M. Tech. (SPDD) Divi Sai Manoj 2013-2015

COGNITIVE ASSESSMENT THROUGH THE ANALYSIS OF EEG SIGNALS

A DISSERTATION
SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE
OF

MASTER OF TECHNOLOGY
IN
SIGNAL PROCESSING AND DIGITAL DESIGN

Submitted by

Divi Sai Manoj

2K13/SPD/04

Under the supervision of

Mr. Rajesh Birok
(Associate Professor)

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

2013-2015

COGNITIVE ASSESSMENT THROUGH THE ANALYSIS OF EEG SIGNALS

A DISSERTATION
SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT
FOR THE AWARD OF THE DEGREE
OF

MASTER OF TECHNOLOGY
IN
SIGNAL PROCESSING AND DIGITAL DESIGN

Submitted by

Divi Sai Manoj

2K13/SPD/04

Under the supervision of

Mr. Rajesh Birok
(Associate Professor)

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

2015

Certificate

This is to certify that the dissertation title ***"Cognitive Assessment through the Analysis of EEG Signals"*** submitted by **Mr Divi Sai Manoj**, Roll. No. *2K13/SPD/04*, in partial fulfilment for the award of degree of Master of Technology in Signal Processing & Digital Design at **Delhi Technological University, Delhi**, is a bonafide record of student's own work carried out by him under my supervision and guidance in the academic session 2014-15. To the best of my belief and knowledge the matter embodied in dissertation has not been submitted for the award of any other degree or certificate in this or any other university or institute.

Rajesh Birok
Supervisor
Associate Professor
Dept. of ECE
Delhi Technological University

Acknowledgement

I am indebted to my thesis supervisor **Mr Rajesh Birok, Associate Professor** Department of Electronics and Communication, for his gracious encouragement and very valued constructive criticism that has driven me to carry out the project successfully.

I am greatly thankful to **Prof. Prem R. Chadda,** Head of Department (Electronics & Communication Engineering), entire faculty and staff of Electronics & Communication Engineering and friends for their continuous support, encouragement and inspiration in the execution of this "**thesis"** work.

I sincerely thank **Dr. Sushil Chandra,** Scientist 'F', HOD, Biomedical Engineering Department, INMAS, DRDO for giving me an excellent opportunity to work as an offline intern. I also thank **Mr. Akash Rao** and other staff of DRDO for their support and guidance through out this work.

Finally I express my deep sense of gratitude to my parents who bestowed upon me their grace and were source of my inspiration and encouragement.

Divi Sai Manoj

M.Tech (SPDD)

2K13/SPD/04

Table of contents

List of Tables

List of Figures

Abbrevations

EEG	ElectroEncephaloGraphy
DWT	Discrete Wavelet Transform
PCA	Principal Component Analysis
ICA	Independent Component Analysis
ERP	Event Related Potential
EMD	Emperical Mode Decomposition
DCT	Discrete Cosine Transform
IMF	Intrinsic Mode Function
VG	Visibility Graph
HVG	Horizontal Visibility Graph
DVG	Difference Visibility Graph
DD	Degree Distribution
MVDR	Minimum Variance Distortionless Response
LDA	Linear Discriminant Analysis
KNN	K Nearest Neighbor
RBFNN	Radial Basis Function Neural Network
ADHD	Attention Deficit Hyperactivity Disorder
MER-MER	Memory and Encoding Related Multifaceted Electroencephalographic Response

Abstract

With the advent of emerging technologies and methodologies of biomedical signal processing, research on the cognitive sciences has become one of the most innovative, enthralling and challenging researches in the field of biomedical engineering.

The problem of cognitive assessment and enhancement has gained major importance amongst the today's cognitive researches, as it aids in identification and treatment of cognitive related disorders like Attention Deficit Hyperactivity Disorder (ADHD), Spatial Navigation, Short Term Memory, Cross Modal Processing etc.

In this work, as a part of Cognitive Assessment, we are concentrated towards the testing of Working Memory and Cognitive Workload through the analysis of the two channel (C_z-P_z) EEG signals obtained from the subjects when they were presented with some familiar stimulus. For this, the subjects were explained about the stimuli related to a situation and later were shown these stimuli and a model had been developed to differentiate these different types of stimuli responses. Several features in conjunction with classifiers have been explored and the corresponding results were analysed to decide the optimum feature and classifier that can be selected for obtaining the optimum classification accuracy.

A maximum classification accuracy of 66.67% had been obtained when tried with the LDA+RBFFNN classifier for P_z-channel when the feature of Hurst Exponent on 3-5 IMFs is used for the task of inter Stimuli Classification and a maximum classification accuracy of 100% had been obtained when tried with the LDA+KNN classifier for C_z and P_z channels when the feature of Hurst Exponent is used for the task of inter Subject Classification and the results show a very good overall Inter subject classification accuracies for the classifier LDA+KNN.

CHAPTER 1

INTRODUCTION

1. INTRODUCTION

1.1 Background:

Cognitive Science is the scientific study of mind and its processes. It relates to the analysis of how information is being represented, processed and transformed inside the brain, during different mental and physical tasks performed by a person. Cognitive science can be viewed as an interdisciplinary branch constituting various disciplines of psychology, artificial intelligence, philosophy, neuroscience, linguistics and anthropology [1].

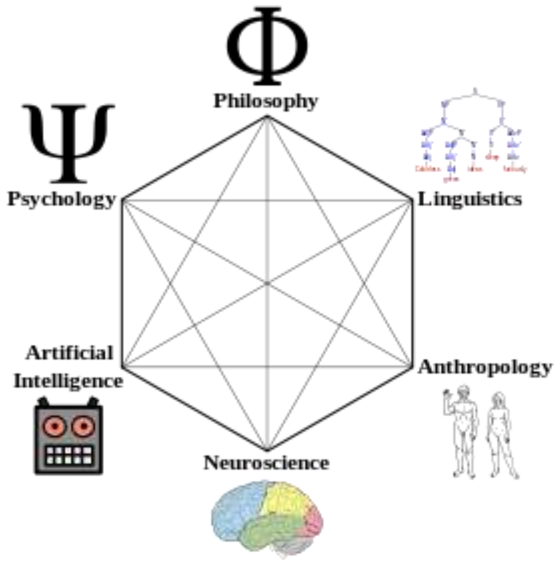


Fig.1.1 Hexagon showing the constituents of cognitive science

The problem of cognitive assessment and enhancement has gained major importance, as it aids in identification and treatment of cognitive related disorders like Attention Deficit Hyperactivity Disorder (ADHD), Spatial Navigation, Short Term Memory, Cross Modal Processing etc. by the analysis of changes that occur in different lobes of the brain. Cognitive Science, is useful not only in the assessment and enhancement of brain capabilities but also, now-a-days it is being used in the field of marketing field, which is known as 'Neuromarketing' [2], [3], [4].

1.2 Human Brain, Lobes and its associated tasks

Human brain is basically divided into four regions called “lobes” namely Frontal, Parietal, Occipital and Temporal and each of these lobes is associated to different skills.

1.2.1 Frontal lobe: It is the region present at the rear of forehead. It is associated with the skills like Organization, Concept Formation, Mental Flexibility, Personality, Execution of behaviour, Abstract Reasoning, Problem Solving, Planning, Judgment, Ethical Behaviour, Inhibition, Expressive Language, Affect, and Attention.

1.2.2 Parietal lobe: It is the lobe that is present next to Frontal lobe and above Temporal lobe. This lobe is responsible for the behaviours like Reading, Calculation, Attention, Short Term Memory, Cross Modal Processing (e.g. listening, writing, and reading), spatial navigation, Visual Perception and Discrimination.

1.2.3 Occipital lobe: Occipital lobe is located at the rear of the brain. It corresponds to activities related to vision like Visual Processing, Visually Perceive, Visual Discrimination, Visual Spatial Skill and Facial Discrimination.

1.2.4 Temporal lobe: This lobe is associated with the critical activities like Memory and new learning, Language Comprehension, Auditory Processing, Spatial Processing, Attention, Spirituality and Emotion.

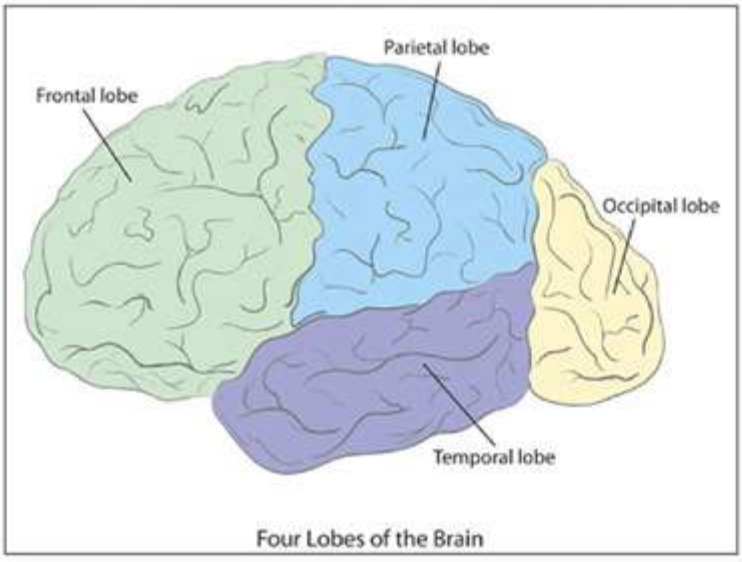


Fig.1.2. Four Lobes of the Brain

1.3 Research Methodologies in Cognitive Sciences

The different research methodologies that we have for the analysis of cognitive sciences can be listed as follows

1.3.1 Behavioural experiments: These experiments involve the analysis of the responses to the stimuli

- P300 of an EEG signal (P300 means the positive component that 250-500ms on the onset of a stimulus)
- Eye tracking in the assessment of person's focus of attention

1.3.2 Brain imaging: The various brain imaging techniques that we have in today's world are

- Single Photon Emission Computed Tomography and Positron Emission Tomography (SPECT and PET)
- Electro-EncephaloGraphy (EEG)
- Functional Magnetic Resonance Imaging (fMRI)
- Optical Imaging
- Magnetoencephalography (MEG)

1.3.3 Computational modelling: These methods try to represent the problem statement in terms of a mathematical model that effectively describes the functional aspects of a cognitive process. Use of Neural networks can be stated as an example for this kind of methodology [5].

1.3.4 Neurobiological methods: These methods include

- Single-Unit recording
- Direct brain stimulation
- Animal models
- Postmortem studies

Among these methods available, Electroencephalography (EEG) signal had been chosen for the task of cognitive assessment. EEG is a non-invasive method to record electrical activity of the brain along the scalp. EEG measures voltage fluctuations resulting from ionic current within the neurons of the brain. In clinical contexts, EEG refers to the recording of the brain's spontaneous electrical activity over a period of time, [6] as recorded from multiple electrodes placed on the scalp. It has several advantages like low cost and portable equipment, spatial localization of information, better temporal resolution when compared to other physiological signals.

1.4 Problem Statement

In this work, we deal with the testing of Working Memory and Cognitive Workload, the lobe that is useful for carrying out the analysis is Parietal lobe, and hence the data from channels C_z, P_z of the 10-20 system EEG setup (to be discussed in detail in Experimental setup section) tends to be effective in this work. Regarding the Cognitive Workload, Features that are useful in exploring the Cognitive workload were discussed while as a part of Working Memory a task of crime investigation had been presented through inter subject and intra subject classifications. This work mainly concentrates on the part, testing of Working Memory.

The rest of the work is divided and organised as per the steps involved in the task, testing of Working Memory, which can be listed as Pre-processing, Feature Extraction, Classification, Results and Discussion.

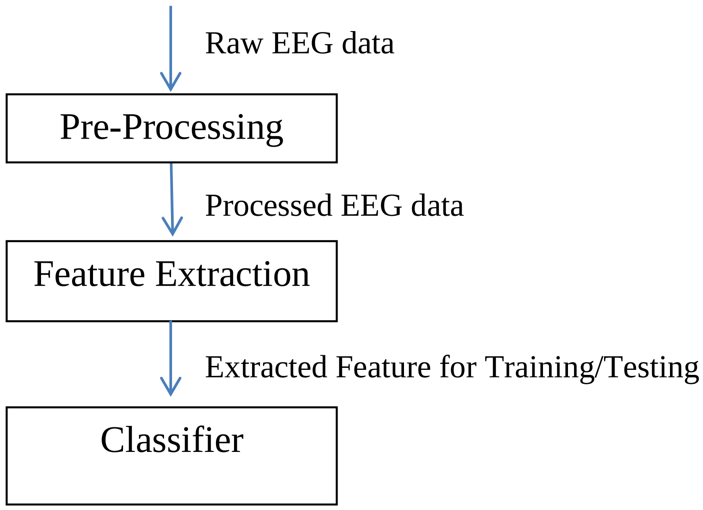


Fig.1.3. Steps involved in the task of testing of Working Memory

CHAPTER 2

PRE-PROCESSING

2. PRE-PROCESSING

In this chapter we discuss various techniques available for denoising the EEG data. Also the techniques that were implemented in this work were also presented.

2.1 Types of EEG Noises

The data obtained from the EEG setup might always be noisy and has to be pre-processed, to avoid any improper analysis of information. The various kinds of noises that we generally find in the EEG data are

2.1.1 Powerline Noise: This noise generally occurs due to the interference of the 50/60Hz line frequency or its higher harmonics from the power sources that are present near the equipment. This noise can easily be eliminated by passing the EEG data through a notch or comb filter designed for a cut-off value as the line frequency.

2.1.2 Baseline Wandering Noise: Baseline interference generally occurs due to the poor contact of the electrodes, respiration of the patient and sometimes due to the temperature variations and change in the bias value of the instrumentation amplifiers in the setup.

