A Brain Computer Interface for Interactive and Intelligent Image Search and Retrieval

Shitij P. Kumar

Follow this and additional works at: http://scholarworks.rit.edu/theses
A Brain Computer Interface for Interactive and Intelligent Image Search and Retrieval

by

Shitij P. Kumar

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Electrical and Microelectronic Engineering

Supervised by

Professor Dr. Ferat Sahin
Department of Electrical and Microelectronic Engineering
Kate Gleason College of Engineering
Rochester Institute of Technology
Rochester, New York
September 2014

Approved by:

Dr. Ferat Sahin, Professor
Thesis Advisor, Department of Electrical and Microelectronic Engineering

Dr. Gill Tsouri, Associate Professor
Committee Member, Department of Electrical and Microelectronic Engineering

Dr. Sildomar T. Monteiro, Assistant Professor
Committee Member, Department of Electrical and Microelectronic Engineering

Dr. Sohail A. Dianat, Professor
Department Head, Department of Electrical and Microelectronic Engineering
Thesis Release Permission Form

Rochester Institute of Technology
Kate Gleason College of Engineering

Title:

A Brain Computer Interface for Interactive and Intelligent Image Search and Retrieval

I, Shitij P. Kumar, hereby grant permission to the Wallace Memorial Library to reproduce my thesis in whole or part.

__________________________________________
Shitij P. Kumar

__________________________________________
Date
Dedication

I dedicate this work to my loving family who has supported and inspired me in everything that I have done, to my closest of friends who have helped and motivated me, my teachers and mentors who have never failed to guide and inspire me and most of all the ’Big Boss - God’ above who gives me the strength and good health while doing this.
Acknowledgments

Through my years here, there are many I would like to thank.

Above all, I am grateful for the guidance and generosity, that has and continues to be provided by my advisor Dr. Ferat Sahin. His words of inspiration and wisdom have found themselves contributing to the basis of my success.

I would like to convey special thanks to Rochester Institute of Technology for the Effective Access Technologies Grant for this work.

I am grateful to all my peers and fellow research assistants at the Multi-Agent Biorobotics Laboratory. Each of you have contributed your unique roles in the completion of this work.

Ryan, thank you for being an excellent mentor and a great friend. The software architecture of this work was only possible because of your guidance, thanks.

Eyup, it was your initial work and help during the Bio-Robotics Class that inspired this work, thanks.

Dan, our discussions about the human-behavior were very helpful while designing the user interface of the proposed Brain Computer Interface, thanks.

For the rest not mentioned I thank you all.
Abstract

A Brain Computer Interface for Interactive and Intelligent Image Search and Retrieval

Shitij P. Kumar

Supervising Professor: Dr. Ferat Sahin

This research proposes a Brain Computer Interface as an interactive and intelligent Image Search and Retrieval tool that allows users, disabled or otherwise to browse and search for images using brain signals. The proposed BCI system implements decoding the brain state by using a non-invasive electroencephalography (EEG) signals, in combination with machine learning, artificial intelligence and automatic content and similarity analysis of images. The user can spell search queries using a mental typewriter (Hex-O-Speller), and the resulting images from the web search are shown to the user as a Rapid Serial Visual Presentations (RSVP). For each image shown, the EEG response is used by the system to recognize the user’s interests and narrow down the search results. In addition, it also adds more descriptive terms to the search query, and retrieves more specific image search results and repeats the process. As a proof of concept, a prototype system was designed and implemented to test the navigation through the interface and the Hex-o-Speller using an event-related potential(ERP) detection and classification system.

A comparison of different feature extraction methods and classifiers is done to study the detection of event related potentials on a standard data set. The results and challenges faced were noted and analyzed. It elaborates the implementation of the data collection
system for the Brain Computer Interface and discusses the recording of events during the visual stimulus and how they are used for epoching/segmenting the data collected. It also describes how the data is stored during training sessions for the BCI. Description of various visual stimuli used during training is also given. The preliminary results of the real time implementation of the prototype BCI system are measured by the number of times the user/subject was successful in navigating through the interface and spelling the search keyword 'FOX' using the mental-typewriter Hex-O-Speller. Out of ten tries the user/subject was successful six times.
List of Contributions

- Design of a software architecture for the proposed Brain Computer Interface. Generation of visual stimuli like the Hex-O-Speller, rapid Serial Visual Presentation for the proposed Brain Computer Interface and other Hex-O-Speller based stimuli, but instead of characters other information is shown.

- Development of an automated stimulus generation and data collection system for the proposed Brain Computer Interface.

- Comparison of results from the Event Related Potential detection and classification system developed for the proposed BCI, by using three different Feature Extraction Methods: Band Powers, Time Segments and Wavelet Decomposition in combination with four different classification methods: Linear Discriminant Analysis (LDA), optimized LDA, Support Vector Machine (SVM) and Neural Networks (NN).

- Publication

# Contents

Dedication ................................................................. iii  
Acknowledgments ......................................................... iv  
Abstract ................................................................ v  
List of Contributions ....................................................... vii  

## 1 Introduction .......................................................... 1  

## 2 Background Literature ............................................. 4  
  2.1 Stimulus Generation ................................................. 5  
      2.1.1 P300-Matrix Speller ........................................... 6  
      2.1.2 Hex-O-Speller .................................................. 7  
      2.1.3 Rapid Serial Visual Presentation ......................... 7  
      2.1.4 Psychotoolbox 3.0 ............................................ 7  
  2.2 Data Collection and Other Analysis Tools ................... 9  
      2.2.1 Basic structure of a Human Brain ........................ 9  
      2.2.2 Electrode Placements and Data Collection Hardware ... 11  
  2.3 Pre-processing of EEG signals .................................. 13  
      2.3.1 Filtering ....................................................... 14  
      2.3.2 Artifact Removal ............................................. 15  
      2.3.3 Pre-processing of the Corrected Artifact Free Data ... 18  
      2.3.4 Feature Extraction Methods ............................... 20  

## 3 Proposed Method .................................................. 25  
  3.1 View .................................................................. 26  
      3.1.1 Training ....................................................... 27  
      3.1.2 Navigation .................................................... 27  
      3.1.3 Hex-O-Speller ............................................... 27  
      3.1.4 RSVP-Rapid Serial Visual Presentation ................ 27
3.2 Model

3.2.1 Motor Imaginary Movement Algorithm

3.2.2 ERP Detection Algorithm

3.2.3 ERP Yes/No Algorithm

3.2.4 ERP Score Generation Algorithm

3.2.5 Content Based Image Similarity Map Generation Algorithm

3.2.6 Image Queuing/Ranking and Search Query Refinement Algorithm

3.3 Controller

3.3.1 Data Collection, Organization and Timing Synchronization during Visual Stimulus

3.4 ERP Detection System and the Yes/No (2-Class Target/non-Target) Classifier

4 Results

4.1 Results of the ERP Detection and Classification on a Standard Dataset

4.1.1 Data Description and Protocol

4.1.2 Pre-processing: Results of ICA Correction

4.1.3 Pre-processing: Spatial and Temporal Selection

4.1.4 Comparison of Results for ERP detection

4.2 Results from the Prototype BCI System

4.2.1 Experiment Methodology and Setup

4.2.2 Results of Simultaneous Data Collection and Stimulus Generation

5 Conclusions

6 Future Work

Bibliography

A Topographic map of an EEG field as a 2-D circular view

B Receiver Operating Characteristic (ROC) Plots

C Channel Numbers
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Independent Components Rejected for Subject A and Subject B</td>
<td>43</td>
</tr>
<tr>
<td>4.2</td>
<td>Channels Selected for Subject A and Subject B</td>
<td>44</td>
</tr>
<tr>
<td>4.3</td>
<td>Time Segments Selected for Subject A and Subject B (from 1 to 168 data points i.e. 0.7 seconds)</td>
<td>45</td>
</tr>
<tr>
<td>4.4</td>
<td>Result Comparison for p300 Matrix speller Dataset for different Feature Extraction Methods and Classifiers.</td>
<td>47</td>
</tr>
<tr>
<td>4.5</td>
<td>Performance ERP Classification on the P-300 sample data-set</td>
<td>48</td>
</tr>
</tbody>
</table>
# List of Figures

2.1 P300-Matrix Speller [1] ................................. 6  
2.2 Hex-O-Speller - Steps 1 and 2 showing the selection of the character 'A' [2]  7  
2.3 Rapid Serial Visual Presentation- RSVP .......................... 8  
2.4 The Human Brain-Cerebrum [3] ............................... 12  
2.5 EEG Electrode Placement 10-20 System [1] ..................... 13  
2.6 20 channel EEG Electrode Placement [4] ........................ 14  
2.7 Raw signal and its Power Spectrum depicting the DC offset noise and 60 Hz noise in a Channel (screen shot from the Bio-Capture Software) [5]  15  
2.8 ICA decomposition [6] ......................................... 17  
2.9 Summed Projection of Selected Components [6] ................. 18  
2.10 Extracted Epochs representing the Stimulus Onset Asynchrony ................................. 19  
2.11 Time Interval Windows Selected based upon $\text{signed} - r^2$ values ................................. 23  

3.2 Data Collection System Block Diagram ............................ 36  
3.3 Timing Diagram .................................................. 37  
3.4 Block Diagram of the ERP Detection and Classification System ................................. 38  

4.1 ERPs in the P300 Matrix Speller Dataset for Subject A and Subject B (in that order) in the mean signal across all data samples/epochs for target and non-target stimuli. ................................. 42  
4.2 Spatial Selection for Subject A ................................. 52  
4.3 Spatial Selection for Subject B .................................... 53  
4.4 Topographic map of BD as a 2-D circular view for Subject A and B ................................. 54  
4.5 Temporal Selection for Subject A ................................. 55  
4.6 Temporal Selection for Subject B ................................. 56  
4.7 Examples of Training Visual Stimuli ............................... 57  
4.8 Sample Raw Data window and Epoch Extraction of a Single channel ................................. 58  
4.9 ERPs in the BCI for self-collected data in the mean signal across all data samples/epochs for target and non-target stimuli. ................................. 59  

A.1 ICA Components for Subject A ................................. 69
Chapter 1

Introduction

Assistive devices and technologies are essential in enhancing the quality of life for individuals. A lot of research in developing assistive smart devices and technologies for disabled individuals that depend on motor movements, speech, touch and bio-signals has been done. Most of these systems for disabled individuals depend on some residual motor movement or speech. Whereas the BCI systems completely bypass any motor-output by decoding the Brain state of an individual, which can be an emotion, attention i.e. an event related potential (ERP) or an imagined movement [7][8][9][10][11]. These can help disabled individuals that have almost none to some voluntary muscle/movement control, thereby attempting to give some autonomy to individuals by providing the brain with alternate ways of communication. Significant EEG based researches that try to classify and understand various brain states like movement of limbs, imaginary or otherwise, emotions, attention etc. have been done but very few have tried to combine and decode these brain states and apply them in daily used applications like a web search.

