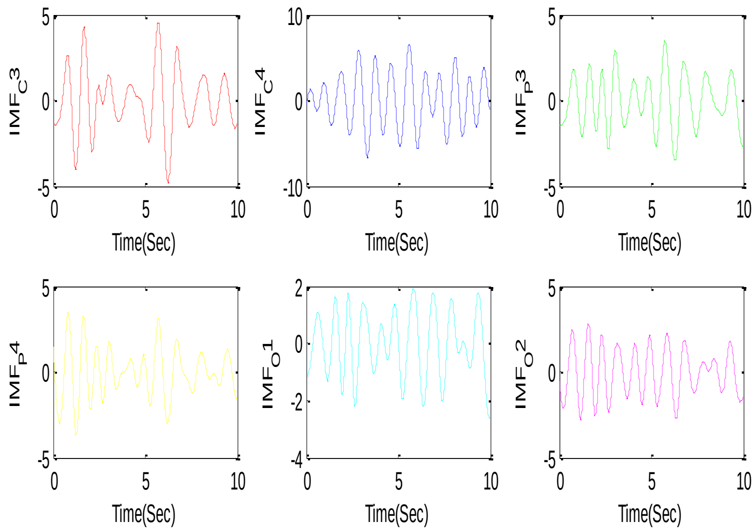
For catering to the need of real world multichannel data, Rehman and Mandic [111] extended the standard univariate EMD algorithm to multivariate data known as MEMD algorithm.In this algorithm, the local mean of n –dimensional signals are computed by the multiple n-dimensional envelopes, which are generated by taking signal projections along different directions in n-variate spaces. For a uniform set of direction vectors to project the signal, low discrepancy Hammersley sequences [113] are used to obtain quasi-uniform points on high dimensional spheres [114]. The IMFs are obtained through an iterative process known as sifting[115]. Finally, the iterative process of this MEMD algorithm is stopped when all the K projected signals fulfill the stopping criterion [116].

The steps of the MEMD algorithm are given in [111].

(a)



(b)



(c)

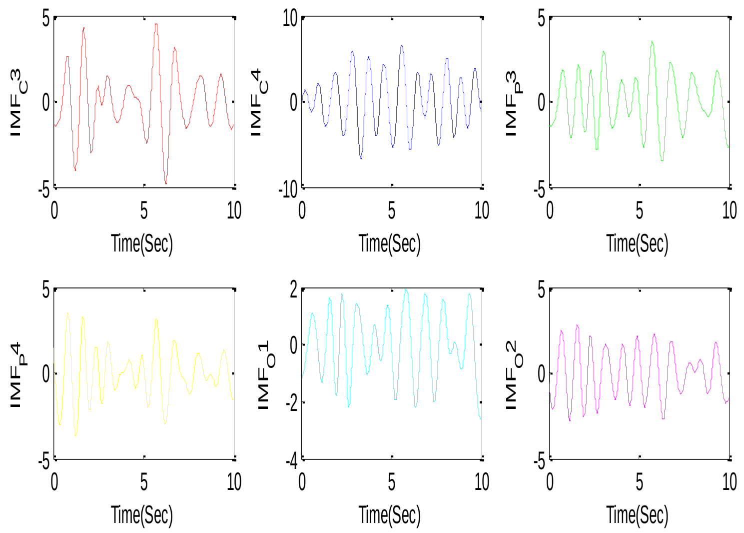


Figure 6.2: Waveforms of six IMFs of sensitive IMF group (a) Base Line Task (b) Mental Multiplication Task (c) Mental Letter Composing Task

6.2.2.2 Multivariate sample entropy

All biological and physiological signals are inherently multidimensional in nature and contain rich information encoded in their long range cross-correlation structure that varies over multiple spatio-temporal scales. Traditional entropy measures i.e., sample entropy, permutation entropy quantifies the complexity of a univariate time series atone scale onlywhereasMulti scale entropy (MSE) is a nonlinear chaotic measure which can quantify the complexity of finite length time series data over a range of scales. The original Multi-Scale Entropy(MSE) algorithm, proposed by Costa et al (2002)[133] computes the sample entropy over a range of scales generated from the coarse-grained time series.The scales can be derived from the original signals through the method of coarse graining or multivariate empirical mode decomposition (MEMD) based decomposition.

The algorithm for Multivariate sample entropy (MSampEn) developedbyAhmed et al, 2011,is described in[132].

6.2.2.3 Multivariate permutation entropy

Multi-scale permutation entropy (MSPE) [134] was developed as multi-scale version of original permutation entropy (C.Bandt and B.Pompe, 2002)[130]. MSPE computes permutation entropy at different temporal scales**.** Multi-Scale permutation entropy (MSPE) can be employed for the complexity analysis of a single channel time series data. Being anunivariate approach, it cannot quantify the long range cross correlation structure present over multiple spatio-temporal scales. But the EEG signals are inherently multi-dimensional in nature due to volume conduction effect. The sequential channel by channel analysis of complexity of multivariate EEG leads to inevitable information loss related to cross-channel dependency. Shaobo He et al, 2016[140] developed multivariate permutation entropy (MvPE) as multivariate version of PE for complexity analysis of chaotic systems.

Multivariate extension of MSPE (Francesco Carlo, Morabito et al, 2012)[137] has been developed like multivariate multi scale sample entropy (MMSE) for capturing the information related to cross-channel dependency.Themultivariate extension of MSPE, termed as MMPEcomputespermutationentropyfor a multivariate signal, andover different time scales.MPE extracts the cross-channel complexity of a multichannel signal by highlighting the long-range spatial (nonlinear) correlations. The algorithm for the computation of multivariate permutation entropy value is given by Shaobo et al (2016)[140] .

6.2.2.4. Multivariate Fuzzy Entropy

Chen et al, (2007) [137] developed a new statisticalmeasureknown asFuzzy entropy (FE). It can evaluate self-similarity and thereby can measure the complexity and irregularity of a signal. InFE, the Heaviside function is replaced with a fuzzy membership function.Due to this, FE is strongly consistent, less dependent on data length, continuous and more resistant to noise. Chen et al reported that FE gives more accurate and reliable result than SampEn for short time series. Since FE computesentropy for univariate signalatone scale,it isnotapplicabletomeasure complexity and irregularity of multivariate signals over multiple scale factors. Zheng et al (2014) [141] developed the multiscale version of Fuzzy Entropy (MFE) through a multiscale framework by combining FE and scale factor. They used Gaussian type fuzzy membership function.Later,Ahmed et al (2017)[13] developed multivariate multiscale fuzzy entropy (MMFE) as fuzzy versionof multivariate multi-scale entropy (MMSE). This MMFE evaluates MFSampEn over different time scales and is able to analyze very short signals. The algorithm for the computation of multivariate fuzzy sample entropy (MvFE) is similar to that of multivariate sample entropy. The only difference arises in replacing the Heaviside function in sample entropy by a fuzzy membership function in fuzzy entropy. The description of the algorithm is given by Ahmed et al (2016)[143].

6.3. RESULTS

6.3.1. Multivariate multi-scale entropy values of six channel EEG data related to non-motor mental task

We decomposed the single trial EEG data into large number of IMFs derived from the data adaptively. But all these IMFs are not useful i.e. contain meaningful information. So, we selected a small set of five IMFs according to their correlation co-efficient values with the original EEG data. This enabled us to get rid of the noise and artifacts present in the signal. We computed multivariate sample entropy, multivariate permutation entropy, and multivariate fuzzy entropy values over five different time scales defined by the five IMFs. The temporal scales for multivariatesample entropy and multivariatefuzzy entropy were defined by five sensitive intrinsic mode functions (IMF_6-IMF_{10}) selected as per their degree of correlation with the original EEG signals.

For the multivariate multi-scale permutation entropy, the scales were obtained from a number of coarse grained time series.

The mean and stdvalues of all the three multivariate multi-scale entropy measures for ten trials of a mental task are presented in Tables 6.1 (a)-6.1(c).

Table 6.1(a): Mean and stdof MMSEfor ten trials of three mental tasks

IMF No	Base line (mean±std)	Mental arithmetic task (mean±std)	Mental letter composing (mean±std)
6	0.8668±0.4464	0.3237±0.0469	0.3528±0.0763
7	0.9248±0.5724	0.4088±0.0741	0.3890±0.1253
8	1.0878±0.5252	0.7139±0.1851	0.5731±0.2228
9	0.9188±0.6046	0.3587±0.1771	0.5261±0.3040
10	0.9830±0.5968	0.4847±0.1703	0.5731±0.3524

Table 6.1(b): Mean and stdof MMPE for ten trials of three mental task

IMF No	Baseline (mean±std)	Mental arithmetic (mean± std)	Mental letter composing (mean±std)
6	3.0505±0.0426	3.0854±0.0239	3.0573±0.0345
7	2.9827±0.0549	3.0351±0.0329	3.0280±0.0155
8	3.0382±0.0492	3.0625±0.0503	3.0662±0.0252
9	3.0648±0.0301	3.0989±0.0378	3.0770±0.0252
10	3.0116±0.0496	3.0792±0.0328	3.0530±0.0292

Table 6.1(c): Mean and stdof MMFE for ten trials of three mental tasks

IMF No	Baseline (mean±std)	Mental arithmetic (mean±std)	Mental letter composing (mean ±std)
6	0.2174±0.0509	0.6288±0.0875	0.2379±0.0640
7	0.2571±0.0593	0.8027±0.0705	0.2924±0.1244
8	0.5125±0.1886	1.1370±0.1652	0.4237±0.1148
9	0.2933±0.0838	0.7653±0.2081	0.4571±0.3642
10	0.4268±0.2293	0.9800±0.1839	0.5216±0.3054

We performed Student's t-test for each pair of entropy values separately for evaluating their discriminatory power. The results of t-test for the three entropy values are presented in Table6. 2(a)-6.2(c).

Table 6.2(a): Probability ('p') values in Student's t-test for MMSE

Pair of Mental Task	IMF6	IMF7	IMF8	IMF9	IMF10
Baseline-Mental arithmetic	0.0061	0.0141	0.0587	0.0153	0.0310
Mental arithmetic-Mental letter composing	0.3885	0.7175	0.1496	0.0743	0.4004
Baseline- Mental letter composing	0.0054	0.0300	0.0202	0.0730	0.1105

Table 6.2(b): Probability ('p') values in Student's t-test for MMPE

Pair of Mental Task	IMF6	IMF7	IMF8	IMF9	IMF10
Baseline-Mental arithmetic	0.0145	0.0165	0.0814	0.0213	0.0008
Mental arithmetic-Mental letter composing	0.1164	0.6094	0.8569	0.2462	0.0896
Baseline- Mental letter composing	0.7652	0.0373	0.1832	0.4214	0.1131

Table 6.2(c): Probability ('p') values in Student's t-test for MMFE

Pair of Mental Task	IMF6	IMF7	IMF8	IMF9	IMF10
Baseline-Mental arithmetic	0.0002	0.0001	0.0231	0.6480	0.2558
Mental arithmetic-Mental letter composing	0.0000	0.0000	0.0000	0.0061	0.0011
Baseline- Mental letter composing	0.5144	0.3466	0.1000	0.2061	0.4902

For binary i.e. pair-wise classification, we constructed the overall feature matrix of twenty trials (ten trials per task), out of which we used 70% data for training and remaining 30% data for testing the LS-SVM classifier. We employed LS-SVM classifier with three different kernel functions using ten-fold cross validation.It uses a set of linear equations instead of nonlinear programming and learns the LS-SVM model considering all the data points.This reduces the computational burden of the LS-SVM classifier significantly.