2.13 Eye and Muscle Artifacts:

- Eye artifact generally occurs due to the eye blinks during the EEG recording. The EEG data affected by the Eye artifact experiences smoothly decreasing EEG spectrum and it influences channels located at the Far-frontal region of the brain.
- Muscle artifacts (EMG noise) generally occur due to the movement of the muscles during the process of EEG recording. These artifacts are spatially localised and show high power at high frequencies (20-50Hz).

2.2 Denoising Techniques

Some of the most widely used techniques for the EEG denoising are filtering technique, wavelet based denoising, PCA based denoising technique and ICA base denoising technique [7], [8].

2.2.1 Filtering: The raw EEG signal taken from any person is always noisy due to various reasons and needs to be filtered to achieve best results. Filtering is achieved by passing the raw EEG data through the series of a bandpass and notch filters (0.1Hz-63.5Hz bandpass+ 60 Hz notch) available in EEG LAB to remove the noise and power line interference.

2.2.2 Wavelet based denoising Technique: In this method, the time-frequency representation of DWT is performed by repeated filtering of the input signal with a pair of filters named as, Low Pass Filter (LPF) and High Pass Filter (HPF), which decompose the signal into different scales. The coefficient of low pass filter is called as Approximation Coefficients (CA) and similarly, high pass filtered coefficients are called as Detailed Coefficients (CD). The CA is consequently divided into new approximation and detailed coefficients. Using these wavelet filters, the raw EEG data is decomposed into various bands namely delta, theta, alpha, beta and gamma [9], [10], [11]. In general the EEG data corresponding to the gamma band is treated as noise. So the wavelet coefficients that are related to these bands are removed by any one of the wavelet shrinkage and threshold techniques. The frequency ranges of these bands, the wavelet decomposition process and the various wavelet denoising techniques available are presented below.

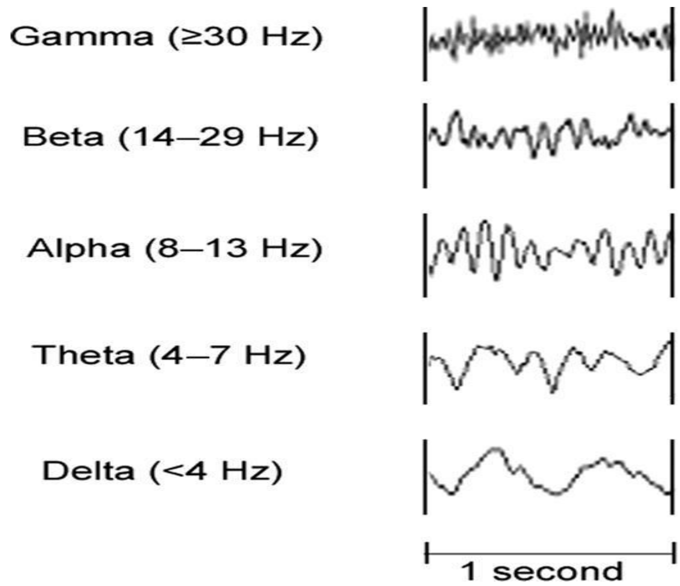


Fig.2.1 Bands of EEG data

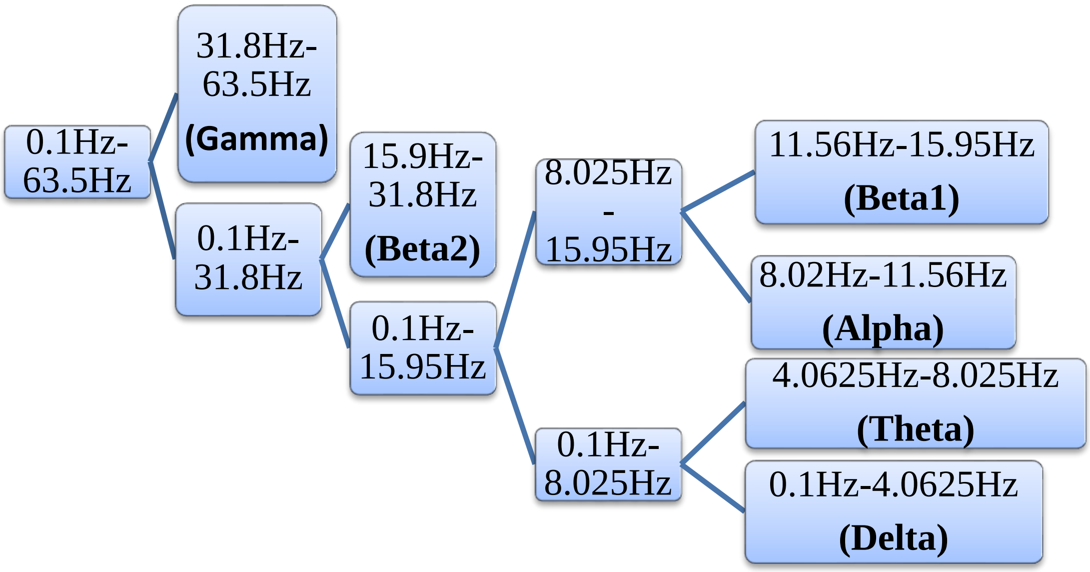


Fig.2.2 Decomposition of EEG data into sub bands using tree structure

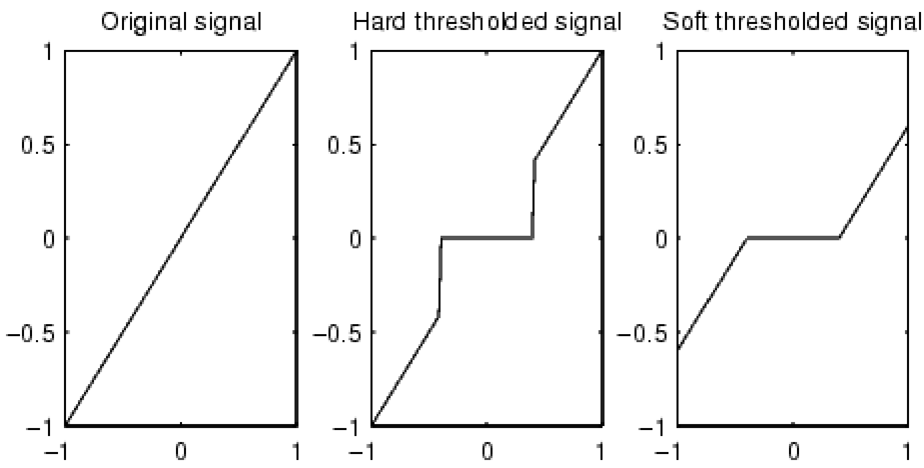


Fig.2.3 Examples of Wavelet tresholding Methods

There are some more tresholding methods available like strain tresholding etc., which can be chosen for better quality of denoising.

2.2.3 Principal Component Analysis (PCA) based denoising Technique

Principal Component Analysis is a linear decomposition technique that transforms a number of (possibly) correlated variables into a (smaller) number of uncorrelated variables called principal components. PCA finds components that are efficient for representing data, useful to represent the data in reduced dimension. Here the technique of Eigen analysis is used. We compute the eigen values and eigen vectors of a square symmetric matrix with sums of squares and cross products. By keeping only the components that are related to high eigen values and rejecting the components related to low eigen values, we will be able to denoise the data. Low eigen value represents low variance. Low variance can often be assumed to represent undesired background noise.

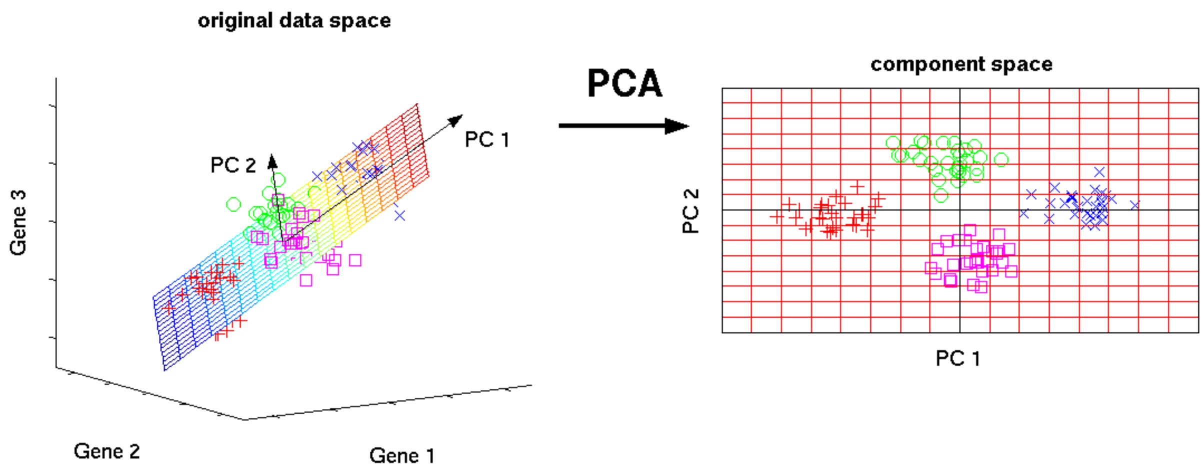


Fig.2.4 Principal Component Analysis

2.2.4 ICA based denoising Technique:

Artifact rejection: The filtered EEG signal still contains various artifacts like eye blink, muscle artifacts etc. These artifacts can be removed using the ICA method available in EEGLAB.

ICA converts the EEG data collected at single scalp channels into spatially transformed ‘virtual channel’ basis (Independent components). So, by removing the components that are associated with the artifacts and by back projecting this data, we can achieve the EEG data that is artifact free (filtered as well) [13].

In contrast to PCA, for which the first component may account for 50% of the data, the second 25%, etc..., ICA component contributions are much more homogeneous, ranging from roughly 5% down to ~0%. This is because PCA specifically makes each successive component account for as much as possible of the remaining activity not accounted for by previously determined components -- while ICA seeks maximally independent sources of activity. PCA components are temporally or spatially orthogonal - smaller component projections to scalp EEG data typically looking like checker boards - while ICA components of EEG data are maximally temporally independent, but spatially unconstrained -- and therefore able to find maps representing the projection of a partially synchronized domain / island / patch / region of cortex, no matter how much it may overlap the projections of other (relatively independent) EEG sources.

The steps for artifact rejection are implemented from the reference tutorial EEGLAB available online [14]. An example of the scalp maps of ICA components of a 14-channel EEG data has been presented in fig.2.5 The artifacts can further be reduced by performing ICA repetitively (multiple times). Original and pruned EEG data are shown in fig. 2.6.

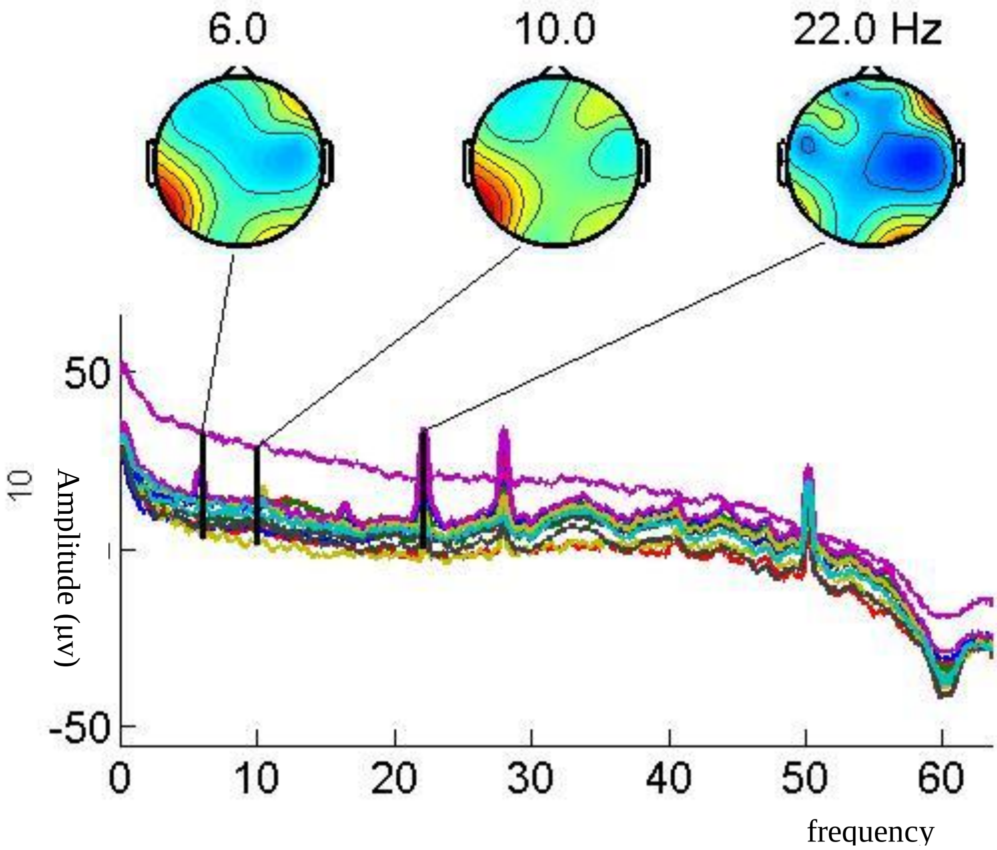
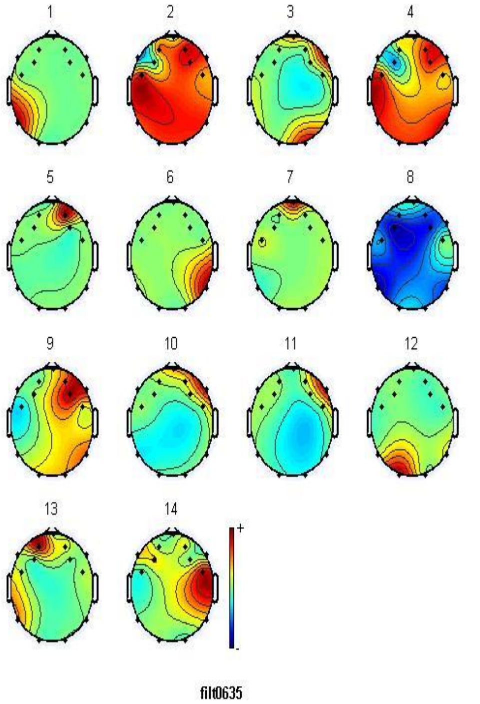


Fig.2.5 Filtered EEG signal (0.1 hz-63.5hz bandpass+ 60 hz notch)



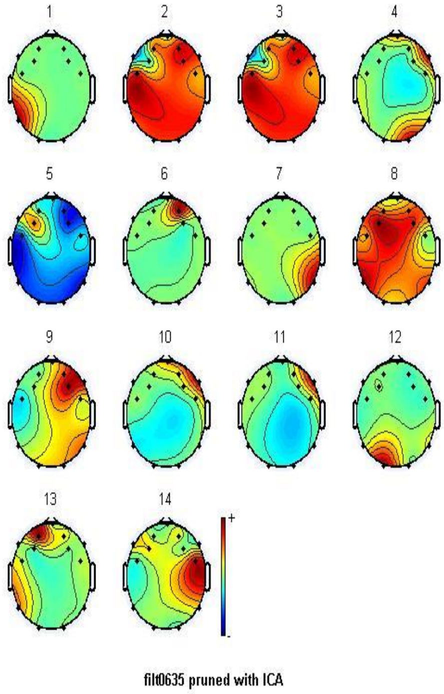


Fig. 2.6 a. ICA components map of Filtered EEG data

b. ICA components map of Pruned EEG data

The red portions in the each component indicate high amplitudes of the voltages in the corresponding areas of brain. The components can be decided to be artifacts based on the region where the level of intensity is high. For example in Fig.2.7.a, the component accounts for eye activity, we may wish to subtract it from the data before further analysis and plotting. Likewise the component shown in Fig.2.7.b appears to be typical muscle artifact component.