This BCI interface will be an attempt to use and augment prior research and understand how it can be applied to make devices (smart-phones, tablets, computers etc.) and technologies smarter and more interactive. Today the dependencies and reliance on smart
devices, computers and applications especially the ones that use the information available on internet are indisputable. The users expect the applications and technologies to be smart and learn from the usage and the choices made by the user. The proposed study will further reinforce this notion, by taking into consideration and understanding their thoughts/actions by decoding their brain state, thus making them more user friendly and intelligent. A lot of studies in the psychophysiology and neuroscience fields have been done to understand the relation of emotions, attention, interest, motor movements etc. to the Brain activity and EEG responses [12][13][14][15][16][17][18][9]. But there are relatively fewer real world applications that use these researches.

Similar studies to the one proposed use EEG signals and image processing to find similar images or specific objects in images [19][20][11][21]. But these only reinforces an image search by leveraging the robust and invariant object recognition capabilities of the human brain. On the other hand, this proposed research attempts to objectively define the subjective nature of a person’s interest in an image. Unlike the above referenced researches, all the images shown would be semantically same, but the choice of the user will be understood not only on recognition but also on emotion and attention. This will result in a more enhanced image browsing, search and retrieval process. Although this proposed research focuses only on the visual stimulus that shows images in a series of rapidly changing presentations (RSVP), it can be further extended to auditory stimuli like music or sounds. The idea of having devices that understand the user, his/her state of mind to a certain extent by decoding the Brain state is quite fascinating. Moreover, this interface will also explore the access and use of existing smart devices for disabled individuals, so that they can have
similar freedom and autonomy as the other capable general population.

The principal challenge with this research is that there is no defined model or relation that relates a user’s interest to the contents of an image. In order to formulate a relation, experiments as proposed in Chapter 6 and [2] can be designed and performed using the proposed BCI. Another challenge is that a large amount of training data is needed to be collected, thereby a longer training time, for getting higher accuracy for a single-trial EEG signal classification [7][8]. Also the EEG recording devices are a bit uncomfortable and less fashionable. However, there have been significant advances in wearable technologies and new devices like the Emotive research headset [22], EEG head bands etc. that collect wireless EEG data are relatively comfortable and fashionable. Nonetheless, this study is step towards better understanding the use of Bio-Signals as a feedback to devices and applications to understand the user better.

The remainder of this document is organized as follows: Chapter 2 describes the background literature and studies used as a reference and part of this this work. Chapter 3 describes the proposed approach for the BCI framework. It describes various aspects of the BCI system, explains the experiment methodology and setup and explains different feature extraction methods and classifiers used for detecting event related potentials on the standard dataset. Preliminary results of the BCI system and comparison of different methods for classifying event related potentials are presented and discussed in Chapter 4. Conclusions are drawn in Chapter 5, and the future work is discussed in Chapter 6.
Chapter 2

Background Literature

Understanding the brain function is a challenge that has inspired a lot of research and insight from researchers and scientists from multiple disciplines. This resulted in the emergence of computational neuroscience, the principle theoretical approach in understanding the workings the nervous system [7]. Part of this research involves decoding the brain-state/human intentions by using the the EEG(electroencephalogram) signals. This branch of the field is greatly influenced by the development of effective communication interfaces connecting the human-brain and the computer [7]. This chapter describes the research, studies and other similar brain-computer interfaces that have been used as a reference and guide for this work.

Generally, in order to develop a BCI (brain computer interface) it involves four major steps. Firstly, the generation of an EEG signal. In order to get a valid EEG signal representing an event of a specific human brain state, the EEG response is evoked by recreating a similar setup which normally would evoke such a response. This is called a ‘Stimulus’. The EEG response can either be evoked using either a visual, auditory, motor, touch or speech stimulus. Here in this research a visually evoked response is used to generate a type of EEG signal called the Event Related Potential (ERP). The EEG response is collected and
stored corresponding to the event that evoked it. Thus the setup and collection of the EEG data is the second step in the development of a BCI. The first two steps occur concurrently, hence synchronization and timing is a crucial part for effecting data collection. Thirdly, the pre-processing and representation of the EEG data. For effective representation of EEG data, the data is cleaned and is represented by a reduced subset of features. The last step involves generation of models that can classify the EEG signals representing different events. Once the EEG response to an event is classified/identified, subsequent action based on the purpose of the BCI can be taken.

2.1 Stimulus Generation

Generally every Brain Computer Interface requires a medium of communication represented by an event that evokes an EEG response. This event can be of the type visual, muscle movement, auditory or touch. Here in this research a visual event is used to generate an EEG response.

In [10] the detection of EEG responses to motor imaginary movements is done. Here the EEG response is evoked by imagining the movement of left and right hand. In an existing software such as the Emotive Education SDK (provided with the Emotive Headset [22]), movement of an animated cube moving left, right, up and down is used to simulate motor imaginary movements from the user/subject. In this research stimulus for generating event related potentials i.e. ERPs that are EEG responses that occur when the subject focuses their attention on symbols or images shown and tries to distinguish between what he/she wants to select. This notion is called the ’oddball paradigm’ [8][7] in which the subject focuses on the target images/symbols and ignores other non-target images.
The most commonly researched example of this stimulus is the attention based mental typewriter. This mental typewriter is used to spell words using the EEG signals. The following subsection explains the visual stimuli used in this research.

2.1.1 P300-Matrix Speller

The so called Matrix Speller consists of a 6x6 character matrix as shown in Figure 2.1 where in characters are arranged within rows and columns [1]. throughout the stimulus generation process the rows and columns are intensified ('flashed') in a random order one after the other. The subject is asked to focuses on the targeted character for the span of stimulus and EEG response corresponding to each intensification is recorded. As the EEG response can be modulated by attention, the response for the targeted character is different from the non-target characters.

Figure 2.1 P300-Matrix Speller [1]
2.1.2 Hex-O-Speller

The Hex-O-Speller is a variant of the Matrix speller [7]. It is represented with six circles placed at the corners of an invisible hexagon as shown in Figure 2.2. Each circle is intensified (‘flashed’) in a random order multiple times. Intensification is realized by up-sizing the circles by 40 percent for 100 milliseconds. In this speller the choice is made in two recurrent steps. At the first step, the group containing the target symbol/character is selected, followed by the selection of the target symbol itself at the second step.

![Hex-O-Speller - Steps 1 and 2 showing the selection of the character ‘A’][2]

2.1.3 Rapid Serial Visual Presentation

Rapid Serial Visual Presentation (RSVP) is an experimental setup used to understand temporal characteristics of attention. In this setup images are flashed or shown for a fixed duration at specific intervals as shown in Figure 2.3.

2.1.4 Psychtoolbox 3.0

The Psychtoolbox 3.0 tool is used to generate the Stimulus in this research [23][24]. It allows control over the graphics for the stimulus. The Stimulus for Hex-O-Speller and
RSVP are generated using the Psychtoolbox. It allows control and flexibility to record the time for events which helps generate decently accurate event markers necessary for data segmentation (epoching or shelving). Every display screen has a display buffer and a 'Vertical Blank Interval' time [25] also known as $V - Blank$, that it takes to completely display a new image. This toolbox gives a two-stage flexibility to change a display on a monitor. The first step is filing the display buffer with the desired image and second is 'Flipping' the screen i.e. clearing the previous display and showing the new display. The refresh rate of a monitor depends on the type of monitor, operating system and the graphics card. This rate remains the same for the stimulus running on a specific system. This refresh rate say $T_{refresh}$, becomes the unit time for change of display on a monitor.

The $V - Blank$ interval for each image varies based upon the size and resolution of the
image. The $V − Blank$ interval is represented as $V − Blank = n \times T_{\text{refresh}}$, where $n$ being the time taken to show the image in the image buffer onto the screen in units $T_{\text{refresh}}$. For all images to be displayed this $V − Blank$ time can be fixed to the maximum possible time taken for the biggest image to be displayed during the stimulus. This results in a uniform screen change rate, thus making the synchronization of stimulus generation and data collection better.

2.2 Data Collection and Other Analysis Tools

Data collection is an integral part of any Brain Computer Interface. Data collection requires proper setup of electrodes and bio-signal collection devices. These devices generally contain buffers and amplifiers to amplify the bio-signal data and transmit it to the computer. It is extremely important for proper synchronization between stimulus generation for the event that evokes the EEG response and its collection, to correctly segregate (also known as epoching or shelving) the EEG response corresponding to the event. An implementation of the simultaneous data collection and stimulus generation tool for the proposed BCI is discussed in detail in Chapter 3.