Table 6.3(a): Classification performance of LS-SVM classifier using MMSE feature

Pair of mental task	Linear Kernel					Polynomial Kernel					Radial Basis Function Kernel				
	ACC	SEN	SPE	PPV	NPV	ACC	SEN	SPE	PPV	NPV	ACC	SEN	SPE	PPV	NPV
Baseline- Mental Arithmetic	83.33	75	66.66	75	100	66.66	60	33.33	60	100	50	50	00	50	00
Mental Arithmetic-Mental LetterComposing	33.33	33.33	33.33	33.33	33.33	33.33	33.33	33.33	33.33	33.33	50	50	66.66	50	50
Baseline - Mental Letter composing	50	50	00	50	00	66.66	66.66	66.66	66.66	66.66	50	50	00	50	00

Table 6.3(b): Classification performance of LS-SVM classifier usingMMPE feature

Pair of mental task	Linear Kernel					Polynomial Kernel					Radial Basis Function Kernel				
	ACC	SEN	SPE	PPV	NPV	ACC	SEN	SPE	PPV	NPV	ACC	SEN	SPE	PPV	NPV
Base Line Vs Mental Arithmetic	83.33	100	100	100	75	66.66	100	100	100	60	100	100	100	100	100
Mental ArithmeticVs Mental LetterComposing	50	50	66.66	50	50	50	50	66.66	50	50	66.66	100	100	100	60
Base Line Vs Mental LetterComposing	66.66	100	100	100	60	83.33	100	100	100	75	66.66	100	100	100	60

Table 6.3(c): Classification performance of LS-SVM classifier usingMMFE feature

Pair of mental task	Linear Kernel					Polynomial Kernel					Radial Basis Function Kernel				
	ACC	SEN	SPE	PPV	NPV	ACC	SEN	SPE	PPV	NPV	ACC	SEN	SPE	PPV	NPV
Base Line VsMental Arithmetic	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Mental ArithmeticVs Mental LetterComposing	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Base Line VsMental Letter Composing	33.33	33.33	33.33	33.33	33.33	16.25	25	00	25	00	16.25	25	00	25	00

6.3.2. Tuning the hyper parameters of the LS-SVM classifier with RBF kernel function

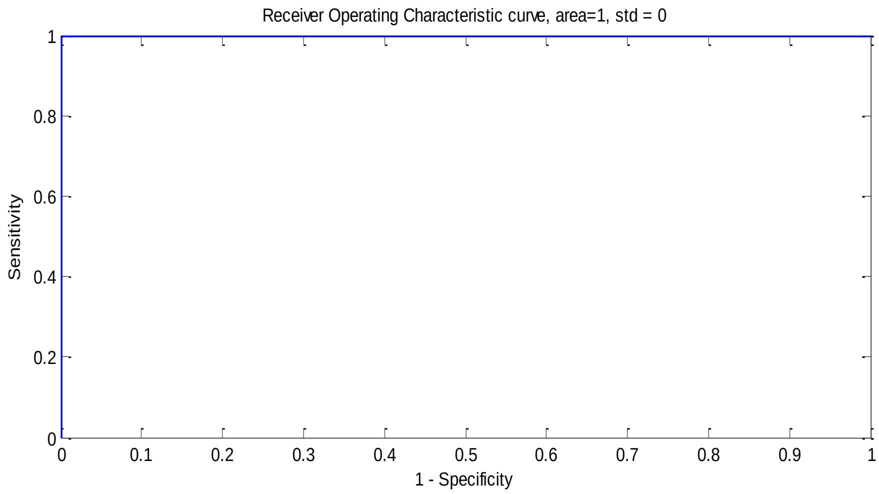
For improving the generalization performance of the LS-SVM classifier with Gaussian RBF kernel, tuning of its two hyper parameters plays important role .These two important hyper parameters should be tuned appropriately for achieving a desirable performance. These tuning parameters are found in a two stage process by using a combination of Coupled Simulated Annealing (CSA) and a standard simplex method. In the first stage, CSA finds the good initial values which are fine-tuned by the grid search or simplex method in the second stage.

Table 6.4: Hyper parameters (gamma, sigma2) values for various cases of binary classification

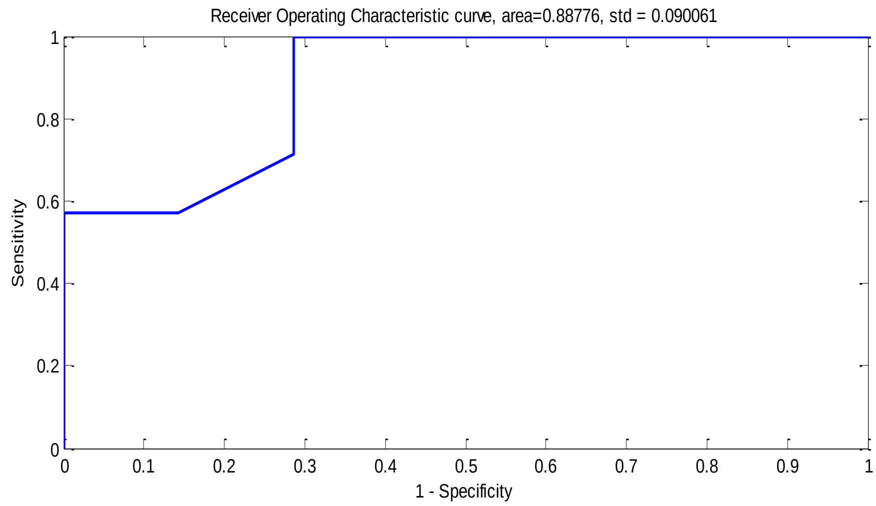
Pair of mental task	MMSE	MMFE	MMPE
Baseline/mental arithmetic	0.2138, 0.1482	2.2262, 1.1524	1.3216, 2.3519
Mental arithmetic/Letter composing	3.8107, 14.8169	1.2798, 2.0423	1.2616, 0.0568
Baseline/ Letter composing	0.6995, 0.2910	1.0125, 2.5320	5.0204,0.0017

These two hyper parameters significantly affect the performance of the LS-SVM. In order to achieve the best results, LS-SVM is trained with different combinations of these parameters. The performance of LS-SVM classifier for the three types of entropies was evaluated separately using various parameters such as ACCURACY, SENSITIVITY, SPECIFICITY ,PPV,NPV and the area under receiver operating characteristic(ROC)curve .Ten-fold cross-validation was employed to ensure better performance of the classifier.

1. **MMSE**



2. MMPE



3. MMFE

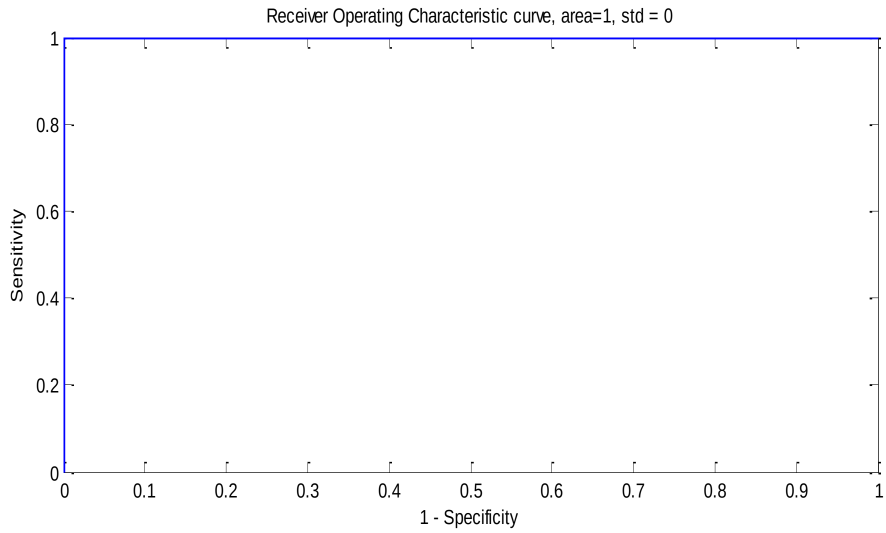


Figure6.3: Receiver operating characteristic (ROC) curve for (a) MMSE (b) MMPE (c) MMFE

6.4 Discussion

In this research, we investigated the applicability of three different multi-scale multivariate entropy (MMSE,MMPE, MMFE) based structural complexity measures as EEG feature for classifying three non-motor cognitive task in EEG based BCI system. We used six channelEEG signals corresponding to one trial (10 second duration) of three non-motor mental task including baseline. EEG signals are inherently multidimensional and shows structural complexity which contains wealth of information. We captured such type of information through computing multi-scale and multivariate version of traditional entropies at IMF defined scales. From the classification results as shown in Tables 6.3(a)- 6.3(c), it is clear that multivariate fuzzy entropy (mvFE) gave highest classification accuracy of 100% for the first two Pairs of mental task (i.e. baseline - mental arithmetic , mental arithmetic- mental letter composing) with all the three kernel functions. The reason of such a high classification accuracy is due to high discriminatory power of MEMD enhanced multivariate fuzzy entropy (mvFE) values as manifested in the low value of Probability ($p<0.05$) as shown in Table 6.2(c). At the same time, the performance of this feature for the third pair(baseline-mental letter composing) is very poor. The reason for their poor classification performance is reflected in their poor discriminatory power as apparent in the test results of paired t-test. The performance of the multivariate permutation (mvPE) values computed over different scale factors was 100% and 83.33% for the first pair (baseline-mental arithmetic) of mental task using RBF and linear kernel functions respectively. But its performance for the other two pairs of mental task was moderate(66.66%) as manifested in the Table6.3(b). The MEMD enhanced multivariate sample entropy (mvSE) values as EEG feature gave highest classification accuracy of 83.33% for the first pair (baseline- mental arithmetic) of mental task using linear kernel function. But it gave moderate performance (66% accuracy) for the other two pairs of mental task. The reason for this moderate value of classification accuracy may be due to selection of improper kernel function and imperfect tuning of the kernel and classifier parameters. Addressing these two issues by automatic selection of kernel functions and optimizing the classifier parameters may enhance the performance of the classifier. From the results of statistical significance test and classification performance, it can be inferred that the similar brain activity patters are induced by both baseline and mental letter composing, while there is significant differences in the activity patterns between baseline and mental arithmetic task. In this way, our findings corroborate the established fact that complex arithmetic calculation requires greater cognitive load than letter composing task. . Higher cognitive load will generate more complex brain activity patterns. Baseline i.e. mentally relaxed state will require least cognitive load generating least complexity dynamics which is reflected in their least value of standard deviation. In this study, we have achieved the highest SEN and SPF of 100% for the first two pairs of mental task for all the three kernel functions using MEMD enhanced multivariate fuzzy entropy values. The highest values achieved for both PPV and NPV are 100% for the first two pairs of mental task using all the three kernel functions. From the performance comparison with the previous studies as shown in Table 6.5, it is clear that our proposed feature extraction approach using multi-scale multivariate version of three different entropies (MMSE, MMPE and MMFE) produced better results. The obtained accuracy of 100% by MMFE and MMPE demonstrate the superiority of our approach over the previous approaches. Besides providing better classification accuracy, our MMFE based approach offers the following advantages:

(i) MMFE is able to analyze very short length time series, where the utility of MMSE is hampered.

(ii) Due to use of fuzzy membership function instead of Heaviside function, MMFE value does not change abruptly; its change is smooth, continuous and robust to small changes in tolerance 'r'.

(iii) As there is no rigid boundary in a fuzzy membership function, MMFE is able to determine the class boundaries when the input pattern belonging to a given class is ambiguous as well as imprecise.

(iv) Since all the computations are in time domain, our approach is suitable to real time implementation

Few of the drawbacks of our approach are following:

(i) The selection of kernel function and its parameters for the LS-SVM classifier are manual.

(ii) Sensitive IMFs for a particular task is based on the degree of correlation between the IMF and the original signal. Considering the non-Gaussian nature of the signals and their IMFs, this criterion may not be appropriate.

Considering the above mentioned drawbacks, future research endeavor should focus on the following:

(i) Selection criterion for sensitive IMFs based on mutual information should be examined.

(ii) Instead of using the three entropy values separately, combining them together and applying feature selection algorithm may be undertaken.

(iii) For computing the entropy values, selection of embedding parameters should be optimized.

(iv) For computing MMFE, a number of fuzzy membership functions, namely, Gaussian, Sigmoid, bell shaped function etc.may be tried for finding the best result.

(v) The kernel function should be selected automatically and its parameters are to be optimized for enhancing performance of the LS-SVM classifier.

