Title
Inverting Visual Stimulus Paradigm and Minimal Electrode Set Improve P300 Speller Performance and Practicality

Author
Deshpande, Aniket

Publication Date
2014

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Inverting Visual Stimulus Paradigm and Minimal Electrode Set Improve P300 Speller Performance and Practicality

A thesis submitted in partial satisfaction

of the requirements for the degree Master of Science in

Bioengineering

by

Aniket Deshpande

2014

ABSTRACT OF THESIS

Inverting Visual Stimulus Paradigm and Minimal Electrode Set Improve P300 Speller Performance and Practicality

by

Aniket Deshpande

Master of Science in Bioengineering

University of California, Los Angeles, 2014

Professor Nader Pouratian, Chair

This thesis describes two studies aiming to enhance the speed, accuracy and practicality of the P300 Speller. The first study introduces the Inverting paradigm, which is based on the hypothesis that a modification in the visual stimulus paradigm in order to increase the stimulus saliency and strength might enhance the physiological response and hence system performance. The second study compares Speller performance using a minimal four electrode set, versus the 32 and six electrode sets previously used and described in literature. The Inverting paradigm

significantly improved Speller speed and accuracy, along with an increase in the physiological response. The minimal electrode set maintained comparable levels of performance and hence increased the system practicality.

The thesis of Aniket Deshpande is approved.

Ian Cook

Dejan Markovic

Nader Pouratian, Committee Chair

University of California, Los Angeles

2014

Dedicated to the most complicated & amazing circuit known to exist in the universe. This circuit lets us derive meaning and beauty from patterns of light and pressure waves; it lets us manipulate limbs to shape the world we live in & feel our own breath; it informs us where the self ends and surroundings begin, and sometimes it doesn't; it lets us sense pleasure & pain and feel emotions; it lets us speak language, have memories, and lets us experience what we call life. Above all, it manages to convince us into believing, that we exist.

This work is also dedicated to my family & friends. Parents, without whose support nothing would have been possible. Grandparents. Aunt, uncle, cousins & their spouses. And the youngest members of our family, my favorite 'humanlings', Aahana & Ayaan.

TABLE OF CONTENTS

LIST OF ACRONYMS

BCI	Brain Computer Interface
EEG	Electroencephalography
ECoG	Electrocorticography
fMRI	functional Magnetic Resonance Imaging
MEG	Magnetic Resonance Imaging
NIR	Near Infrared Spectroscopy
EMG	Electromyography
ADC	Analog to Digital Convertor
ECG	Electrocardiography
SNR	Signal to Noise Ratio
MRI	Magnetic Resonance Imaging
LFP	Local Field Potential
fNIR	functional Near Infrared Spectroscopy
ERD	Event Related Desynchronization
ERS	Event Related Synchronization
ERP	Event Related Potential
EP	Evoked Potential
ALS	Amytrophic Lateral Sclerosis
RC	Row-Column

SWLDA	Stepwise Linear Discriminant Analysis
FLD	Fischer's Linear Discriminant
CBP	Checkerboard Paradigm
SBP	Sub-matrix Based Paradigm
ITR	Information Transfer Rate
SR	Selection Rate

ACKNOWLEDGEMENTS

Study One from Chapter Two is a version of Deshpande A, Speier W and Pouratian N. Inverting visual stimulus paradigm improves P300 Speller performance. 2013. Submitted to Journal of Neural Engineering.

William Speier is a Biomedical Engineering graduate student at UCLA and a member of the Neurosurgical Brain Mapping and Restoration Laboratory. Nader Pouratian is an affiliate faculty in UCLA Bioengineering and assistant professor in the Department of Neurosurgery at UCLA. Their guidance and advice has been very valuable while conducting this research, and I would like to thank them for their support.

CHAPTER ONE: AN INTRODUCTION TO BRAIN-COMPUTER INTERFACES

I. Brain Computer Interfacing

A Brain Computer Interface (BCI) uses signals from the human brain as control signals for controlling an external computer or an actuator. These signals may be acquired using various modalities – the most common ones being EEG, ECoG or single unit recordings. Some other modalities that are also used are fMRI, magnetoencephalography (MEG), or Near Infrared Spectroscopy (NIR).

The overview of a BCI system can be summarized using the following block diagram (Fig. 1).

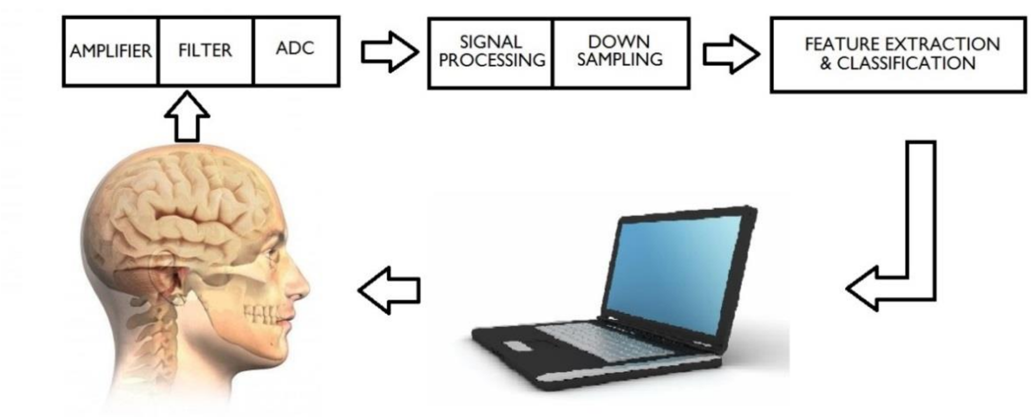

Fig. 1: Block Diagram of a BCI system

The first stage of the system typically consists of signal acquisition. The signals usually have very low amplitudes, to the order of tens to hundreds of microvolts. The thermal noise in the system also has amplitudes comparable to the signals of interest. This necessitates amplification and filtering of the signal acquired. Amplification may be done in two stages, with the first stage meant for providing a reasonable buffering between the electrodes and the filtering stage, and the second stage for providing a large gain to amplify the signals to a range such that the signals may be appropriately passed onto an analog to digital convertor, without any wastage of resolution or possible saturation. Filtering is used to get rid of high frequency noise from thermal sources, EMG and activity from other devices, power line noise, and low frequency drift from the electrode electrolyte contact, movement artifacts and eye movements. The ADC converts these analog signals to digital values, which can be passed on to a computer.

Since brain signals are collected from several channels, and at a considerable length of time, it is often necessary to down sample the data with respect to time or channels. This, possibly along with further filtering for any new noise introduced by the system, such as quantization noise, is done in the digital domain.

The third stage of the system includes feature extraction and classification. This enables the system to make decisions based on the signal acquired. In the feature extraction step, those features or patterns within the signal that help in making a distinction between two or more of several end goals, are extracted. The classification step makes use of these features to come to a decision as to which particular class the signal belongs to, based on how similar the features are to a particular class.

The output obtained from the feature classification process initiates an appropriate action, which may be the motion of a robotic arm or typing a character on a computer screen. This also serves as a feedback mechanism for the user, who can then make any modifications if required.

II. Acquisition Modalities:

a. Electroencephalography (EEG):

EEG detects summed electrical activity along the scalp, by measuring the voltage difference between any two points on the scalp. Such two points, between which the voltage difference is measured, are referred to as a channel, and an EEG may be comprised of as many as 256 such channels. Usually, the voltage difference is measured with respect to a common ground which is located on the ear lobe or mastoids, and re-referenced to another point on the scalp or to the spatial average of all the voltages. The ground is not connected to the mains ground, for safety reasons. Hence, it has to be on the body. Also, it is kept on the head, either at the earlobes or mastoids to avoid noise from any other electrophysiological sources, such as ECG or EMG. Often, a right leg driver circuit is used that drives any common mode noise back into the ground on the body. The electrodes are placed in a standardized pattern over the scalp, usually in the International 10-20 system. Fig. 2 shows a subject using EEG.

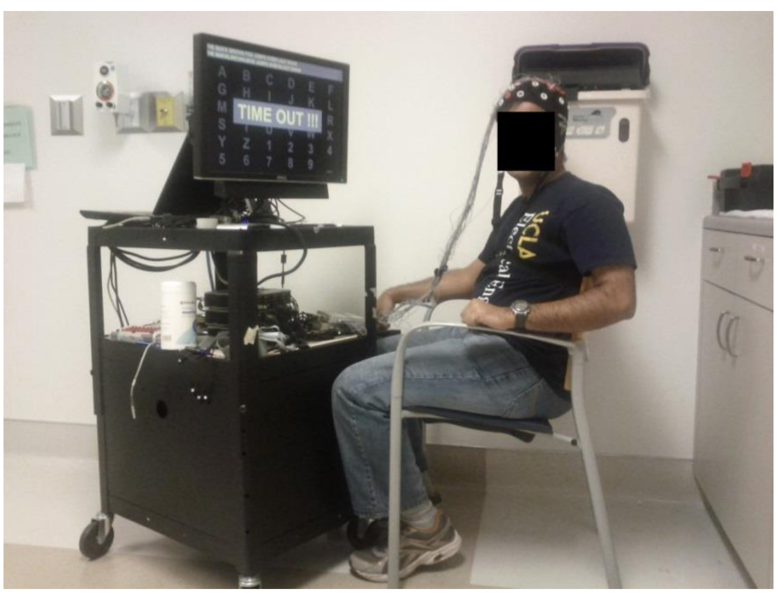

Fig. 2: Subject using EEG

There are several reasons behind the popularity of EEG as a signal acquisition modality for BCI systems. Mainly, it is the fact that EEG is non-invasive. Being non-invasive vastly widens its scope of use. It is also relatively more convenient to use, not being as bulky as other systems, even though it does require a considerable setup time. Also, compared to other modalities, EEG is cost effective. Since electrical potentials are being measured, EEG has a very good temporal resolution. It is subject friendly in the sense that it does not expose the subject to large magnetic fields or radiation.