2.2.1 Basic structure of a Human Brain

In order to understand the EEG signals, a basic understanding of different areas of the human brain is important. The brain is made of three main parts: the forebrain, midbrain, and hindbrain [3]. The forebrain consists of the cerebrum, thalamus, and hypothalamus (part of the limbic system). The midbrain consists of the tectum and tegmentum. The hindbrain is made of the cerebellum, pons and medulla. Often the midbrain, pons, and
medulla are referred to together as the brainstem. Here, the part where the EEG signals are collected from is known as the Cerebrum. The cerebrum or cortex is the largest part of the human brain, associated with higher brain function such as thought and action. The cerebral cortex is divided into four sections, called "lobes": the frontal lobe, parietal lobe, occipital lobe, and temporal lobe. The visual representation of the cortex is shown in Figure 2.4. Each lobe of the brain has a specific function.

- **Frontal Lobe** is associated with reasoning, planning, parts of speech, movement, emotions and problem solving.

- **Parietal Lobe** is associated with movement, orientation, recognition and perception of stimuli.

- **Occipital Lobe** is associated with visual processing.

- **Temporal Lobe** is associated with perception and recognition of auditory stimuli, memory and speech.

The EEG signals or brain waves have specific bands of frequencies that are most active corresponding to a brain state. As explained in [26] EEG has four main bands $delta(\delta)$, $theta(\theta)$, $alpha(\alpha)$ and $beta(\beta)$.

- $beta(\beta)$ waves are small, faster brainwaves (13 Hz - 30 Hz) associated with a state of mental, intellectual activity and outwardly focused concentration. This is a "bright-eyed, alert and focused" state of awareness. This is very wide range of frequency and has been broken down into sub bands of frequency.
- *beta*($\beta_1$) brainwaves (13-15 Hz) are associated with being in a physically relaxed and mentally alert state of mind. These waves are often associated with peak performance training, e.g., professional athletes.

- *beta*($\beta_2$) brainwaves (16-18 Hz) are typically associated with performing mental tasks such as reading, mathematics and problem solving.

- *beta*($\beta_3$) brainwaves (19-26 Hz) are also associated with problem solving and thinking in general, however, there is also some association with worry or anxiety.

- *alpha*($\alpha$) waves are slower, larger brainwaves (8 Hz - 12 Hz) than the beta waves and are associated with the state of relaxation. Their presence represents the brain in a relaxed, somewhat disengaged state, waiting to respond if needed.

- *theta*($\theta$) brainwaves (4 Hz - 8 Hz) represent a daydream-like, rather spacey state of mind that is associated with mental inefficiency. At very slow levels, theta wave activity is associated with a very relaxed state and often represents the twilight zone between wakefulness and sleep.

- *delta*($\delta$) brainwaves (.50 Hz - 3.5 Hz) are the slowest, highest amplitude (magnitude) brainwaves, and are what a person experiences when asleep.

### 2.2.2 Electrode Placements and Data Collection Hardware

The non-invasive EEG signals data collection is done by placing electrodes on the skin. For that various EEG caps, headsets and head bands are available. The EEG caps generally use the standard 10-20 placement system as shown in Figure 2.5. These caps can have 14, 20, 32, 64 or 128 channels depending on the EEG cap. Some headsets like the Emotive
Figure 2.4 The Human Brain-Cerebrum [3]

have 14 channels, and some headbands have only four. For this work an EEG cap with 20 channels is used. Figure 2.6 shows the placement of electrodes. These electrodes through cable are connected to data-collection hardware that generally have two-stage amplifiers as the magnitude of EEG signals ranges in micro volts along with analog-to-digital converters.

There are various data-collection hardware available such as Bio-Radio 150 [5], Bio-Semi [27], B-Alert[28], Emotive head set [22] that collect the data. Here Bio-radio 150 is used. According to the bio-radio manual [5], the data is collected and wirelessly transmitted via blue-tooth in serialized data packets. The Bio-Radio 150 provides libraries so that its functions can be called and the data collected parsed and put into a proper matrix

\[ X_{\text{bioradio}} \] of dimensions \((c \times p)\) where \(c\) is the Number Of Channels/electrodes connected and \(p\) Number of Data Points in the time series of the signal received at each
It also allows the system to configure the data sampling rate and resolution of the data collected. A maximum of eight channels can be connected to a single Bio-radio 150. As more than 8 channels are required for getting sufficient information to represent an Event Related Potential (ERP), programs for recording EEG signals from multiple Bio-Radio 150 modules simultaneously are implemented.

2.3 Pre-processing of EEG signals

The EEG signals recorded using a non-invasive procedure are riddled with noise. Before generation of features that would represent the characteristics of an EEG response, the raw data signal is to be cleaned, corrected and segmented (epoched/shelved). The types of noise present in an EEG response are skin impedance noise, data collection instrument noise, eye
movement and blink artifacts i.e. Electrooculogram (EOG) (if they are not considered as useful information for the objective of the Brain Computer Interface classification), Electrocardiogram (ECG) noise, and other muscle movement noise (voluntary or involuntary).

2.3.1 Filtering

The skin impedance noise can generally be removed by ensuring proper conductivity between the scalp and the electrode. As mentioned in Section 2.2.2 the EEG cap used here has passive electrode i.e. it requires conductive gel to be inserted as cushion between the scalp and the electrodes in the EEG cap. For that the Bio-Capture Software [5] is used along with a testing program implemented in this system to ensure that minimum skin impedance (also known as ‘contact noise’) is present. Generally this noise affects the DC offset i.e. the zero Hz frequency of the raw EEG signal.
Instrument noise present in the raw signal can be either because of the 60 Hz power noise, noise over the wireless communication or if the electrode cable is not connected properly to the Bio-radio 150 (see Figure 2.7). As described in Section 2.2 the main bands of EEG signal generally range from 0.5 Hz to 30 Hz. In order to remove the DC noise and the 60 Hz noise, the raw signal can be filtered using a band-pass filter. Here in this work a 6th order Butterworth band pass filter with cut-off frequencies as 0.1 Hz and 30 Hz is used [8].

**Figure 2.7** Raw signal and its Power Spectrum depicting the DC offset noise and 60 Hz noise in a Channel (screen shot from the Bio-Capture Software) [5]

2.3.2 Artifact Removal

The raw EEG data is corrupted by other bio-signals that are generated involuntarily by the subject during data collection and are undesirable. These signals are called ‘Artifacts’. The biggest contributor of these artifacts are eye movements and blinks. For removal of these artifacts the most commonly used method is Independent Component Analysis as explained in [6] and [29]. The Winsorizing method (also known as ‘Clipping’) is also used to reduce
the effects of large amplitude outliers which are representative of these artifacts. In [8] this method is used to dampen the effects of these artifacts. Both methods are described briefly below:

**Independent Component Analysis (ICA)**

ICA based artifact correction can separate and remove wide variety of artifacts from EEG data by linear decomposition. However, there are certain assumptions that are made [6][29]. They are:

- **(1)**: spatially stable mixtures of activities of temporally independent brain and artifact sources
- **(2)**: summation of potentials from different channels/sensors is linear at the electrodes
- **(3)**: propagation delays from the source of the electrodes is negligible

For analyzing the EEG signals, let $X$ be the input response matrix where

- $X$ of dimensions $(N \times P)$, where $N$ is the number of channels and $P$ is the number of data points.

The ICA finds an unmixing matrix $W$ which linearly decomposes the multi-channel EEG data into sum of temporally independent and spatially fixed components.

$$U = W \cdot X$$

(2.1)

where the rows of $U$ are the time courses of activation of ICA components. The columns of the inverse matrix, $W^{-1}$ gives the relative projection strengths of the respective components.
at each electrode. Columns of $W^{-1}$ are used to determine which components are to be selected. Say a set of $a$ components are selected, then clean data can be represented as

$$CleanData = W^{-1}(; , a) \ast U(a, :)$$

(2.2)

, where the dimensions of each matrices are

- $CleanData : (c \times N)$
- $W^{-1}(; , a) : (c \times \text{length of } a)$
- $U(a, :) : (\text{length of } a \times N)$

Figures 2.8 and Figure 2.9 show the decomposition and summed projection of the selected ICA components as described in [6].

**Figure 2.8 ICA decomposition [6]**

![ICA decomposition](image)

**Winsorizing Data- Clipping**

Winsorizing is a method used for attenuating and reducing the affect of outliers. A percentile of data say $p^{th}$ percentile is to be attenuated. Then for a data vector say of size $N$,  

...
the first and last \((0.5 \times p)^{th}\) percentile of the sorted values of the data vector are replaced by the value nearest to the side of the sorting order.

For a sorted data vector say \(X = X_1, X_2, X_3, X_4, X_5, ..., X_{99}, X_{100}\), 10th percentile win-
sorizing will results in \(X_{clipped} = X_6, X_6, X_6, X_6, X_6, X_7, ..., X_{93}, X_{94}, X_{94}, X_{94}, X_{94}, X_{94}\), where \(X\) and \(X_{clipped}\) have the same length.

### 2.3.3 Pre-processing of the Corrected Artifact Free Data

The corrected data is segmented (epoched/shelved) corresponding to the response of each stimulus. In order to keep the ratio of meaningful signal to that of noise high, multiple trials of the same stimulus are performed and data collected. These trials are averaged in order to attenuate the noise. Each data sample has data points corresponding to time interval \(T_{SOA}\) prior to the display of the stimulus and data points after the display of the stimulus. This signal collected for the time interval \(T_{SOA}\) is called Stimulus Onset Asynchrony (SOA). The mean of the data points corresponding to the time course of \(T_{soa}\) is subtracted from the
rest of the data sample. This is called 'Baseline Correction’. As the data collection for a stimulus is continuous during a trial baseline correction is very important to remove any change of offset due to the response to the previous stimulus in a trial.

