HYBRID AND MODEL BASED APPROACHES FOR NEW BCI SPELLERS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

By
Suleman Aijaz Memon
July 2019

Hybrid and Model Based Approaches for New BCI Spellers
By Suleman Aijaz Memon
July 2019

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Yusuf Ziya Ider(Advisor)

Hacı Hulusi Kafalıgönül

Yeşim Serinağaoğlu Doğrusöz

Approved for the Graduate School of Engineering and Science:

Ezhan Karaşan
Director of the Graduate School

ABSTRACT

HYBRID AND MODEL BASED APPROACHES FOR NEW BCI SPELLERS

Suleman Aijaz Memon
M.S. in Electrical and Electronics Engineering
Advisor: Yusuf Ziya Ider
July 2019

Electroencephalography (EEG) based brain-computer interfaces (BCIs), due to their non-invasive, portable and temporal resolution properties, are widely used in the field of neural engineering. In order to make BCI paradigms more practical and feasible for real life applications, new approaches are being tested such as hybrid BCIs and model based BCIs. In the first phase of this study, a novel hybrid speller BCI is proposed, incorporating P300 and code-modulated visual evoked potential (c-VEP) paradigms, with the objective of improving the spelling accuracy and information transfer rate (ITR), compared to individual P300 and c-VEP paradigms. Moreover, fusion techniques have been applied in order to effectively combine the information of P300 and c-VEP at the score level. We have implemented and compared two different approaches, linear discriminant analysis (LDA) and maximum probability estimation (MPE), in order to identify which one works best for this hybrid BCI. The proposed BCI consists of 36 targets presented as 6x6 matrix on screen with a refresh rate of 120 Hz. Seven healthy subjects participated in experiments where each subject performed a training session followed by five test sessions. The P300 and c-VEP signals are obtained by using bandpass filters of 0.5-6 Hz and 6-41 Hz respectively, on the raw hybrid data. For P300, stepwise linear discriminant analysis (SWLDA) is performed on training data from all the 10 EEG channels to obtain the feature vector. For c-VEP, canonical correlation analysis (CCA) is performed on training data to obtain the reference templates for all 36 symbols. In comparison with the accuracy and ITR values of c-VEP alone, that is without simultaneously making use of the P300 data obtained during the hybrid experiments, MPE-based hybrid improved only by 1.1% and 2.1 bits/min, on average, respectively, whereas the values worsened by 12.3% and 19.8 bits/min in the case of LDA-based hybrid. Moreover, the statistical tests on the mean accuracy and ITR values of all the subjects showed that the results of MPE-based hybrid and of c-VEP alone are

not statistically different (p=0.293). Although the MPE-based hybrid is not statistically better than the c-VEP alone, it can be highly effective if the primary goal is to only increase the accuracy, using a range of improvements in P300 methods as discussed in conclusion. However, it would not be useful if the purpose is to increase the speed of the speller since the individual c-VEP paradigm, when optimized for timing, has the capability of giving an average ITR of 114.9bits/min or higher, on its own. In the second phase of this study, model based c-VEP BCI is implemented, aimed at improving the training time compared to the case where all the targets are assigned arbitrary pseudorandom binary sequences and training is required for all the symbols separately. For this purpose, moving average model has been implemented to simulate the responses for c-VEP visual stimulation patterns, for 60Hz and 120Hz monitor refresh rate respectively. The average of the correlation between measured response and modeled response for 60Hz and 120Hz is 0.357 and 0.396 respectively. The average accuracy and ITR obtained for model based c-VEP BCI is 87.1% and 76.4 bits/min for 60Hz respectively and 82.1% and 72.4 bits/min for 120Hz respectively. Modeling results suggest that it is possible to perform a training on a single visual stimulus pattern and achieve a good fit model.

Keywords: Brain-computer interface (BCI), Electroencephalogram (EEG), P300, code-modulated visual evoked potential (c-VEP), stepwise linear discriminant analysis (SWLDA), canonical correlation analysis (CCA), maximum probability estimation (MPE), modeling, moving average..