EEG also has several limitations. Since the electrical signals have to pass through the cranial thickness, the signal strength obtained is very low. Signal amplitudes are to the order of tens of microvolts, making them comparable to the ambient noise levels. Further, the skull filters out frequencies higher than about 100 Hz. EEG has a poor spatial resolution, in centimeters [1]. It is difficult to localize the signal source to a narrow region of interest. Some degree of spatial averaging can be said to occur inherently, due to the spread of electric potential to neighboring electrodes. Since the brain surface is not flat, but is lined with gyri and sulci, only some patches on the brain surface may be parallel to the scalp surface. These surfaces have neurons that are perpendicular to the scalp surface. Only such surfaces can produce dipole moments between two electrodes over the scalp surface, further reducing the scope of activity visible to EEG [2].

b. Electrocorticography (ECoG):

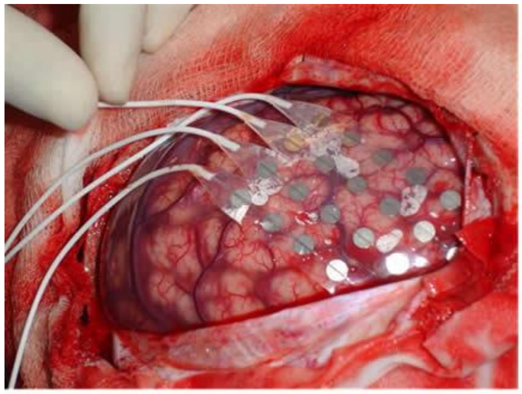

Fig. 3: ECoG [3]

ECoG (Fig. 3) measures electrical potentials from the brain by directly placing electrodes on the cerebral cortex. A part of the skull is surgically opened for this purpose. The use of ECoG is primarily limited because of this invasive nature; however, it is comparatively less invasive than insertion of electrode arrays into brain tissue which penetrate the blood brain barrier. In terms of signal quality, it is clearly better than EEG in terms of SNR – since the signal amplitude is about 50-100μV compared to 5-20μV offered by EEG [1], and spatial resolution which is in millimeters for ECoG [1]. Higher frequencies are not lost since there is no filtering through the skull and a bandwidth of up to 200Hz is achieved [1]. EEG or MRI may often be used to narrow down to a certain region of the brain, before craniotomy is performed to expose that area of the cortex. This also implies the specificity of the tasks for which ECoG is used.

c. Intracortical recordings:

Intracortical recordings employ electrodes implanted up to a small depth into the brain cortex. This offers better spatio-temporal resolution compared to the semi-invasive and non-invasive methods discussed earlier [4], due to the fact that the electrodes being closer to individual neurons averages signals from fewer number of neurons in space and at any time. Apart from being able to record better resolved local field potentials, these electrodes can record single unit activity. Single unit activity is nothing but the recording of individual action potentials from single neurons. Intracortical recordings have a frequency bandwidth of up to 1 kHz for LFPs whereas individual action potentials can be above 1 kHz [4]. The rate of firing of these action potentials is used by the brain to encode various features, for example in the motor cortex it may be used to encode velocity, force, etc. These firing patterns can be used as a control signal for external applications with a greater level of detail than non-invasive methods which utilize the overall activity of a large group of neurons [4].

Despite its advantages, intracortical recording has several drawbacks as well. It requires surgery to be able to access the cortex and electrodes that could perform chronically. However since the electrodes are inserted in brain tissue, the tissue response causes the electrodes to fail eventually due to lack of electrical contact caused by deposition of glial cells. The figure below (Fig. 4) shows the three different recording sites described thus far.

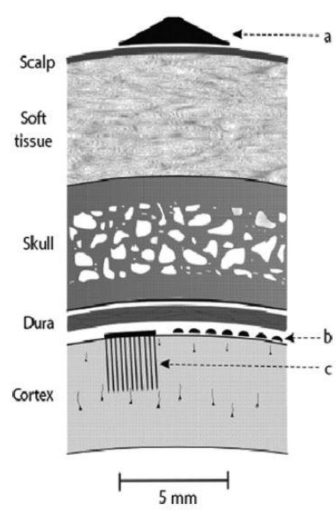

Fig. 4: a. Scalp (EEG), b. Subdural (ECoG) and c. Intracortical recording sites [4]

d. Other Modalities:

A few other modalities, which are based on the hemodynamic response of the brain, such as fMRI (functional Magnetic Resonance imaging) (Fig. 5b) or fNIR (functional Near Infrared Spectroscopy), may also be used for brain signal acquisition. Whenever a certain population of neurons is active, more amount of oxygen is demanded by that portion of the brain. This results in a greater blood flow to that area, which can be detected using these modalities. These modalities have the advantage of being non-invasive. Although MRI is bulky, expensive, time consuming and exposes the patient to high magnetic fields, fNIR has the advantages of being portable, affordable and safe [3]. However, the hemodynamic response is a slow one and hence these techniques have poor temporal resolution. Further, these techniques only tell us about how much activity is present in the brain and where. It does not inform about the nature of this activity, for example the frequency content of the signal, whether the neuronal population is firing synchronously or asynchronously, occurrence of event related potentials or other temporal events of interest. For these reasons, their use for BCI applications is very limited.

MEG (Fig. 5a) or magnetoencephalography may also be used to acquire information about the magnetic activity of the brain. However, the instrumentation required for this purpose is highly non-portable and expensive. Also, it is much more susceptible to noise, and its use in a BCI system is not very practical.

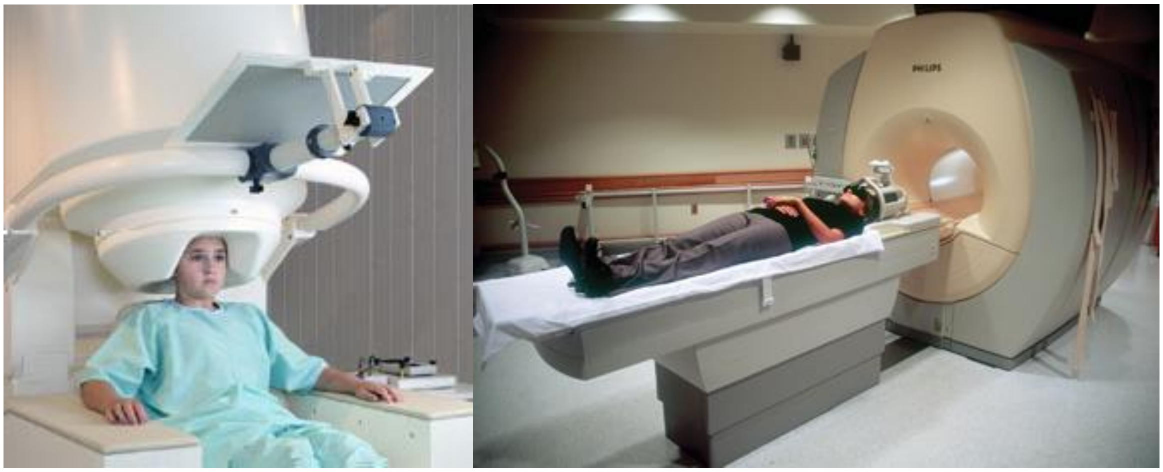

Fig. 5: a. MEG [5] and b. fMRI [6]

III. Use of EEG signals in BCI applications:

We discussed electroencephalography, as a technique to acquire brain signals for BCI. Let us now elaborate upon the electroencephalogram, or the EEG signal itself and its components which can be used as control signals for a BCI system. EEG has a frequency range of 0.1 Hz to 100 Hz. Certain frequency bands within this range are biologically relevant and hence have special nomenclatures for identification.

Human EEG was invented by the German scientist, Hans Berger (1873–1941) [7], [8]. Berger recorded the first EEG in the mid-1920s, and also was the first to describe alpha waves, now also known as Berger waves [8]. Fig. 6 shows the first recorded human EEG by Hans Berger. Alpha waves are associated with mental and physical relaxation during wakefulness, and tend to appear on the closure of eyes. These waves lie in the frequency band of 8Hz to 13Hz, over occipital sites [9]. Alpha rhythm is suppressed during attention, and visual or mental effort [9]. Mu waves are nothing but alpha waves, observed in the sensorimotor cortex when limbs are relaxed. These may be thought of as the alpha waves corresponding to motor function. Similar to the alpha waves, mu waves are suppressed during movements.

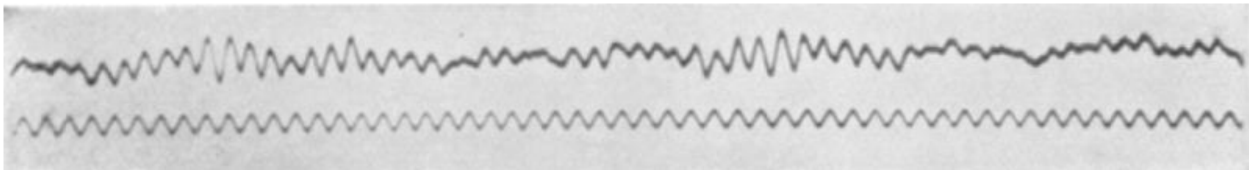

Fig. 6: First recorded human EEG. The lower graph is a 10Hz calibration signal, while the upper one is a recording from Hans Berger's young son. [8]

As the wave frequency increases, the amplitudes tend to decrease; thus the lowest frequency bands have highest amplitudes. Fig. 7 shows the frequency spectrum of an average EEG signal.

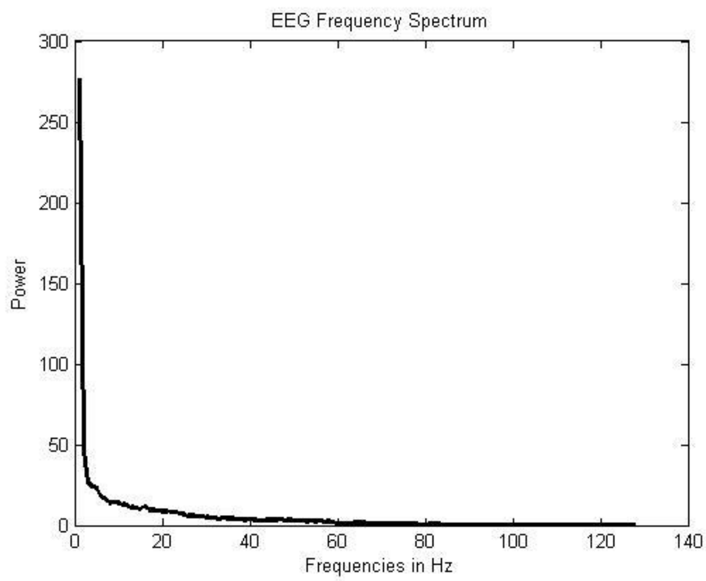

Fig. 7: Frequency Spectrum obtained from average EEG signal

Theta waves, observed from 4Hz to 7Hz are indicative of relaxation, meditation, drowsiness, and sleep. Delta waves, between 0 to 4Hz, are observed during slow wave sleep, otherwise called deep sleep in adults and also normally seen in babies.

Beta waves, 13Hz to 30Hz are observed during alertness or thinking and mental concentration. Gamma band consists of frequencies between 30Hz to 100Hz, and is known to occur during activation of motor functions along with a suppression of beta.

Apart from these frequency bands, there are other signals encountered in the EEG, such as Event Related Desynchronizations (ERDs), Event Related Synchronizations (ERSs), and Event Related Potentials (ERPs), which includes Evoked Potentials (EPs). These are utilized for BCI applications.