Following the guidelines of recording EEG signals and ERP data as given in [30], the EEG signal data is re-referenced and baseline corrected. Re-referring EEG data means that reference channels are to be subtracted from all other channels. Generally the reference channels selected are the signals from mastoids i.e. Channels A1 and A2 as shown in 2.6. Another re-referencing method used is subtracting the mean of all channels at a given data point from all the channels. Here the later is used as the re-referencing procedure.

**Figure 2.10** Extracted Epochs representing the Stimulus Onset Asynchrony
2.3.4 Feature Extraction Methods

After the raw EEG data is corrected, segmented/epoched, re-referenced, averaged and baseline corrected feature extraction is performed. Here event-related potentials (ERPs) in the collected EEG signals are to be identified and classified. The property of an ERP lies in both spatial and temporal domain. Spatial domain is the representation of the signal generated at different channels (electrodes/sensors) and the contribution of each channel to successfully characterize an ERP. Temporal domain represents the generation of ERP after the stimulus is shown. This response time varies from subject to subject. Some subjects might respond faster or slower than other subjects. However, it has been noted that for an average person the ERP response, specifically the P-300 response happens 300 milliseconds after the stimulus event occurs [1][8]. Hence the nomenclature of P-300 Matrix Speller, as it tries to identify this specific data point response in an EEG response to a stimulus.

Spatial- Channel Selection

Not all channels contribute equally to the distinguishable characteristics of an ERP response to target and non-target stimuli. In [8] the channel selection is done by sorting the channels using Bhattacharyya Distance. The efficiency of each channel is measured by the ability to discriminate target and non-target patterns in the training dataset of EEG responses. For that at each data point in the time series of the response a statistical measure i.e. the Bhattacharyya Distance ($BD$) is used between the two respective target and non
target patterns. This real-valued scaler $BD$ can be defined as:

$$BD = \frac{1}{8} (m_1 - m_2)^T \left( \frac{C_1 + C_2}{2} \right)^{-1} (m_1 - m_2) + \frac{1}{2} \ln \left( \frac{\sqrt{C_1} \times \sqrt{C_2}}{\sqrt{|C_1| \times |C_2|}} \right) \quad (2.3)$$

where $\left| . \right|$ denotes the determinant of a matrix, $m_1$ is the mean vector of target pattern signals, $m_2$ is the mean vector of non-target pattern signals, and $C_1$ and $C_2$ are the corresponding co-variance matrices. Results of using this statistical measure on the standard dataset provided in [1] are shown in Chapter 4.

**Temporal- Time Intervals/Window Selection**

Similar to the spatial selection certain time intervals in the EEG response from each selected channels contain more discriminative information for target patterns and non target patterns. This approach is used in [7] to select clusters of timing data points for generation of features. The statistical measure used in [7] is a biserial correlation coefficient $r$. This correlation coefficient is defined as

$$r(x) = \left( \frac{N_1 \times N_2}{N_1 + N_2} \right) \left( \frac{\text{mean}(x_i \mid y_i = \text{target}) - \text{mean}(x_i \mid y_i = \text{non-target})}{\text{std}(x_i)} \right) \quad (2.4)$$

where $N_1$ and $N_2$ are the number of samples for target and non-target patterns, $x_i$ is the data point in the time series at time $t_i$ of the signal for a selected channel, $y_i$ is the classification of the pattern i.e. either target or non-target, $\text{mean}(.)$ is the mean of the vector of the data points across samples and $\text{std}(.)$ is the standard deviation of the vector of data points across all samples, target and non-target. The discriminative measure used is signed $- r^2$ value which is

$$\text{signed} - r^2 := \text{sign}(x) \times r(x)^2 \quad (2.5)$$
Results of using this statistical measure on the standard dataset provided in [1] are shown in Chapter 4.

Following the identification of discriminative spatial channels and temporal data points, features representing them can be generated. The temporal characteristics of the EEG signal can be identified and represented, both in time and frequency domain. Here in this research three different methods for feature selection were used for comparison.

**Band Powers: Sum of Fast Fourier Transform (FFT) Coefficients for specific EEG bands**

This feature extraction method explores the frequency domain features of the signal. For each channel of an EEG response the signal is passed through various band-pass filters with cut-off frequencies representing different bands of the EEG signal (Section 2.2.1) and the corresponding FFT is taken and the absolute values of the FFT coefficients are used to represent each channel. These coefficients are summed at fixed equally spaced intervals and vectorized to represent a feature vector for all channels.

**Procedure 2.1 FFT Feature Extraction**

**Input:** data Sample $x_i$

**Input:** cut-off frequencies for EEG bands

**Input:** Sampling Rate of Data Collection

**Output:** Feature Vector Representing the $x_i$

---

**Sum of Timing Segments/Windows**

This feature extraction method explores the time domain features of the signal. For each channel of an EEG response the discriminating timing intervals are identified, and summed for each channel. For $n$ number of timing intervals identified for $c$ number of channels, a feature vector of size $(n \times c)$ is generated for each data sample. Let a data sample be
• $x_i$ of dimensions $(c \times p)$ for $i \in [1, N_1 + N_2]$, where $c$ is the Number of selected channels and $p$ is the Number of data points in the signal and $N_1$ and $N_2$ are the number of samples in the target and non-target pattern signals respectively.

• Let $\tau_{ijk}$ represent a set of time interval where $j \in [1, c]$ and $k \in [1, m]$, and $m$ is the Number of Time Intervals identified. Figure 2.11 shows an example of timing intervals being selected.

• Let $X$ be the feature set representing $x$. Then $X_i = \{X_{ijk} \forall j \in [1, c], k \in [1, m]\}$, where $X_{ijk}$ is the sum of all data points in the time interval $\tau_{ijk}$ for channel $j$ of data sample $i$.

• $X_{ijk} = \sum \tau_k x_{ij}$, where $x_{ij}$ is the signal of data sample $i \in [1, N_1 + N_2]$ from channel $j \in [1, c]$, and $\tau_k$ is the $k^{th} \in [1, m]$ time interval.

**Figure 2.11** Time Interval Windows Selected based upon signed $- r^2$ values
Wavelet Decomposition

This feature extraction method explores both the time and frequency domain features of the signal. Wavelet analysis can be performed either in continuous mode (CWT) or in discrete mode (DWT). Discrete Wavelet Transforms have been used in [8] for representing the EEG data sample. In [8] the author performs wavelet decomposition for each channel using the Daubechies family of wavelets up to six levels and records the approximation and detail coefficients for each level. These coefficients are then used to represent the feature vector for an EEG data sample.
Chapter 3

Proposed Method

The proposed BCI [2] allows the user to navigate, and spell search queries in the user interface by detecting ERPs and/or motor imaginary movements. In order to perform a web image search, an initial search query spelled by the user using a Hex-O-Speller [7] is used and the resulting images are retrieved. A content based image similarity map is generated that defines the relation between all the images with shape, color, edge, texture and quality as the criteria (see Section 3.2). These images are then displayed in a Rapid Serial Visual Presentation (RSVP) and EEG response is recorded corresponding to each image shown. This data is then passed to the ERP interest score generator that takes into account the interest/attention shown in an image by the user. According to studies an interest/attention can be a result of recognition, emotion or both [14][15][17] and can be measured by analyzing the EEG spatio-temporal responses and their ERP components [12][13] elicited by the images. By combining the EEG score and the similarity map by using an approach similar to [20][11], the images are ranked and selected. The names of the images as obtained from the web generally contain more descriptive terms than the initial query and using that, the search query is refined in order to retrieve more pertaining images, thereby narrowing the relevant image search results. After the set of images that
interest the user have been identified, the user can choose the image or sub-set of images that he/she likes the most.

The proposed BCI system is modeled as a Model View Controller (MVC) architecture as shown in Figure 3.1. The working of the system is described in detail as follows:

### 3.1 View

The View controls the display to the user, hence is responsible for the user interface of the system that generates the visual stimulus with the pertaining information (explained in detail below) for the user to choose. The visual stimulus used is a hexagonal placement of six circles that contain information, which are intensified by up-sizing in a random order for a short amount of time [7]. There are four different types of visual stimuli; Training,
Navigation, Hex-O-Speller and RSVP. Their function is explained in detail below.

### 3.1.1 Training

The Training generates different visual stimuli for ERP detection, motor imaginary movements and eye movements during the training of the system for a specific user.

### 3.1.2 Navigation

The Navigation displays different options for easy navigation through the interface either by using ERP detection or through motor imaginary movements.

### 3.1.3 Hex-O-Speller

The Hex-O-Speller is used to type search queries which are then passed to the Controller for image retrieval from the web (see Section 2.1.2).

### 3.1.4 RSVP-Rapid Serial Visual Presentation

After the images are retrieved and processed these images are shown to the user in a rapid serial visual presentation (RSVP) (see Section 2.1.3).

The stimulus generator handles the change and the placement of proper information in all of these displays. The rate of change and timing of the visual stimulus is synchronized with the Controller (Section 3.3). The Stimulus Generator also provides the necessary information about the visual stimulus to the Controller for tracking and organizing the data representing the EEG response corresponding to the visual stimulus.