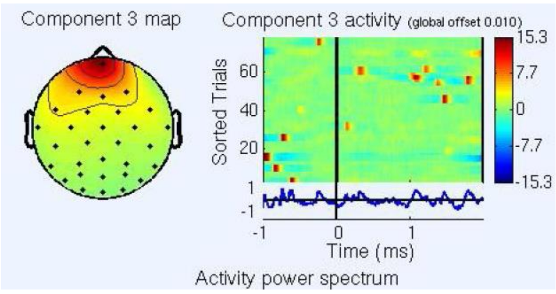
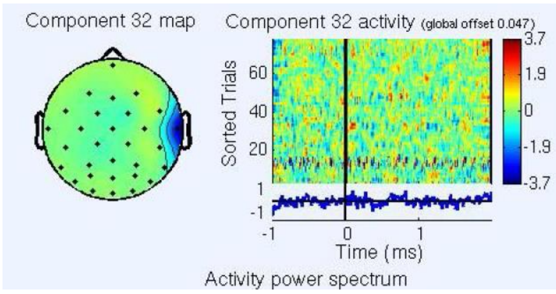


Fig. 2.7 a. Component map related to Eye artifact



b. Component map related to Muscle artifact

CHAPTER 3

FEATURE EXTRACTION

3. FEATURE EXTRACTION

In this section all the features that are explored in this work are discussed in detail.

The processed EEG data is very large in size, highly random and nonlinear in nature. Hence, we need to identify and extract the unique property of the signal so that we would be able to distinguish different kinds of EEG signal inputs using this unique property. This unique property which is used for the classification is known a 'Feature'.

Feature extraction step can be viewed as core part of any classification problem. The performance of any classifier is decided by the Feature that is selected. In this work, a number of features in conjunction with different classifiers were explored and the feature that yields the best results is found. The various features which are computed and analysed are discussed below one by one.

3.1 Power spectral density

Power Density Spectrum of Power Signal x(t) is given by,

$$S_x(w) = lim_{T\to\infty} \frac{(E[X(w)^2])}{2T} \tag{3.1}$$

3.2 Spectral Centroid

Spectral centroid is used to find the dominant spectral energy from the power spectrum. It is usually used to detect the dominant frequency from signals.
The equation for computing centroid is given by,

$$SC = \frac{\sum_{k=1}^{N} KX(K)}{\sum_{k=1}^{N} X(K)} \tag{3.2}$$

Where X[K] is the N-point DFT for a signal x(t).

3.3 Spectral Energy

The spectral energy distribution in a given EEG signal is obtained through the formula given by,

$$SE = \sum_{k=1}^{N} |X(K)|^2 \tag{3.3}$$

The first three features namely Power Spectral Density (PSD), Spectral Centroid, Spectral Energy are calculated in by decomposing the EEG signal into each of the EEG bands (delta, theta, alpha, beta and gamma). The decomposition of the EEG signal is done using the wavelet packet analysis toolbox available in MATLAB. The wavelet function used for the decomposition is Symlets 4.

3.4 ERP (Event Related Potential) Response

It is the response that is observed in the recorded EEG signal after the onset of stimuli (termed as event) to a subject. ERP components are often represented using the letters P/N followed by the time after which the component is occurring after the onset of a stimuli. For e.g. N1 represents that it is the first negative peak after the onset of stimuli, it can also be represented as N100 which is the negative peak occurring after 100ms after the onset of a stimulus. Similarly P300 is the positive peak occurring after 300ms after the onset of a stimulus shown in fig. 3.1.

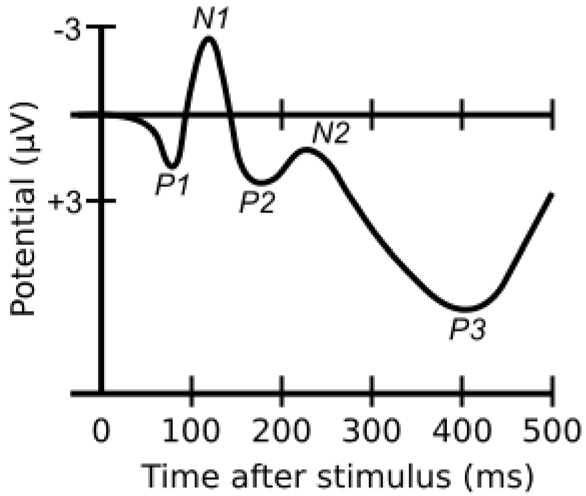


Fig.3.1 ERP response

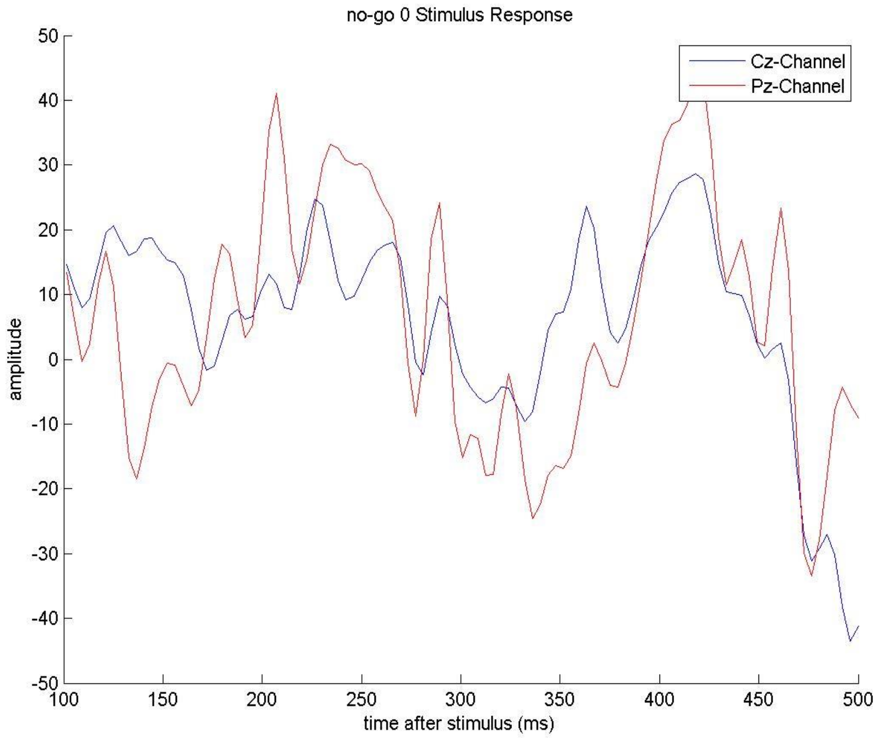


Fig.3.2 ERP response of Cz, Pz channels from a random subject

The ERP response for a stimulus computed for the channels Cz, Pz channels for a subject has been presented in fig.3.2. Research papers in [15], [16] show the use of ERP response used as a feature for the analysis of EEG signals.

3.5 Hurst exponent

It is used to find the rate at which the autocorrelations of a discrete signal is varying as the time difference between two instants of the signal is increased, and is represented using the letter 'H'

Its value generally lies between 0 and 1.

The value of H lying in the range 0-0.5 indicates the values in the signal are switching between high and low values, value of H between 0-1 indicates that a value in a signal is followed by a value that is higher than its previous value.

And H=0.5 represents an uncorrelated signal.

The procedure for calculating the Hurst Exponent has been presented in the form of algorithm as shown below.

Let the partial signal be, x= [$x_1, x_2, x_3 \dots \dots \dots, x_n$] of length n

a. Calculate the mean of the signal 'm', given by

$$m = \frac{1}{n} \sum_{i=1}^{n} x_n \tag{3.4}$$

b. Adjust the signal such that it is zero centred,

$$y_{t=}x_t - m \text{ , for t=1, 2,..., n} \tag{3.5}$$

c. Find the Cumulative deviate of the zero centred data

$$z_t = \sum_{i=1}^{t} y_i \text{ , for t=1, 2,..., n} \tag{3.6}$$

d. Find the range R,

$$\mathrm{R(n)} = max(z_1, z_2, \dots, z_t)\text{-}min(z_1, z_2, \dots, z_t) \tag{3.7}$$

e. Find the standard deviation S,

$$S(n) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_{i-}m)^2)} \tag{3.8}$$

f. Compute the rescaled range, $\frac{R(n)}{S(n)}$, and average W over the length n.

$$W = E\left[\frac{R(n)}{S(n)}\right] \tag{3.9}$$

Hurst exponent can be found from the expression,$W = C\, n^H$, which is the slope of the linear equation with log(W) taken along y-axis and log(n) along x-axis. D.A.J. Blythe et al. [17] discussed the use of the Hurst exponent feature for the analysis of EEG signals. Sriram et al.[18] used the feature of Hurst Exponent for the detection of sleep and wake stages.

3.6 Emperical Mode Decomposition

Emperical Mode Decomposition [19] is designed to decompose a non-linear and non-stationary signal s(n) into finite and small number of components. These are known as Intrinsic Modal Functions (IMFs), each one containing a spectrally independent oscillatory mode. This method is adaptive and highly efficient. Since the decomposition is based on the local characteristic time scale of the data, it can be applied to nonlinear and non-stationary processes. The decomposition is based on the direct extraction of the energy associated with various intrinsic time scales, the most important parameters of the system. Mohammad Zavid Parvez et al. [20] performed the analysis of EEG signals using the feature EMD and DCT, for the Classification of Ictal (state of seizure, stroke or headache) and Interictal (time period between Ictals) EEG Signals. Lisheng Chen et al. [21] shows the extraction of features from the EEG signals based on the EMD decomposition and Gabor Transform.

An IMF is defined as a function that satisfies the following requirements:

1. In the whole data set, the number of extrema and the number of zero-crossings must either be equal or differ at most by one.
2. At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

The algorithm for the extraction for IMF is presented below,

IMF Extraction:

- The process of IMF extraction is also known as Sifting process. It is described as follows :
- Find local minima and maxima of s(n).
- Form upper, e_u(n), and lower, e_l (n),
- Find the mean, m(n) = $[e_u(n) + e_l(n)]/2$
- If h(n) = s(n)−m(n) is not an IMF, go to step 1 using h(n) instead if s(n). Else, h(n) = IMF_1 (n).
- If the residue, r_1 (n) = s(n)− IMF_1 (n) has more than a zero cross, go to step 1 and find next IMF.

Once IMF_k(n) are extracted, the signal can be expressed as:

$$\mathrm{S}(n) = \sum_{k=1}^{M} IMF_k(n) + r(n) \tag{3.10}$$

EMD is performed i times, each one decomposing

$$S_i(\mathrm{n}) = \mathrm{s(n)} + w_i(\mathrm{n}) , \tag{3.11}$$

Where w_i(n) = different realizations (i = 1, .., I) of white noise

ε = variance of white noise

The stoppage criterion determines the number of sifting steps to produce an IMF. Following are the four existing stoppage criterion:

3.6.1 Standard Deviation

It similar to the Cauchy convergence test, and we define a sum of the difference, SD, as

$$SD_k = \sum_{t=0}^{T} \frac{|h_{k-1(t)} - h_{k(t)}|^2}{{h_{k-1(t)}}^2} \tag{3.12}$$

Then the sifting process stops when SD is smaller than a pre-given value. A typical value for SD can be set between 0.2 and 0.3. As a comparison, the two Fourier spectra, computed by shifting only five out of 1024 points from the same data, can have an equivalent SD of 0.2-0.3 calculated point-by-point. Therefore, a SD value of 0.2-0.3 for the sifting procedure is a very rigorous limitation for the difference between siftings.

3.6.2 S Number Criterion

This criterion is based on the so-called S-number, which is defined as the number of consecutive siftings for which the number of zero-crossings and extrema are equal or at most differing by one. Specifically, an S-number is pre-selected. The sifting process will stop only if, for S consecutive siftings, the numbers of zero-crossings and extrema stay the same, and are equal or at most differ by one.

3.6.3 Threshold Method

Proposed by Rilling, Flandrin and Goncalves, threshold method set two threshold values to guaranteeing globally small fluctuations in the mean while taking into account locally large excursions.