ERDs and ERSs, referred together as EEG reactivity, may be understood as the variation (decrease or increase respectively), in percent power of a particular frequency band over a certain time period, related to an internal or external event. This variation is with respect to a non-event time period as a reference. It is temporally, and usually spatially limited; and may occur in anticipation of or in response to an event.

An example of EEG reactivity would be pre-movement alpha ERD and post-movement beta ERS. Patients with neurological disorders may show differences in their EEG reactivity to certain tasks, compared to normal subjects [10]. Thus, EEG reactivity may be used in neuroprosthetic applications or in the diagnosis of neurological disorders.

Event Related Potentials are temporally bound activities of neuronal populations, seen in the form of positive or negative going peaks, occurring in response to an external stimulus or internal event at a particular amplitude and latency. The latency of each kind of ERP is more or less fixed, and hence it is convenient to name these potentials based in their polarity and latency. For example, a negative going potential observed 100ms after a stimulus would be called N100. Different types of ERPs respond to different kinds of stimuli. ERPs have small amplitudes, which are comparable to the background activity. However, because of their fixed latencies, averaging can be performed over several stimuli to extract the ERPs from the signal.

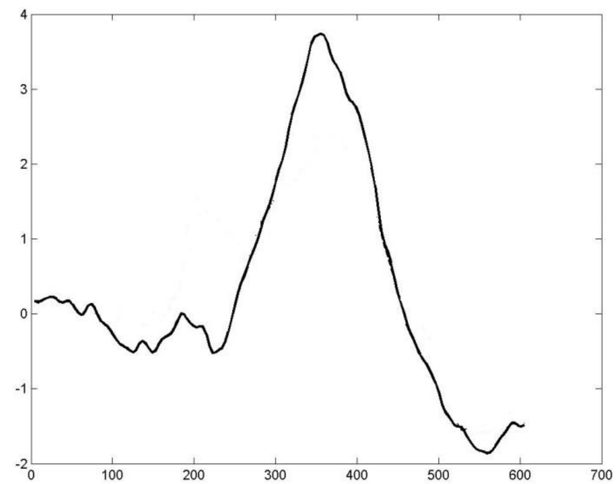

Fig. 8: Average P300 signal in microvolts versus time in milliseconds

IV. P300 Signal:

The P300, as the name suggests, is an ERP with a positive going peak at around 300ms following a stimulus. It was first reported in November 1965, by Sutton et al [11]. P300 is elicited in response to an oddball stimulus, which is a target stimulus which occurs with a low probability amongst other non-target high probability stimuli. An average P300 signal is shown in Fig. 8. P300 amplitude varies directly with stimulus relevance and attentional resource allocation, while and inversely with stimulus probability, and is usually observed with a stronger amplitude at parietal sites.

The P300 was popularized since the 1980s for possible use in lie detection. In a typical interrogation scenario, say for example a crime related investigation; a suspect would be presented with an oddball task. This oddball task would consist of low probability stimuli, highly relevant to the specific crime, interspersed between generic high probability stimuli, relatively irrelevant to the crime. This way, only a subject with knowledge about the specific details of a crime scene would be able to distinguish between the two kinds of stimuli, and attach special relevance to those related to the crime, thus eliciting P300 responses upon the presentation of such stimuli. This test is also known as the guilty knowledge test.

In 1988, Lawrence Farewell and Emanuel Donchin first described the P300 speller as a communication aid for ALS patients [12]. The P300 speller that was first described by Farewell & Donchin, and has been conventionally used, consists of a 6x6 matrix of gray characters on a black screen. The user attends to a specific character that is to be typed, as all rows and columns flash in random order. This character is called the attended or target character. Whenever a row or column flashes, all the characters in that particular row or column change color from gray to white for a short duration and then change back to gray. When the attended character flashes, it acts as a low probability or oddball stimulus amongst several other high probability group stimuli and is also relevant to the user. This causes a P300 response to be elicited, whenever the attended character flashes. The visual interface described here is known as the Row-Column (RC) flashing paradigm and is most commonly used in P300 speller systems. It is often used as the golden standard while performing comparative studies for new paradigms. Fig. 9 shows a screenshot of the P300 Speller user interface with the RC flashing paradigm.

The appropriate character can be selected by detecting which row and column elicited the P300 response. However, in order to be able to do so, the P300 response has to be averaged across several flashes to have a good signal to noise ratio. Also, all other rows and columns must flash every time the row and column of interest have flashed, to maintain the low probability of the target stimulus. This makes the process of typing something useful with the P300 speller, very time consuming and tedious. Further, as the time required to type something is increased, it results in more fatigue for the user possibly causing a poorer performance.

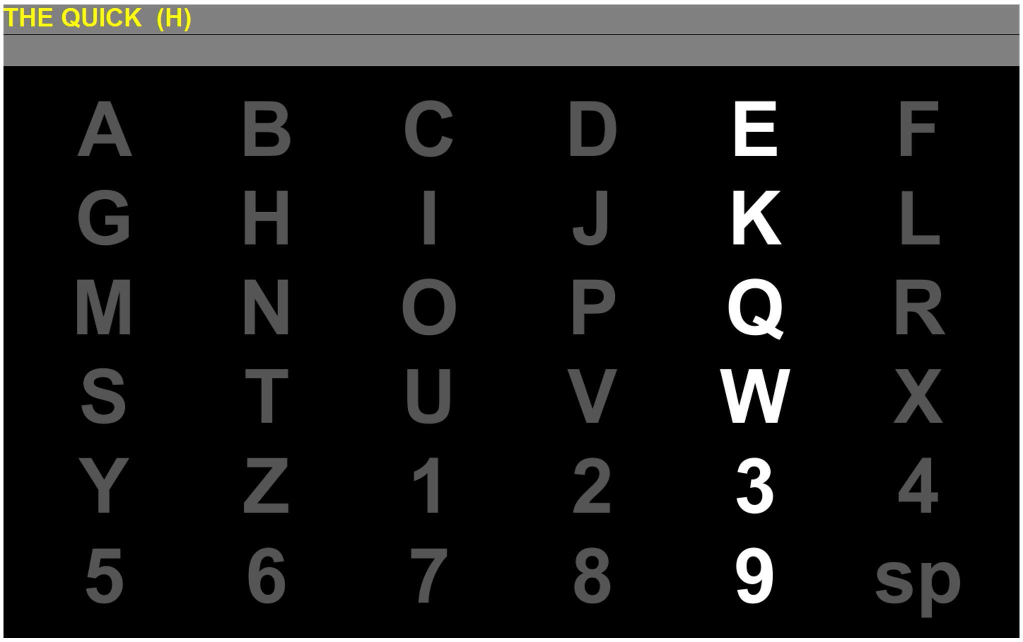

Fig. 9: Screenshot of a P300 Speller User Interface

Several techniques may be used to detect the P300 wave from the signal. These techniques may or may not require prior training data. One such example of detection techniques is cross-correlation. Cross-correlation uses a generalized template which has characteristics similar to the expected wave. This template is compared to the signal at all points across the entire length of the signal to obtain correlation values. In case the expected feature – P300 wave in this case, is present in the given signal, the correlation values will increase for particular signal duration. The template for the purpose of comparison may be produced by using any conventional shape similar to the wave, such as half a sine or triangular wave, or it may be produced by averaging time windows of a signal known to contain the feature of interest. Here, the first case would not require any prior training set but the latter would probably give better results. The Woody adaptive filter draws a middle path between these methods. It uses a half sine or triangular wave as a template to begin with, but adapts the template to be the average of detected features using the first iteration and continues this process till there is no further increase in the correlation.

Another technique used very often is a classifier based on discriminant analysis. A classifier is an algorithm that uses a set of training data consisting of pre-defined target characters to extract features from the EEG signal. Features are sets of data points from the EEG signal within a fixed time window, and from a fixed set of scalp sites. It then assigns weights to these features based on which set of features distinguish the best between target and non-target characters. When detection or training is performed, the new EEG signal is compared against such distinguishing features to indicate whether the signal corresponds to a target or non-target stimulus. This is known as supervised classification, since the classifier is already presented with the response a target stimulus is expected to produce in a particular subject. Thus, a classifier training session needs to be performed prior to the actual use of the P300 speller for communication purposes. This further multiplies the overall time required for the use of the system.

An example of discriminant analysis commonly used for P300 Speller applications [13] is the Stepwise Linear Discriminant Analysis (SWLDA). SWLDA progressively adds most significant features to the feature set and eliminates the least significant ones, using a process of forward selection and backward elimination. Another procedure used for feature selection is the Pearson Correlation method . In this method, features are observed independently for correlation with the labeled outputs and weights are assigned to the features based on the value of the correlation coefficient. This study also utilizes the Naïve Bayes classifier apart from SWLDA, which takes a probability based approached to classification and selects an output class once the probability of that class crosses a threshold value. A few other techniques used for classification are Support Vector Machines, Neural Networks, and Fischer's Linear Discriminant.

V. Amyotrophic Lateral Sclerosis and the use of BCI:

Brain Computer Interfaces provide an alternate way of communication, ambulation and interaction with the physical environment to patients suffering from neuromuscular diseases such as stroke, paralysis, and Amyotrophic Lateral Sclerosis.

At any given point in time, it is estimated that as many as 30,000 Americans may be suffering with Amyotrophic Lateral Sclerosis; and approximately 5,600 are diagnosed with it each year [14].

Amyotrophic Lateral Sclerosis (ALS), also known as Lou Gehrig's disease is a motor neuron disease causing muscle weakness and atrophy due to degeneration of upper and lower motor neurons. Degeneration of the motor neurons causes inability to control muscles, and this in turn weakens the muscles further, due to lack of use. Progressively, the ability to initiate and control all voluntary movement is lost [15].

As a result, there are several patients suffering from ALS, who have lost the ability to control their physical environment, actuate physical movements or communicate with the outside world. However, they usually have normal cognitive function and are fully conscious.

Apart from ALS, there could be patients suffering from stroke, paralysis, or spinal cord injury, whose ability to make voluntary movements is either lost completely or compromised. In case of patients who have some degree of residual functionality, control signals may be tapped from spinal cord, nerves, muscles, eye movements, breathing, etc. Several prosthetic devices utilize muscle signals or peripheral nerve signals for amputees. However, in case of patients with complete loss of voluntary movement, accessing brain signals directly becomes necessary.

CHAPTER TWO: IMPROVEMENTS TO THE P300 SPELLER

I. Introduction:

The field of Brain Computer Interfacing has greatly progressed in the last fifteen years. This progress has largely been because of the surge of cost effective and increasingly powerful computer hardware, and advances in software that enable making sense of vast amounts of EEG data in real time [16].