### 3.2 Model

The Model contains different classifiers giving the classification needed as per the stimuli provided and thus updating the View. It provides the information needed by the stimulus
generator to change and update the display based on the EEG response corresponding to the previous stimuli. The model contains sub-systems for classifying motor imaginary movements, detecting ERPs, generating ERP scores, content based image similarity maps and ranking/queuing the images and also refining the search queries. There are two different types of ERP classifiers; one classifies a Target/Non-target i.e. a Yes/No choice, used in Navigation and the Hex-O-Speller; and the other generates an ERP interest score for the images displayed during the RSVP. The score along with the similarity map is passed to the Image Queuing/Ranking sub-system that determines the subset of images that the user has shown interest in. Using the names of the images, it finds similar keywords and adds/refines the initial search query.

3.2.1 Motor Imaginary Movement Algorithm

This sub-system generates the training model for classifying left,right,up and down movements as shown in [9][10]. Using this classification the user can control and navigate through the interface, emulating the control paradigm for a Mouse or a Joystick. The Motor Imaginary Movement Classifier is an alternate way of controlling the interface other than the ERPs.

3.2.2 ERP Detection Algorithm

This sub-system identifies the event related potentials (ERPs) in a EEG response. It filters, removes the artifacts and noise using Independent Component Analysis (ICA) and generates the features for representing the components of ERPs (see Section 2.3.2).
3.2.3 **ERP Yes/No Algorithm**

This subsystem generates the training model for classifying the ERPs as Target and Non Target, thus emulating a ‘Yes/No’ choice from the user. This classification is used in Navigation to emulate a click, and in Hex-O-Speller to select the alphabets to type the search query [7].

3.2.4 **ERP Score Generation Algorithm**

This subsystem uses the EEG response to define an objective measure for the interest shown by the user in an image during the Rapid Serial Visual Presentation (RSVP). It generates the features needed by the Image Queuing/Ranking subsystem to select the subset of images representing the user’s interest.

3.2.5 **Content Based Image Similarity Map Generation Algorithm**

This subsystem compares and analyzes all the images retrieved from the web in terms of similar content. The criteria of similar content can be both local and global features such as shapes, color, texture, edges or any other information that can be retrieved from the images [31][32][33][34][11]. References [32][11] show how a graph based representation can be achieved based on the similarity of these images.

3.2.6 **Image Queuing/Ranking and Search Query Refinement Algorithm**

This subsystem combines the results of Similarity map generation and the ERP score and selects the subset of images representing the user’s interest. The metadata i.e. the name of an image from the web can be used to refine the search query by matching similar keywords.
3.3 Controller

The Controller acts as an intermediary between the View and the Model. It collects the EEG data and also synchronizes the rate of change of the display in the View to that of the data collection. It also organizes the data corresponding to each stimulus and retrieves images from the web, thus providing the necessary data to the Model and also chooses the sub-system or classifier to be used. The data is represented as a structure that contains the information of the visual stimuli and the data collected corresponding to it. This is explained in detail in the following Section 3.3.1.

3.3.1 Data Collection, Organization and Timing Synchronization during Visual Stimulus

The data collection is needed to be done simultaneously and synchronously with the visual stimulus shown. During training or testing the data is to be collected and stored along with the information corresponding to the visual stimulus. Moreover, the information of event timings; here the intensification of circles containing the pertinent information in a visual stimulus, is also stored along with the data. The event information is necessary in order to extract the EEG response corresponding to each event in a visual stimulus. The data sample for each stimulus is stored and organized in a data structure.

The Block Diagram representing the flow of information and working of the View and Controller for data collection during Training is shown in Figure 3.2. The timing diagram representing the tasks performed during Data Collection is shown in Figure 3.3. Task1 represents the buffering of data, Task2 shows the display of the visual stimulus as controlled by the Stimulus Generator (see Section 3.1), Task3 records the events taking place
during a visual stimulus and Task 4 organizes and stores the data collected for the visual stimulus. The workings of the Controller subsystems i.e. the data buffering, organization and synchronization is discussed in detail below:

**Data Collection - Buffering**

This sub-system interfaces with the data-collection hardware i.e. Bio-Radio 150 [5] and reads the bio-signal data. This data is stored in a buffer of a fixed size. The size of the buffer can be controlled and is determined by the time taken by a single visual stimulus.

- The buffer, say rawWindow is a defined matrix of size \((c \times p)\), where \(c\) are the Number of Channels and \(p\) is the Collection Interval, i.e.

- Collection Interval \(p = F_s \times T_{collection}\), where \(F_s\) is the Sampling Rate and the \(T_{collection}\) is the duration of collection. Here in the given system,

- \(F_s = 960\) samples/second, \(T_{collection} = 3\) seconds. Therefore \(p = 960 \times 3 = 2880\) data points.

The buffer is constantly updated and the data appended as long as the entire BCI system is running. The rawWindow is transferred to the Data Organizer (explained in detail below) at the end of the stimulus when the Time Synchronizer invokes the Read Event(explained below). Hence, the continuous bio-signal response for the entire duration of the visual stimulus is recorded and transferred to the Data Organizer. This process is represented as Task 1 in the Timing Diagram(refer Figure 3.3).
Timing Synchronization

This sub-system records the times of the events in the visual stimulus. These times of the events are crucial for epoching/segmenting the bio-signal data collected for the entire duration of the stimulus. The timings and the working of different tasks performed by the subsystems are represented in Figure 3.3. This process is represented as Task3 in the Timing Diagram.

There are eight different events tracked during a visual stimulus. The significance of each event is explained as follows:

- **Start Event** This event signifies the start of the visual stimulus. This event is important as the data from the start event to the beginning of the first up-sizing/intensification of the circle i.e. Event 1 in the rawWindow can be used for baseline correction [30].

  In the Timing Diagram (refer Figure 3.3) the times shown for a single visual stimulus (Task2) span the DurationOfCollection i.e. 3seconds, and the Start Event happens at 0.4seconds. For a duration of 0.2seconds after the Start Event there is no up-sizing/intensification in the visual stimulus.

- **Events 1 to 6** These Events represent the up-sizing/intensification of the six circles. These event times are important for identifying the epochs i.e. the EEG response for each up-sizing circle that may contain event related potentials (ERP). The stimulus onset i.e. the time between each upsizing/intensification here is 0.25seconds. The up-sizing lasts for 0.1seconds and 0.15seconds is the time taken before the next up-sizing. The stimulus onset can be controlled and modified as per the need of the
experiment during training or the reaction time of the subject.

- **Stop Event** This event signifies the end of the visual stimulus. This event happens 0.8 seconds after the last up-sizing of circles. In order to include the epoch response of the last event, delay of 0.8 seconds is given. In the Timing Diagram (refer Figure 3.3) the time for the stop event shown is at 2.9 seconds. Immediately after, the Time Synchronizer triggers the Read Event asking the Data Collector to transfer the buffer to the Data Organizer.

- **Read Event** This event signifies the exact time of transfer of the raw data from the buffer to the Data Organizer. The reason for separating the Stop Event and the Read Event is to compensate for any processing delay that could happen, as these processes are controlled by different subsystems. In the Timing Diagram (refer Figure 3.3) the time for the read event shown is at 3 seconds, i.e. the last data point of the rawWindow. This is used as the reference for the times of the previous events.

**Data Organization**

This sub-system organizes each data sample for a single visual stimulus and stores it into a data structure. The fields of this structure are described as follows:

- **Order** This field contains information of the sequence in which the six circles in a visual stimulus are up-sized. In case of RSVP it has the order/queue of the images.

- **Association** This field contains the information to be displayed in the visual stimulus. This information can be the options during Navigation, Alphabets in Hex-O-Speller, image pointers during RSVP or sequences of a Training trial. During training the
association is predetermined.

- **Raw EEG Data** contains the raw data received from the data collection system. Epoching/Segmentation is performed on this data and is stored in a 4-Dimensional matrix say `rawData`.

  - `rawData` is of dimensions \((c \times w \times e \times t)\), where \(c\) is the Number of Channels, \(w\) is the Size of Epoch Window, \(e\) is the Number of Events, and \(t\) is the Number of Trials.

  - **Number of Channels** \(c\) : number of locations of the electrodes from where the data is recorded.

  - **Size of Epoch Window** \(w\) : The Epoch Window represents the set of data points to be selected from the `rawWindow` representing each epoch i.e. the range of data points that will contain the response for each up-sizing of circle in a visual stimulus. The position of the epoch window is determined by the Events i.e Event 1 to 6 (see Figure 4.8). Here the size is for a duration is \(1\) seconds \(= 960\) datapoints. The Epoch Window is placed such that all the data points representing \(0.2\) seconds prior and \(0.8\) seconds after the Events \(i : (i = 1 to 6)\).

  - **Number of Events** \(e\) : for 6 events, as there are 6 circles up-sizing in a single visual stimulus.

  - **Number of Trials** \(t\) : In case of ERP recording, multiple trials of the same visual stimulus are taken and averaged after artifact removal to suppress noise \([30][8]\).
• **Trigger information** contains information of the time of the events in a visual stimulus (see Figure 3.3).

• **Label** represents the result of a classifier on the data. For Navigation or Hex-O-Speller, it is labeled as Up/Down/Left/Right or Yes/No depending on the classifier used. In case of RSVP, an ERP score is assigned. This field is predetermined during training.

• **Inference** Based on the Labels determined after the classification results from the Model, the action or the choice of the user is inferred. For example, in case of a Hex-O-Speller or Navigation, the choice of the user; and during RSVP the ranking of the image. The Stimulus Generator generates the next visual Stimulus based on this inference. During Training this attribute is predetermined, and is used to instruct the subject what option to choose (as shown in Figure 4.7).