3.6.4. Energy Different Tracking

Proposed by Cheng, Yu and Yang, energy different tracking method utilized the assumption that the original signal is a composition of orthogonal signals, and calculate the energy based on the assumption. If the result of EMD is not a orthogonal basis of the original signal, the amount of energy will be different from the original energy.

Once a stoppage criterion is selected, the first IMF, c_1, can be obtained. Overall, c_1 should contain the finest scale or the shortest period component of the signal. We can, then, separate c_1 from the rest of the data by X(t)-c_1= r_1., Since the residue, r_1, still contains longer period variations in the data, it is treated as the new data and subjected to the same sifting process as described above. Here Standard Deviation is chosen as the stopping criterion.

In this work first 9 IMF's are extracted and the some of the statistical properties of these IMF's were computed and analysed. The statistical properties computed are discusses below

a. **Hurst Exponent of IMF's**:

The first feature that is explored is the Hurst Exponent of the IMF's. It has been observed that the results are more satisfactory by taking hurst exponent computed on the IMF's, rather than taking the hurst exponent of the original EEG signal as a feature vector. Different combinations of IMF's (First 3 or IMF 3, 4, 5 etc.) have been analysed to achieve the optimum classification accuracy and are discussed in the results section.

b. Energy, Standard Deviation of amplitude, Standard Deviation of phase of Hilbert Transform applied on IMFs:

Research paper [22] used the concept of Empirical Mode Decomposition along with Hilbert Transform for the assessment of the power quality. In this work a nine component vector is used as a feature vector. This feature vector is composed of the energy, standard deviations of amplitude and phase for Hilbert transformed data applied of first 3 IMFs (3 components per IMF) is used as a feature vector.

Hilbert transform tries to extract the analytic function of any given signal. It can be calculated as described below.

i. Calculate the Discrete Fourier Transform (DFT) of the discrete signal x(n), n=1,2,….N. where 'N' is the length of the signal.

The DFT of a discrete signal is computed using the formula,

$$X[K]=\sum_{n=0}^{N-1} x(n)\, e^{-2\pi ikn/N} \quad k=1,2,\ldots N \tag{3.13}$$

ii. Multiply the computed DFT matrix X with the mask M1 defined by,

$$Y=X*M \tag{3.14}$$

M1= {0, j, j,…..,0, -j, -j……..,-j}, if n is even

M1= {0, j, j,…..,0, j, -j, -j……..,-j}, if n is odd

iii. Compute the inverse DFT of the matrix Y to obtain, $\widehat{x(t)}$.

$$z(t)= x(t)+ \widehat{x(t)} \tag{3.15}$$

The amplitude a(t) can be derived from the expression

$$a(t)= \sqrt{x(t) + \widehat{x(t)}} \tag{3.16}$$

$$\theta(t) = \arctan(\widehat{x(t)}/x(t)) \tag{3.17}$$

Then the features like standard deviations of amplitude and phase are computed which are used as feature vectors.

The component, Energy of the IMF is found using the formula,

$$E = \sum_{t=0}^{N} a(t)^2 . \tag{3.18}$$

3.7 Difference Visibility Graph

Many of the brain researchers have started exploring the Graph based features of the physiological signals, and the results have been found promising, since the data is analysed epoch by epoch.

Lucasa et al. [23] first proposed the theory of visibility graph for the analysis of EEG signals. Research papers [24], [25] discuss the use of visibility graph for the diagnosis of the diseases like epilepsy and Alzheimer's disease (AD). In [27] a feature of difference visibility graphs has been used, for the analysis and classification of different sleep stages of a human being from the EEG signal data.

An undirected graph G(V,E) is basically a set of nodes and edges, where each point in the time series is taken to be a vertex 'V' and an edge 'E' can be a line joining any two vertices. The degree sequence k for the Graph is given by the sum of all the edges that are connected to the node.

Given a time series $\{x_t\}$, t= 1,…,n, the steps for transforming the data into difference visibility graph have been depicted below.

1. Form all the vertices of the Graph, $v_i = x_i \ \forall i = 1,2, \dots, n$.

2. Construct the Visibility Graph(VG) by finding all the edges that are possible between any two nodes of the graph based on the rule proposed by Lacasa et al. [23] given by the condition and find the degree sequence k of VG.

$$\frac{(x_j - x_k)}{j - k} > \frac{(x_j - x_i)}{j - i} \;;\; \forall k \in (i, j)$$

3. Now construct a Horizontal Visibility Graph (HVG) proposed by Luque et al. [26]., based on the condition for forming an edge as given below,

$$x_j > x_k \;\; and \; x_i > x_k; \; \forall k \in (i, j)$$

Find the degree sequence of the HVG.

4. Now Subtract the degree sequence of the HVG form VG to get the degree sequence of the Difference Visibility Graph (DVG).

The construction of the Visibility Graph and the Horizontal Graph and thereby the Difference Visibility Graph has been elicited through an example given below.

Let us assume the time series to be x= (7.3, 5.0, 6.2, 6.6, 5.7, 5.0, 9.1). Fig.3.3 and 3.4 show the visibility graph and horizontal visibility graphs constructed for the series x. The degree sequences for the corresponding VG and HVG's are found to be $k_{vg} = (4, 2, 4, 5, 3, 3, 5)$ and $k_{hvg} = (4, 2, 3,4 , 3, 2, 4)$ respectively. The degree sequence for the DVG is obtained as $k_{dvg} = k_{vg} - k_{hvg}$ which comes out to be $k_{dvg} = (0, \;\; 0, 1, 1, \;\; 0, 1, 1)$

The concept of DVG is introduced in order to eliminate the drawback involved in the HVG, where many of the nodes would be having the degree as 2, which makes the degree sequence to be absurd especially for a chaotic time series.

The feature vector chosen is of eleven components among which two are the mean of degree sequences of the DVG and HVG and the rest nine are the Degree Distributions (DD's) of the DVG, $p_{dvg(0)\;to}\, p_{dvg(8)}$. The number of degree distributions to be taken is decided intuitively. DD of the DVG is given by,

$$p(k) = p_{dvg(k)=} \frac{Number\ of\ times\ the\ degree\ k\ occuring\ in\ a\ degree\ sequence}{Total\ number\ of\ elements\ in\ the\ degree\ sequence} \tag{3.19}$$

Hence the degree distribution of the difference visibility graph is found to be $p_{dvg(0\,)=\frac{3}{7}}$, $p_{dvg(1)=\frac{4}{7}}$, Since the degree values that we get are 0 and 1.

The feature vector can finally be represented as,

$$F= [\overline{k_{dvg}}, \overline{k_{hvg}}, p_{(0)}, p_{(1)}, p_{(2)}, p_{(3)}, p_{(4)}, p_{(5)}, p_{(6)}, p_{(7)}, p_{(8)}]. \tag{3.20}$$

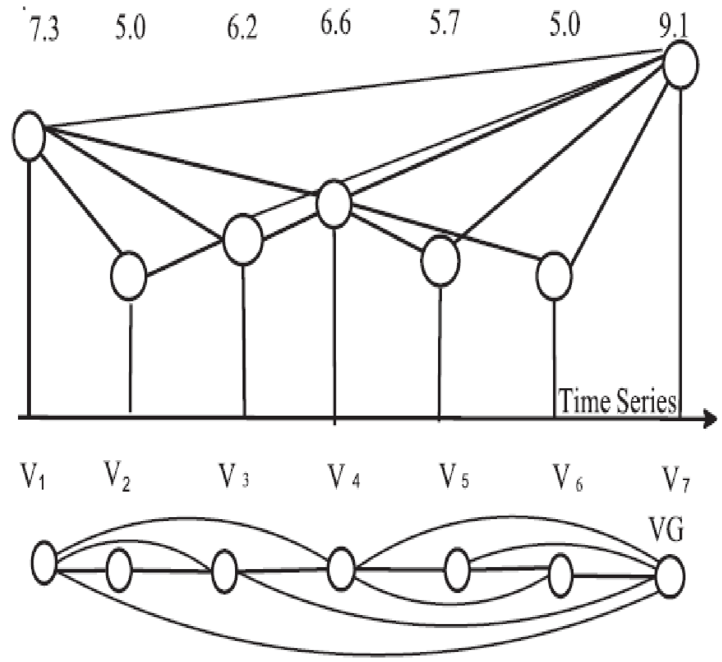


Fig. 3.3 Visibility Graph

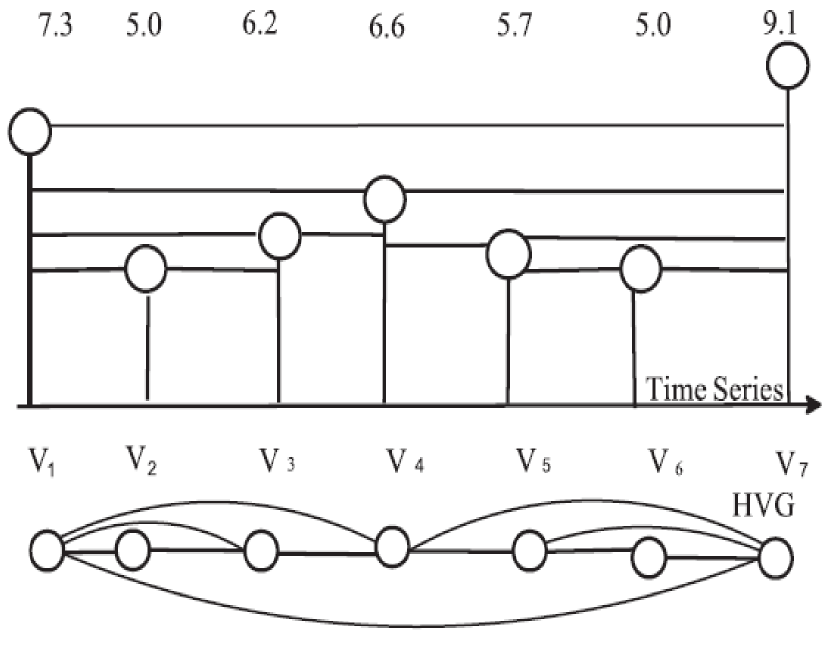


Fig. 3.4 Horizontal Visibility Graph

Each epoch (either GO or NOGO stimuli response) serves as an ERP which is the response to the stimuli. Here each epoch is the data extracted after 100-500ms on the onset of the stimuli.

3.8 Coherence Analysis

A coherence function can be used to find common frequencies of two signals and to evaluate the similarity of these signals. A wavelet coherence function, finds not only common frequencies of two signals, but also gives information when these frequencies appear. Jacob Benesty et al. [28] introduced the use of Minimum Variance Distortionless Response (MVDR) for finding the coherence frequencies present between two signals.

Coherence function is defined as

$$C_{xy} = \frac{P_{xy}(w)}{\sqrt{P_{xx}(w)\, P_{yy}(w)}} \qquad (3.21)$$

where Pxx and Pyy are power spectra of signals x and y, Pxy is cross-power spectrum for these signals. In case, when Pxx(w) = 0 or Pyy(w) = 0, then also Pxy(w) = 0 and we assume, that value Cxy(w) is zero.

3.9 Matching Pursuit Algorithm

Searching over an extremely large dictionary for the best matches is computationally unacceptable for practical applications. In 1993 Mallat and Zhang proposed a greedy solution that is known from that time as Matching Pursuit. The algorithm iteratively generates for any signal f and any dictionary D a sorted list of indexes and scalars which are the sub-optimal solution to the problem of sparse signal representation. The residual after calculating γ_n and a_n is denoted by R_{n+1}.

Algorithm:

- Input: Signal: f(t),dictionary D.
- Output: List of coefficients: (a_n,g_{γ_n}).
- Steps:
  a. Initialization:
     - $R_1 \leftarrow f(t)$
     - n←1
  b. Repeat:
     - Find $g_{\gamma_n} \in D$ *with maximum inner product* $| < R_n\, , g_{\gamma_n} > |$

- $a_n \leftarrow |< R_n, g_{\gamma_n} >|$
- $R_{n+1} \leftarrow R_n - a_n g_{\gamma_n}$
- $n \leftarrow n + 1$

Until stop condition (for example: $||R_n||$< treshold)

← represents shorthand for “changes to”

Matching Pursuit Algorithm finds the list of coefficients that will fit the given data. In this work it is shown how to calculate these weights. Further, statistics of these weights can be explored for the analysis of workload. Zin Mar Lwin et al. [29] presented the use of Gabor based Matching Pursuit for the classification of mental tasks.

CHAPTER 4

CLASSIFICATION

4. CLASSIFICATION

In this section we discuss concepts of classification, classifier, and the three types of classifiers that are used in this work.

A Classifier is a mathematical model that is used for distinguishing two or more objects represented as set of features (where objects can be images, biomedical signals, etc.). The entire set of objects is divided into two, namely training and testing sets. The division is made such that their intersection is a null set (training set ∩ testing set= Ø). The training set is taken to be the set of objects/examples that are known to the classifier. Testing set is taken to be set of objects/examples that are unknown. The ability to categorize correctly new examples that differ from those used for training determines the efficiency of the classifier.

4.1 Types of Classifiers:

Classifiers can be broadly categorised into two

4.1.1 Statistical Schemes: These schemes use the probabilistic models for the purpose of classification.

E.g.: Bayes Classifiers, Hidden Markov Models (HMM)

4.1.2 Learning Schemes: These schemes use mathematical models that require some kind of learning strategies for performing the classification task. In general, we find three kinds of learning strategies.

4.1.2.1 Supervised Learning: If both the input and the output to a classifier are known it is called supervised learning. An example of this kind of learning strategy is training a child with the alphabets, by inputting a picture of an alphabet and specifying its name.

4.1.2.2 Unsupervised Learning: If only the inputs to the classifier are known, then it is termed as unsupervised learning. An example of this type of strategy is classification of two animals (Buffaloes, Elephants) which doesn't require any prior knowledge of their corresponding outputs (exact categories). Here the classification is done just by differentiating into clusters based on some metric (Size in the above example mentioned).