ÖZET

HIBRIT VE MODEL BAZLİ YENI HECELEYICI BEYIN-BILGISAYAR ARAYÜZLERI

Suleman Aijaz Memon
Elektrik ve Elektronik Mühendisliği, Yüksek Lisans
Tez Danışmanı: Yusuf Ziya İder
Temmuz 2019

Elektroansefalografi (EEG) tabanlı beyin-bilgisayar arayüzleri (BBA) müdahalesiz olması, taşınabilir olması ve zamansal çözünürlüğünün iyi olması sebebiyle nöral mühendislik alanında sıkça tercih edilmektedirler. BBA paradigmalarını günlük uygulamalar için daha pratik ve uygulanabilir hale getirmek amacıyla, hibrit BBA'lar ve model tabanlı BBA'lar gibi yeni yaklaşımlar denenmektedir. Bu çalışmanın ilk aşamasında, tek başına P300 ve tek başına kod modülasyonlu görsel uyarılmış potansiyel (KMGUP) paradigmalarının elde edebileceğinen daha yüksek doğruluk oranı ve bilgi aktarım hızı (BAH) geliştirmek amacıyla geleneksel P300 ve KMGUP paradigmalarını birleştiren yeni bir hibrit heceleyici BBA önerilmiştir. Buna ek olarak, P300 ve KMGUP verilerini en son karar aşamasında skor düzeyinde etkili bir şekilde birleştirmek için füzyon teknikleri uygulanmıştır. Geliştirilen hibrit BBA için iki farklı yaklaşım, doğrusal ayrımcı analizi (linear discriminant analysis (LDA)) ve maksimum olasılık tahmini (maximum probability estimation (MPE)), uygulanmıştır ve performansları karşılaştırılmıştır. Önerdiğimiz BBA, 120 Hz yineleme hızına sahip bir ekranda 6x6 matris dizilimine sahip 36 hedeften oluşmaktadır. Her bir deneğe bir eğitim seansından sonra beş test seansının uygulandığı deneylere 7 sağlıklı denek katılmıştır. Deneylerde kaydedilen ham EEG'den P300 ve KMGUP sinyalleri sırasıyla 0.5-6 Hz bant geçiren filtre ve 6-41 Hz bant geçiren filtre uygulanarak elde edilmiştir. P300 için, öznitelik vektörünü elde etmek amacıyla 10 EEG kanalından alınan eğitim verileri üzerinde aşamalı olarak lineer ayırma analizi (stepwise linear discriminant analysis (SWLDA)) uygulanmaktadır. KMGUP için ise 36 hedef için şablon elde etmek amacıyla eğitim verisi üzerinde kanonik korelasyon analizi (KKA) uygulanmaktadır. Hibrit deneylerinde elde edilen veriden P300 verisini kullanmayıp sadece kmGUP verisini kullanarak yapılan KMGUP ile karşılaştırıldığında MPE-tabanlı hibrit paradigma, doğruluk ve BAH değerlerini