Since Farewell & Donchin first described the P300 Speller in 1988, there have been several attempts at improving the P300 Speller. These attempts have largely focused on overcoming various drawbacks of the original system, such as low speed and accuracy, adjacency distraction, double flash errors [17], fatigue effects, etc. The main objective of these efforts however, is to improve the speed and accuracy of the system. There have been different approaches taken towards this goal, and several groups have implemented and tested modifications to the speller paradigm. Some groups have worked towards improved classifier algorithms. Krusienski et al [18] compared five established classifiers for offline performance and practical concerns. They reported Fisher's Linear Discriminant (FLD), and Stepwise Linear Discriminant Analysis (SWLDA) to have the best overall performance and implementation characteristics, with SWLDA having an added advantage by eliminating insignificant features [18]. Speier et al [13], integrated language information by adding Natural Language Processing with a Naïve Bayes Classifier and a trigram model and showed significant improvements.

With respect to the visual interface design, one approach has been to alter the arrangement of flashing combinations in order to overcome some of the system's drawbacks or classify targets with less input and hence increased speed. In 2010, Townsend et al [19] introduced the Checkerboard paradigm (CBP) in order to prevent errors due to distraction caused by the simultaneous flashing of adjacent characters and consecutive flashing of the same character. The checkerboard avoids flashing rows and columns, but flashes pseudo-randomly arranged flash groups [20] such that directly adjacent characters would never be flashed simultaneously. It also ensures that there are at least six flashes between two repeating flashes of the same character [20]. The Five Flash paradigm was developed by the same group based on the previously described CBP, in order to reduce the number of sequences required for character detection. Shi et al [17], describe the sub-matrix based paradigm (SBP). The SBP divides the matrix into four sub-matrices, and one letter from each sub-matrix is randomly selected as a flashing group during one flash, independent of previous groups. A sequence is completed when each character has flashed once. Not only does this method prevent adjacency errors, but it also reduces target character probability.

Another approach is to make changes to the flashing paradigm that would enhance the signal-to-noise of the P300 waveform. This is different from altering the flashing combinations, in that the basic RC paradigm may remain the same, however the visual parameters of the interface may be changed. Yang Liu et al [21], describe several paradigms involving variations of visual parameters specific to the stimulus presentation. They compared the traditional RC paradigm with novel paradigms such as translation or rotation motion, zoom in or zoom out, and transformation of texture or sharpness, of each character during its presentation as a stimulus. Further, these different paradigms were tested using two transformation modes – the impulse mode, in which each character immediately returns to its original standard state after stimulus presentation and the step mode, in which each character stays in its oddball or stimulus state until the next one acts as stimulus. No particular paradigm studied by Liu et al was found to be consistently better across all subjects, and they recommended using subject specific stimuli. Takano et al [22] compared the RC paradigm to two other paradigms consisting of blue colored characters that turned green during stimulus. Since the RC paradigm in which gray characters turn white during stimulus, it may be viewed as an increase in luminance of the characters rather than transformation of color. Takano and colleagues have thus referred to the RC paradigm as the luminance paradigm, and the other two are - the isoluminance chromatic paradigm which changes color from blue to green during stimulus without any change in the luminance, and the luminance and chromatic paradigm which increases luminance along with changing color from blue to green. They observed significantly better results using the luminance and chromatic paradigm and suggested that neurons in certain brain regions specialized to process color and luminance information might be responsible for the improvements. Salvaris & Sepulveda [23] varied and compared several aspects of stimulus presentation, such as keeping a static white background for the entire grid for the normal as well as stimulus state, small and large inter-symbol distance paradigms, small and large symbol paradigms, and so on. Although the static white background paradigm consistently outperformed other paradigms, the improvement was statistically significant compared only to the worst performing paradigm.

There may be further enhancements possible to the signal to noise ratio of the ERPs, which may reduce the number of flashing sequences required for classification, making communication faster. The strength of the P300 ERP increases with a decrease in the probability of occurrence of the stimulus. Thus, increasing the saliency of the stimuli might enhance the P300 response. Similarly, increasing the strength or intensity of the stimuli may also cause the P300 response to be enhanced. Sugg and Polich [24] investigated the use of varying intensities of auditory stimuli for eliciting a P300 response, and reported significant increases in the P300 amplitude with an increase in the intensity of the auditory stimuli. Li and colleagues [25] reported increases in mean P300 amplitude and significant increase in speller accuracy with the use of greater luminosity contrast stimuli. Thus, modifications made to the visual paradigm, with increased saliency and strength of the stimuli as objectives may improve the P300 response.

Likewise, increasing the saliency, strength and luminance of the stimuli may improve the visual evoked component of the ERP as well. Krusienski et al [26] performed an online study to describe the use of additional scalp locations, other than the standard P300 scalp locations (Fz, Cz, Pz) exclusively used until then. Other groups had previously shown improved classification accuracy using posterior locations with offline studies [27]. They found, that using the occipital sites PO7, PO8 and Oz along with the traditional scalp locations resulted in significant improvement in the classifier performance. Their study highlights the importance of the visual component in the classification process of the P300 speller.

It is implicit from the designs of the novel paradigms introduced by the previous studies mentioned here, that an attempt was made towards increasing the saliency and strength of the stimuli, possibly so with an end goal of enhancing the P300 response and the visual component. For example, Liu et al [21] described stimuli involving character motion and character size transformation. These stimuli might have had greater saliency and strength since the oddball state is more obviously different from the normal state. It also might have increased the luminance of the stimuli in increasing the character size. The paradigms suggested by Salvaris & Sepulveda also indicate similar changes that might improve the salience, strength and luminance of the signal.

A. STUDY ONE: MODIFYING THE VISUAL INTERFACE

I. Methods

a. The Inverting Paradigm:

We hypothesized that increasing the intensity, luminosity and salience of the stimulus, thereby increasing the strength of the visual component of the system and enhancing the P300 response, might increase the speed and accuracy of the P300 speller system. A novel flashing paradigm is described here, in which the black background around the symbol changes to white and the gray symbol turns black. Thus, the stimulus is more intense, luminous and salient than before. In order to test this paradigm, it was compared against the traditional RC paradigm reported by Farewell & Donchin [12]. The system was tested online, and further offline testing was performed using two different classifiers to verify the results [13].

Chapter One introduced the conventionally used RC paradigm. The RC paradigm only consists of the intensification of the row/column which is flashing, from gray to white and back. Thus, the interface background remains as is and is not involved in stimulus presentation. However, since the background comprises of a relatively large area over the screen compared to the characters, it may be utilized in order to produce a stronger stimulus. The method proposed is the RC paradigm with an inversion in the grayscale of the letter versus its background instead of the simple highlighting of the letter in the traditional RC paradigm (Fig. 10a). Henceforth, these will be referred to as Non-Inverting and Inverting paradigms, respectively. In the inverting paradigm, as long as a row/column is not flashing, it has a black background and a gray letter – exactly the same as in the case of non-inverting paradigm. However, when a row/column flashes, its background changes to white and the letter becomes black (Fig. 10b). Hence the background intensifies instead of the characters, giving greater luminance and a more salient stimulus than before since the variation from the normal state to oddball state is larger.

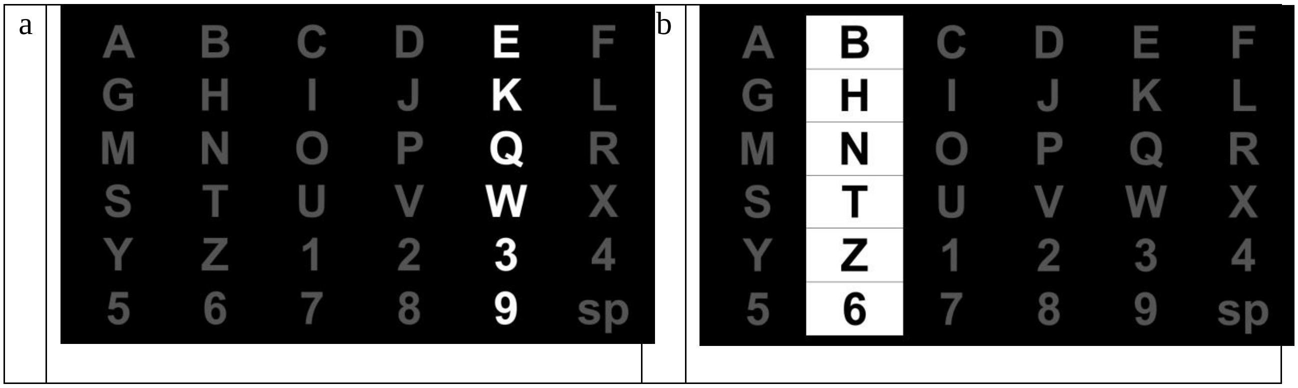

Figure 10: Screenshots of a stimulus presentation using Non-Inverting (a) and Inverting Paradigms (b)

b. Data Collection:

Data was collected from 15 healthy volunteers, consisting of 11 male and four female undergraduate and graduate students (age range 17 to 30 years), having perfect or corrected to perfect vision. Recording was done using 32 electrodes (Fpz, Fz, FC1, FCz, FC2, FC4, FC6, C4, C6, CP4, CP6, FC3, FC5, C3, C5, CP3, CP5, CP1, P1, Cz, CPz, Pz, POz, CP2, P2, PO7, PO3, O1, Oz, O2, PO4, PO8) with the left or right earlobe as ground. A pair of 16 channel g.tec USB biosignal amplifiers was used for amplification, digitized at 256Hz and filtered between 0.1 and 60Hz, while BCI2000 software system for BCI research was used for data acquisition and stimulus presentation. The number of sequences used was ten that is, for every character there were 12 flashes (six rows and six columns) repeated ten times. An ISI of 125ms was used between two stimuli.

Experiments were performed with the objective of comparing the effects and results produced by the Inverting paradigm to the Non-Inverting paradigm, comprising of a training session to train the classifier algorithm and testing in which subjects used the trained classifier to type series of characters with the speller. An experiment with one subject thus consisted of ten words of training, with five words for each paradigm, and then ten words of testing with five for each paradigm. All of these were randomly chosen five lettered words (Table 1). Which paradigm was presented first, was alternated for every subject. Given this method, the set of words for training and evaluation sessions remained constant for all subjects, while the words within these sessions were interchanged between flashing paradigms.

AVOID	JUICE	UNITS	MINUS	NOTED
MIXED	NIGHT	DAILY	SCORE	GIANT
BEING	MAJOR	HOURS	SHOWN	PANEL
FIRST	BLOCK	CLEAR	GIVEN	AFTER

Table 1: Target words used in the experiment

The experiment consisted of copy spelling and visual feedback during testing. Copy spelling implies that the subjects were presented with words on the screen, which they were expected to type rather than having them choose the words for themselves. In cases where the classifier could not generate feature weights for online testing, the experiment was continued without any visual feedback so as to complete data collection for a set of twenty words.

The fraction of letters correctly detected during this testing was the basis of the comparative performance analyses performed later. This is referred to as the online testing and was performed on two thirds of the subjects.