4.1.2.3 Recurrent Learning: In this kind of learning the network adjusts by itself through its experience. The inputs are specified in terms of YES/NO i.e., an example is given to a network and network adjusts itself until it yields a correct output which is specified as YES when correctly classified and NO when wrongly classified.

Here in this work LDA based Radial basis neural network, LDA based k nearest neighbour (KNN) classifiers are used for the purpose of classification task.

4.2 Linear Discriminant Analysis:

Sometimes the feature vectors found for the examples of different classes might lie in such a way that they are hard to discriminate one amongst the other. Linear discriminant analysis is a technique that seeks the directions that are efficient for discrimination.

LDA transforms the training data such that for the transformed training data, the scatter within a class is minimised and the scatter between the classes is maximised as depicted in the fig.4.1

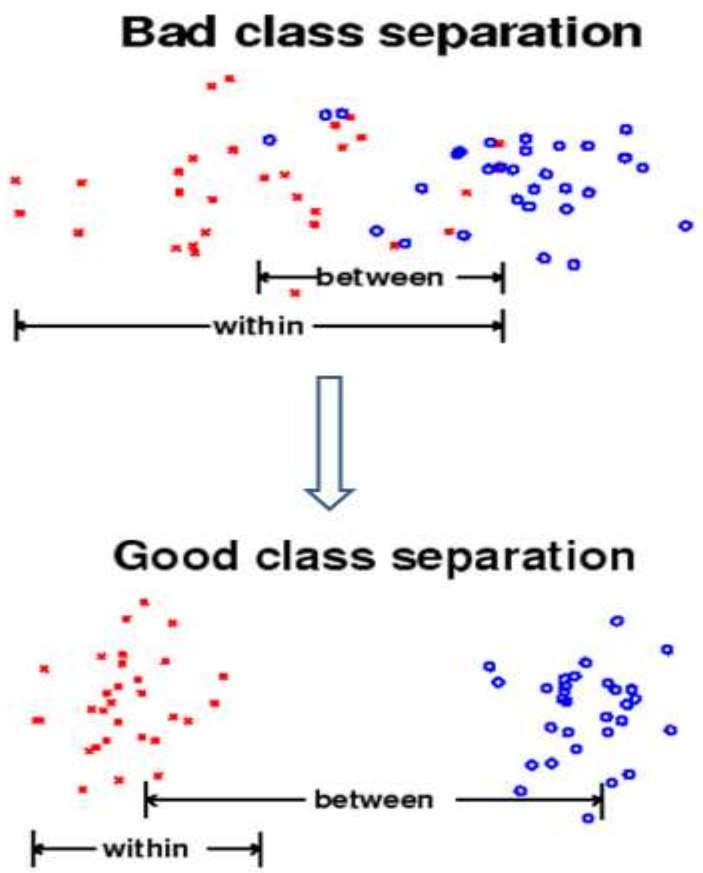


Fig.4.1. Linear Discriminant Analysis

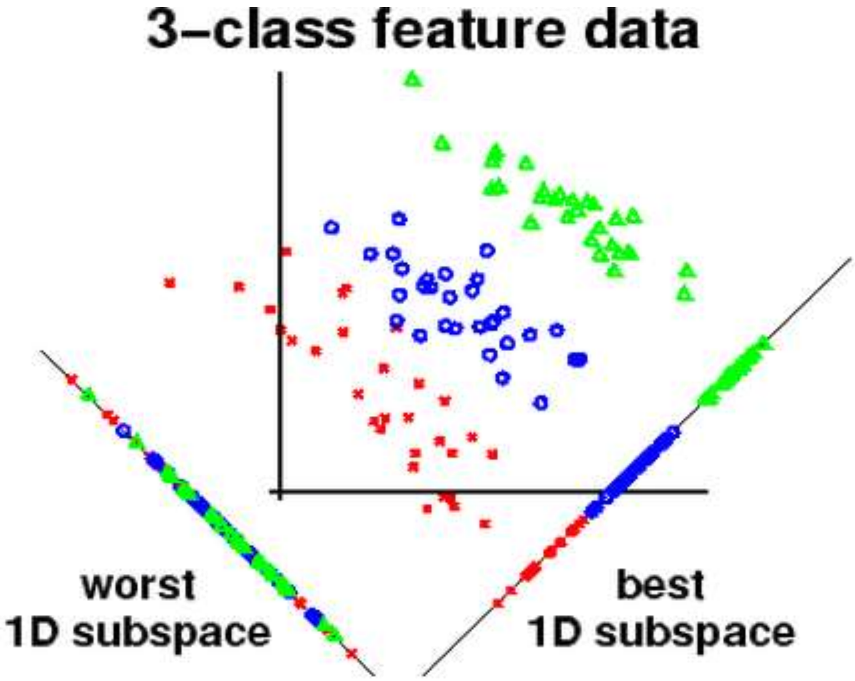


Fig.4.2. Application of LDA on a 3-class Feature data

It had been observed that the performance of the classifier when tested on the LDA transformed data was better than when it was done on the non-transformed data. Mohammad Shahin Mahanta et al. [30] shows the use of LDA for the extraction of the features from the EEG signals.

4.3 K-NN Classifier

It is a simplest, nonparametric and robust technique that can be used for the purpose of classification. Classification is done by relating the unknown to the known according to some metric. One of the most popular metrics chosen is distance (Euclidean distance, mahalanobis distance). The algorithm works as follows,

1. Training
   - In the training phase the feature vectors and their corresponding classes with which they are associated specified.

2. Testing
   - For any new example X= $(x_1, x_{2,} \dots \dots, x_n)$, find the Euclidean distance d(X, Y) between X and all the available training examples. The Euclidean distance is given by,

$$d(X, Y) = \sqrt{\sum_{j=1}^{n}(x_j - y_j)^2} \qquad (4.2)$$

- Find the k-nearest neighbours to X among all the available training examples.
- Determine the class values of k nearest neighbours found.
- The example X is assigned a class based on the majority of the classes among the class values of K nearest neighbours.

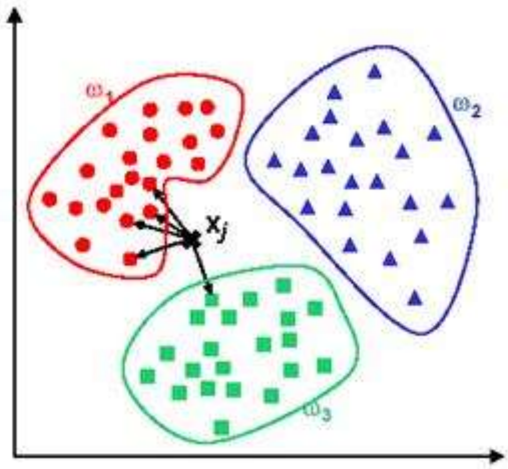


Fig.4.3. Illustration of Euclidean distance

Yazdani et al. [31] shows the use of KNN classifier for the classification of EEG signals.

4.4 Radial Basis Function Neural Network Classifier:

A Neural network is defined as an interconnected network of perceptrons. Where a perceptron is a linear classifier that maps the given input feature vector onto an output value using activation function. The activation function basically maps the values that are the linear combination of the feature vector with some predefined weights on to output values. The linearity/non-linearity of the classifier is decided by the activation function. Hard limiting function is an example of linear activation function, whereas a sigmoid function is used as an activation function for non-linear classifier.

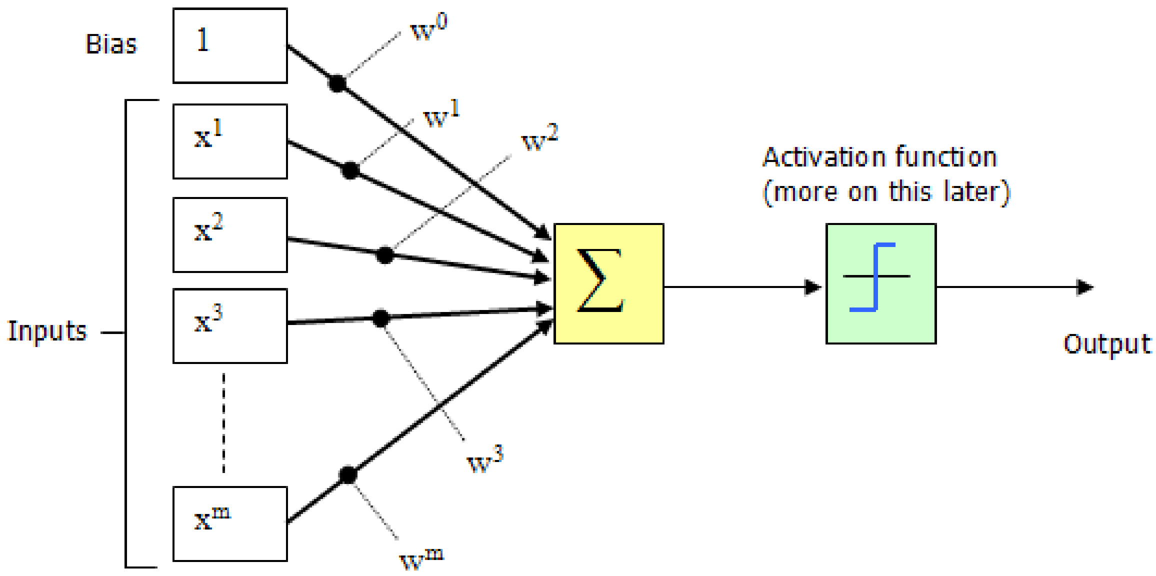


Fig.4.4. Perceptron

Hard limiting function is given by,

$$f(x)=\begin{cases} 1 & if\ x > 0 \\ -1 & if\ x < 0 \end{cases} \qquad (4.3)$$

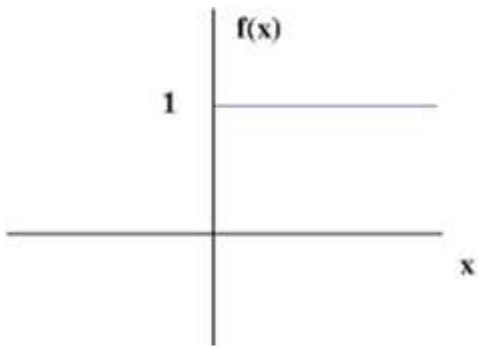


Fig.4.5. Hard Limiting function

The sigmoid function is given by,

$$f(s) = \frac{1}{1+e^{-s}},\quad s \in R . \qquad (4.4)$$

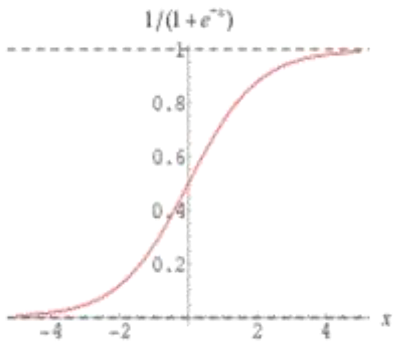


Fig.4.6. Sigmoid function

A neural network is made up of three layers namely the input layer, hidden layer and output layer as shown in fig.4.6. The number of neurons in the input layer N_i is determined by the number of components of the input vector. In the hidden layer the number (N_h) is taken roughly between $\frac{N_i}{3}$ or $\frac{N_i}{2}$. In the output layer the neuron number (N_o) is decided by the number of classes present in the classification problem.

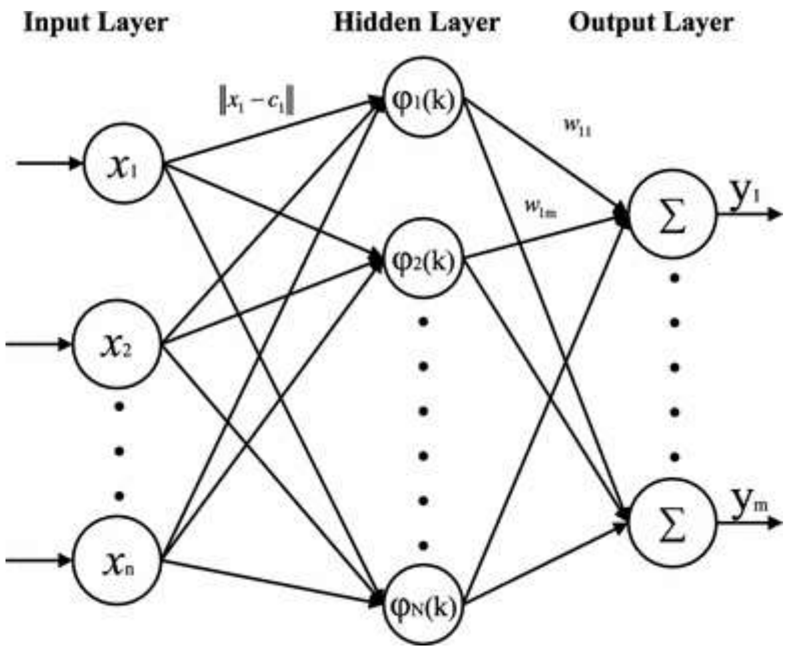


Fig. 4.7 Radial Basis Functional Neural Network (RBFNN)

There are some modifications that are made in RBFNN thus making it advantageous, when compared to Multi-Layer Perceptron (MLP) networks.

- Unlike the MLP networks, RBFNN uses a clustering algorithm for the adjustment of weights in the hidden layer, where as in the MLP networks use the concept of back propagation for weights adjustment which requires more time for training/learning.
- The learning methodology in input-hidden layer is unsupervised in RBFNN and it is supervised in the hidden-output layer. It is supervised in both, input-hidden and hidden-output layers.
- MLPBNN use linear summation functions for the neurons both in hidden and output layers. The hidden layer neurons use the set of radial basis

functions and the output layer neurons use linear summation functions in RBFNN.

The advantage of using RBFNN is it helps in reducing the learning time and it provides better interpolation of the given data. The algorithm for the RBFNN is given below.

1. **Weight Initialization:**

Weights in the hidden layer are determined by the clustering algorithm, and the weights in the output layer are initialized to small random values.