Table6.5: Comparison of classification results of the proposed multivariate multi scale entropy with existing features

Authors	Features used	Classifier used	Accuracy (%)
Keirn and Anon (1990)	FFT based band power Scalar AR model coefficients Asymmetry ratio	Quadratic Bayesian classifier	81 -87.5
Charles W Anderson, Z Latko,Sijercie (1996)	AR model co-efficient	Back Propagation Neural Network (BPN)	71
		Probabilistic Neural Network (PNN)	38
Charles W .Anderson , Erik A. Stolz , SanyogitaShamsunder (1998)	Multivariate AR model, Scalar AR coefficients, Eigen values of Correlation matrix	Feed forward neural network trained via error back propagation algorithm	86.1 to 91.4
RamaswamyPalaniappan,RaveendranParamesran, Shogo Nishida, Naoki Saiwai (2002)	Power Spectral Density (PSD)	Fuzzy ARTMAP	83.5
Deon Garret, David A .Peterson, Charles W Anderson, Michael H.Thaut, (2003)	AR Model Co-efficient Power Spectra	SVM/ANN/LDA	76(highest)
CK HO, M Sasaki(2005)	Wavelet transform	Neural Network	75
Nan-Ying Liang, ParamasivamSaratchandran, Guang-Bin-Huang NarasimhanSundarajan (2006)	AR Model Co-efficient	Extreme learning Machine	50
Diez et al (2009)	RMS, Variance, Shanon Entropy, Lempel-Ziv complexity etc in EMD domain	Linear Discriminant Analysis(LDA)	91.17
Martina Tolic, FranjoJovik (2013)	DWT based features	Back Propagation Neural Network	90.75(for multi-class) 99.87(for binary)
Zin Mar Lwin(2015) Mie Mie Thaw	Gabor based Matching Pursuit(MP)	Multiclass SVM Classifier	a) 9.73 (OneVsOne) b)52.43 (OneVsAll)
S.K. Agarwal(2015) Saatvik Shah Rajesh Kumar	Power Spectral Density by Welch Periodogram	Backtracking search optimization based neural classifier (BSANN)	83.32 (highest)
MouniaHendel, AbdelkaderBenyetton, FatihaHendel (2016)	-	Hybrid Classifier using Self Organizing Map and Probabilistic Quadratic Loss Multiclass SVM	81.73 to 91.9
Our previous work	1. Largest singular values in the phase space of IMFs	(i)LS-SVM	83.33
	2. Eigen values from the MVAR model in the MEMD domain	(ii) LS-SVM	94.43
Our current work	MMSE	LS-SVM classifier	83.33
	MMPE		100
	MMFE		100

6.5 Conclusion

Human brain is a complex adaptive system (CAS) showing structural complexity manifested in the long range cross correlation structure across the data channels of multichannel EEG data. In this study, we captured this structural complexity through multivariate multi-scale version of different entropy measures and demonstrated their use in binary i.e. pair-wise classification of three different pairs of non-motor cognitive task in multitask BCI systems . So far these structural complexity measures have been used for examining the complexity loss in EEG due to aging, Alzheimer's (AD) and other neurological disorders like epilepsy, cerebral palsy (CP), Parkinson's disease etc. The application of these entropy based complexity measures for classification of mental task in EEG based BCI system is novel. For the case of MMFE, we obtained highest classification accuracy of 100% using LS-SVM classifier for two pairs of mental task (mental arithmetic-baseline, mental letter composing-mental arithmetic). Hence, we can conclude that MEMD enhanced multivariate fuzzy entropy has better discretionary power than the others two entropies. Our proposed multivariate approach for capturing structural complexity provides a fertile ground for taking up a number of further research through exploring many new features and improving the computation of the existing feature values. The obtained test result validates the applicability of these multi-scale multivariate entropy values as BCI features and opens up a number of direction for further analysis of EEG signals in diverse application domains such as clinical diagnosis of neurological disorders like epilepsy, cerebral palsy etc.

CHAPTER7

SINGULAR VALUE FEATURES IN THE RECONSTRUCTED PHASE SPACE

In this chapter, we present singular values (SVs) in thephase spaceof EEG as novelfeature forclassifyingnon-motormental tasksinMT based BCI**.** In our approach, first we reconstructed the phase space of single trial, six channel non-motors EEG signals. In the second stage, the SVs of the covariance matrix corresponding to the phase space trajectory matrix of one trial EEG were used for constructing the feature vectors .Feature matrix corresponding to a specific mental task was formed from the feature vectors of ten trials of that particular mental task. Overall feature matrix composing of three different mental tasks was formed from these three individual task specific feature matrices. Finally, 70% data from this feature matrix i.e. seven trials were employed for training and remaining 30% i.e. three trials for testing the performance of a set of classifiers. Their performances were assessed for three class classification as well as binary classification using various parameters.We tested our approach on five subjects and three mental tasks, namely, baseline,mental arithmetic andletter composing task. We achieved highest value of average classification accuracy of 75.54% for three class classification using linear discriminate analysis (LDA) and 96.66% for binary classification between baseline and mental letter composing task usingLS-SVM. The singular values in the RPS of EEG as feature provide us a new perspective to view the non-linear dynamics of cognitive brain activity.

7.1. Introduction

Fast, reliable and accurate classification of these EEG signals related to performance of different cognitive task is the central challenge for designing real time BCI systems. Keirn and Aunon (1990)[22,75,76] proposed that EEG signals could distinguish between various mental tasks accurately. They designed the experimental protocol and acquired EEG signals related to five non-motor imagery mental tasks which invoke hemispheric brainwave asymmetry. They used AR model coefficients and band power asymmetry ratio as features with quadratic Bayesian classifier for classifying five non-motor mental task EEG signals which was acquired by them. Later, Anderson et al(1998)[78-81] employed MVAR model of EEG for the classification of same set of five non-motor mental task(1990). They continued the work of Keirn and Aunon and derived both scalar AR model as well as multivariate AR model from the raw EEG signals without employing any decomposition.The classification performances of these previous works are sub optimal since these employed either FFT based non-parametric method orparametric andlinearmethods of time series analysis like autoregressive (AR) model.Due to its complex origin, they show extremely complex and subtle patterns containing wealth of information. Decoding this extremely complex patterns and extracting information from its noisy background is a formidable task. The conventionalFFTbased non-parametric andlinear parametric auto-regressive (AR) model of time series analysis are inadequate to capture the extremely complex patterns of brain electrical activity. Combination of best features with optimum machine learning algorithm will lead to better decoding of these complex patterns which is required for enhanced performance of BCI system. As a result, search for better EEG features having more discriminatory power is yet an open research problem for improved classification of these EEG signals. Though it is not universally accepted till date, with the development of nonlinear dynamics in recent years, "more and more evidences indicate that human brain is a nonlinear dynamic system and EEG signal can be regarded as its output"(P.Faure and H.Kom,2001; H.Kom and P.Faure,2003)[124,125] . In 1985, A. Babloyantz et al [126] first put forward that "II and IV stage EEG signals of human sleep cycle are chaotic i.e. a kind of nonlinear dynamics." Hereafter, a large number of study results reported that the EEG was derived from chaotic systems (D.Duke et al, 1991; W.S.Pritchard et al, 1995; A.Babloyantz et al, 1986) [122,123, 126]. Therefore, people startedanalyzingEEG signals through a number of nonlinear approaches including time-frequency, time-scale transforms, higher-order statistics,chaos theory to get new knowledge of the brain function. In this research, we adopted the dynamical systems and chaos theory based approach for capturing the nonlinear aspects of the cognitive brain dynamics induced by non-motor cognitive task. The basis of ourapproach is to embed a signal intohigh dimensional reconstructed phase space (RPS). The topological structure of RPS is equivalent to the original phase spaceof the system generating the signal. Due to its powerful representing capacity, the attractor and its associated RPS offers a fertile ground for investigating new features required for signal analysis and classification. The prime motivation of our work stems from exploiting the rich information contained in the geometrical structure of the RPS. In our previous research, we used the largest singular values in the RPS of the task sensitive multivariate IMFs as feature for classifying same set of non-motor mental task [106] .Our current research using SVs in the RPS of the original EEG signal as new feature for classifying the same set of non-motor EEG signals can be viewed as continuation of our earlier work[106]. In this work, we intend to demonstrate the applicability of singular valuesas discriminating feature for classifying non-motor cognitive task. Our main contribution comes in demonstrating the applicability of the singular values in the RPS as new feature for classifying a small a set of non-motor EEG signals such as mental arithmetic, mental letter composing etc. in a new BCI paradigm ,called multi task BCI system.

7.2 Methodology

Our proposed methodology employing multivariate phase space reconstruction and singular value decomposition of the single trial, six channel EEG data is described through the following block diagram in Figure7.1.

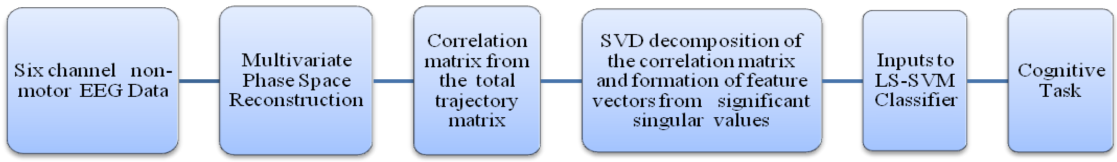


Figure 7.1: Block diagram of the proposed methodology

From the aboveblock diagram ,it is clear that the dynamics of thesingle trial, six channel EEG time series data is first reconstructed in its reconstructed phase space(RPS) by employing the technique of multivariate phase space reconstruction .After phase space reconstruction, covariance matrix was computed from the of multivariate phase space trajectory matrix generated through this procedure .SVD was applied on the covariance matrix to find out the set of singular values for constructing the feature vector .Finally, LS-SVMwas employed as a nonlinear classifier for classifying the feature vectors into a class of mental task.

7.2.1 Phase Space Reconstruction

Chaotic dynamics, a case of nonlinear dynamics is represented by phase space trajectory,known as strange attractor. Phase space reconstruction put forward by Packard et al, 1980 [139] is a powerful technique for detecting the hidden structure of such dynamics.Strange attractor or the phase space trajectory in the reconstructed phase space provides a very useful framework for extracting and visualizing the nonlinear dynamical behavior of the signal over time known as phase space representation (PSR). The mathematical basis of this practical technique lies in the time delay embedding theorem by F.Taken, 1981[140]. This technique is applied for accurate analysis of chaotic time series without increase in computational complexity. In this technique, the phase space of ‘d’ dimensions of a scalar time series:{x_i: i= 1,2,3,4,……N} can be represented by : $X_i = x_i, x_{i+t}, x_{i+2t},$

In this high dimensional RPS, the signal becomes a trajectory of unknown dimension and shape. The points in the reconstructed phase space usually converge to a manifold or some other subspace of the n-dimensional Euclidean space or attractors of fractal dimension. Phase space reconstruction based on Takens' embedding theorem (1981) requires two variables- the time delay, 'τ' and the embedding dimension'd' (Jaskowski, 1995). As stated by Takens(1981),embedding dimension'd'should be higher than attractor dimension ' D' according to the relationship : $d \geq 2D + 1$ .The proper choice of these parameters has important impact on the quality of the obtained results in applications . The embedding theorem by Takens(1981) guarantees the preservation of topological properties of the attractor. For a nonlinear system like human brain, the exact value of'd' i.e.embedding dimension is not known. Due to finite number of samples in a practical biological time series (including noise), the attractor reconstructed in the phase space is not ideal. This makes determination of time delay so important in the phase space reconstruction of the original signal(Klikova et al, 2011)[141] . In the literature, several methods by which the value of time delay can be selected are mentioned. These include: autocorrelation, mutual information, higher order statistics, fill factor and wavering product. In this work, we reconstructed the dynamics of the six channel EEG data in the RPS by employing procedure of multivariate phase space reconstruction. For this procedure, we used the value of embedding dimension m=2 and time delay τ=3. We could have used different values of m like m =3,4, 5 etc. The trajectory matrix of the six channel EEG data is the concatenation of the component trajectory matrices, computed for each component.The 2D PSR for a component phase space trajectory corresponding to three mental task is shown in Figure 7.4