The Data Organizer performs different tasks during Training and Testing modes of the system. During Training the data structures are stores and written to a file corresponding to each visual stimulus. Where as during Testing the data structure is passed to the Model for classification. During Training the Labels and Inference are pre-determined and the Model reads these files and generates the training model needed for classification.

### 3.4 ERP Detection System and the Yes/No (2-Class Target/non-Target) Classifier

In order to test and understand the behavior of event related potentials the, ERP Detection System and the Yes/No (2-Class Target/non-Target) classifier were implemented for a
standardized data set of P-300 Speller responses [1]. Figure 3.4 shows the block diagram of the implemented system. The data from the data set is first segmented/epoched into a 4-Dimensional matrix, say $D$.

- $D$ is of size $(p \times c \times r \times t)$, where $p$ is the Number of data points in a signal per channel i.e. time/sampling rate, $c$ is the number of channels, $r$ is the number of responses depending on the stimuli (P-300 has twelve, while a Hex-O-Speller has six, see Section 2.1) and $t$ is the number of trials (see Section 3.3.1).

Each data sample for a response and trial is filtered using a band pass filter with 0.1 Hz and 30 Hz cut-off frequencies. This data sample is then winsorized with a 90th percentile. The training model is generated for four different classifiers: Linear Discriminant Analysis
(LDA), optimized LDA, Support Vector Machine (SVM) and Neural Networks (NN), for three different types of features: Band Power (BP), Timing Intervals/Segments (TS) and Wavelet Decomposition (WD). This results in 12 different classifiers for a sample data set.

After the data is shelved in a structure Target and Non-Target samples are extracted. For the all the data samples for a character , the data matrix $D_k$ for $k^{th}$ character, is concatenated across all channels, resulting in a matrix say $M$.

- $M$ is of size $(c \times m)$, where $m$ is the length of the concatenated points, which is $m = p \times r \times t$.

Following this ICA is performed on $M_{avg}$ where

- $M_{avg} = mean(M_k)$ : $\forall$ Characters, resulting in the unmixing matrix $W_{avg}$. 

---

**Figure 3.3 Timing Diagram**

![Timing Diagram](image)
Then a subset of all

- $M_j : j \in \text{Random Set of Characters}$, is chosen and ICA is performed resulting in $W_j$.

For all selected $W_j$ correlation with $W_{avg}$ is performed and the components and component selected using a threshold. Following which the components that are to be kept are identified and the resulting truncated $W_{avg}^{truncated}$ is stored, and as the ICA parameter and the rest of the data is corrected as explained in Section 2.8. the corrected data is the re-referenced, averaged and base-line corrected. Spatial and temporal features are identified, features extracted, features reduced using Principal Component Analysis and the classifier models are
generated. The results of this system are presented in the following Chapter 4.
Chapter 4

Results

This chapter discusses the results obtained from the ERP detection and the Yes/No (2-Class Target/non-Target) classifiers described in Section 3.4. It explains the experiment setup and methodology used in the collection of data. It describes the training stimulus and navigation through the BCI. It also discusses the results obtained from simultaneous data collection and stimulus generation for the implemented prototype BCI system.

4.1 Results of the ERP Detection and Classification on a Standard Dataset

The ERP detection and the classifier were tested on the P300 Matrix Speller standard dataset from the BCI Competition 2005 [1] for two subjects. The P300 Matrix Speller stimulus results in finding two target responses from the twelve responses to the stimulus. Each target response corresponds to a row and column in the matrix of symbols, thus identifying the targeted symbol. This setup is different from the Hex-O-Speller, as it requires one target response for 6 stimuli. In P300 speller the goal is to identify 1 character out of 36 characters from the 6x6 matrix, whereas for Hex-O-Speller the goal is to identify group of characters or character out of 6 choices, depending on the step (Section 2.1.2).
4.1.1 Data Description and Protocol

The P300 speller paradigm (Section 2.1.1) of BCI Competition 2005 [1] displays a 6 X 6 matrix of characters to each subject. Each row and column in the display are flashed at random, and the subject’s task is to focus on characters in a given word, one character at a time. Two out of the 12 flashed rows or columns contain the desired letter (intersecting row or column). An ERP is generated when either the target row or column is flashed.

The dataset is recorded for two subjects, ‘Subject A’ and ‘Subject B’. For each subject EEG responses from 64 channels/sensors are collected at the rate of $F_s$ is 240 samples/second for 15 trials per character. The EEG electrodes are placed using the 10-20 electrode placement system as shown in Figure 2.6. The recorded EEG is band-passed from 0.1Hz to 60Hz. As the ERP relevant information can be found in the frequency range of less than 30Hz, this data is band passed further during epoching/segmentation. The training and testing datasets contain 85 and 100 characters, respectively for each subject.

Hence the total number of data samples/epochs, say $N_{train}$ in training is

- $N_{train} = k \times r \times t$, where $k$ is the Number of Characters $r$ is the Number Of Responses in 1 trial and $t$ is the Number of Trials.

- Hence, $N_{train} = 85 \times 12 \times 15 = 15300$; and testing, say $N_{test} = 100 \times 12 \times 15 = 18000$.

Figure 4.1 gives an idea of how the ERP signal looks like and behaves over the time course for a data sample/epoch. The graph in Figure 4.1 shows the time course of the ERPs at channel Cz and the average of channels Po7 and Po8. The baseline is taken from -100ms to 0ms as the Stimulus Onset Asynchrony i.e the time between each intensification (row or
Figure 4.1 ERPs in the P300 Matrix Speller Dataset for Subject A and Subject B (in that order) in the mean signal across all data samples/epochs for target and non-target stimuli.

column stimulus) is 100ms.
4.1.2 Pre-processing: Results of ICA Correction

Each ICA component is a representation of the contribution of all the channels. The artifacts that are introduced in the signal are highly correlated and generally contributed from the extreme frontal and temporal region channels of the brain. In the Figure A.1 and A.2 all the 64 components of the signal are shown for Subject A and Subject B respectively. For Subject A, 42 independent components were selected, and for Subject B, 54 components were selected. The threshold for correlation for identifying the artifact channels was chosen as 0.5 where 0 and 1 represent highly uncorrelated and correlated respectively. Table 4.1 shows the list of rejected components. In order to perform the ICA the FastICA algorithm [35] was used.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Independent Components Rejected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject A</td>
<td>3, 5, 7, 9, 10, 11, 15, 18, 20, 22, 26, 27, 28, 33, 34, 36, 37, 50, 57, 59, 60, 62</td>
</tr>
<tr>
<td>Subject B</td>
<td>3, 20, 32, 34, 39, 43, 49, 50, 54, 62</td>
</tr>
</tbody>
</table>

It can be seen that Subject B data is less riddled with noise compared to Subject A. This can also be observed in the final classification results in Table 4.4.

4.1.3 Pre-processing: Spatial and Temporal Selection

For feature extraction, spatial and temporal discriminating criterion is identified using Bhattacharyya Distance (BD) and signed-$r^2$ respectively (see Section 2.3).

The results for the channel selection using BD are shown in Figure 4.2 and Figure 4.3. In Figure 4.2 and Figure 4.3, the figure on the left column shows the discriminating behavior for each data point as a 2-D color mapped image, and the bar graph on the right shows the
mean BD for the entire time course of a data sample/epoch. For the 2D-color map the x-axis is the index of each channel as shown in Appendix C, and the y-axis is the data points ranging from 1 to 168, that correspond to 0ms- 700ms given \( F_s = 240 \) samples/second ,the data sampling rate . Table 4.2 lists the selected 32 channels sorted in descending order of the BD. Figure 4.4 shows the topographic map for the channels selected for Subject A and B. It can be seen that the most significant channels are present in the parietal region.

The results for the selection of time segments using signed-r\(^2\) values are shown in Figure 4.5 and Figure 4.6. In Figure 4.5 and Figure 4.6, figure on the left column shows the discriminating behavior for each channel as a 2-D color mapped image, and the bar graph on the right shows the mean signed-r\(^2\) for the selected channels shown in Table 4.2. For the 2D-color map the y-axis is the index of each channel as shown in Appendix C, and the x-axis is the data points ranging from 1 to 168, that correspond to 0ms- 700ms given \( F_s = 240 \) samples/second ,the data sampling rate. Table 4.3 lists the selected 8 time ranges. The intervals of the time segments are selected by finding the peaks and valleys in the averaged signed-r\(^2\) values for all the selected channels. It can be seen that the most prominent time segment lie between 0.2 second (48 data point)and 0.6 second (144 data point).