This was followed by offline analysis combining the training and testing data, in order to find trends in the data that were not apparent in the online results.

c. Data Analysis:

BCI2000 was used for online testing and MATLAB (version 7.10, MathWorks, Inc., Natick, MA) was used for offline testing, which was done using 10-fold cross-validation. BCI2000 used Stepwise Linear Regression for classification, and a fixed value of 10 trials was used. Offline analysis was performed using the SWLDA and Naïve Bayes classifiers [13].

c.1 SWLDA Classifier:

Classification was performed using stepwise linear discriminant analysis (SWLDA) to identify differences in response to attended vs. non-attended stimuli [13]. The signal from a particular electrode is composed of voltage level values at several time points. Each of these different time points is treated as features or variables, with the voltages as values. SWLDA extracts features from the signal, which help in discriminating the signals for target stimuli from non-target stimuli. Extraction of the most significant features also serves to reduce the data size. It does so by a process of forward selection, in which most significant features are added to the classifier and backward elimination, in which the least significant features are removed. Thus, the entire data stream is not used in decision making unlike in other methods such as covariance. It also outputs a feature weight vector, which suggests the relative importance of the signal features in discrimination of the targets and non-targets. The feature vector and feature weight vector are created using the training data set. Each new flash is assigned a score, which is indicative of its proximity to the attended class. This score y_t^i is calculated as the dot product of the feature weight vector, w, with the feature vector, x_t^i (for every flash *i* and every target character *t*),

$$\boldsymbol{y_t^i} = \boldsymbol{w} \cdot \boldsymbol{x_t^i} \tag{1}$$

This function, which is used to calculate the score, is known as the discriminant function. In order to validate the discriminant function, a portion of the data is used to generate the function and the other portion is used to test it. This is known as cross-validation. Similar testing is done during an online session to determine which character is being attended.

At the end of a flashing sequence, each character is assigned a score, which is the sum of all scores for flashes containing that character. The character corresponding to the highest score is selected. Thus, the score for each possible next character is given as the sum of scores for flashes containing that character:

$$\boldsymbol{g}(\boldsymbol{x_t}) = \sum_{\boldsymbol{i}:\boldsymbol{x_t} \in \boldsymbol{A_t^i}} \boldsymbol{y_t^i} \tag{2}$$

c.2 Naïve Bayes Classifier:

While the SWLDA classifier assigns scores to each character and selects the character with the highest score by the end of total number of sequences, the Naïve Bayes classifier takes a different approach. It still assigns scores to each character after every round of a sequence (a sequence is when all rows and all columns have flashed once, and consists of 12 flashes in all), but does not wait for the total number of sequences to exhaust. Instead, the Naïve Bayes classifier assigns probabilities to the signal based on the score and selects the character once a threshold probability is reached. These probabilities represent the chance that the stimulus is an attended one. Probabilities are assigned to each character, and the character that reaches threshold first is selected. The scores were assumed to have a normal distribution and hence the probability density functions for the score given an attended or unattended character were calculated as:

$$f\left(y_t^i \middle| x_t\right) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma_a^2}} e^{\frac{-1}{2\sigma_a^2}(y_t^i - \mu_a)^2} \text{ if } x_t \in A_t^i \\ \frac{1}{\sqrt{2\pi\sigma_a^2}} e^{\frac{-1}{2\sigma_a^2}(y_t^i - \mu_n)^2} \text{ if } x_t \text{ is not } \in A_t^i \end{cases} \quad (3)$$

where, $\mu_a, \sigma_a^2, \mu_n, \sigma_n^2$, are the means and variances respectively for attended and non-attended stimuli.

Thus, we know the probability of a particular score given an attended or non-attended character. Hence, using Bayes theorem we can calculate the probability of a particular character being attended given its score. Assuming conditional independence between flashes given the current attended character, the posterior probability is given as,

$$\begin{aligned} P(x_t|y_t, x_{t-1}, \dots, x_0) &= \frac{P(x_t|x_{t-1}, \dots, x_0) P(y_t|x_t, \dots, x_0)}{P(y_t|x_{t-1}, \dots, x_0)} \\ &= \frac{1}{Z} P(x_t|x_{t-1}, \dots, x_0) \textstyle\prod_i f(\, y_t^i|x_t) \end{aligned} \quad (3)$$

Once a predetermined threshold probability is reached or the maximum number of flashes is reached, the character with maximum probability is selected [13].

Further, this classifier used a corpus of words from the English language in order to assign bias probabilities to letters depending on the previous letters chosen and letters most likely to make a word. This way, knowledge about the language is utilized for making more efficient predictions about the characters to be typed. Sets of three letters are compared with each word in the corpus. The probability of the next character being ‘x ‘ is the number of times it appears following the previous two characters divided by the number of times the previous two characters appear consecutively in the corpus. In this case, the prior probabilities are calculated as,

$$P(x_t|x_{t-1}, \dots, x_0) = \frac{c(x_{t-2}, x_{t-1}, x_t)}{c(x_{t-2}, x_{t-1})} \tag{4}$$

Where, $c(x_{t-2}, x_{t-1}, x_t)$ is the number of occurrences of the string '$x_{t-2}x_{t-1}x_t{}'$ in the corpus [13].

For the offline analysis, this threshold probability was varied and the information transfer rate and accuracy calculated for each threshold. This was done for every subject and the threshold with the highest ITR was selected for that particular subject.

d. Evaluation:

The information transfer rate (ITR) is derived from the bits per symbol, B, which is calculated as:

$$B = \log_2 \mathrm{N} + \mathrm{P} \log_2 \mathrm{P} + (1 - \mathrm{P}) \log_2 \frac{1 - \mathrm{P}}{\mathrm{N} - 1} \tag{5}$$

Where N is the total number of characters in the matrix, P is the selection accuracy, given by the ratio of correct selections to total trials. Also, the selection rate (in selections per minute) is given as the inverse of the average selection time,

$$\mathrm{SR} = \frac{60}{3.5+0.125*12*a} \tag{6}$$

Where, 3.5s is the gap between flashes, 0.125s is the time taken by every flash, 12 is the number of flashes in each set, and 'a' is the average number of sets of flashes. The information transfer rate, in terms of bits per minute, can be calculated as the product of bits per symbol, B and the selection rate, S. The ITR was obtained separately for inverting and non-inverting paradigms, in order to use it as a comparison metric.

Paired t-tests are used to evaluate differences in the performance between paradigms.

III. Results:

a. Offline Results:

Offline analysis was performed on the data using SWLDA and Naïve Bayes classifiers. The Inverting paradigm showed significantly improved results independent of which classifier was used (Table 2).

Using the SWLDA (original) classification methodology, ITRs significantly improved when using Inverting vs. Non-Inverting paradigm (27.59 vs. 21.23, p=0.0002). Likewise, using Naïve Bayes classification, use of the inverting flashing paradigm significantly improved performance compared to non-inverting paradigms (41.70 vs. 34.57, p=0.0002)). Overall, using a Naïve Bayes classifier and an inverting paradigm on average doubled ITRs compared to the original SWLDA non-inverting paradigms traditionally used in p300 spellers (41.70 vs. 21.23). Only one subject (J) showed a decrease in ITR from non-inverting to inverting using SWLDA classifier by 0.77%. The same subject showed an increase in ITR by 7.52% using the Naïve Bayes classifier.

As mentioned earlier, the order in which inverting and non-inverting trials were performed was randomized in order to minimize the effect of a practice effect. Out of the 15 subjects, 7 performed inverting before non-inverting, while 8 performed non-inverting first. The average increase in ITR for inverting versus non-inverting was 7.12 bits per minute for the Naïve Bayes classifier. For those who performed inverting first, the average ITR increase was 6.49 bits per minute (standard deviation = 6.16); while for those who performed non-inverting first, it was 7.68 bits per minute (standard deviation = 5.46). The difference in the ITR increases was not significant between the subjects who performed non-inverting first versus those who performed inverting first in either of the classifiers (p value = 0.69). Thus, the order of paradigm presentation did not have a significant effect on the amount of ITR improvement.

Subject	SWLDA Non-Inverting	SWLDA Inverting	Naïve Bayes Non-Inverting	Naïve Bayes Inverting
A	29.02	32.92	45.69	52.47
B	18.05	20.72	27.88	37.02
C	17.12	31.59	26.11	42.83
D	12.85	16.98	24.44	29.81
E	16.57	20.62	32.09	35.26
F	25.67	35.70	45.76	46.21
G	3.83	6.28	8.77	10.14
H	16.10	25.67	27.86	41.33
I	31.59	32.92	48.89	52.57
J	25.38	24.61	37.12	39.92
K	20.89	30.29	33.98	46.11
L	31.41	48.30	44.33	62.75
M	37.15	43.59	54.90	57.87
N	17.83	21.77	33.38	37.97
O	15.04	21.95	27.34	33.18
Average	21.23	27.59	34.57	41.70
p value	0.000187		0.000228	

Table 2: Offline results for inverting and Non-Inverting trials using the SWLDA and Naïve Bayes classifiers.

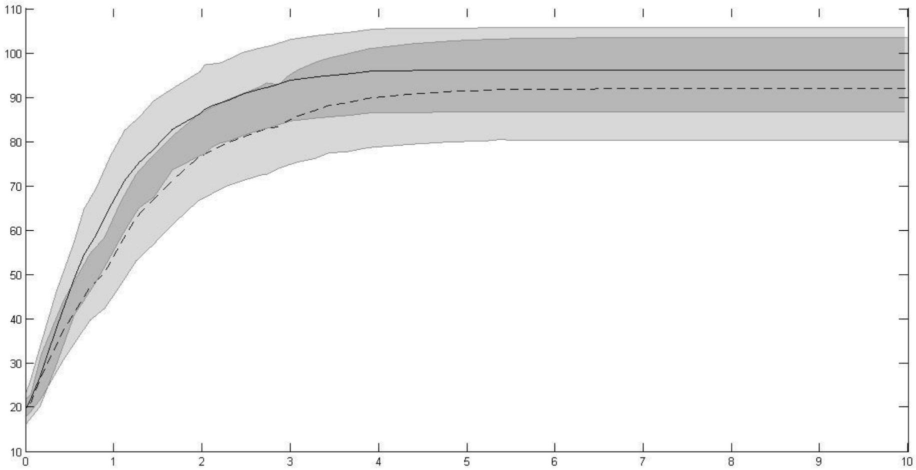

Figure 11: Accuracy versus number of sequences for Non-Inverting (dashed lines) and Inverting (solid lines) using offline analysis.

Subjects using the Naïve Bayes classifier with inverting flashing show a rise in average accuracy to approximately 95% in about 4 sequences (Fig. 11). Non-inverting flashing using Naïve Bayes achieves less than 90% accuracy in the same number of flashes.