2. **Calculation of activation functions:**

- For the j^th neuron in the hidden layer the activation function is given by,

$$\varphi_j(x) = exp\left(\frac{-\left\|(x-c_j)^2\right\|}{2\sigma_j{}^2}\right) \tag{4.5}$$

$where\ x \in (x_{1,}\ x_{2,\dots,}x_n),\ c_j\ is\ the\ centre\ of\ jth\ neuron, \sigma_j is\ the\ spread\ width$

- For the j^th neuron in the output layer the activation function is given by,

$$y_j = \sum_{i=1}^{N_h} w_{ji}\, y_i \tag{4.6}$$

$where\ y_i is\ the\ output\ of\ the$ ith $neuron\ in\ the\ hidden\ layer, w_{ji}\ is\ the\ weight$ $between$ ith $neuron\ in\ the\ hidden\ layer\ and$ jth $neuron\ in\ the\ output\ layer$

3. **Weight Learning:**

- In the hidden layer, use k-means clustering
  - (i) Find the index of centre closer to x,

$$k(x) = \arg min_k \|x(t) - w_k(t)\| \tag{4.7}$$

(ii) $$w_k(t+1) = \begin{cases} w_k(t) + \eta\big(x(t) - w_k(t)\big)\ if\ k = k(t) \\ w_k(t) \qquad\qquad\qquad\qquad otherwise \end{cases} \tag{4.8}$$

- In the output layer,

$$w_{ji}(t+1) = w_{ji}(t) + \Delta w_{ji} \tag{4.9}$$

$$\Delta w_{ji} = \eta\ \delta_j\ o_i\,, \delta_j\ = (d_j\ - y_j\,) \tag{4.10}$$

Where $\eta\ is\ the\ learning\ rate.$

Repeat the above two steps until convergence.

In our work we used the three classifiers. Firstly a Radial Basis Function Neural Network (RBFNN), RBFNN on LDA transformed data (LDA+RBFNN) and K-Nearest Neighbour Classifier (LDA+KNN Classifier). These three classifiers are analysed in conjunction with the different features discussed in the previous section and the results are presented in the next section. Stuti Shukla et al. [21] shows the use of RBFNN for the classification of PQ events.

The complete classification process is summarized into a Flowchart as shown in Fig.4.7

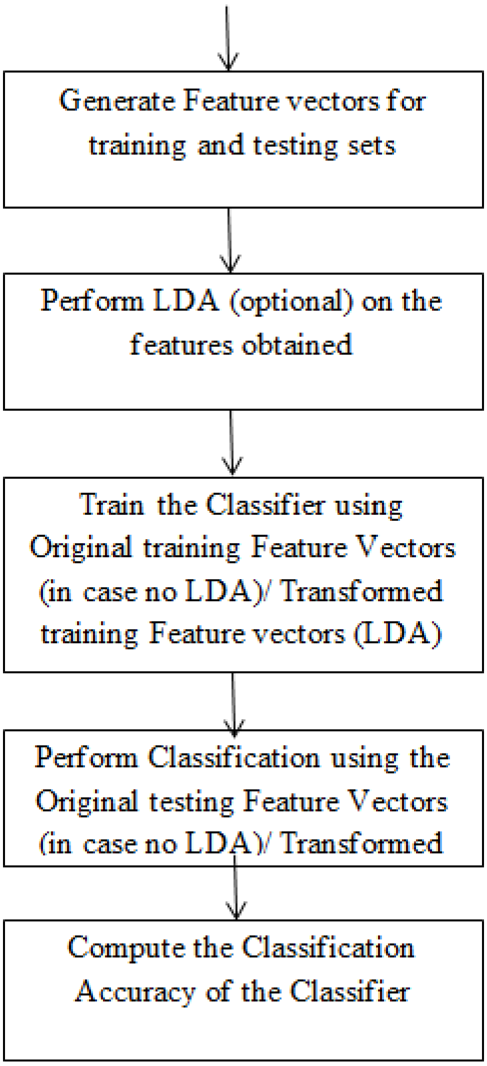


Fig.4.8 Flowchart for classifier

CHAPTER 5

EXPERIMENTAL SETUP

5. Experimental Setup

In this work, for the acquisition of the EEG data, 10-20 system (shown in fig.5.1) has been adopted. The electrode placement procedure can be referred from the manual by Trans Cranial Technologies [32]. The EEG data has been acquired using the equipment, NeXus-10 MKII equipment [33] shown in fig.5.2 which has 10 channels that can be used for simultaneous acquisition of various physiological signals like EEG, Heart rate, relative blood flow, skin conductance, temperature etc. Out of the 10 channels available, 4 channels can be used for the EEG data acquisition. An EXG sensor shown in fig.5.2 is used for the acquisition of the data from two channels namely C_z-P_z. 2 of the 4 connectors available in EXG sensor are connected to Central (C_z) and Parietal zones (P_z) of the human scalp, the other two connectors act as ground reference for the two electrodes. The ground reference channel for C_z electrode is taken as right mastoid (the bone which sits behind the ear) of the subject. The ground reference channel for P_z electrode is taken as left mastoid of the subject, and right ear lobe (the soft part of the ear) has been chosen as a common ground. The data is acquired at a sampling rate of 256 Hz (samples/sec).

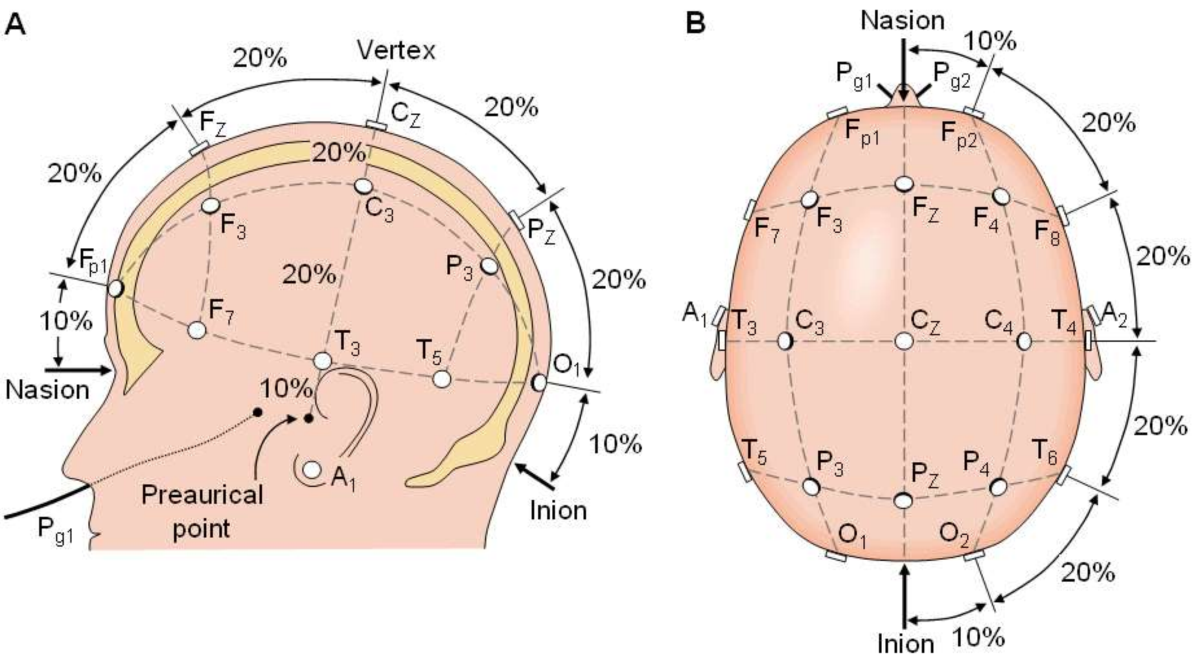


Fig.5.1. 10-20 Positioning System

Fig.5.2 NeXus-10 MKII equipment

Fig.5.3 EXG Sensor for EEG data Acquisition

The equipment comes with a software package BioTrace+, which is a versatile and highly customizable software that runs on Windows 8.1/8/7 (32/64-bit machines). The software can be used for several Bio/Neuro-feedback applications like

- Migraine and tension headache reduction
- Abdominal breathing, respiration training
- Stress test, stress management/reduction, relaxation training, anxiety training
- SMR, SMR & Theta, Theta/Beta, Gamma Band, Alpha/Theta Training
- Attention training for ADHD
- Slow Cortical Potentials (SCP)
- Heart Rate Variability (HRV) training
- Neuromuscular re-education
- Pelvic floor training for urinary incontinence
- Event-Related Potentials (ERP) and more.

Here we are using this software for the application of Event-Related Potentials (ERP). The subjects were asked to remain in isolation for one day to avoid any deviations to the experiment and also to imitate the way criminals do. After the setup of the equipment the subjects were asked to remain calm and stable, in order to remove any type of emotional stress upon them. Then they were shown presentations (Stimulus) of different categories (classes) which were made using the software and the response is recorded.

The presentation starts with a baseline of 10 seconds followed by a series of 'NOGO' and 'GO' stimulus. A 'NOGO' stimulus represents some irrelevant pictures, whereas a 'GO' stimulus represents a relevant picture of that category. A 'GO' signal is actually sandwiched between every four to five 'NOGO' signals. This is basically done to divert the attention of the suspect before the presentation of actual relevant (GO) signal, so that he/she would not be able to adjust the neuronal circuit present inside the brain to evade the stimulus under interrogation.

There are actually three categories of stimulus that are used in this work as a part of investigation. First one is the 'instrument' that was used for committing the crime, second the 'place' at which the crime had been committed, third the 'victim' of the crime.

The subjects are divided into two sets. First type is of actual and the second type are of control subjects. The actual subjects consist of persons who are actual or formal suspects of the crime, i.e. the actual suspects are the persons who suspects in real, whereas formal suspects are the made up ones. These persons were actually trained or explained about all the details pertaining to the crime (instrument used in the crime, place of crime and the victim of the crime). The details are so specific in nature. For instance, if we take the case of instrument used, a presentation had been made that consisted of the 10 seconds baseline followed by some irrelevant pictures (NOGO's) followed by some detail of the instrument (GO). All the pictures are displayed at a resolution of 640x480 pixels, which is an effective resolution that suits the angle of view of a human eye. Secondly the control subjects are the persons who are completely innocent and ignorant of any of these stimuli.

According to the hypothesis, the ERP response for the actual subjects and the control subjects varies since the control subjects are no way correlated to these crime incidents. The presentation was made after the careful of the methodologies described in the research studies [34], [35], [36], [37], [38].

For testing the Cognitive Workload, the subject was asked to perform a task of Multiple Object Tracking, a mind game with varying levels of difficulty to track some balls among a group of balls moving at random.

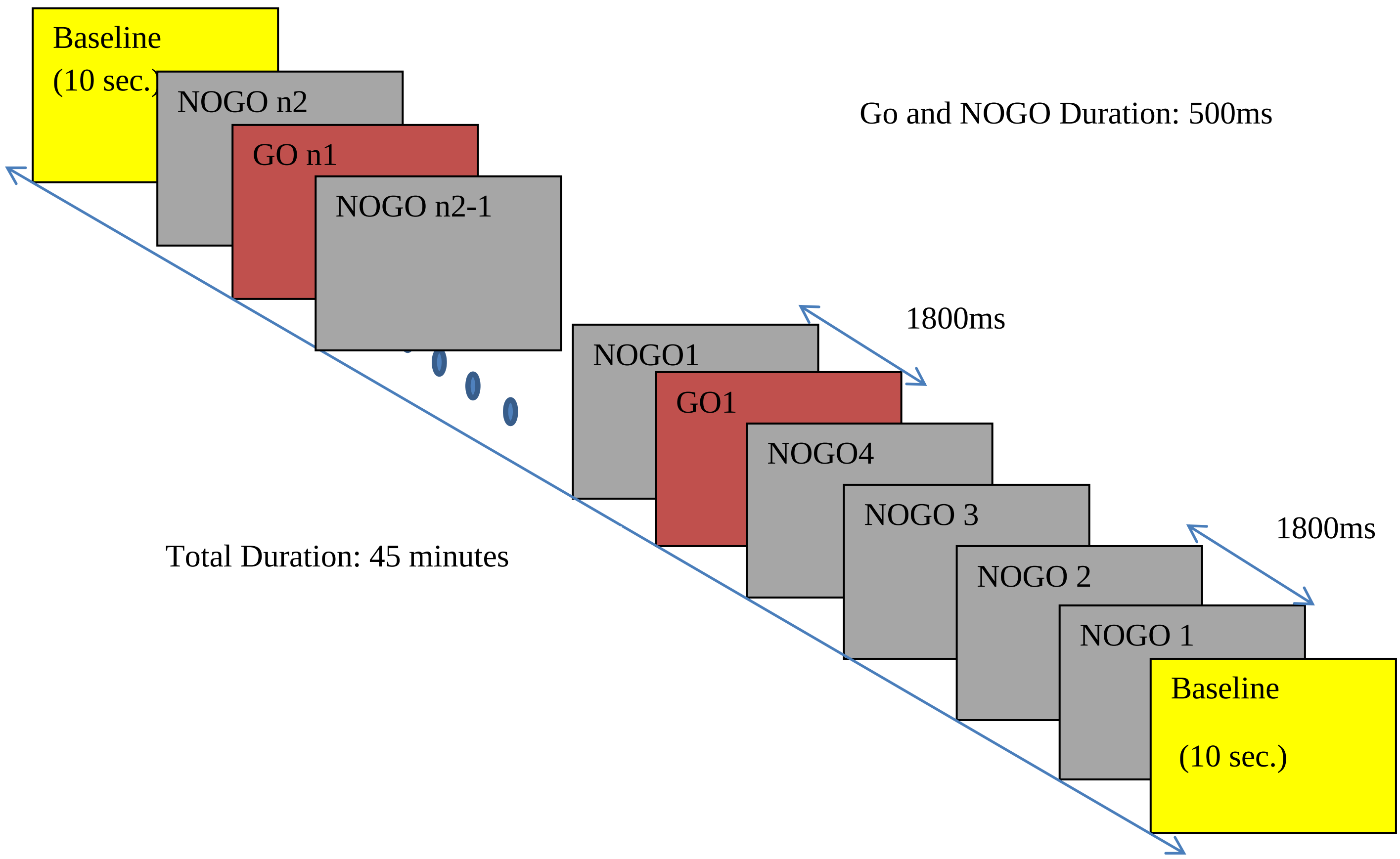


Fig.5.4 Stimulus Presentation For the task of Crime Investigation (Working Memory)

CHAPTER 6

RESULTS AND DISCUSSION

6. RESULTS AND DISCUSSION

This work, as a part of Cognitive Assessment, concentrates on the Working Memory and Cognitive Workload. The features useful for the analysis of cognitive workload have been explored and the results are presented. The concept of working memory is used for the purpose of crime investigation. Inter Stimuli and Intra subject classifications have been performed and the results are presented.