<table>
<thead>
<tr>
<th>Table 4.2 Channels Selected for Subject A and Subject B</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Sorted in descending order based on their discriminative characteristics)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject</th>
<th>Channels Selected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject A</td>
<td>{Po7..Fc1..Iz..O1..P7..Cz..Fc2..F4..Oz..T9..Tp7..Fz..P1..}</td>
</tr>
<tr>
<td></td>
<td>{Po8..C2..O2..F1..F2..Af3..P5..Cp2..Cp1..C4..Po3..P3..Tp8..T10..P8..Fc4..}</td>
</tr>
<tr>
<td>Subject B</td>
<td>{Po8..Po7..Cpz..Cz..O1..P04..P8..C2..P6..Fc2..O2..Fc2..Fc2..P03..Iz..C1..P5..}</td>
</tr>
<tr>
<td></td>
<td>{Cp2..F7..Fc4..Oz..Fc1..Cp1..F2..T10..Tp8..Pz..Fc6..F4..P1..Fz..T8..C4}</td>
</tr>
</tbody>
</table>
Table 4.3 Time Segments Selected for Subject A and Subject B (from 1 to 168 data points i.e. 0.7 seconds)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Time Segments/Intervals Selected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject A</td>
<td>( \tau_1 = 12 ) to ( 28 )</td>
</tr>
<tr>
<td></td>
<td>( \tau_2 = 30 ) to ( 52 )</td>
</tr>
<tr>
<td></td>
<td>( \tau_3 = 53 ) to ( 70 )</td>
</tr>
<tr>
<td></td>
<td>( \tau_4 = 70 ) to ( 90 )</td>
</tr>
<tr>
<td></td>
<td>( \tau_5 = 92 ) to ( 106 )</td>
</tr>
<tr>
<td></td>
<td>( \tau_6 = 106 ) to ( 125 )</td>
</tr>
<tr>
<td></td>
<td>( \tau_7 = 126 ) to ( 136 )</td>
</tr>
<tr>
<td></td>
<td>( \tau_8 = 136 ) to ( 168 )</td>
</tr>
<tr>
<td>Subject B</td>
<td>( \tau_1 = 12 ) to ( 29 )</td>
</tr>
<tr>
<td></td>
<td>( \tau_2 = 30 ) to ( 39 )</td>
</tr>
<tr>
<td></td>
<td>( \tau_3 = 40 ) to ( 63 )</td>
</tr>
<tr>
<td></td>
<td>( \tau_4 = 63 ) to ( 77 )</td>
</tr>
<tr>
<td></td>
<td>( \tau_5 = 78 ) to ( 108 )</td>
</tr>
<tr>
<td></td>
<td>( \tau_6 = 109 ) to ( 133 )</td>
</tr>
<tr>
<td></td>
<td>( \tau_7 = 134 ) to ( 150 )</td>
</tr>
<tr>
<td></td>
<td>( \tau_8 = 151 ) to ( 168 )</td>
</tr>
</tbody>
</table>

4.1.4 Comparison of Results for ERP detection

Three different feature extraction methods were used to generate features for the given P300 sample dataset. They were Band Powers (BP) i.e. sum of FFT coefficients for the EEG bands, sum of data points over the selected Time Segments/Intervals (TS) and Wavelet Decomposition (WD) (Section 2.3). Each of these features were classified using three different classifiers: Linear Discriminant Analysis (LDA), Support Vector Machine (SVM) and Neural Networks (NN). An optimized variation of LDA (oLDA) was also used as a classifier. The objective of the optimized LDA was to find the best subset of samples for each class in training using the scatter matrices as the optimization criterion (see Section 3.4). A five-fold cross-validation with 50% of the training data set was performed and the mean performance parameters were noted. Table 4.4 list all the combinations of the results for both subjects. The performance parameters with the target (positive) and non-target
(negative) class are true positives (TP), false positives (FP), false negatives (FN) and true negatives (TN). The Receiver Operating Characteristics for LDA and NN are shown in Appendix B.1. Other derived performance parameters were calculated as follows [8]:

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.1}
\]

\[
\text{TruePositiveRate(TPR)} = \frac{TP}{TP + FN} \tag{4.2}
\]

\[
\text{FalsePositiveRate(FPR)} = \frac{FP}{TN + FP} \tag{4.3}
\]

\[
\text{TrueDetectionRate(TDR)} = \frac{TP}{TP + FP + FN} \tag{4.4}
\]

The training models subsequently generated were used to classify the test data sets of Subjects A and B that contained data epochs for 100 characters. The second last column of Table 4.4 lists the number of characters accurately detected. As there are two ERP responses for every 12 response correct classification of two responses simultaneously was requires to accurately detect the target character. The last column lists the number of correctly detected row or column. If the row was detected correctly, 0.5 was added to the total of the target detected.

### 4.2 Results from the Prototype BCI System

As described in Chapter 3 a part of proposed BCI system was implemented. The Stimulus Generation controlled by the 'View', timing synchronization for data collection and stimulus generation controlled by the 'Controller' and ERP detection and the Yes/No(2 class target and non-target) classifier subsystems for the 'Model' were implemented. In order to test the workings of the 'Controller', initially electromyogram (EMG) signals from
Table 4.4 Result Comparison for p300 Matrix speller Dataset for different Feature Extraction Methods and Classifiers.

<table>
<thead>
<tr>
<th>Sub</th>
<th>FtExt.</th>
<th>Clf.</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>Acc</th>
<th>TPR</th>
<th>TDR</th>
<th>Char</th>
<th>Tar</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP</td>
<td>LDA</td>
<td>83</td>
<td>39</td>
<td>87</td>
<td>811</td>
<td>811</td>
<td>0.876</td>
<td>0.49</td>
<td>0.39</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>oLDA</td>
<td>82</td>
<td>44</td>
<td>88</td>
<td>806</td>
<td>806</td>
<td>0.870</td>
<td>0.48</td>
<td>0.39</td>
<td>13</td>
<td>31.5</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>97</td>
<td>86</td>
<td>73</td>
<td>764</td>
<td>764</td>
<td>0.844</td>
<td>0.57</td>
<td>0.38</td>
<td>5</td>
<td>32.5</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>103</td>
<td>20</td>
<td>67</td>
<td>830</td>
<td>830</td>
<td>0.91</td>
<td>0.61</td>
<td>0.55</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>A</td>
<td>LDA</td>
<td>128</td>
<td>32</td>
<td>42</td>
<td>818</td>
<td>818</td>
<td>0.92</td>
<td>0.76</td>
<td>0.65</td>
<td>79</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>oLDA</td>
<td>125</td>
<td>54</td>
<td>45</td>
<td>796</td>
<td>796</td>
<td>0.91</td>
<td>0.74</td>
<td>0.57</td>
<td>48</td>
<td>69.5</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>133</td>
<td>31</td>
<td>37</td>
<td>819</td>
<td>819</td>
<td>0.93</td>
<td>0.78</td>
<td>0.66</td>
<td>32</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>141</td>
<td>15</td>
<td>29</td>
<td>835</td>
<td>835</td>
<td>0.95</td>
<td>0.83</td>
<td>0.77</td>
<td>67</td>
<td>67</td>
</tr>
<tr>
<td>TS</td>
<td>LDA</td>
<td>126</td>
<td>44</td>
<td>44</td>
<td>806</td>
<td>806</td>
<td>0.91</td>
<td>0.75</td>
<td>0.59</td>
<td>78</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>oLDA</td>
<td>122</td>
<td>51</td>
<td>48</td>
<td>799</td>
<td>799</td>
<td>0.90</td>
<td>0.72</td>
<td>0.56</td>
<td>60</td>
<td>64.5</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>151</td>
<td>84</td>
<td>19</td>
<td>766</td>
<td>766</td>
<td>0.89</td>
<td>0.89</td>
<td>0.59</td>
<td>46</td>
<td>68.5</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>130</td>
<td>13</td>
<td>40</td>
<td>837</td>
<td>837</td>
<td>0.94</td>
<td>0.77</td>
<td>0.71</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>WD</td>
<td>LDA</td>
<td>148</td>
<td>4</td>
<td>22</td>
<td>846</td>
<td>846</td>
<td>0.97</td>
<td>0.88</td>
<td>0.86</td>
<td>91</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>oLDA</td>
<td>143</td>
<td>40</td>
<td>27</td>
<td>810</td>
<td>810</td>
<td>0.94</td>
<td>0.85</td>
<td>0.70</td>
<td>72</td>
<td>76.5</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>142</td>
<td>6</td>
<td>28</td>
<td>844</td>
<td>844</td>
<td>0.97</td>
<td>0.83</td>
<td>0.80</td>
<td>67</td>
<td>82.5</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>160</td>
<td>3</td>
<td>10</td>
<td>847</td>
<td>847</td>
<td>0.94</td>
<td>0.95</td>
<td>0.88</td>
<td>88</td>
<td>88</td>
</tr>
<tr>
<td>B</td>
<td>LDA</td>
<td>145</td>
<td>18</td>
<td>25</td>
<td>832</td>
<td>832</td>
<td>0.95</td>
<td>0.85</td>
<td>0.77</td>
<td>85</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>oLDA</td>
<td>150</td>
<td>25</td>
<td>20</td>
<td>825</td>
<td>825</td>
<td>0.95</td>
<td>0.88</td>
<td>0.77</td>
<td>65</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>153</td>
<td>46</td>
<td>17</td>
<td>804</td>
<td>804</td>
<td>0.93</td>
<td>0.9</td>
<td>0.70</td>
<td>48</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>155</td>
<td>7</td>
<td>15</td>
<td>843</td>
<td>843</td>
<td>0.97</td>
<td>0.91</td>
<td>0.88</td>
<td>70</td>
<td>70</td>
</tr>
</tbody>
</table>

*where Sub - Subject, FtExt.-Feature Extraction Method, Clf.-Classifier, TP- average true positives, FP - average false positives, FN - average false negatives, TN- average true negatives, Acc- mean Accuracy, TPR- average true positive rate, TDR - true detection rate, Char- characters detected correctly, Tar - targets detected correctly
BP - Band Powers, TS- Time Segments, WD - Wavelet Decomposition, LDA- Linear discriminant Analysis, oLDA - optimized Linear discriminant Analysis, SVM -Support Vector Machine and NN -Neural Network.

the wrist are used. This is followed by collecting EEG signals from 10 channels using the electrode placement shown in Figure 2.6 is done. A classifier model using Time Segments/Intervals (TS) as the feature extraction method and Linear Discriminant Analysis (LDA) as the classifier is generated. Navigation though the interface is tested in real time.
Table 4.5 Performance ERP Classification on the P-300 sample data-set

<table>
<thead>
<tr>
<th>Algorithm Used</th>
<th>Mean Classification Accuracy for Subject A and B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algorithm as used in [8]</td>
<td>93-98%</td>
</tr>
<tr>
<td>Winner of the BCI Competition III [1]</td>
<td>96.5%</td>
</tr>
<tr>
<td>Best result of the Algorithm used here (Time Segments and LDA)</td>
<td>85%</td>
</tr>
</tbody>
</table>

* In [8], the author compares the results by using different sets of EEG channels and features representing the ERPs to identify the ones giving the highest accuracy. Binary Particle Swarm optimization is used to search for the best combination of wavelet features and channels.