For the inverting paradigm using Naïve Bayes, nine subjects reached 100% accuracy. On an average, these nine subjects took 2.56 flashes to reach 100% accuracy, the lowest number of flashes being 0.98 (subject L). In case of the non-inverting paradigm, four subjects reached 100% accuracy, taking an average of 2.45 flashes. The lowest number of flashes in this case were 1.43 (subject M).

b. Online Analysis:

Online analysis was performed on 10 out of the 15 subjects. There was an increase in the average accuracies between non-inverting and inverting from 82% to 91.2%, which approached statistical significance (p=0.078). One subject (J) performed poorly with Inverting versus Non-Inverting. Six subjects performed better with Inverting compared to Non-Inverting, while three others remained constant. Two of these three, reached 100% accuracy with both the paradigms. Non-Inverting paradigm had only these two subjects reaching 100%, while Inverting had three others as well. The worst performer achieved accuracies of 40% and 72% respectively with Non-Inverting and Inverting paradigms.

	Non-Inverting		**Inverting**	
	Accuracy (%)	ITR	Accuracy (%)	ITR
A	88	13.05	92	14.13
B	80	11.10	80	11.10
C	40	3.64	72	9.33
F	84	12.04	100	16.77
I	92	14.13	100	16.77
J	92	14.13	76	10.20
L	100	16.77	100	16.77
M	100	16.77	100	16.77
N	76	10.20	100	16.77
O	68	8.51	92	14.13
Average	82	12.03	91.2	14.27

Table 3: Accuracies and Information Transfer Rates for online trials.

c. Waveform Analysis:

The EEG data was also analyzed to be able to visualize the average physiological responses observed for each paradigm. Windows of 600ms following target and non-target stimuli were separately averaged across trials and subjects, and such averages were obtained for each electrodes site. The average P300 responses across subjects show an increase in amplitude with the Inverting paradigm versus the Non-inverting in all electrodes (Fig. 12).

Five subjects with a large amount of noise were excluded from the averaging process. This exclusion was done prior to taking the grand averages of the data, by visually inspecting similar time-windowed averages and standard deviations for individual subjects, across all trials. Two criteria were used for exclusion, presence of unusually high values of DC drifts and standard deviation in the averaged data. The DC drift was observed in all channels within the five excluded subjects, although in varying amounts.

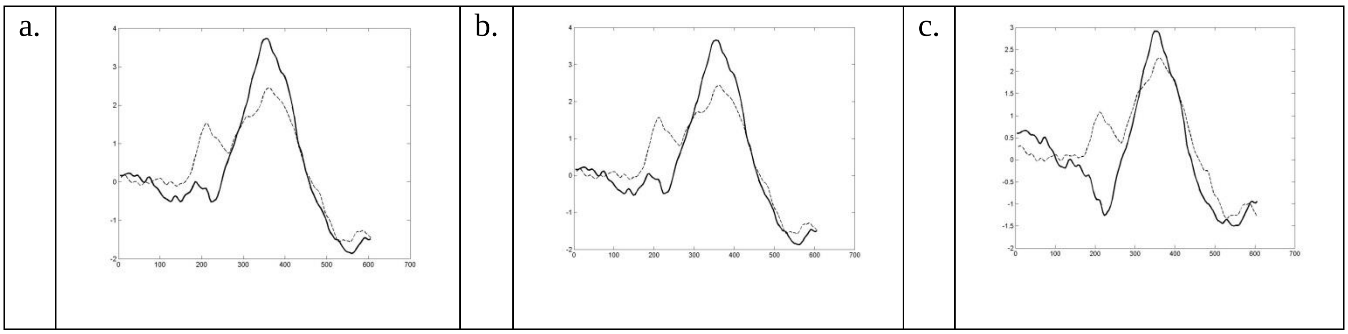

Figure 12: Average differences between target and non-target responses for Inverting (solid) and Non-Inverting (dashed) signals at the Pz, Cz, and Oz sites (a, b, and c respectively).

Peaks in average amplitude at roughly 300ms in response to target characters were observed at all electrode sites. The difference in amplitude in response to target versus non-target characters for the Inverting paradigm shows a roughly 1µV higher amplitude compared to those for Non-Inverting (Fig. 12). The occipital electrode shows a distinct negativity at about 200ms for the Inverting paradigm.

d. Classifier Analysis:

The offline results were also analyzed with respect to the scores assigned to target and non-target stimuli by the Naïve Bayes classifier for non-inverting versus inverting paradigms. These scores are indicative of how close a new signal is to the attended (target) class; the higher the score, the greater is the certainty of it belonging to the attended class. A paired t-test was performed for the differences between average target and non-target scores, for non-inverting versus inverting paradigms, across all fifteen subjects. The inverting paradigm was found to have a significantly higher score (p = 0.00094) difference between target and non-target as opposed to non-inverting paradigm.

Fig. 13 shows the normal distributions obtained by averaging the normal distributions for the scores of each subject for the target stimuli, using the Non-Inverting and Inverting paradigms. As can be seen from the figure, the average normal distribution for the target scores with the inverting paradigm is farther away from zero, than that for the target scores with the non-inverting paradigm.

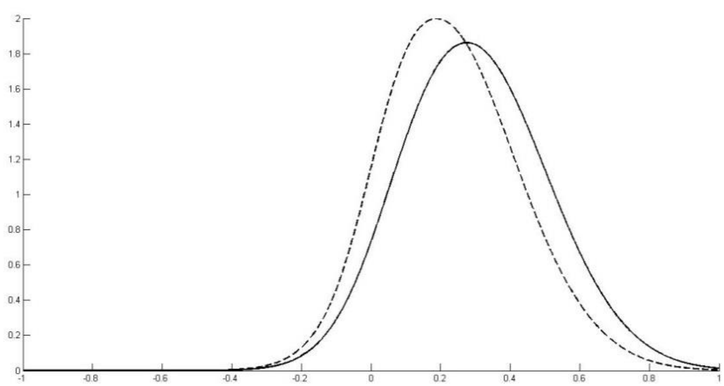

Fig. 13: Averaged normal distributions of target scores across all subjects for non-inverting (broken lines) versus inverting (solid lines) paradigms

Further, analysis was also performed on the feature weight vectors generated by the BCI2000 online classifier for non-inverting and inverting paradigms. Feature weight vectors were averaged across subjects and all electrode sites. The averaging was performed across electrode sites and not just subjects, because the classifier discards redundant weight information between electrodes and hence looking at individual electrodes might not have revealed the complete picture. Naturally, only those subjects were included for which the classifier could generate feature weight vectors. These averaged vectors indicate which features were mainly used by the classifier in discriminating between target and non-target stimuli.

The vectors generated for each of the paradigms were similar –showing a peak at 300ms, except the one for the Inverting paradigm shows a negative peak at about 200ms (Fig. 14).

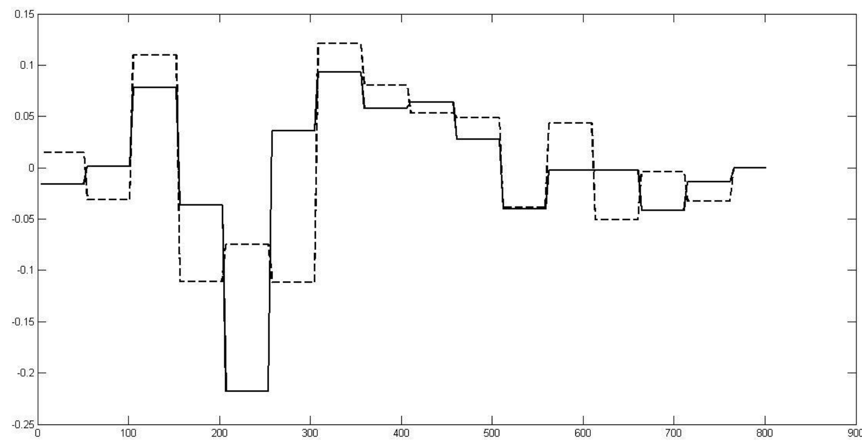

Figure 14: Feature weight vectors from the BCI2000 online classifier averaged across subjects and electrode sites for inverting (solid) and non-inverting (dashed) paradigms

IV. Discussion:

Using inverting flashing with the P300 speller as opposed to a non-inverting paradigm, results in practical improvements in ITR, which are further amplified by using a Naïve Bayesian classifier. Improved ITRs are accomplished by eliciting higher amplitude evoked response which improved classification accuracy and ITRs. The average ITR observed using the conventional method that is SWLDA classifier with Non-Inverting paradigm, almost doubled with the use of the Naïve Bayes Classifier used with the inverting paradigm. Each of the classifiers showed increases in the average ITR of 30.18% for SWLDA, and 20.59% for the Naïve Bayes classifier.

Online testing was only performed on ten of the fifteen subjects, since the BCI2000 classifier could not generate feature weight vectors for the other five. In the case of these subjects, data was acquired without providing visual feedback. Several of the subjects on whom online testing was performed, reached accuracies close to or up to 100% with both the paradigms. Half the subjects reached 100% accuracy using the Inverting paradigm, while two of those reached 100% using Non-Inverting. This might have caused a ceiling effect, which could have been the reason behind not reaching statistical significance for performance differences in the online analysis. Such a ceiling effect was not present in the offline analysis because the number of sequences that yielded the maximum ITR was selected for each subject regardless of the accuracy provided.

The ITRs and accuracies achieved in this study, using the conventional non-inverting paradigm, have been comparable to those of similar studies in the past, using the traditional row column paradigm introduced by Farewell & Donchin [12]. The average ITRs using SWLDA and Naïve Bayes classifiers offline respectively, with the non-inverting paradigm were, 21.23 bits/minute and 34.57 bits/minute using ten sequences of flashes. Speier et al. [13], reported average ITRs for SWLDA and Naïve Bayes to be 22.07 bits/minute and 33.15 bits/minute respectively, using 15 sequences of flashes for every character. They also reported an offline accuracy of 82.97% for the SWLDA classifier, which was 82% for non-inverting online in this study. Townsend et al. [19], reported an ITR of 19.85 bits/minute for the RC paradigm which is similar to the non-inverting paradigm used here. Fazel-Rezai et al. [28] reported an accuracy of 85% using the RC paradigm with 12 sequences of flashes.

The occipital electrode Oz shows a distinct negative peak at about 200ms for the inverting paradigm (Fig. 12c). A similar peak was observed with the inverting paradigm for other occipital sites as well. It is possible that this peak might be an N200 ERP. N200 is a negative going ERP elicited by an oddball stimulus, often observed together with P300 [29]. In case of visual stimuli, it is known to occur in posterior sites [29]. Kaufmann et al. observed a third of subjects controlling a P300 BCI achieve higher discrimination with an N200 at occipital sites than with P300 at centro-parietal sites [30]. Liu et al [31] also reported negative going dips at about 200ms in the SWLDA classifier for certain paradigms such as translation and rotation of characters. The feature weights found by the SWLDA classifier in our study show a similar negative going peak at about 200ms for the inverting paradigm and not for the non-inverting paradigm (Fig. 14). Apart from the increased amplitudes of the P300 response, this could be a possible explanation for the improved performance with the inverting paradigm.