6.1 Cognitive Workload

Features like matching pursuit analysis and coherence analysis discussed in the above sections were found and the results are presented.

6.1.1 Coherence Analysis Results

Coherence analysis is used to find the common frequencies present in the C_z and P_z channels. From these frequencies found are helpful in the assessment of workload. It is analysing the correlation between different parts of the brain at various workloads.

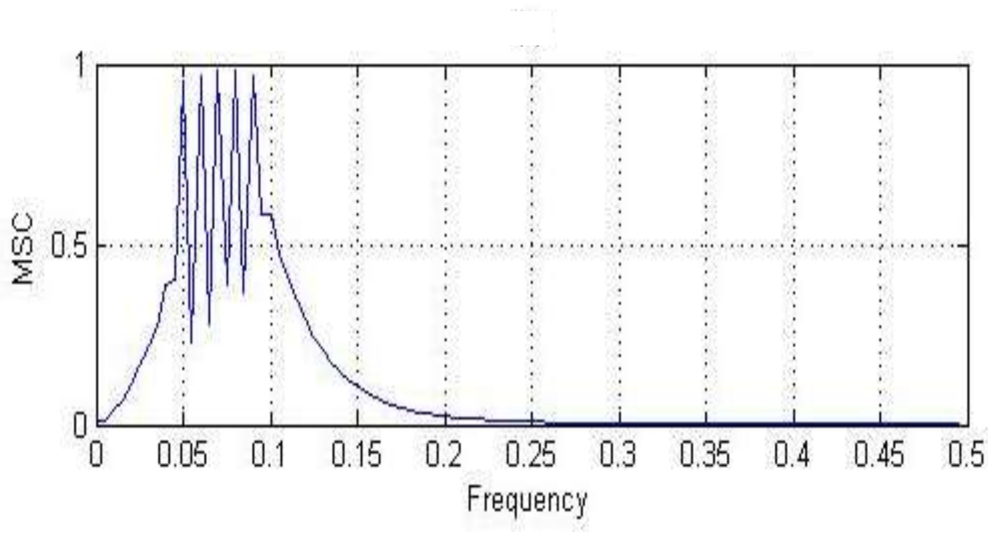


Fig.6.1 Coherence Analysis performed on C_z and P_z channel data

6.1.2 Matching Pursuit Algorithm Results

Matching Pursuit analysis finds the list of coefficients that will fit the given data. In this work it is shown how to calculate these weights. Further, statistics of these weights can be explored for the analysis of workload.

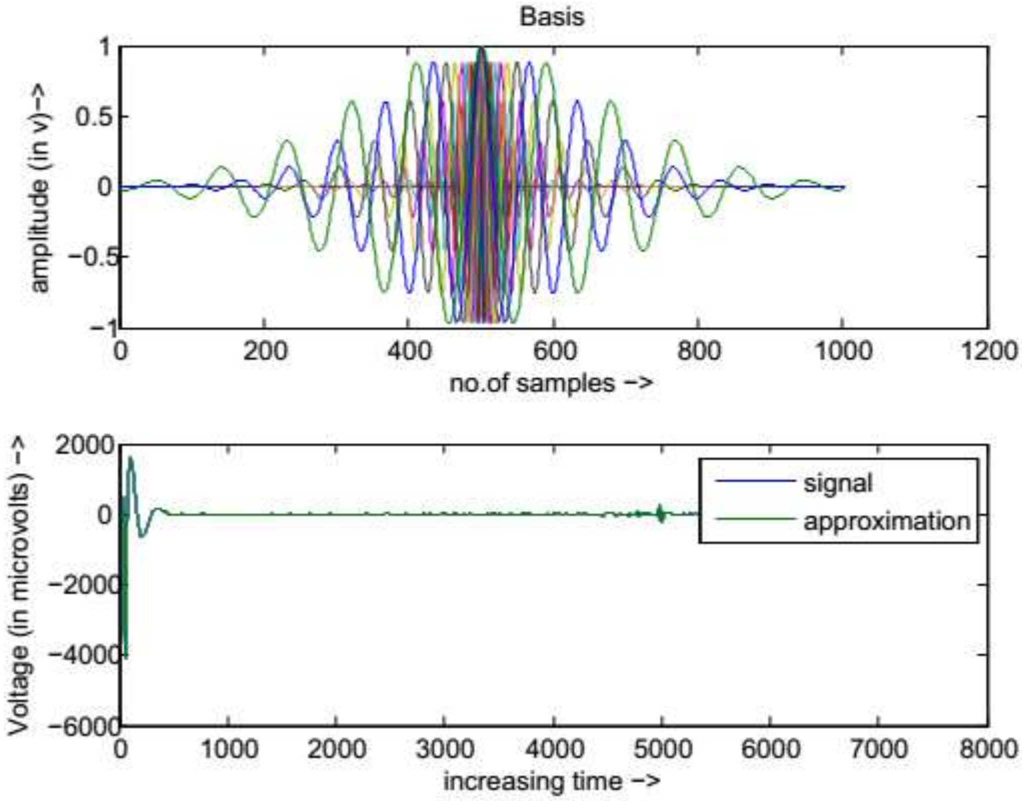


Fig.6.2 Matching Pursuit Analysis performed for channel C_z data

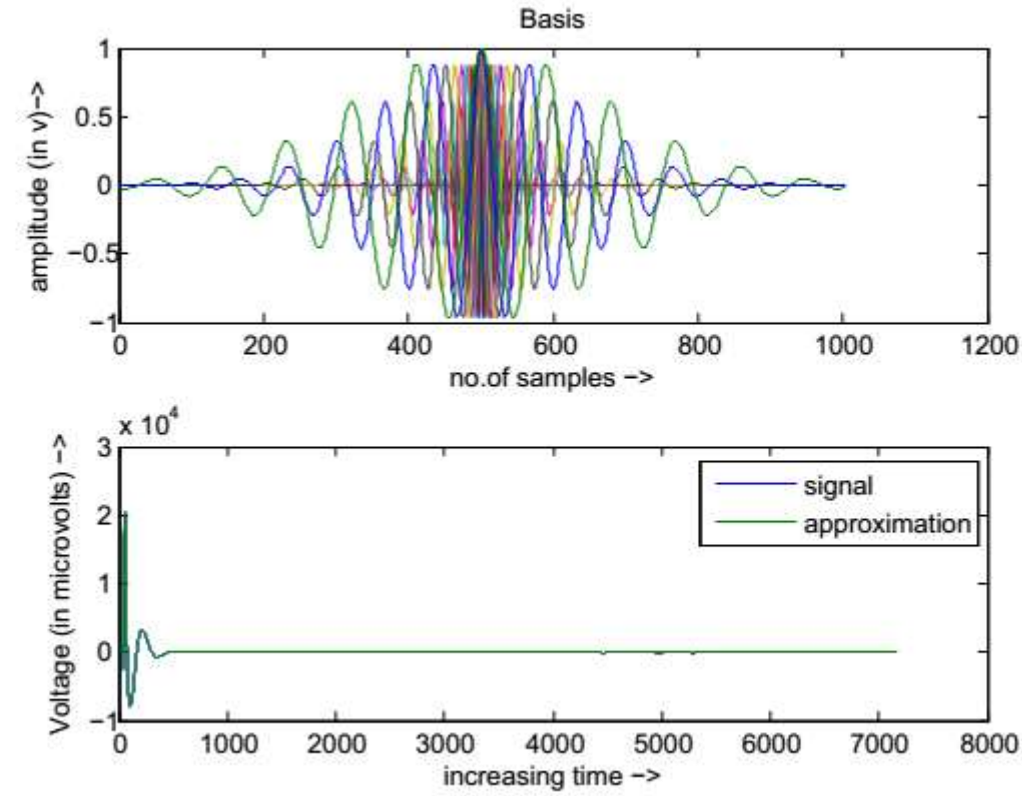


Fig.6.3 Matching Pursuit Analysis performed for channel P_z data

6.2 Working Memory

As a part of working memory, Crime investigation can be done by performing Inter Stimuli and Inter Subject classification.

6.2.1 Inter Stimuli Classification

Inter Stimuli Classification can be used to categorize a given data from the subject into one of the three categories that represent different kinds of stimuli pertaining to the crime. This gives out the affinity of the subject towards the stimuli pertaining to the crime.

As mentioned earlier, every subject was trained with three different kinds of stimuli namely instrument, victim, place. In this work, data from five subjects were taken. For each subject and for each category of stimuli, five examples were shown. They were given as follows

Category I: Instrument used in the crime: inst1, inst2, inst3, inst4, inst5

Category II: Place of crime: Place 1, Place 2, Place 3, Place 4, Place 5

Category III: Victim of crime: Victim1, Victim2, Victim3, Victim4, Victim5

As far as Inter Stimuli Classification is concerned, inst1-3, Place1-3, Victim1-3 of all the five subjects were placed in the training set. Inst4-5, Place4-5, Victim4-5 of all the five subjects were placed in the testing set.

So, near about 40 examples were placed in the training set, and about 28 examples were placed in the testing set.

The classification results for different features are as follows,

FEATURE	C_z-Channel Class. Accuracy (%)			P_z-Channel Class. Accuracy (%)		
	CLASSIFIER			CLASSIFIER		
	RBFNN	LDA +RBFNN	LDA+ KNN	RBFNN	LDA +RBFNN	LDA+ KNN
Feature I	28.5714	28.5714	32.1428	50	37.5	37.5
Feature II	35.7143	32.1429	42.85714	33.33	50	41.667
Feature III	32.1429	42.85714	32.1428	29.1667	37.5	33.33
Feature IV	35.7143	32.1429	28.5714	33.33	66.67	33.33
Feature V	17.857	39.2857	42.8571	29.1667	25	16.67

Table.6.2.1.1 Inter Stimuli Classifictaion Accuracies

Feature I: Hurst Exponent

Feature II: EMD (Hurst Exp. On first 3 IMFs)

Feature III: EMD (Energy, Standard deviation of amplitude, Standard deviation of phase of Hilbert transform applied on first 3 IMFs) [E, Sa, Sp]

Feature IV: EMD (Hurst Exp. On 3-5 IMFs)

Feature V: EMD (Energy, Standard deviation of amplitude, Standard deviation of phase of Hilbert transform applied on 3-5 IMFs) [E, Sa, Sp]

RBFNN: Radial Basis Function Neural Network Classifier

LDA+RBFNN: Linear Discriminant Analysis + RBFNN Classifier

LDA+ KNN: Linear Discriminant Analysis + K-Nearest Neighbor Classifier

The above features used were calculated on the whole EEG signal (like inst1). Another feature was also extracted using the concept of Difference Visibility Graph (DVG). Instead of extracting the feature from the entire data inst1, this deals the data epoch by epoch i.e., we try to convert the GO and NOGO stimuli of inst1 to Difference Visibility Graphs and try to extract the feature vector which is a combination of mean of degrees and degree distributions. Here the classifier used was Radial Basis Function Neural Network Classifier.
The classification accuracies for C_z and P_z channels were as follows,

FEATURE	C_z-Channel Class. Accuracy (%)	P_z-Channel Class. Accuracy (%)
DVG	55	55

Table.6.2.1.2 Inter Stimuli Classifictaion Accuracies using DVG Feature

6.2.2 Inter Subject Classification

Inter Subject Classification is performed to segregate the actual and control subjects i.e., this classification method is useful for finding out the actual criminals. This is a more useful method than the existing methods for crime investigation like lie detection and polygraph tests which rely on the heart rate variability, pulse variations wherein the major crimes hard-core criminals are quite trained to beat these tests. But this method discussed in our work suits better for crime investigation as the subject won’t be able to manipulate the brain’s response that occurs for a period of milliseconds.

In this classification task we tried to perform one against one type of classification i.e., data of one actual subject and one control subject are classified. For doing this task, the inst1,2,3; Place1,2,3; Victim1,2,3 of both the control and actual subjects were placed in training set and inst4,5; Place4,5; Victim4,5 of both the control and actual subjects were placed in testing set.