7.2.2 Singular Value Decomposition

The SVD ofamatrix X is given by

$$\boldsymbol{X} = \boldsymbol{U} \sum \boldsymbol{V}^{\boldsymbol{T}} \quad \textbf{(8)}$$

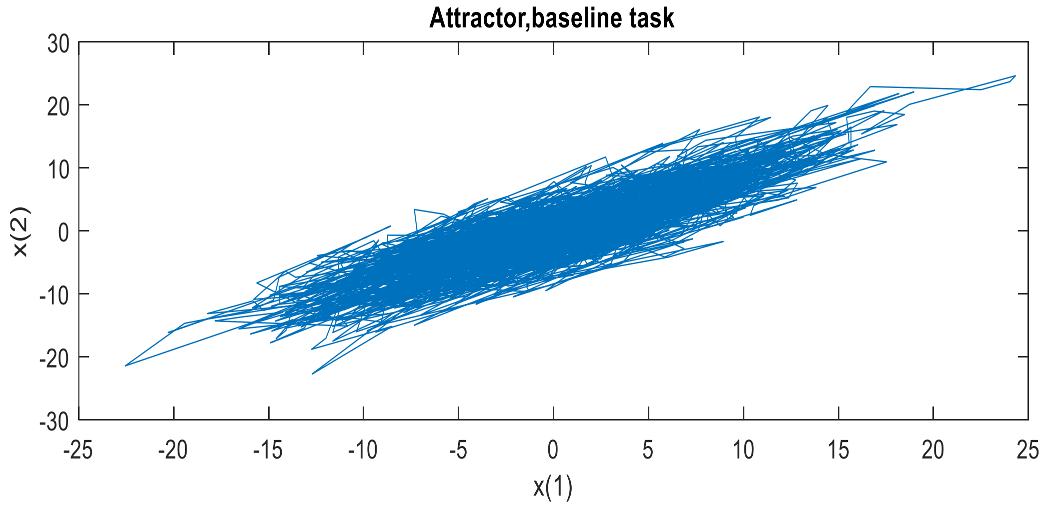
The singular values (σ_{ii}) represent the importance of individual singularvectorsin the composition of the matrix. SVs corresponding to the larger singular values have more information about the structure of patterns embedded in the matrix than the other SVs. A set of statistically significant singular values of the covariance matrix associated with the total trajectory matrix of a multichannel EEG signal cancharacterize the signal in the RPS.

In this study, we used six channel (C_3, C_4, P_3, P_4, O_1, O_2) EEG signals corresponding to single trial performance (10 second duration) of three non-motor mental tasks, namely, mental arithmetic, mental letter composing and baseline. We reconstructed the dynamics of these multichannel EEG signals by multivariate phase space reconstruction based on Taken's embedding theorem. The multivariate phase space trajectory matrix generated from such reconstruction is concatenation of the component trajectory matrices. Through repeated trial, Xingyuan Wang et al (2010) [142] found that the value of time delay (τ=3) and data point N=2000 reconstructs the attractor quite well. Assuming these values for time delay and embedding dimension, they reconstructed the attractors of all five kinds of non-motor cognitive tasks. Getting inspiration from their results, we used the value of embedding dimension, D= 2 and time delay (τ=3) for constructing the EEG attractors of baseline and two non-motor cognitive tasks as shown in Figure3. We adopted the procedure of multivariate phase space reconstruction. The temporal dynamics of a signal is best represented by its phase space representation (PSR) in the RPS. The covariance matrix of the multivariate phase space trajectory matrix represents the spatio-temporal dynamics of the original multivariate signals. The geometrical structure such spatio-temporal covariance matrix encodes wealth of information regarding the coupled spatio-temporal dynamics. We applied SVD on the covariance matrix and computed singular values as these control the geometrical structure of this spatio-temporal co-variance matrix.

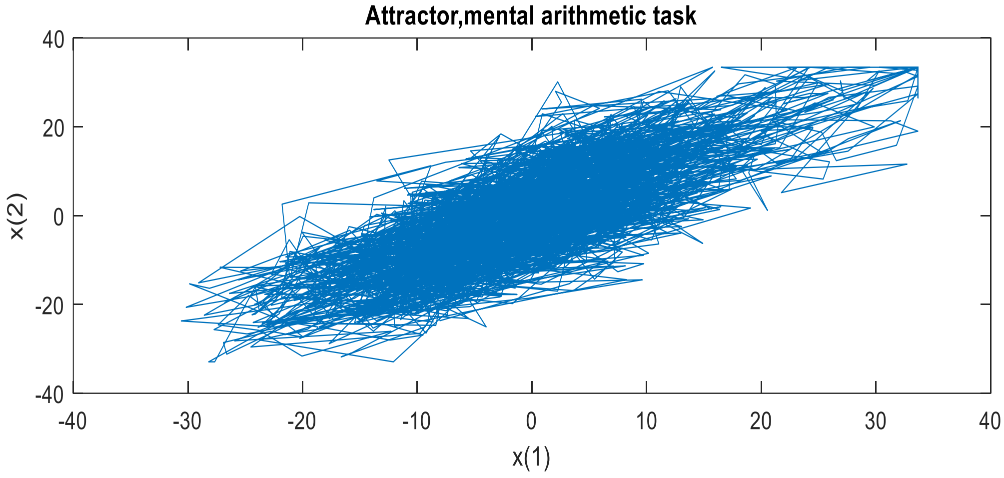
7.3 Results

Visualization of attractors for single channel EEG data

(a)



(b)



(c)

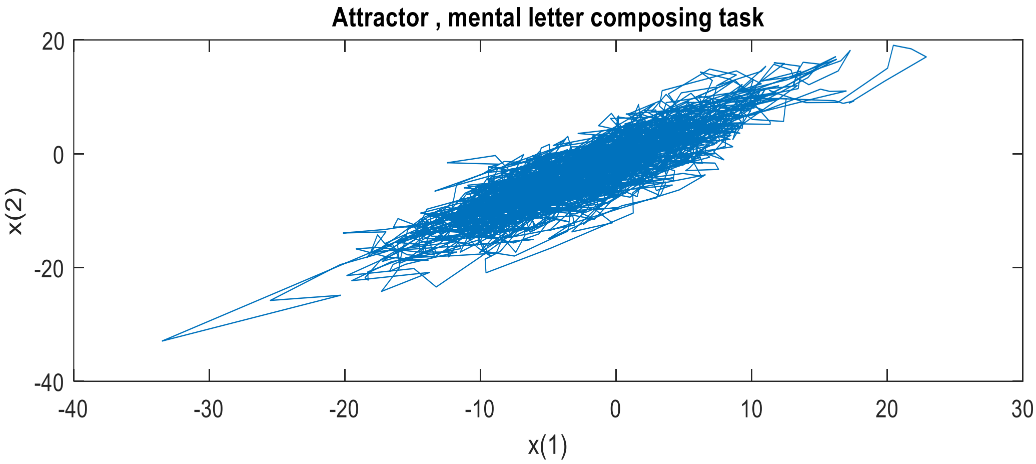


Figure 7.2: EEG attractors for (a) baseline (b) mental arithmetic task (c) mental letter composing task.

Figure7.2 visualizes the two dimensional EEG attractors of six channel EEG signals corresponding to single trial performance of three non-motor mental tasks. From the figure, it is clear that there is significant differences in the geometrical and topological properties of the attractors reflecting the underlying nonlinear dynamics induced under a particular mental task. The geometrical and topological structure of this attractor is significant as the same characterizes the true nonlinear dynamics of brain activity .This result supports that that the changes induced in the brain activity pattern during mental letter writing task is not so significant as compared to mental arithmetic task.

The feature vector was formed for each trial of a mental task using the singular values of the covariance matrix and feature matrix were formed from the ten trials of each mental task

Table 7.1: Mean andstdof the singular values for ten trials of three mental tasks for five subjects

(a) Subject1

Singular value no	Baseline (mean±std)	Mental arithmetic (mean±std)	Mental letter composing (mean ± std)
1	107.6517±21.2825	140.7809±39.2751	109.5785±25.2020
2	23.2641±7.8081	26.7266±5.0228	23.4549±4.7779
3	10.3239±2.0310	11.8959±1.9152	10.6025±1.0790
4	4.4633±1.3960	5.6904±0.9129	5.9165±0.9107
5	2.9504±0.9906	3.1957±0.8485	3.2440±0.7069
6	1.6454±0.9918	1.8528±1.1773	1.4085±0.6819

(b) Subject3

Singular value no	Baseline (mean±std)	Mental arithmetic (mean±std)	Mental letter composing (mean ± std)
1	597.3024±276.4733	625.3489±210.4848	582.1427±272.9761
2	103.8268±36.7456	77.7080±22.3690	92.3831±16.4185
3	65.3166±13.1991	55.1532±16.1894	66.8112±8.6582
4	21.0941±6.5521	14.2881±3.9870	17.5940±4.9154
5	10.4121±1.7970	10.3493±2.2907	12.5852±2.1301
6	4.1019±0.8973	3.9022±0.8259	3.3892±1.4282

(c) Subject4

Singular value no	Baseline (mean±std)	Mental arithmetic (mean±std)	Mental letter composing (mean ± std)
1	204.5474±25.6243	377.1392±133.1890	383.2305±173.6161
2	33.4138±22.2140	78.6034±139.5386	61.6456±51.1597
3	20.2159±12.9099	27.9439±17.2187	17.2855±2.6222
4	7.8422±5.4362	10.5338±14.3584	6.2513±2.3111
5	4.0517±2.0821	4.6272±1.9816	3.3775±0.5811
6	1.8191±0.7940	2.6692±1.1098	2.0271±0.3987

(d) Subject5

Singular value no	Baseline (mean±std)	Mental arithmetic (mean±std)	Mental letter composing (mean ± std)
1	244.9918±58.5428	429.3409± 199.8600	349.1693±241.1841
2	57.6701±18.0279	66.4582±18.9781	79.9099±37.1848
3	31.8460±12.1290	45.0232±13.4008	43.4685±14.6270
4	16.2812±8.8338	20.6235±11.5865	21.0266±8.8616
5	6.3572±2.0174	6.9314±2.3526	7.8171±2.1434
6	2.6642±0.6040	2.4995±0.5593	3.4499±0.9148

(e) Subject6

Singular value no	Baseline (mean±std)	Mental arithmetic (mean±std)	Mental letter composing (mean ± std)
1	525.2842±192.1669	735.3513±255.2634	530.2355±115.0922
2	93.7031±25.2136	121.3603±22.0858	133.1679±37.3172
3	61.9057±13.1736	86.0398±7.8742	81.5469±9.3988
4	17.3072±3.5170	22.9010±5.8551	21.0417±3.2730
5	12.9850±2.3821	13.1869±2.2161	13.8116±1.7813
6	5.5946±1.4540	6.2285±1.7433	6.0756±1.0449

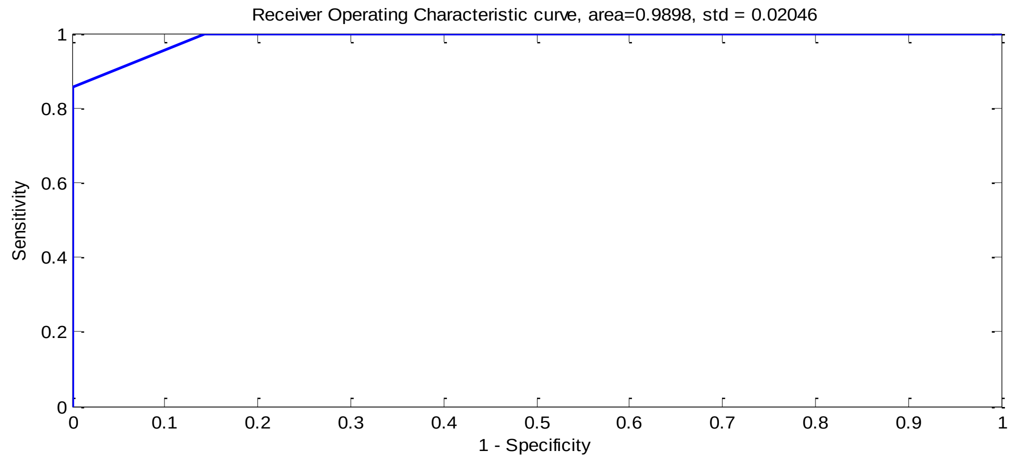
Table 7.2: Comparison of classifier performance (three class classification)