There was only one subject who showed decreased performance with the Inverting paradigm. This decrease was by 28% in ITR for the online method; whereas it was 0.78% in ITR for offline analysis using the SWLDA classifier. This subject had long and dense hair, making electrode contact difficult especially at posterior sites. During the experiment, the raw signal showed increasing levels of noise in parietal and occipital electrodes compared to frontal and central ones. Poor contact at occipital sites might have hurt the Inverting paradigm more because of the hypothesized larger visual component. The greater decrease in online compared to offline, may be because online has a smaller sample size, of 25 letters (five words of five letters each) – giving an accuracy of 4% for every correct letter. Offline analysis on the other hand, had a larger sample size due to cross validation.

The threshold probabilities used for the Naïve Bayes classifier were optimized for every user individually, by selecting the threshold with the highest ITR for each subject. Thus, the threshold used for every subject was different. Regardless of this, the average accuracies achieved with the inverting paradigm were still higher for every different number of sequences and hence for every threshold probability as opposed to those for non-inverting (Fig. 11).

a. Prior Similar Work:

Previous studies have tried to change the stimulus presentation or graphical display, but could not show either significant improvements or consistency across subjects. Nevertheless, these studies showed that changes to the visual stimulus can produce considerable increments in the ITRs and accuracies. Liu et al [31] experimented with several paradigms, and reported improved results with rotation of characters for some subjects and zoom in or zoom out for others. Any single paradigm was not found to be consistently better across subjects. They suggested the necessity to explore more paradigms in an online system. Salvaris and Sepulveda [23] also demonstrated various paradigms, and consistent best performance was achieved using a paradigm in which the entire background covering the grid was white and kept static as the symbols changed from gray to black during stimulus. None of the paradigms described in the study, showed performances significantly different from each other except for the best (white background) and the worst (small symbol size) performers.

b. Limitations and Future Directions:

While the improvement described here are practical and applicable, further improvements are still possible. Combining some other visual paradigms with the Inverting paradigm, such as the zoom in, zoom out or the color method shown by Liu et al [31] can possibly result in superior results. Lu and colleagues [32] had shown optimal inter stimulus intervals and flashing timings, which were used in this study. These same parameters may or may not be optimal for the inverting paradigm, and studies done by Lu may be repeated with the inverting paradigm.

Using Inverting paradigm with colors other than black and white during the flashing of stimuli may add to the level of contrast and hence further increase the visual aspect of the paradigm. Paradigms using gray scale values intermediate to black and white may be considered and compared for performance.

This study is limited in that all the subjects were healthy individuals, with normal to corrected vision. Further, the users only perform copy spelling, in which words are presented to be typed. Using copy spelling, subjects don't have the freedom to type words of their choice. If given such a choice, the word selection could be better randomized and the subject might be more motivated to type the word; also, it would better approximate the practical application of the speller. Subject gaze was not monitored for motion in order to focus on the target character, nor were the subjects instructed not to alter gaze. In practical use of the speller, the patient may not be able to alter gaze and any dependency of the paradigm on gaze would be problematic. Future work may include gaze monitoring to ensure that consistent results are still obtained without needing the user to shift gaze towards the target character.

B. STUDY TWO: MODIFYING THE ELECTRODE SET

I. Introduction:

The signal acquisition for implementing the P300 Speller BCI system using EEG as a modality is done with 32 electrodes over the scalp (Fpz, Fz, FC1, FCz, FC2, FC4, FC6, C4, C6, CP4, CP6, FC3, FC5, C3, C5, CP3, CP5, CP1, P1, Cz, CPz, Pz, POz, CP2, P2, PO7, PO3, O1, Oz, O2, PO4, PO8), in several studies including study one. Electrode gel has to be applied or injected into each of these electrode locations for conduction. Further, each of the electrodes may have to be connected over an electrode cap individually prior to each experiment. This adds greatly to the total experiment time including the volunteer's time, not just the experimenter's. Some of this time may be contributing towards fatigue and general lowering of interest of the volunteer. The computational time and effort for selecting and weighting features from the electrode set is also increased. Overall, the use of a large number of electrodes is cumbersome, causes practical difficulties, increases experiment time, and may be prohibitive to the development of wearable, portable headsets.

Krusienski et al [26] have studied the use of four different electrode set layouts and compared their performances. This study recommended a six electrode set consisting of central and posterior sites – Fz, Cz, Pz, Oz, PO7 and PO8. However, the sets that were compared in this study were chosen empirically and were not optimized or minimalized by comparing with other sets. Speier and Pouratian [33] conducted a more systematic study which utilized Gibbs sampling in order to narrow down on to electrode sets found to best classify the EEG signals. Electrode sets starting from one and up to five electrodes were compared to the six electrode set proposed by Krusienski et al and also to the full 32 electrode set. One, two and three electrode sets produced ITR values significantly lower compared to the 32 electrode set. Using the four electrode set (PO7, PO8, POz, Oz) also produced ITR values significantly lower compared to 32 electrode set, however it was still effective. The five electrode set (PO8, PO7, POz, Oz, CP2) produced ITR values not significantly different from either the 32 electrode set or the six electrode set. The study concluded that the electrode set can be reduced to four electrodes while still achieving clinically significant results. Neither of these studies performed online analysis to verify offline results.

The present study is further incremental to the previous studies and includes online experiments to support the offline results obtained by comparing electrode sets generated using Gibbs sampling [33]. The figure below (Fig. 15) shows the three electrode sets compared in this study.

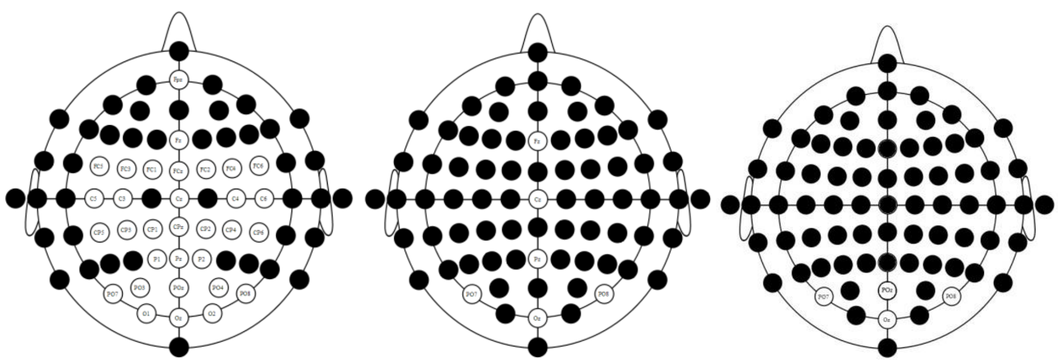

Fig. 15: The 32, six and four electrode sets [33]

II. Data Collection:

Data was collected from six healthy volunteers, males between the ages of 22 to 30 years with perfect or corrected to perfect vision. Each experiment consisted of training and testing. Training consisted of the standard sentence, ‘The quick brown fox jumps over lazy dogs’ broken into three parts with breaks in between. Testing was done in free mode, where the subjects were allowed to use a sentence of their choice. Three testing sessions of five minutes each were performed for the three electrode sets in random order with the same sentence. All sentences were either terminated at five minutes or repeated in order to run until five minutes. The hardware used for the study was the same as the Inverted paradigm study. The Inverting paradigm was used for the visual stimuli and Naïve Bayes classifier was used for classification.

III. Results:

The four electrode set achieved a higher average accuracy of 70.36% compared to the six and 32 electrode sets which achieved accuracies of up to 62.62% and 60.02% respectively. The 32 electrode set had the fastest average selection rate (7.19 selections per minute) compared to the six electrode set at 6.32 selections per minute and four electrode set at 5.83 selections per minute. The ITR values for the 32, six and four electrode sets were found to be as 18.75, 16.97, and 18.72 respectively.

Subject	**Selection Rate**			**Accuracy**			**ITR**		
	32	**6**	**4**	**32**	**6**	**4**	**32**	**6**	**4**
1	9.90	7.44	6.84	68.75	56.76	41.18	26.45	14.62	8.03
2	8.13	8.59	8.80	75.00	82.93	87.36	25.01	31.23	34.98
3	10.33	5.08	7.32	91.11	88.00	100.00	44.23	20.44	37.82
4	8.32	8.76	6.88	43.90	62.79	70.59	10.84	20.23	19.17
5	5.95	2.66	3.35	72.41	15.38	75.00	17.30	0.56	10.31
6	6.67	4.51	3.21	18.18	68.18	62.50	1.93	11.90	7.35
7	3.87	3.52	3.46	40.00	30.77	46.15	4.35	2.56	4.88
8	6.99	5.49	4.83	47.83	25.00	41.18	10.45	2.81	5.67
9	6.04	6.76	5.49	75.00	88.00	100.00	18.59	27.19	28.40
10	4.80	4.87	3.92	61.11	33.33	53.33	10.61	4.05	6.98
11	6.65	8.20	5.72	64.00	87.50	77.27	15.83	32.70	18.47
12	8.66	9.98	10.21	94.12	81.58	89.74	39.37	35.30	42.55
Average	7.19	6.32	5.83	62.62	60.02	70.36	18.75	16.97	18.72

Table 4: Selection Rates, Accuracies and ITR values for online trials

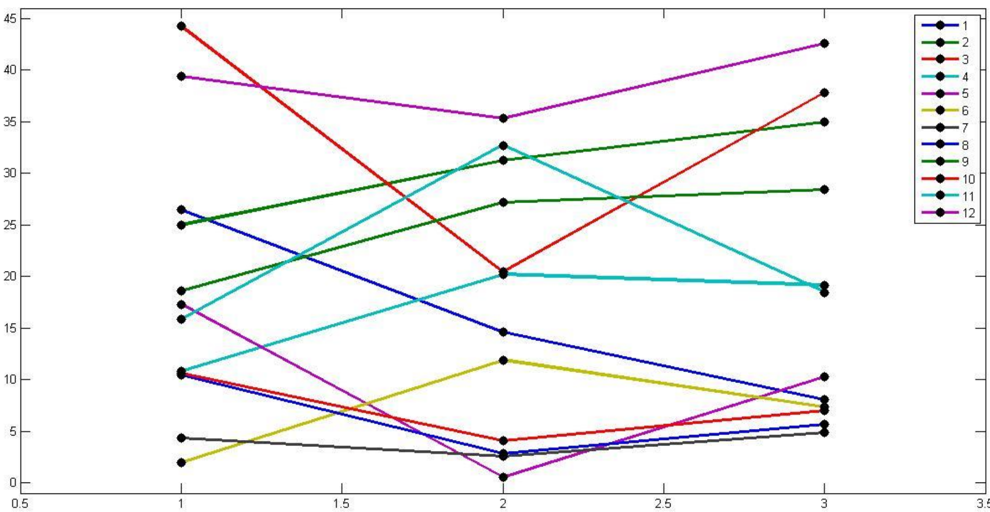

Fig. 16: ITR values for individual subjects for 32, six and four electrode sets

IV. Discussion:

There were differences in the performances achieved using the three different electrode sets with respect to the accuracy and ITR values, however no statistical significance was found. P-values were calculated for each pair of the three electrode sets – 32 and six (0.62) six and four (0.47) and 32 and four (0.99). This indicates that using four electrodes might be clinically effective, without compromising on the performance.