The classification results of three actual subjects against a control subject were obtained as follows

FEATURE	C_z-Channel Class. Accuracy (%)			P_z-Channel Class. Accuracy (%)		
	CLASSIFIER			CLASSIFIER		
	RBFNN	LDA +RBFNN	LDA+ KNN	RBFNN	LDA +RBFNN	LDA+ KNN
Feature I	50	50	58.333	83.333	25	66.667
Feature II	50	41.667	66.667	50	41.667	41.667
Feature III	50	33.333	41.667	33.333	50	41.667

Table.6.2.2.1 Inter Subject Classifictaion Accuracies for SUBJECT1 vs SUBJECT4

FEATURE	C_z-Channel Class. Accuracy (%)			P_z-Channel Class. Accuracy (%)		
	CLASSIFIER			CLASSIFIER		
	RBFNN	LDA +RBFNN	LDA+ KNN	RBFNN	LDA +RBFNN	LDA+ KNN
Feature I	100	100	100	100	100	100
Feature II	50	75	91.667	50	91.667	83.333
Feature III	25	66.667	66.667	50	66.667	75

Table.6.2.2.2 Inter Subject Classifictaion Accuracies for SUBJECT2 vs SUBJECT4

FEATURE	C_z-Channel Class. Accuracy (%)			P_z-Channel Class. Accuracy (%)		
	CLASSIFIER			CLASSIFIER		
	RBFNN	LDA +RBFNN	LDA+ KNN	RBFNN	LDA +RBFNN	LDA+ KNN
Feature I	91.667	91.667	100	58.333	58.333	83.333
Feature II	50	66.667	83.3330	50	75	91.667
Feature III	58.333	58.333	75	66.667	66.667	66.667

Table.6.2.2.3 Inter Subject Classifictaion Accuracies for SUBJECT3 vs SUBJECT4

Feature I - Hurst Exponent

Feature II - EMD (Energy, Standard deviation of amplitude, Standard deviation of phase of Hilbert transform applied on first 3 IMFs) [E, Sa, Sp]

Feature III - EMD (Hurst Exp. On first 3 IMFs)

Here Subjects 1, 2, 3 are actual and subject 4 is a control subject.

Table 6.2.1, Table 6.2.2, Table 6.2.3 show a very good overall Inter subject classification accuracies for the classifier LDA+KNN.

A maximum classification accuracy of 66.67% had been obtained when tried with the LDA+RBFFNN classifier for P_z-channel when the feature of Hurst Exponent on 3-5 IMFs is used for the task of inter Stimuli Classification and a maximum classification accuracy of 100% had been obtained when tried with the LDA+KNN classifier for C_z and P_z channels when the feature of Hurst Exponent is used for the task of inter Subject Classification and the results show a very good overall Inter subject classification accuracies for the classifier LDA+KNN in comparison to the other two classifiers namely RBFNN, LDA+RBFNN.

CHAPTER 7

CONCLUSION & FUTURE SCOPE

7. CONCLUSION & FUTURE SCOPE

In this work, there were two things that are discussed. Firstly, analysis of some of the most commonly used features for the analysis of Cognitive Workload. Secondly, a task of crime investigation as a part of Memory testing had been discussed.

It was shown through the analysis of features described in the previous sections, that how we can assess the workload of the brain. Under Memory testing, two tasks namely, Inter Stimuli Classification and Inter subject Classification were performed. Inter Subject Classification is performed to segregate the actual and control subjects i.e., this classification method is useful for finding out the actual criminals. Inter Stimuli Classification can be used to categorize a given data from the subject into one of the three categories that represent different kinds of stimuli pertaining to the crime, which tells us, to which stimuli presented in the crime is the criminal more connected with.

Also, through the analysis it is evident that the inter subject classification of the EEG data can be used an effective method when compared with any of the existing methods for the task of crime investigation.

The classification accuracy of this experiment is solely dependent on the quality of the data acquired from the subjects. Quality of the data is dependent on the various factors such as up to what level the subjects are made familiar with the crime scenarios, subject's memory power etc. Proper care must also be taken such that same stimulus is not being repeated in the presentations of different sets of a single category (inst1, 2, 3, 4, 5) for the same subject.

Through the results we can also derive a conclusion that which region of the brain is more useful for the analysis and assessment of the cognitive related tasks. In our case the Pz channel data seems to be more useful for the task of Inter Stimuli Classification achieving a maximum classification accuracy of 66.67% shown in the results.

Another notable thing of this work is that the data used for the analysis of the tasks is only of two channels (Cz, Pz). As a whole, it can be stated that this method used for the task of crime investigation is the simplest, fastest and highly accurate method when compared to any other existing tasks in present.

Boosting up the classification accuracy by the analysis of some more algorithms for the task of Inter Stimulus classification is recommended to be an extension to this work.

Regarding the task of Inter Subject classification, there is also a limitation that lies in this work to judge the certainty of the results obtained for the task of crime investigation. There may be some special cases, where a person undergoing the test is no way related to the crime but someway has the knowledge of the stimulus used for the investigation, can be mistaken to be a criminal. Few people like Journalists, Crime investigators might not actually be connected to the crime but can be familiar with the specific information related to the crime which is being used as a stimulus.

To overcome this limitation we can include analysis of P300 MER-MER (Memory and Encoding Related Multifaceted Electroencephalographic Response) [39], through which we would be able to know whether the information obtained, is it actually stored in the brain or not. This can be viewed as the future scope to this work.

References

1. Simpson, G.V., Connectivity and dynamic function in the brain. Biomedical Imaging, 2002. Proceedings. 2002 IEEE International Symposium on, 2002, pp. 399 – 400.

2. Fabio Babiloni, Consumer nueroscience: a new area of study for biomedical engineers, IEEE PULSE, 31 May 2012, pp. 2154-2287.

3. Du-Jian Gang, Wang Lin, Zheng Qi, Liu-Li Yan, Neuromarketing: Marketing through Science, Service Sciences (IJCSS), 2012 International Joint Conference on, 24-26 May 2012, pp. 285 –289.

4. Laura Astolfi, F. De Vico Fallani, F. Cincotti, D. Mattia, L. Bianchi, M. G. Marciani, S. Salinari, A. Colosimo,A. Tocci, R. Soranzo, and F. Babiloni, Neural Basis for Brain Responses to TV Commercials:A High-Resolution EEG Study, Ieee Transactions On Neural Systems And Rehabilitation Engineering, December 2008, Vol. 16, No. 6, pp. 522 – 531.

5. Dahal, N, Nandagopal N, Nafalski A, Nedic, Z, Modeling of cognition using EEG: A review and a new approach, TENCON 2011 - 2011 IEEE Region 10 Conference, 21-24 Nov. 2011, pp. 1045 – 1049.

6. Niedermeyer E. and da Silva F.L. (2004). Electroencephalography: Basic Principles, Clinical Applications, and Related Fields. Lippincot Williams & Wilkins. ISBN 0781751268.

7. Simranpreet Kaur, Sheenam Malhotra, Various Techniques for Denoising EEG signal : A Review, International Journal Of Engineering And Computer Science, August, 2014, ISSN:2319-7242, Volume 3, Issue 8, pp. 7965-7973.

8. Priyanka Khatwani, Archana Tiwari, A survey on different noise removal techniques of EEG signals, International Journal of Advanced Research in Computer and Communication Engineering, February 2013, Vol. 2, Issue 2, pp. 1091-1095.

9. Liwei Zhang, Guozhong Liu, Ying Wu, Wavelet and Common Spatial Pattern for EEG Signal Feature Extraction and Classification, Computer, Mechatronics, Control and Electronic Engineering (CMCE), 2010 International Conference on, 24-26 Aug., Vol 5, 2010, pp. 243 – 246.

10. Prochazka A, Kukal J, Vysata O ,Wavelet Transform Use for Feature Extraction and EEG Signal Segments Classification, Communications, Control and Signal Processing, 2008. ISCCSP 2008. 3rd International Symposium on, 12-14 March 2008, pp. 719 – 722.

11. Michahial S, Ranjith Kumar R, Hemath Kumar P, Puneeth Kumar A, Hand rotate EEG signal feature extraction by second order Daubechies wavelet transform (DWT), Computing Communication & Networking Technologies (ICCCNT), 2012 Third International Conference on, 26-28 July 2012, pp. 1-6.

12. Arjon Turnip, Dwi Esti Kusumandari, Artifacts Removal of EEG Signals using Adaptive Principal Component Analysis, 3rd International Conference on Computation for Science and Technology (ICCST-3), 30 November 2014.

13. Chun-Hsiang Chuang, Li-Wei Ko, Yuan-Pin Lin, Tzyy-Ping Jung, Chin-Teng Lin Independent Component Ensemble of EEG for Brain–Computer Interface, Volume:22, March24, Issue: 2, pp. 230 – 238.

14. EEGLAB, an open source environment for electrophysiological signal processing http://sccn.ucsd.edu/eeglab/index.html

15. Nathalie Fritsch, Lars Kuchinke, Acquired affective associations induce emotion effects in word recognition: An ERP study, Brain and Language Volume 124, Issue 1, January 2013, pp. 75–83.

16. Julie Onton, Scott Makeig, Information-based modeling of event-related brain dynamics, Progress in Brain Research, 2006, Vol.159, pp. 99-120.

17. Duncan A.J. Blythe, Stefan Haufe, Klaus-Robert Muller, Vadim V. Nikulin, The effect of linear mixing in the EEG on Hurst exponent estimation, NeuroImage, October 2014, Volume 99, 1 ,pp. 377–387

18. N.Sriraam, Purnima, B.R, Uma K , Padmashri T.K., Hurst Exponents Based Detection Of Wake-Sleep- A Pilot Study, Circuits, Communication, Control and Computing (I4C), 2014 International Conference on, Nov. 2014, pp. 118 – 121.

19. Norden E. Huang, Zheng Shen, Steven R. Long, Manli C. Wu, Hsing H. Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, Henry H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, The Royal Society, March 1998, Volume: 454, Issue: 1971, pp. 903-995.

20. Mohammad Zavid Parvez, Manoranjan Paul, Features Extraction and Classification for Ictal and Interictal EEG Signals using EMD and DCT, Computer and Information Technology (ICCIT), 2012 15th International Conference on, 22-24 Dec. 2012, pp. 132 – 137.

21. Lisheng Chen, Erbo Zhao, Dahui Wang, Zhangang Han, Shouwen Zhang, Cuiping Xu, Feature Extraction of EEG Signals from Epilepsy Patients Based on Gabor

Transform and EMD Decomposition, Natural Computation (ICNC), 2010 Sixth International Conference on (Volume:3), 10-12 Aug. 2010, pp. 1243 – 1247.

22. Stuti Shukla, S. Mishra, Bhim Singh, Empirical-Mode Decomposition With Hilbert Transform for Power-Quality Assessment, IEEE transactions on power delivery, october 2009, vol. 24, no. 4, pp. 2159 - 2165.

23. L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nu˜no, "From time series to complex networks: The visibility graph," in Proc. Nat. Academy Sci., Apr. 1 2008, vol. 105, pp. 4972–4975.

24. X. Tang, L. Xia, Y. Liao, W. Liu, Y. Peng, T. Gao, and Y. Zeng, "New approach to epileptic diagnosis using visibility graph of high-frequency signal," Clinical EEG Neuroscience, Mar. 17, 2013, vol. 44, pp. 150–156.

25. M. Ahmadlou, H. Adeli, and A. Adeli, "New diagnostic EEG markers of the Alzheimer's disease using visibility graph," J. Neural Transmiss.,2010 Sep., vol. 117, pp. 1099–1109, 2010.

26. B. Luque, L. Lacasa, F. Ballesteros, and J. Luque, "Horizontal visibility graphs: Exact results for random time series," Phys. Rev. E, vol. 80, pp. 046103-1–046103-11, 2009.

27. Guohun Zhu, Yan Li, Peng (Paul) Wen, "Analysis and Classification of Sleep Stages Based on Difference Visibility Graphs From a Single-Channel EEG Signal" IEEE Journal of Biomedical and health informatics, november 2014, vol. 18, no. 6, pp. 1813 – 1821.

28. Benesty, J.INRS-EMT, Jingdong Chen, Yiteng Huang, Estimation of the Coherence Function with the MVDR Approach, Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on, Vol 3, May 2006, pp III.

29. ZinMarLwin, MieMie Thaw, Mental Tasks Classification from Electroencephalogram (EEG) Signal Using Gabor Based Matching Pursuit (MP), International Journal of Computer Science And Technology, Jan-March 2015, Vol. 6, Issue 1, pp. 22-26.

30. Mohammad Shahin Mahanta, Amirhossein S. Aghaei, Konstantinos N. Plataniotis, A Bayes Optimal Matrix-variate LDA for Extraction of Spatio-Spectral Features from EEG Signals, 34th Annual International Conference of the IEEE EMBS , 28 August - 1 September, 2012., pp. 3955-3958.

31. Yazdani A, Ebrahimi T, Hoffmann U, Classification of EEG signals using Dempster Shafer theory and a k-nearest neighbor classifier, Neural Engineering, 2009. NER '09. 4th International IEEE/EMBS Conference on, May 2009, pp. 327-330.

32. 10/20 system Positioning Manual by Trans Cranial Technologies, https://www.trans-cranial.com/local/manuals/10_20_pos_man_v1_0_pdf.pdf

33. http://www.mindmedia.info/CMS2014/products/systems/nexus-10-mkii, NeXus-10 MKII equipment by MINDMEDIA,

34. Missonnier P, Herrmann FR, Richiardi J, Rodriguez C, Deiber MP, Gold G, Giannakopoulos P, Attention-related potentials allow for a highly accurate discrimination of mild cognitive impairment subtypes, Neurodegenerative Diseases, Year 2013,Vol. 12, No. 2, pp. 59-70.

35. Etienne Sallard, Jerome Barral, Camille F. Chavan, Lucas Spierer, Early attentional processes distinguish selective from global motor inhibitory control: An electrical neuroimaging study, Feb. 2014, NeuroImage, pp. 183–189.

36. Thomas S. Bearden, Jeffrey E. Cassisi, and J. Noland White, Electrophysiological Correlates of Vigilance During a Continuous Performance Test in Healthy Adults, Applied Psychophysiology and Biofeedback, September 2004, Vol. 29, No. 3, pp. 175-188.

37. Alicia Callejas, Juan Lupianez, Maria Jesus Funes, Pio Tudela, Modulations among the alerting, orienting and executive control networks, Experimental Brain Research November 2005, Volume 167, Issue 1, pp 27-37.

38. Luis J. Fuentes , Guillermo Campoy, The time course of alerting effect over orienting in the attention network test, Experimental Brain Research, April 2008, Volume 185, Issue 4, pp 667-672.

39. Lawrence A. Farwell, Drew C. Richardson, Graham M. Richardson ,Brain fingerprinting field studies comparing P300-MERMER and P300 brainwave responses in the detection of concealed information, Cognitive Neurodynamics August 2013, Volume 7, Issue 4, pp 263-299.