Classifier	Subject1	Subject3	Subject4	Subject5	Subject6	Average
	Accu (%)	Accu (%)	Accu (%)	Accu (%)	Accu (%)	Accu (%)
LS-SVM	88.88	55.55	77.77	66.66	55.55	68.88
LDA	88.88	77.77	55.55	77.77	77.77	75.54

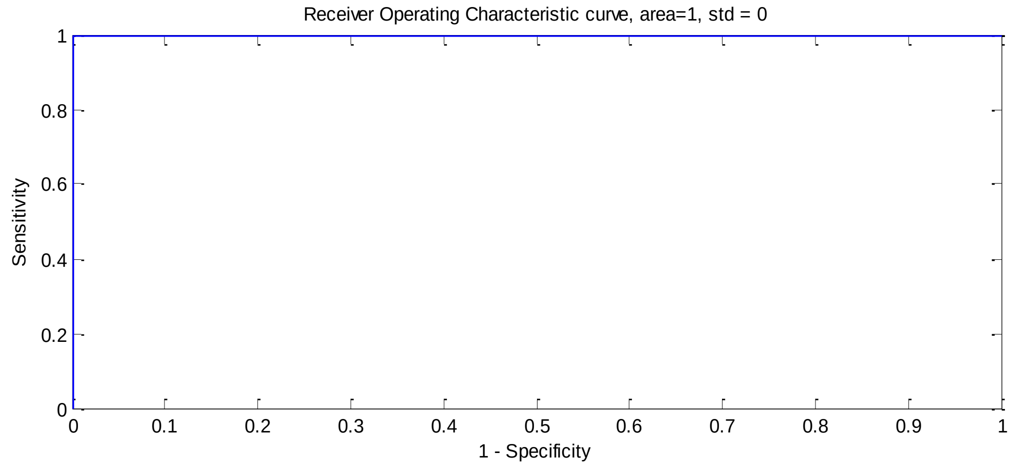
Table 7.3: Comparison of classifier performance (binary classification)

CLASSIFIER	Pair of mental task	Subject1			Subject3			Subject4			Subject5			Subject6			Avg
		ACC (%)	SEN (%)	SPE (%)	ACC (%)	SEN (%)	SPE (%)	ACC (%)	SEN (%)	SPE (%)	ACC (%)	SEN (%)	SPE (%)	ACC (%)	SEN (%)	SPE (%)	ACC (%)
LS-SVM	Base Line Vs Mental Arithmetic	83.33	75.00	66.66	66.67	66.67	66.67	100.00	100.00	100.00	83.33	100.00	100.00	100.00	100.00	100.00	86.86
	Mental Arithmetic Vs Mental Letter Composing	83.3	100.00	100.00	83.33	100.00	100.00	50.00	00.00	100.00	66.67	100.00	100.00	100.00	100.00	100.00	76.66
	Base Line Vs Mental Letter Composing	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	88.83	100.00	100.00	100.00	100.00	100.00	96.66
LDA	Base Line Vs Mental Arithmetic	83.33	75.00	66.66	66.66	66.66	66.66	100.00	100.00	100.00	66.66	100.00	60.00	100.00	100.00	100.00	83.33
	Mental Arithmetic Vs Mental Letter Composing	83.33	100.00	75.00	83.33	100.00	100.00	50.00	00.00	100.00	100.00	100.00	100.00	66.66	60.00	100.00	76.96
	Base Line Vs Mental Letter Composing	100.00	100.00	100.00	83.33	100.00	100.00	100.00	100.00	100.00	66.66	100.00	60.00	100.00	100.00	100	90.00
SVM	Base Line Vs Mental Arithmetic	50.00	50.00	00.00	50.00	50.00	50.00	83.33	75.00	100.00	83.33	100.00	75.00	100.00	100.00	100.00	73.33
	Mental Arithmetic Vs Mental Letter Composing	50.00	00.00	50.00	66.66	100.00	60.00	100.00	100.00	100.00	100.00	100.00	100.00	50.00	100.00	00.00	73.33
	Base Line Vs Mental Letter Composing	100.00	100.00	100..00	33.33	40.00	00.00	100.00	100.00	100.00	50.00	00.00	50.00	100.00	100.00	100.00	76.66

(a)



(b)



(c)

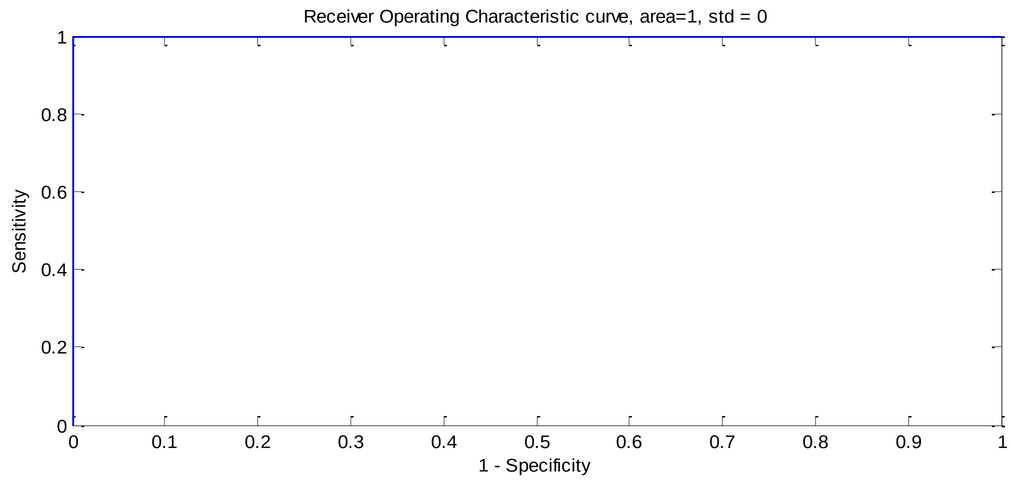


Figure7.3: ROC curves for (a) Baseline- Mental Arithmetic
(b) Mental Letter Composing -Mental Arithmetic
(c) Baseline-Mental Letter Composing

Table 7.3 describes the performance of two classifiers for binary classification through different statistical measurements, such as SEN, SPE, CAand ROC curve.The above Figure 7.3visualizes the ROC curves for binary i.e. pair wise classification of three pairs of mental task.

ROC curve: It visualizes the performance of a binary classifier.When the area under the ROC curve is 1, it indicates perfect discrimination ability. When the area equals 0.5,it no discriminative power at all.The 1or near 1 value of this area under the ROC curves for three cases of binary classifications as shown in Figure7. 5. This Figure7.5, indicates that the LS-SVM classifier model with very high accuracy.

7.4 Discussion

In this research, we employed multivariate phase space reconstruction technique with singular value decomposition (SVD) for feature extraction of EEG signals in the reconstructed phase space (RPS).The nonlinear temporal dynamics of a signal is best represented in its RPS. The covariance matrix of the multivariate phase space trajectory matrix represents the spatio-temporal dynamics of the original multivariate signals. The geometrical structure such spatio-temporal covariance matrix encodes wealth of information regarding the coupled spatio-temporal dynamics. Due to this, the phase space trajectory and the RPS provide a fertile ground for extracting new features from different perspectives. In our previous research [106], we used singular values from the RPS of the task sensitive intrinsic mode functions (IMFs) and achieved highest classification accuracy of 83.33% for binary classification. In this study, we computedsingular values from the covariance matrix ofphase space trajectory matrix corresponding to single trial, six channel EEG data for defining new feature space. The feature vector was formed for each trial of a mental task using these singular values and feature matrices were formed from the ten trials of each mental task .We performed three class classification as well as binary *i.e*. pair wise classification of three non-motor mental task using a set of classifiers. From the feature matrix of a task, we used 70% data i.e. seven trials for training and remaining 30% data i.e. three trials for testing classifiers for binary as well as three class classification. For three class classification, we employed one linear classifier, namely linear discriminate analysis (LDA) and one nonlinear classifier, namely, LS-SVM classifier (‘OneVsAll’) and achieved average classification accuracy of 75.54% and 68.88% respectively. For binary classification, we employed three different classifiers, namely, LS-SVM, LDA and SVM and assessed their performance through parameters like accuracy, sensitivity,and specificity. For binary classification, this new feature space enabled us to achieve highest average classification accuracy of 96.66% for binary classification between baseline and mental letter composing. This experimental findings corroborate the established fact that the performance of complex cognitive tasks require significant cognitive effort compared to the baseline where no cognitive effort is required ideally. The details of classification results for three class as well as binary classification are presented in Table 7.2 and Table7.3 respectively.

The advantages of our proposed feature extraction approach are:

(a) Multivariate phase space reconstruction of the six channel EEG signal enabled us to reconstruct the signal dynamics more accurately in the reconstructed phase space (RPS).
(b) This RPS provides a fertile ground which can be explored for constructing many new features
(c) The computational complexity of our technique is relatively lower than that of conventional techniques
(d) It is not strongly dependent on data length.
(e) It does not require the mathematical model of the system under study.
(f) Being in timedomain, ourapproach issuitable for real time implementation of BCI system.

Though we achieved satisfactory classification performance, our proposed approach has few challengesalso, such as (a) accurate determination of the optimum values of the embedding parametersfor accurate reconstruction of the signals (b) automatic selection of kernel function and its parameters for the LS-SVM classifier.(c) The method is susceptible to noise since it is based on time delay embedding (d) accuracy of the method is hard to guarantee.

Table 7.4: Performance comparison of the proposed approach with existing approaches

Authors	Features used	Classifier used	Classification accuracy (%)
Keirn and Anon(1990)	FFT based band power Scalar AR model coefficients Asymmetry ratio	Quadratic Bayesian classifier	81 -87.5
Charles W Anderson, Z Latko,Sijercie (1996)	AR model co-efficient	Back Propagation Neural Network (BPN)	71
		Probabilistic Neural Network (PNN)	38
Charles W .Anderson , Erik A. Stolz , SanyogitaShamsunder (1998)	Multivariate AR model, Scalar AR coefficients, Eigen values of Correlation matrix	Feed forward neural network trained via error back propagation algorithm	86.1 to 91.4
RamaswamyPalaniappan,RaveendranParamesran, Shogo Nishida, Naoki Saiwai (2002)	Power Spectral Density (PSD)	Fuzzy ARTMAP	83.5
Deon Garret, David A .Peterson, Charles W Anderson, Michael H.Thaut, (2003)	AR Model Co-efficient Power Spectra	SVM/ANN/LDA	76(highest)
CK HO, M Sasaki(2005)	Wavelet transform	Neural Network	75
Nan-Ying Liang, ParamasivamSaratchandran, Guang-Bin-Huang NarasimhanSundarajan (2006)	AR Model Co-efficient	Extreme learning Machine	50
Diez et al (2009)	RMS, Variance, Shanon Entropy, Lempel-Ziv complexity etc in EMD domain	Linear Discriminant Analysis(LDA)	91.17
Martina Tolic, FranjoJovik (2013)	DWT based features	Back Propagation Neural Network	90.75(for multiclass) 99.87 (for binary)
Zin Mar Lwin(2015) Mie Mie Thaw	Gabor based Matching Pursuit(MP)	Multiclass SVM Classifier	89.73 (OneVsOne) 52.43(OneVsAll)
S.K. Agarwal(2015) Saatvik Shah	Power Spectral Density by Welch Periodogram	Backtracking search optimization based neural classifier (BSANN)	83.32(highest)
MouniaHendel, AbdelkaderBenyetton, FatihaHendel (2016)	-	Hybrid Classifier using Self Organizing Map and Probabilistic Quadratic Loss Multiclass SVM	81.73 to 91.9
Our previous work	Singular values in the reconstructed phase space of the IMFs	LS-SVM	83.33(binary)
Our current work	Singular values in the reconstructed phase space of the original signals	LS-SVM/LDA/SVM	83.33(binary) 88.89(3 class)

7.5 Conclusion

In this study, we demonstrated the applicability of singular values in the reconstructed phase space (RPS) of the EEG signals as new features for classifying three non-motor mental tasks in EEG based BCI system. Our proposed feature extraction approach is novel as it exploits the geometrical structure of the RPS .We tested our approach for three mental states on five subjects and achieved highest average classification accuracy of 75.54% for three class classification using LDA classifier and 96.66% for binary classification using LS-SVM classifier. The singular values have dominant control over the geometrical structure of the RPS. This motivated us to employ the singular values of the covariance matrix in the reconstructed phase space as characterizing features of the signals. TheRPS is able to capture the nonlinear dynamical aspects of underlying cognitive process which cannot be fully captured by traditional linear methods. The accuracy and consistency of our test results corroborate the nonlinear dynamic characteristic of human brain and offers us a new avenue from different perspective which may encourage for further investigations. From our study, it is clear that the study of brain dynamics under profound cognitive task requires multidisciplinary approach integrating the field of neuroscience, modern mathematics, physics i.e. nonlinear dynamics and computer to reach new heights.Though, we computed the proposed feature values from the original EEG signals, computing these features from the IMFs in the multivariate EMD (MEMD) domain and evaluating their performance may form the basis of future work in diverse application domains of EEG analysis such as detection ofepilepsy, cerebral palsy, Alzheimer's disease, Parkinson's disease , migraine, sleep disorder etc.