The approximate time taken to apply the four electrodes was about three to four minutes, for six electrodes it was about five minutes, and 20 to 25 minutes for the 32 electrode set including ground and reference. Thus a considerable amount of time was saved using the reduced electrode sets. There was however, no measure of how the reduced set up time might affect subject performance.

This study further emphasizes the importance of the visual component in the P300 speller. The four electrode set consists of PO7, PO8, POz and Oz, all of which are posterior electrodes while the central electrodes from the Krusienski electrode set, Pz, Cz and Fz are excluded. In spite of this, a comparable performance was achieved. It remains to be seen whether this affects the dependence of the speller on gaze.

The average ITR values obtained in this study are comparable to the one obtained in the Inverting paradigm study, with online analysis – 14.27 bits/minute, as well as to other online studies reported in literature. The average ITRs achieved in this study using the 32, six and four electrode sets respectively were 18.75 bits/minute, 16.97 bits/minute and 18.72 bits/minute respectively.

Krusienski et al [26], who described the six electrode set first, reported average accuracies of about 90% at a total of 15 flashes. In order to be able to fairly compare their results to those obtained in this study; we must first convert the percentage accuracy into an ITR value so that the number of flashes would also be taken into consideration. Krusienski et al used an intensification period of 100ms and an inter-stimulus interval of 75ms, with a 6x6 grid of characters. Thus, the 90% accuracy translates into an ITR of 13.9 bits/minute, which is close to the ITR achieved in this study using the six electrode set – 16.97 bits/minute.

Townsend et al conducted an online study to compare the RC paradigm to other novel paradigms, and reported 19.85 bits/minute using the RC paradigm. Ryan et al [34] also performed online analysis in their study using the RC paradigm as a baseline and achieved an ITR value of 19.39 bits/minute.

Subjects 5 and 6 performed poorly on electrode sets six and 32 respectively, possibly due to hardware issues. Except for these two, all other subjects were pretty consistent in performance across the different electrode sets, although there were a few subjects who performed relatively poorly overall such as subjects 7, 8 and 10. Subject 3 was the best performer for 32 electrode sets (44.23 bits/min) while subject 12 was the best performer for the four and six electrode set (42.55 bits/min and 35.3 bits/min respectively).

CHAPTER THREE: CONCLUSION

The current P300 speller system suffers from several drawbacks which might be limiting its use and adherence by patients suffering from ALS and other neuromuscular disorders. The two studies performed and presented here sought to reduce a few important limitations of the system, namely the low speed and accuracy as well as the electrode requirements. Poor speed and accuracy is responsible for making the system slow and is self-propagating in the sense that the slow system speed might cause fatigue and loss of attention in the user causing more errors, while the need for correcting errors might make the process of typing slower. These issues can be a cause of frustration for the user whose sole means of communicating with the outer world is the speller system. Further, the requirement of a large number of electrodes can make the system cumbersome and difficult to use. The hardware setup time and maintenance can be inconvenient for both the caretakers and patients and may limit the system's use.

The first chapter introduced the Inverting paradigm which was a novel flashing paradigm meant to increase the saliency, strength and luminance of the system's visual interface. Using the Inverting paradigm yielded significant improvements in the speed and accuracy of the system. This was possibly due to the brighter and more obvious flashing eliciting an enhanced P300 response as seen from the average responses at all electrode sites. Also, an analysis of the classifier revealed an increase in the target scores for the Inverting paradigm and a change in the feature weight vector obtained using this paradigm.

The cumbersomeness and setup time of the current system necessitated the use of a reduced set of electrodes. However, it was important to verify that the reduction in electrode number does not heavily compromise on the system performance. The 32, six and four electrode sets were compared for performance in the second study. Reducing the number of electrodes to a four electrode set maintained comparable levels of speed and accuracy, while greatly reducing the setup time, effort and hardware requirement.

The improvements achieved in these studies were on two fronts – enhancing the system performance by means of greater SNR and ITR, and increasing convenience of use for patients and caretakers. These improvements have largely enhanced the practicality of the system. Future studies may involve using the two systems together and also observing the gaze dependence of the system.

REFERENCES:

[1] Laureys S 2006 *The Boundaries of Consciousness: Neurobiology and Neuropathology* Elsevier

[2] Nunez P and Srinivasan R 2005 *Electric Fields of the Brain: The Neurophysics of EEG*, Oxford University Press

[3] http://jonlieffmd.com/blog/brain-electricity-and-the-mind

[4] Wolpaw J and Boulay C Brain Signals for Brain-Computer Interfaces. In Graimann B, Allison B and Pfurtscheller G *Brain Computer Interfaces: Revolutionizing Human-Computer Interaction* 29-46

[5] http://thebrain.mcgill.ca/flash/capsules/outil_bleu13.html

[6] http://news.psu.edu/story/141593/2006/10/02/research/moral-brain

[7] Haas L F 2003 Hans Berger (1873-1941), Richard Caton (1842-1926), and electroencephalography *Journal of Neurology, Neurosurgery & Psychiatry* **74** (1): 9

[8] Niedermeyer E and Lopes Da Silva F *Electroencephalography: Basic Principles, Clinical Applications and Relaed Fields*

[9] Gerrard P and Malcolm R 2007 Mechanisms of modafinil: A review of current research *Neuropsychiatr Dis Treat* **3** (3): 349–364

[10] Neiman E and Hanna P EEG in Dementia and Encephalopathy http://emedicine.medscape.com/article/1138235-overview

[11] Sutton S, Braren M, Zubin J and John ER 1965 Evoked-Potential correlates of stimulus uncertainty *Science* **150** (3700):1187-8

[12] Farwell L A and Donchin E 1988 Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials *Electroencephalography and Clinical*

Neurophysiology **70** (6): 510-23

[13] Speier W, Arnold C, Lu J, Taira R and Pouratian N 2012 Natural language processing with dynamic classification improves P300 speller accuracy and bit rate *J. Neural Eng.* 9 016004

[14] http://www.alsa.org/about-als/facts-you-should-know.html

[15] Dugdale DC, Hoch DB and Zieve D 2010 Amyotrophic lateral sclerosis *A.D.A.M. Medical Encyclopedia*

[16] Reza Fazel-Rezai 2011 Recent Advances in Brain-Computer Interface Systems

[17] Shi J, Shen J, Ji Y and Du F 2012 A submatrix-based P300 brain-computer interface stimulus presentation paradigm *J. Zhejiang Univ. Sci. C* 13 452-9

[18] Krusienski D, Sellers E, Cabestaing F, Bayoudh S, McFarland D, Vaughan T and Wolpaw J 2006 A Comparison of Classification Techniques for the P300 Speller *Journal of Neural Engineering* **3**(4):299-305

[19] Townsend G, LaPallo B, Boulay C, Krusienski D, Frye G, Hauser C, Schwartz N, Vaughan T, Wolpaw J, Sellers E 2010 A novel P300-based brain–computer interface stimulus presentation paradigm: moving beyond rows and columns *Clinical Neurophysiology* **121**:1109–1120

[20] Townsend G, Shanahan J, Ryan DB and Sellers EW 2012 A general P300 brain-computer interface presentation paradigm based on performance guided constraints *Neurosci Lett* **531**(2):63-8

[21] Liu Y, Zhou Z and Hu D 2010 Comparison of Stimulus Types in Visual P300 Speller of Brain-Computer Interfaces *Proc. 9th IEEE Int. Conf. on Cognitive Informatics* 273-9

[22] Takano K, Komatsu T, Hata N, Nakajima Y, Kansaku K 2009 Visual stimuli for the P300 brain–computer interface: A comparison of white/gray and green/blue flicker matrices *Clin. Neurophysiol* **120**:1562-6

[23] Salvaris M and Sepulveda F 2009 Visual modifications on the P300 speller BCI paradigm *J. Neural Eng* **6** 046011

[24] Sugg M and Polich J 1995 P300 from auditory stimuli: intensity and frequency effects *Biol. Psychol.* **41**:255-69

[25] Li Y, Bahn S, Nam C, Lee J 2013 Effects of Luminosity Contrast and Stimulus Duration on User Performance and Preference in a P300-Based Brain-Computer Interface (BCI) *Int. J. Hum-Comput. Int.* Taylor & Francis

[26] Krusienski D, Sellers E, McFarland D, Vaughan T and Wolpaw J 2008Toward enhanced P300 speller performance *J. Neurosci. Methods* **167**:15-21

[27] Vaughan TM, McFarland DJ, Schalk G, Sellers E, Wolpaw J R 2003 Multichannel data from a brain–computer interface (BCI) speller using a P300 (i.e., oddball) protocol. *Soc Neurosci Abs*

[28] Reza Fazel-Rezai, Scott Gavett, Waqas Ahmad, Ahmed Rabbi and Eric Schneider 2011 A Comparison among Several P300 Brain-Computer Interface Speller Paradigms *Clin EEG Neurosci* **42**: 209

[29] Patel S and Azzam P 2005 Characterization of N200 and P300: Selected Studies of the Event-Related Potential *Int. J. Med. Sci.* **2**:147-54

[30] Kaufmann T, Hammer E and Kübler A 2011ERPs contributing to classification in the P300 BCI *5th International Brain-Computer Interface Conference Graz,* Austria: University of Technology

[31] Liu Y, Zhou Z and Hu D 2010 Comparision of Stimulus Types in Visual P300 Speller of Brain-Computer Interfaces *Proc. 9th IEEE Int. Conf. on Cognitive Informatics* 273-9

[32] Lu J, Speier W, Hu X and Pouratian N 2013 The effects of stimulus timing features on P300

speller performance *Clin. Neurophysiol***.124**:306-14

[33] Speier W and Pouratian N 2013 Optimizing the EEG Electrode Configuration for Signal Acquisition in P300 Speller Systems. The 6th International IEEE EMBS Neural Engineering Conference

[34] Ryan D, Fyre G, Townsend G, Berry D, Mesa-G S, Gates N and Sellers E 2010 Predictive Spelling with a P300-Based Brain Computer Interface: Increasing the Rate of Communication *International Journal of Human-Computer Interaction* **27**(1):69-84