4.2.1 Experiment Methodology and Setup

The EEG data is collected using a standard 10/20 placement EEG cap and multiple Bio-Radio 150 data collection systems [5]. Along with EEG data, EOG (electrooculogram) can also be collected for artifact and noise removal. The ERP training of the system involves collection of EEG data samples for the visual stimuli with feedback from the user (if not physically disabled) as he/she concentrates on a choice and reaffirms it using the Spacebar key on the keyboard. The tools used for making the prototype are MATLAB for collecting, analyzing and processing the data and the Psych Toolbox [23] for the generation and synchronization of the visual stimulus. For faster processing, the Parallel Processing Toolbox in MATLAB is used sparsely. The system can be made faster by making efficient use of this toolbox.

The generation of visual stimulus during training can be tailored. Figure 4.7 shows two different types of visual stimulus used to collect ERP related data. Figure 4.7(a) shows a choice based stimulus where a polygon is positioned in one of the circles. The position of the polygon displayed changes in random order during multiple trials and the ERP response is collected and stored. The Hex-o-Speller based stimulus instructs the subject to spell predefined words and the responses are stored. As shown in Figure 4.7(b) in order to
spell the letter ’A’, the Hex-O-Speller displays the choice to be made at the top-left corner of the screen. Before the beginning of each trial in training, a fixation time is allotted to the user to concentrate on the desired area of the screen. During that time, the training system also makes the user aware of the character to be spelled and its corresponding choice by displaying the character for a very brief time in the center of the screen (before moving to the top-left corner) and flashing/up-sizing the related choice. This information regarding the choice is stored in the Inference and Label fields of the data structure. During testing these fields are assigned after classification by the Model (see Section 3.2).

4.2.2 Results of Simultaneous Data Collection and Stimulus Generation

Collecting EMG signals

Here as a proof of concept for testing the workings of the data collection system EMG signals from the wrist were used and the corresponding data collected and segmented. Figure 4.8 shows the raw data collected and the epochs extracted. The red line represents the start of the visual stimulus, the green lines represent the end of up-sizing of a circles in the stimulus and the black line that almost coincides with end of the raw data window represents the stop event. Here the green lines that represent the events are spaced on an average time of 0.22 seconds. The Timing Diagram shown in Figure 3.3 the events are spaced at 0.25 seconds, which is the maximum time allowed between the up-sizing of the circles. Ideally the spacing between the up-sizing events should be 0.2 seconds, but a delay error of range 0.05 seconds happens due to processing delays. Hence, the stimulus onset varies from 0.20 – 0.25 seconds. But as the epoching/segmentation is done with respect to event times, it is safe to assume that the epochs extracted adequately represent the response
for each up-sizing of the circle with minimal delays.

This data collection system allows the user to control and change various parameters during data collection. As this system can interface with multiple Bio-Radio 150s, the Number of Channels can be increased. The size of the Collection Interval can be changed based on the time taken by a single visual stimulus. The user can also add events and vary their stimulus onset during a visual stimulus. The size and placement of the Epoch Window can be changed as per the need of the experiment or reaction time of the subject to be tested. The number of repeated trials needed for averaging to suppress ERP related noise can be set.

**Collecting EEG signals**

The EEG signals are collected using the 2 Bio-Radio 150 from 10 channels at 960 samples/second data sampling rate. The channels used are P3, P4, C3, C4, F3, F4, O1, O2, Cz and Pz (see Figure 2.6). For each stimulus 10 trials are recorded of the same epoch. As shown in Figure 4.7(a), three different polygons were randomly placed for 3 repetitions and 10 trials of each repetition were recorded. This generated 9 selection responses. This was followed by the Hex-O-Speller, where the text 'MABLRT' was spelled. This text is relatively small as compared to the dataset used in [7]. This generates 16 responses as each letter is spelled in two recurrent steps. Ideally the following text will be used: 'PACKMYBOXWITHFIVEDOZENLIQUORJUGS' that contains all the alphabet. This visual stimuli is followed by a Rapid Serial Visual Presentation, where the subject tries to identify a 'red car'. The RSVP is repeated for 10 times, in order to get 10 trials for the EEG response corresponding to an image.
After collection of the EEG responses, followed by pre-processing and data correction, target and non-target epochs were extracted. Figure 4.9 shows the mean ERP response. The graph in Figure 4.9 shows the time course of the ERPs at channel Cz and the average of channels P3 and P4. The data shown is base-line corrected.

The collected EEG data is then passed to the ERP detection and classification system (Section 3.4) for the generation of a training model. The classifier was integrated with the system to detect ERPs in real time. In order to measure the success of the working of the BCI in real time, the user/subject was asked to perform a task of navigating through the interface and spelling the word 'FOX'. The task was achieved 6 times out of 10 tries.
Figure 4.2 Spatial Selection for Subject A

Data Points: 0.7 seconds i.e. 168 data points

Bhattacharyya Distance (BD)
Figure 4.3 Spatial Selection for Subject B.
Figure 4.4 Topographic map of BD as a 2-D circular view for Subject A and B
Figure 4.5 Temporal Selection for Subject A
Figure 4.6 Temporal Selection for Subject B
Figure 4.7 Examples of Training Visual Stimuli

(a) Training for ERP i.e. Choice related Stimuli

(b) Training for Hex-O-Speller Stimuli
Figure 4.8 Sample Raw Data window and Epoch Extraction of a Single channel
Figure 4.9 ERPs in the BCI for self-collected data in the mean signal across all data samples/epochs for target and non-target stimuli.
Chapter 5

Conclusions

The data collection system for the proposed BCI framework ([2]) is adequately able to collect data in sync with the visual stimulus. However, due to the lack of data from more than one user/subject, comment cannot be made on the accuracy of the Timing Synchronization. It allows some flexibility to control various aspects and parameters of data collection. It also organizes the data during training and records the related visual stimulus and event information in a well defined data structure. This system can be used for performing future experiments and collecting data for various subjects; both healthy and disabled.

The results obtained from classification of ERP’s on the standardized data set show that the data obtained from Subject B was less noisy, and it resulted in higher accuracy compared to Subject A. For feature extraction, the temporal domain information for ERP’s is important and is preserved in the Time Segment feature extraction method. Band powers resulted in the worst performance for ERP classification. Wavelet Decomposition results were relatively better compared to the Band Powers. Even though the results for performance during validation was high, it did not necessarily give good results during testing. In comparison to the Wavelet Decomposition the Time Segments were more consistent. The training model with Linear Discriminant Analysis in combination with Time Segment
features resulted in the best classification accuracy for both the subjects.

However, the training model generated from the real-time implementation of the proposed BCI system data did not yield decent results. One of the possible reasons could be the smaller size of the data set.
Chapter 6

Future Work

Approval from the Institutional Review Board (IRB) to perform data collection using the proposed BCI system on a group of 15-30 subjects; both healthy and disabled is being sought. Pending approval this system will be tested on the participating subjects. Using this system training models for these subjects will be generated. The User Interface will be optimized and changed based on the feedback from the subjects regarding the ease of navigation, representation of information and the options available. Also an analysis on time synchronization of the EEG data collected to the visual stimuli and the processing time of the system will be done.

Other Classifier models for detecting motor imaginary movements, eye movements, etc. will be implemented. Training of the user for detecting eye blinks and movements will be done in order to remove artifacts or noisy sample data. For training of the motor imaginary movement, an object is animated to move left, right, up and down and the EEG data is collected. In order to understand the relation of the EEG scores to the selected images and define a formulation, data from multiple subjects (healthy and disabled) will be collected and analyzed. Each subject will be shown a series of semantically similar images and the EEG data will be recorded. The subject will then be asked to rank the images according
to his/her liking. They will also give a valence score from 1 to 5 for each image shown to determine the criteria and the reason for their interest. They will score an image in terms of it being pleasant, unpleasant, and neutral; or was it recognized as something similar or familiar [15][18][12].
Bibliography


Appendix A

Topographic map of an EEG field as a 2-D circular view

(* looking down at the top of the head)
Figure A.1 ICA Components for Subject A
IC25  IC26  IC27  IC28

IC29  IC30  IC31  IC32

IC33  IC34  IC35  IC36

IC37  IC38  IC39  IC40

IC41  IC42  IC43  IC44

Department of Electrical and Microelectronic Engineering
Figure A.2 ICA Components for Subject B

IC1  IC2  IC3  IC4

IC5  IC6  IC7  IC8

IC9  IC10  IC11  IC12

IC13  IC14  IC15  IC16

IC17  IC18  IC19  IC20

IC21  IC22  IC23  IC24
Appendix B

Receiver Operating Characteristic (ROC) Plots

Figure B.1 Subject A Band Powers : sum of FFT Coefficients
Figure B.2 Subject A Band Powers: sum of FFT Coefficients

Figure B.3 Subject A - Time Segments: sum of data points in selected time intervals
**Figure B.4** Subject B - Time Segments: sum of data points in selected time intervals

![Diagram showing Time Segments LDA and Time Segments NN]

**Figure B.5** Subject A: Wavelet Decomposition

![Diagram showing Wavelets LDA and Wavelets NN]
Figure B.6 Subject B : Wavelet Decomposition
Appendix C

Channel Numbers
Figure C.1 Channel Numbers and its corresponding names and positions