CHAPTER 8

CONCLUSION & FUTURE WORKS

8.1Conclusions

In this dissertation, we studied and developed signal processing methods for feature extraction of EEG signals in general mental task (MT) based BCI paradigm where there are no known control signals that are consistent across tasks, subjects, participants and sessions. As a result, general mental task BCI para digmsrely more heavily on accurate signal representations and machine learning methods to identify relevant patterns on a subject and persession basis. In order to enhance the performance of classification techniques, we studied and developed three novel EEG feature extraction algorithms for supervised classification of mental tasks in BCI applications. Our proposed approaches are novel since these exploit the nonlinear, non-stationary and multi-dimensional nature of the signals from the RPSof the IMFsin the MEMD domain. MEMD based adaptive decomposition of the multi-channel signals and subsequents election of a number of sensitive IMFs offers USA new function space. The RPS of these sensitive IMFs enables us to fathom deep into the subtle structure of the signal dynamics. We tested our approach on a bench mark EEG data set and evaluated the results.The accuracy and consistency of our test results across the subjects corroborate the chaotic nonlinear dynamic characteristic of human brain and confirms our belief that the RPS of a signal is a fertile ground for exploring nonlinear features which are novel and offers us a different perspective which may encourage for further investigations. From our study, it is clear that there is correlation between nonlinear brain dynamics and performance of complex cognitive task. The study of this correlation requires multidisciplinary approach integrating the field of neuroscience, modern mathematics, physics i.e. nonlinear dynamics, signal processing.We strongly believe that the EEG feature extraction approaches presented in this dissertation will be suitable in addressing the needs BCI system.

In the first approach, we investigated the applicability of a new frame-work for EEG feature extraction based on the combination of two powerful methods, namely, MEMD based decomposition and PSR. The phase space of the sensitive IMFs were reconstructed based on the fixed values of embedding parameters (τ = 1, m=2). The largest singular values computed from the phase space trajectory matrices of the sensitive IMFs enabled us to generate very useful feature space for the classification of non-motorEEG signals.The extracted features are in time domain which causes our approach suitable for real time implementation.The extracted feature vectors had been classified using a LS-SVM classifier with different kernel functions and the results were compared with LDA and KNN classifiers.

In second approach, we explored the applicability of eigen values of the covariance matrix corresponding to co-efficient matrix of the multivariate AR model as novel features for classifying three non-motor cognitive taskin EEG based BCI system. The coefficient matrix of the sixth order multi variate AR model corresponding to the most sensitive IMF group generates feature space which is very useful for the classification of EEG signals related to different mental task. The extracted features are in time domain and this makes our approach suitable forreal time implementation. The extracted feature vectors were classified by LS-SVM classifier with three different kernel functions. The highest value of average classification accuracy of 94.3% obtained for binary classification exceeds 91.4% of the same in previous works. Besides, we obtained average classification accuracy of 77.7% for three class classification. Our proposed approach for analysis of EEG signals using MVAR modeling the MEMD domain is novel and quite encouraging. It opens up a number of directions for further analysis of EEG signals in diverse application domains. Future research may involve extract ingeigen value features from the 2 D or 3 DPSR of the most sensitive IMF or a number of sensitive IMFs and the results should be compared with the current research approach. Finally, it would be of great interest to try the proposed methodology for classification of other biomedical signals like ECG, EMG etc. in the context of developing medical diagnostic expert system software.

In the third approach, we explored the application of multi variate multi-scale entropy based structural complexity measures (MMSE, MMPE, MMFE) as features for classifying non-motor mental tasks in mental task (MT) based BCI system. Human brain is a complex adaptive systems howing structural complexity which is manifested in the multi channel EEG data. In this study, we captured this structural complexity through multivariatemulti scale version of different entropy measures and demonstrated their use in binary i.e. pair-wise classification of three different pairs of non-motor cognitive task in MT based BCI. So far, these structural complexity measures have been used for examining the complexity loss in EEG due to aging, Alzheimer’s (AD) and other neurological disorders like epilepsy, cerebral palsy(CP), Parkinson’s disease etc. The application of these entropy based complexity measuresfor classification of mental taskin EEG based BCI system is novel. For the case of MMFE, we obtained highest CA of 100% using LS-SVM classifier for two pairs of mental task (mental arithmetic-baseline, mental letter composing-mental arithmetic). Hence, we can conclude that MEMD enhanced multi variate fuzzy entropy has better discretionary power than the others two entropies. The obtained test result validates the applicability of the semulti-scale multi variate entropy valuesas BCI features and opens up a number of directions for further analysis of EEG signals in diverse application domains such as epilepsy, cerebral palsy, Alzheimer's disease, Parkinson's disease, migraine, sleep disorder etc.

Overall, it is concluded that the next generation of EEG analysis and classificational gorithms will not be based on hand crafted features characterizing the previously known patterns only. These should be based on detecting the hidden unknown patterns without making any prior assumptions. We should focus on developing deep learning based solutions capable of filtering, identifying and exploiting a wide variety of patterns automatically.

8.2Future work

We believe that this new EEG based BCI paradigm based on non-motor mental task offers greater potential than motor imagery (MI) based BCI in enhancing the quality of life of a physically disabled person. But developing such type of MT based BCIs is a challenging task since the neurological control signals driving these BCIs lack clearly defined patterns. At present, this MT based multi command BCI is relatively under explored. The spatiotemporal patterns contained in the neurological control signal vary over multiple scales. We do believe that the feature extraction methods proposed in this dissertation would provide promising outcome in the area of EEG signal classification. Extensive future work is to be under taken to establish the mass value blemethods in diverse domains of EEG analysis. However, to facilitate the further development of the proposed methods, we have highlighted a few key issues which provide a venues for future research.

The MT related EEG control signals likely contain a number of different types of sophisticated patterns that have not yet been identified. Current feature extraction algorithms may discard information and rely on prior assumptions about the patterns contained in the data. This may limit the ability of algorithms to identify new types of patterns. So future works should focus on developing feature extraction algorithms which can identify new and previously unknown patterns in the control signal.

EEG signals are known to have characteristics like non linearity, non-stationary and transient. The exact types of patterns that are relevant for classification are not well understood. So future works should focus on developing algorithms that do not rely on our prior knowledge about the types of patterns contained in the signal.

Besides, the use of algorithms that rely on our prior knowledge may also limit our in sights about the types of patterns contained in EEG signals. Besides, these signals vary considerably across subjects and even across sessions for the same subject. Robust feature extraction and machine learning based classification algorithms are required for identifying the relevant patterns in the control signal induced by a non-motor mental task. The existing feature extraction algorithms based on advanced signal processing techniques capture patterns which are known to exist.

As a result, these hand crafted features are not adaptive to capture the multi scale spatiotemporal patterns of the control signal. Improved machine learning algorithms like deep learning algorithms may achieve better classification performance which further leads to better BCI systems.

REFERENCES

[1] Berger, H, Über das elektrenkephalogramm des menschen. *Archivfürpsychiatrie und* nervenkrankheiten, *87*(1), 527-570, 1929.

[2] Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., & Muller, K. R. (2008). Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal processing magazine*, *25*(1), 41-56.

[3] Grosse-Wentrup, M., Liefhold, C., Gramann, K., & Buss, M. (2009). Beamforming in noninvasive brain–computer interfaces. *IEEE Transactions on Biomedical Engineering*, *56*(4), 1209-1219.

[4] Vidal, J. J. (1977). Real-time detection of brain events in EEG. *Proceedings of the IEEE*, *65*(5), 633-641.

[5] Wolpaw, J. R., Birbaumer, N., Heetderks, W. J., McFarland, D. J., Peckham, P. H., Schalk, G., & Vaughan, T. M. (2000). Brain-computer interface technology: a review of the first international meeting. *IEEE transactions on rehabilitation engineering*, *8*(2), 164-173.

[6] Wolpaw, J., &Wolpaw, E. W. (Eds.). (2012). *Brain-computer interfaces: principles and practice*. OUP USA.

[7] Nicolas-Alonso, L. F., & Gomez-Gil, J. (2012). Brain computer interfaces, a review. *Sensors*, *12*(2), 1211-1279.

[8] Dornhege, G. (Ed.). (2007). *Toward brain-computer interfacing*. MIT press.

[9] Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical neurophysiology*, *113*(6), 767-791.

[10] Murguialday, A. R., Hill, J., Bensch, M., Martens, S., Halder, S., Nijboer, F., & Gharabaghi, A. (2011). Transition from the locked in to the completely locked-in state: a physiological analysis. *Clinical Neurophysiology*, *122*(5), 925-933.

[11] Plum, F., & Posner, J. B. (1982). *The diagnosis of stupor and coma* (Vol. 19). Oxford University Press, USA.

[12] Bauby, J. D. (1998). The Diving Bell and the Butterfly: A Memoir of Life in Death, trans. Jeremy Leggatt.

[13] Sellers, E. W., Vaughan, T. M., &Wolpaw, J. R. (2010). A brain-computer interface for long-term independent home use. *Amyotrophic lateral sclerosis*, *11*(5), 449-455.

[14] Owen, A. M., & Coleman, M. R. (2008). Detecting awareness in the vegetative state. *Annals of the New York Academy of Sciences*, *1129*(1), 130-138.

[15] Birbaumer, N., Kubler, A., Ghanayim, N., Hinterberger, T., Perelmouter, J., Kaiser, J., &Flor, H. (2000). The thought translation device (TTD) for completely paralyzed patients. *IEEE Transactions on rehabilitation Engineering*, *8*(2), 190-193.

[16] Donchin, E., Spencer, K. M., &Wijesinghe, R. (2000). The mental prosthesis: assessing the speed of a P300-based brain-computer interface. *IEEE transactions on rehabilitation engineering*, *8*(2), 174-179.

[17] Kennedy, P. R., Bakay, R. A., Moore, M. M., Adams, K., &Goldwaithe, J. (2000). Direct control of a computer from the human central nervous system. *IEEE Transactions on rehabilitation engineering*, *8*(2), 198-202.

[18] Kostov, A., &Polak, M. (2000). Parallel man-machine training in development of EEG-based cursor control. *IEEE Transactions on Rehabilitation Engineering*, *8*(2), 203-205.

[19] Pfurtscheller, G., Neuper, C., Guger, C., Harkam, W. A. H. W., Ramoser, H., Schlogl, A., & Pregenzer, M. A. P. M. (2000). Current trends in Graz brain-computer interface (BCI) research. *IEEE Transactions on Rehabilitation Engineering*, *8*(2), 216-219.

[20] Wolpaw, J. R., McFarland, D. J., & Vaughan, T. M. (2000). Brain-computer interface research at the Wadsworth Center. *IEEE Transactions on Rehabilitation Engineering*, *8*(2), 222-226.

[21] Penny, W. D., Roberts, S. J., Curran, E. A., & Stokes, M. J. (2000). EEG-based communication: a pattern recognition approach. *IEEE transactions on Rehabilitation Engineering*, *8*(2), 214-215.

[22] Keirn, Z. A., &Aunon, J. I. (1990). A new mode of communication between man and his surroundings. *IEEE transactions on biomedical engineering*, *37*(12), 1209-1214.

[23] J. C., Ornstein, R., &Galin, D. (1974). Lateral specialization of cognitive mode: II. EEG frequency analysis. *Psychophysiology*, *11*(5), 567-578.

[24] Ehrlichman, H., & Wiener, M. S. (1980). EEGasymmetry during covert mental activity. *Psychophysiology*, *17*(3), 228-235.

[25] Gevins, A. S., Zeitlin, G. M., Doyle, J. C., Yingling, C. D., Schaffer, R. E., Callaway, E., & Yeager, C. L. (1979). Electroencephalogram correlates of higher cortical functions. *Science*, *203*(4381), 665-668.

[26] Osaka, M. (1984). Peak alpha frequency of EEG during a mental task: Task difficulty and hemispheric differences. *Psychophysiology*, *21*(1), 101-105.

[27] Rappelsberger, P., &Petsche, H. (1988). Probability mapping: power and coherence analyses of cognitive processes. *Brain topography*, *1*(1), 46-54.

[28] Pfurtscheller, G. (1992). Event-related synchronization (ERS): an electrophysiological correlate of cortical areas at rest. *Clinical Neurophysiology*, *83*(1), 62-69.

[29] Pfurtscheller, G., &Neuper, C. (1997). Motor imagery activates primary sensorimotor area in humans. *Neuroscience letters*, *239*(2-3), 65-68.

[30] Pfurtscheller, G., Neuper, C., Flotzinger, D., &Pregenzer, M. (1997). EEG-based discrimination between imagination of right and left hand movement. *Electroencephalography and clinical Neurophysiology*, *103*(6), 642-651.

[31] Neuper, C., &Pfurtscheller, G. (2001). Event-related dynamics of cortical rhythms: frequency-specific features and functional correlates. *International journal of psychophysiology*, *43*(1), 41-58.

[32] Friedrich, E. V., Scherer, R., &Neuper, C. (2013). Stability of event-related (de-) synchronization during brain–computer interface-relevant mental tasks. *Clinical Neurophysiology*, *124*(1), 61-69.

[33] Kayikcioglu, T., &Aydemir, O. (2010). A polynomial fitting and k-NN based approach for improving classification of motor imagery BCI data. *Pattern Recognition Letters*, *31*(11), 1207-1215.

[34] Thomas, K. P., Guan, C., Lau, C. T., Vinod, A. P., &Ang, K. K. (2009). A new discriminative common spatial pattern method for motor imagery brain–computer interfaces. *IEEE Transactions on Biomedical Engineering*, *56*(11), 2730-2733.

[35] Friedrich, E., Scherer, R., Faller, J., &Neuper, C. (2011). Do user-related factors of motor impaired and able-bodied participants correlate with classification accuracy? In *Proc. of the 5th International Brain-Computer Interface Conference 2011*.

[36] Lulé, D., Noirhomme, Q., Kleih, S. C., Chatelle, C., Halder, S., Demertzi, A., & Thonnard, M. (2013). Probing command following in patients with disorders of consciousness using a brain–computer interface. *Clinical Neurophysiology*, *124*(1), 101[35]Farwell, L. A., &del

[37] R Millan, J., Mouriño, J., Franzé, M., Cincotti, F., Varsta, M., Heikkonen, J., &Babiloni, F. (2002). A local neural classifier for the recognition of EEG patterns associated to mental tasks. *IEEE transactions on neural networks*, *13*(3), 678-686.

[38] Millán, J. D. R., Renkens, F., Mouriño, J., & Gerstner, W. (2004). Brain-actuated interaction. *Artificial Intelligence*, *159*(1-2), 241-259.

[39] Millán, J. R., Renkens, F., Mourino, J., & Gerstner, W. (2004). Noninvasive brain-actuated control of a mobile robot by human EEG. *IEEE Transactions on biomedical Engineering*, *51*(6), 1026-1033.

[40] Galán, F., Nuttin, M., Lew, E., Ferrez, P. W., Vanacker, G., Philips, J., &Millán, J. D. R. (2008). A brain-actuated wheelchair: asynchronous and non-invasive brain–computer interfaces for continuous control of robots. *Clinical neurophysiology*, *119*(9), 2159-2169.

[41] Millan, J. R. (2004, July). On the need for on-line learning in brain-computer interfaces. In *Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on* (Vol. 4, pp. 2877-2882). IEEE.

[42] Millán, J. D. R., Ferrez, P. W., &Buttfield, A. (2007). The IDIAP brain–computer interface: an asynchronous multiclass approach. *Toward Brain–Computer Interfacing, The MIT Press, Cambridge, MA*, 103-110.

[43] Carlson, N.R.(2002a)Foundations ofphysiologicalpsychology, 5thed, Boston, MassLondon: Allyn and Bacon.

[44] Siuly, S., Li, Y., & Zhang, Y. (2016). Electroencephalogram (EEG) and Its Background. In *EEG Signal Analysis and Classification* (pp. 3-21). Springer, Cham.

[45] Gray, F. J. (2002). *Anatomy for the medical clinician*. FJ Gray.

[46] Atwood, H. L., & MacKay, W. A. (1989). *Essentials of neurophysiology*. BC Decker.

[47] Sanei, S., & Chambers, J. A. (2013). *EEG signal processing*. John Wiley & Sons.

[48] Niedermeyer, E., & da Silva, F. L. (Eds.). (2005). *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins.

[49] Nunez, P. L., &Srinivasan, R. (2006). *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, USA.

[50] Collura, T. F. (1993). History and evolution of electroencephalographic instruments and techniques. *Journal of clinical neurophysiology*, *10*(4), 476-504.

[51] Hazarika, N., Chen, J. Z., Tsoi, A. C., &Sergejew, A. (1997). Classification of EEG signals using the wavelet transform. *Signal processing*, *59*(1), 61-72.

[52] Adeli, H., Zhou, Z., &Dadmehr, N. (2003). Analysis of EEG records in an epileptic patient using wavelet transform. *Journal of neuroscience methods*, *123*(1), 69-87.

[53] Singh, M. (2014). *Introduction to biomedical instrumentation*. PHI Learning Pvt. Ltd.

[54] Jasper, H. (1958a). The ten-twenty electrode system of the Intenational Federation. *Electroencephclin Neutrophysiol*, *10*, 367-380.

[55] Jasper, H. (1958b). Report of the committee on methods of clinical examination in electroencephalography. *Electroencephalogr Clin Neurophysiol*, *10*, 370-375.

[56] Sharbrough, F. (1991). American Electroencephalographic Society guidelines for standard electrode position nomenclature. *J clin Neurophysiol*, *8*, 200-202.

[57] Fisch, B. J., &Spehlmann, R. (1999). *Fisch and Spehlmann's EEG primer: basic principles of digital and analog EEG*. Elsevier Health Sciences.

[58] Libenson, M. H. (2012). *Practical Approach to Electroencephalography E-Book*. Elsevier Health Sciences.

[59] Vaughan, T. M., Heetderks, W. J., Trejo, L. J., Rymer, W. Z., Weinrich, M., Moore, M. M., & Wolpaw, E. W. (2003). Brain-computer interface technology: a review of the Second International Meeting.

[60] Bashashati, A., Fatourechi, M., Ward, R. K., & Birch, G. E. (2007). A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals. *Journal of Neural engineering*, *4*(2), R32.

[61] Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., &Arnaldi, B. (2007). A review of classification algorithms for EEG-based brain–computer interfaces. *Journal of neural engineering*, *4*(2), R1

[62] Wu, W., Gao, X., Hong, B., &Gao, S. (2008). Classifying single-trial EEG during motor imagery by iterative spatio-spectral patterns learning (ISSPL). *IEEE Transactions on Biomedical Engineering*, *55*(6), 1733-1743.

[63] Li, Y., Long, J., Yu, T., Yu, Z., Wang, C., Zhang, H., & Guan, C. (2010). An EEG-based BCI system for 2-D cursor control by combining Mu/Beta rhythm and P300 potential. *IEEE Transactions on Biomedical Engineering*, *57*(10), 2495-2505

[64] Jain, A. K., Duin, R. P. W., & Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on pattern analysis and machine intelligence*, *22*(1), 4-37.

[65] Duda, R. O., Hart, P. E., & Stork, D. G. (2001). Pattern classification, edition wileyinterscience. *New York*.

[66] Brunelli, R. (2009). *Template matching techniques in computer vision: theory and practice*. John Wiley & Sons.

[67] McFarland, D. J., Anderson, C. W., Muller, K. R., Schlogl, A., &Krusienski, D. J. (2006). BCI meeting 2005-workshop on BCI signal processing: feature extraction and translation. *IEEE transactions on neural systems and rehabilitation engineering*, *14*(2), 135-138.

[68] J. H. K. Friedman. On bias, variance, 0/1-loss, and the curse-of-dimensionality. Data Mining andKnowledge Discovery, 1(1), 1997.

[69] Raudys, S. J., & Jain, A. K. (1991). Small sample size effects in statistical pattern recognition: Recommendations for practitioners. *IEEE Transactions on pattern analysis and machine intelligence*, *13*(3), 252-264.

[70] Jain, A. K., & Chandrasekaran, B. (1982). 39 Dimensionality and sample size considerations in pattern recognition practice. *Handbook of statistics*, *2*, 835-855.

[71] Breiman, L. (1998). Arcing classifier (with discussion and a rejoinder by the author). *The annals of statistics*, *26*(3), 801-849.

[72] Ng, A. Y., & Jordan, M. I. (2002). On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. In *Advances in neural information processing systems* (pp. 841-848).

[73] Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257-286.

[74] Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE transactions on neural networks*, *10*(5), 988-999.

[75] Keirn, Z. A. (1988). Alternative modes of communication between man and machine. *Master's thesis, Purdue University*, 35-56.

[76] Keirn, Z. A., &Aunon, J. I. (1990). Man-machine communications through brain-wave processing. *IEEE Engineering in Medicine and biology magazine*, *9*(1), 55-57.

[77] Fernández, T., Harmony, T., Rodríguez, M., Bernal, J., Silva, J., Reyes, A., &Marosi, E. (1995). EEG activation patterns during the performance of tasks involving different components of mental calculation. *Electroencephalography and clinical Neurophysiology*, *94*(3), 175-182.

[78] Anderson, C. W., Stolz, E. A., &Shamsunder, S. (1995, September). Discriminating mental tasks using EEG represented by AR models. In *Engineering in Medicine and Biology Society, 1995., IEEE 17th Annual Conference* (Vol. 2, pp. 875-876). IEEE.

[79] Anderson, C. W., Devulapalli, S. V., &Stolz, E. A. (1995). EEG signal classification with different signal representations. In *Neural Networks for Signal Processing [1995] V. Proceedings of the 1995 IEEE Workshop* (pp. 475-483). IEEE.

[80] Anderson, C. W., &Sijercic, Z. (1996, June). Classification of EEG signals from four subjects during five mental tasks. In *Solving engineering problems with neural networks: proceedings of the conference on engineering applications in neural networks (EANN'96)* (pp. 407-414). Turkey.

[81] Anderson, C. W., Stolz, E. A., &Shamsunder, S. (1998). Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks. *IEEE Transactions on Biomedical Engineering*, *45*(3), 277-286.

[82] Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical Neurophysiology*, *70*(6), 510-523.

[83] Donchin, E., Spencer, K. M., &Wijesinghe, R. (2000). The mental prosthesis: assessing the speed of a P300-based brain-computer interface. *IEEE transactions on rehabilitation engineering*, *8*(2), 174-179.

[84] Palaniappan, R., Paramesran, R., Nishida, S., &Saiwaki, N. (2002). A new brain-computer interface design using fuzzy ARTMAP. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *10*(3), 140-148.

[85] Palaniappan, R. (2005, March). Brain computer interface design using band powers extracted during mental tasks. In *Neural Engineering, 2005. Conference Proceedings. 2nd International IEEE EMBS Conference on* (pp. 321-324). IEEE.

[86] Garrett, D., Peterson, D. A., Anderson, C. W., &Thaut, M. H. (2003). Comparison of linear, nonlinear, and feature selection methods for EEG signal classification. *IEEE Transactions on neural systems and rehabilitation engineering*, *11*(2), 141-144.

[87] Huan, N. J., &Palaniappan, R. (2004). Neural network classification of autoregressive features from electroencephalogram signals for brain–computer interface design. *Journal of neural engineering*, *1*(3), 142.

[88] Huan, N. J., &Palaniappan, R. (2005, March). Classification of mental tasks using fixed and adaptive autoregressive models of EEG signals. In *Neural Engineering, 2005. Conference Proceedings. 2nd International IEEE EMBS Conference on* (pp. 633-636). IEEE.

[89] Liang, N. Y., Saratchandran, P., Huang, G. B., &Sundararajan, N. (2006). Classification of mental tasks from EEG signals using extreme learning machine. *International journal of neural systems*, *16*(01), 29-38.

[90] Anderson, C. W., Knight, J. N., O'Connor, T., Kirby, M. J., &Sokolov, A. (2006). Geometric subspace methods and time-delay embedding for EEG artifact removal and classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *14*(2), 142-146.

[91] Palaniappan, R. (2006). Utilizing gamma band to improve mental task based brain-computer interface design. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *14*(3), 299-303.

[92] Anderson, C. W., Knight, J. N., Kirby, M. J., & Hundley, D. R. (2007). 15 Classification of Time-Embedded EEG Using Short-Time Principal Component Analysis. *Toward Brain-Computer Interfacing*, 261.

[93] Zhiwei, L., &Minfen, S. (2007, August). Classification of mental task EEG signals using wavelet packet entropy and SVM. In *Electronic Measurement and Instruments, 2007. ICEMI'07. 8th International Conference on* (pp. 3-906). IEEE.

[94] Diez, P. F., Mut, V., Laciar, E., Torres, A., & Avila, E. (2009, September). Application of the empirical mode decomposition to the extraction of features from EEG signals for mental task classification. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE* (pp. 2579-2582). IEEE.

[95] Kaleem, M. F., Sugavaneswaran, L., Guergachi, A., & Krishnan, S. (2010, August). Application of empirical mode decomposition and Teager energy operator to EEG signals for mental task classification. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE* (pp. 4590-4593). IEEE.

[96] Chai, Rifai, et al. "Mental non-motor imagery tasks classifications of brain computer interface for wheelchair commands using genetic algorithm-based neural network." *Proceedings of the International Joint Conference on Neural Networks, (IJCNN), Brisbane, Queensland, Australia, 10-15 June 2012*. 2012.

[97] Gupta, A., &Agrawal, R. K. (2012, May). Relevant feature selection from EEG signals for mental task classification. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (pp. 431-442). Springer, Berlin, Heidelberg.

[98] Vidaurre, D., Bielza, C., &Larranaga, P. (2013). Classification of neural signals from sparse autoregressive features. *Neurocomputing*, *111*, 21-26.

[99] Tolić, M., &Jović, F. (2013). Classification of wavelet transformed EEG signals with neural network for imagined mental and motor tasks. *Kinesiology: International journal of fundamental and applied kinesiology*, *45*(1), 130-138.

[100] Hariharan, M., Vijean, V., Sindhu, R., Divakar, P., Saidatul, A., &Yaacob, S. (2014). Classification of mental tasks using stockwell transforms. *Computers & Electrical Engineering*, *40*(5), 1741-1749.

[101] Li, X., Chen, X., Yan, Y., Wei, W., & Wang, Z. J. (2014). Classification of EEG signals using a multiple kernel learning support vector machine. *Sensors*, *14*(7), 12784-12802.

[102] Gupta, A., Agrawal, R. K., &Kaur, B. (2015). Performance enhancement of mental task classification using EEG signal: a study of multivariate feature selection methods. *Soft Computing*, *19*(10), 2799-2812.

[103] Lwin, Z. M. (2015). Analysis of Matching Pursuit Features of EEG Signal for Mental Tasks Classification. *World Academy of Science, Engineering and Technology, International Journal of Computer and Information Engineering*, *2*(1).

[104] Hendel, M., Benyettou, A., &Hendel, F. (2016). Hybrid self organizing map and probabilistic quadratic loss multi-class support vector machine for mental tasks classification. *Informatics in Medicine Unlocked*, *4*, 1-9.

[105] Gupta, A., & Kumar, D. (2017). Fuzzy clustering-based feature extraction method for mental task classification. *Brain informatics*, *4*(2), 135-145.

[106] Dutta, S., Singh, M., & Kumar, A. (2018a). Classification of non-motor cognitive task in EEG based brain-computer interface using phase space features in multivariate empirical mode decomposition domain. *Biomedical Signal Processing and Control*, *39*, 378-389.

[107] Dutta, S., Singh, M., & Kumar, A. (2018b). Automated classification of non-motor mental task in electroencephalogram based brain-computer interface using multivariate autoregressive model in the intrinsic mode function domain. *Biomedical Signal Processing and Control*, *43*, 174-182.

[108] Rahman, Mohammad Mahinur, M. A. Chowdhury, and ShaikhAnowarul Fattah. "An efficient scheme for mental task classification utilizing reflection coefficients obtained from autocorrelation function of EEG signal." *Brain informatics* 5.1 (2018): 1.

[109] Nuamah, J. K., and YounhoSeong. "Support vector machine (SVM) classification of cognitive tasks based on electroencephalography (EEG) engagement index." *Brain-Computer Interfaces* 5.1 (2018): 1-12.

[110] Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., ... & Liu, H. H. (1998, March). 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* (Vol. 454, No. 1971, pp. 903-995). The Royal Society.

[111] Rehman, N., Mandic, D.P.Multivariate empirical mode decomposition, Proc.R.Soc. A-Math, Phys. Eng.Sci. 466(2117) (2010)1291

[112] Lee, S. H., Lim, J. S., Kim, J. K., Yang, J., & Lee, Y. (2014). Classification of normal and epileptic seizure EEG signals using wavelet transform, phase-space reconstruction, and Euclidean distance. *Computer methods and programs in biomedicine, 116*(1), 10-25.

[113] Faure, H., &Pillichshammer, F. (2009). L p discrepancy of generalized two-dimensional Hammersley point sets. *MonatsheftefürMathematik*, *158*(1), 31-61.

[114] Cui, J., &Freeden, W. (1997). Equidistribution on the sphere. *SIAM Journal on Scientific Computing*, *18*(2), 595-609.

[115] Deléchelle, E., Lemoine, J., &Niang, O. (2005). Empirical mode decomposition: an analytical approach for sifting process. *IEEE Signal Processing Letters*, *12*(11), 764-767.

[116] Junsheng, C., Dejie, Y., & Yu, Y. (2006). Research on the intrinsic mode function (IMF) criterion in EMD method. *Mechanical systems and signal processing*, *20*(4), 817-824.

[117] Kay, S. M. (1988). *Modern spectral estimation*. Pearson Education India.

[118] Priestley, M. B. Spectral Analysis and Time Series (Academic, London, 1981). *Google Scholar*, 397.

[119] Swami, A., Giannakis, G., &Shamsunder, S. (1994). Multichannel ARMA processes. *IEEE Transactions on Signal Processing*, *42*(4), 898-913.

[120] Hannan, E. J. (1976). The identification and parameterization of ARMAX and state space forms. *Econometrica: Journal of the Econometric Society*, 713-723.

[121] Strand, O. (1977). Multichannel complex maximum entropy (autoregressive) spectral analysis. *IEEE Transactions on Automatic Control*, *22*(4), 634-640.

[122] Pritchard, W. S., & Duke, D. W. (1992). Measuring chaos in the brain: a tutorial review of nonlinear dynamical EEG analysis. *International Journal of Neuroscience*, *67*(1-4), 31-80.

[123] Pritchard, W. S., & Duke, D. W. (1995). Measuring chaos in the brain-a tutorial review of EEG dimension estimation. *Brain and cognition*, *27*(3), 353-397.

[124] Faure, P., &Korn, H. (2001). Is there chaos in the brain? I. Concepts of nonlinear dynamics and methods of investigation. *ComptesRendus de l'Académie des Sciences-Series III-Sciences de la Vie*, *324*(9), 773-793.

[125] Korn, H., & Faure, P. (2003). Is there chaos in the brain? II. Experimental evidence and related models. *Comptesrendusbiologies*, *326*(9), 787-840.

[126] Babloyantz, A., Salazar, J. M., &Nicolis, C. (1985). Evidence of chaotic dynamics of brain activity during the sleep cycle. *Physics letters A*, *111*(3), 152-156.

[127] Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences*, *88*(6), 2297-2301.

[128] Grassberger, P., &Procaccia, I. (1983). Estimation of the Kolmogorov entropy from a chaotic signal. *Physical review A*, *28*(4), 2591.

[129] Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology*, *278*(6), H2039-H2049.

[130] Bandt, C., &Pompe, B. (2002). Permutation entropy: a natural complexity measure for time series. *Physical review letters*, *88*(17), 174102.

[131] Chen, W., Wang, Z., Xie, H., & Yu, W. (2007). Characterization of surface EMG signal based on fuzzy entropy. *IEEE Transactions on neural systems and rehabilitation engineering*, *15*(2), 266-272.

[132] Ahmed, M. U., &Mandic, D. P. (2011). Multivariate multiscale entropy: A tool for complexity analysis of multichannel data. *Physical Review E*, *84*(6), 061918.

[133] Costa, M., Goldberger, A. L., &Peng, C. K. (2002). Multiscale entropy analysis of complex physiologic time series. *Physical review letters*, *89*(6), 068102.

[134] Aziz, W., &Arif, M. (2005, December). Multiscale permutation entropy of physiological time series. In *9th International Multitopic Conference, IEEE INMIC 2005* (pp. 1-6). IEEE.

[135] He, S., Sun, K., & Wang, H. (2016). Multivariate permutation entropy and its application for complexity analysis of chaotic systems. *Physica A: Statistical Mechanics and its Applications*, *461*, 812-823.

[136] Morabito, F. C., Labate, D., La Foresta, F., Bramanti, A., Morabito, G., &Palamara, I. (2012). Multivariate multi-scale permutation entropy for complexity analysis of Alzheimer's disease EEG. *Entropy*, *14*(7), 1186-1202.

[137] Zheng, Jinde, Junsheng Cheng, Yu Yang, and SongrongLuo. A rolling bearing fault diagnosis method based on multi-scale fuzzy entropy and variable predictive model-based class discrimination. *Mechanism and machine theory*78 , 187-200, 2014.

[138] Ahmed, M. U., Chanwimalueang, T., Thayyil, S., &Mandic, D. P. (2016). A multivariate multiscale fuzzy entropy algorithm with application to uterine EMG complexity analysis. *Entropy*, *19*(1), 2.

[139] Packard, N. H., Crutchfield, J. P., Farmer, J. D., & Shaw, R. S. (1980). Geometry from a time series. *Physical review letters*, *45*(9), 712.

[140] Takens, F. (1981). Detecting strange attractors in turbulence. In *Dynamical systems and turbulence, Warwick 1980* (pp. 366-381). Springer, Berlin, Heidelberg.

[141] Kliková, B., &Raidl, A. (2011). Reconstruction of phase space of dynamical systems using method of time delay. In *Wds* (Vol. 11, pp. 134-137).

[142] Wang, X., Meng, J., Tan, G., & Zou, L. (2010). Research on the relation of EEG signal chaos characteristics with high-level intelligence activity of human brain.