ortalamada sırasıyla %1.1 ve 2.1 bit/dk kadar arttırmıştır. LDA-tabanlı hibrit paradigma kullanıldığında ise bu değerlerin %12.3 ve 19.8 bit/dk kadar azaldığı gözlemlenmiştir. Bunlara ek olarak, bütün deneklerin doğruluk ve BAH değerlerinin ortalaması üzerinden yapılan istatistiki testler, MPE-tabanlı hibrit paradigma ve hibrit-tabanlı KMGUP paradigmasının istatistiki olarak farklı olmadığını göstermektedir (p=0.293). MPE-tabanlı hibrit paradigma her ne kadar istatistiki olarak daha iyi olmasa da temel amaç doğruluk oranını arttırmak ise bu çalışmanın sonuç bölümünde tartışıldığı üzere P300 yöntemlerinde yapılacak geliştirmeler ile beraber bu hibrit paradigmanın kullanılması oldukça etkili olacaktır. Öte yandan, amaç heceleyicinin hızını arttırmak ise bu yöntem faydalı olmayabilir, çünkü KMGUP paradigması, zamanlama açısından optimize edildiğinde, kendi başına ortalamada 114.9 bit/dk veya daha fazlası kadar bir BAH elde edebilmektedir. Çalışmanın ikinci aşamasında ise, heceleyicideki her hedefe atanan kod sekansının rastgele seçildiği ve bu yüzden bütün hedefler için ayrı ayrı eğitim gerektiği durumda, eğitim için gereken sürenin azaltılması amacıyla model-tabanlı bir KMGUP BBA geliştirilmiştir. Bu amaçla, sırasıyla 60 Hz ve 120 Hz monitör yineleme hızı için KMGUP görsel stimülasyonlarına verilen tepkeleri simüle etmek için hareketli ortalama modeli uygulanmıştır. Ölçülen tepkeler ile modelden elde edilen tepkeler arasındaki ortalama korelasyon 60 Hz için 0.357, 120 Hz için ise 0.396'dır. Model-tabanlı KMGUP BBA'nın ortalama doğruluk ve BAH değerleri 60 Hz için sırasıyla %87.1 ve 76.4 bit/dk, 120 Hz için sırasıyla %82.1 ve 72.4 bit/dk'dır. Modelleme sonuçları, tek bir görsel uyaran düzeninde bir eğitim gerçekleştirdikten sonra elde edilen veriler kullanılarak uygun bir model elde etmenin mümkün olduğunu göstermektedir.

Anahtar sözcükler: Beyin-Bilgisayar Arayüzü (BBA), Elektroansefalografi (EEG), P300, kod-modülasyonlu görsel uyarılmış potansitel (KMGUP), stepwise linear discriminant analysis (SWLDA), kanonik korelasyon analizi (KKA), maksimum olasılık tahmini (MPE), modelleme, ridge regresyonu.

Acknowledgement

First and foremost, I want to express eternal gratitude to my parents for their love and faith in me. They may not be here with me today, but remembering them gives me the strength I need to work harder and chase my dreams. Thank you, Ami and Abu.

I would like to thank my supervisor, Professor Yusuf Ziya İder of the Department of Electrical and Electronics Engineering, for his guidance, encouragement, and constant feedback. His constructive criticism as well as praise always motivated me to push my limits, and I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions so promptly.

I also want to express my gratitude to Assistant Professor Hacı Hulusi Kafalıgönül as his course on neuroscience proved to be very helpful for me during my research.

I am also grateful to the BCI Group members, Abdul Waheed, Muhammad Nabi, and Toygun Başaklar, who have been great friends and mentors, for their valuable support, constant guidance, and smart ideas in the organization and operation of the experiments. My thesis would not have been possible without some fun and they made sure I had it.

I gratefully acknowledge the funding received from the Scientific and Technological Research Council of Turkey (TÜBİTAK) under Grant 116E153 during my MSc studies.

I also want to thank my friend, Faria, who proofread my thesis several times, and for her continued support and encouragement.

Finally, completing this work would have been all the more difficult were it not for my loving sisters, Shafaque, Laila, and Sara, and guardians, Abujan and Amijan, whose presence in my life means everything to me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

1.1.1 Brain-Computer Interfaces (BCIs)

BCI technology is a powerful communication tool between users and systems. It uses the brain signals, generated as a result of a certain stimulus, to estimate the user's intent to communicate with his surroundings [1], [2]. BCI systems have numerous applications and they are commonly used in controlling a computer cursor; operating real-time drowsiness detection system for drivers; recognizing users' mental states and measuring attention levels based on the electrophysiological signals [3], [4], [5]. EEG-based BCIs, due to their non-invasive, portable and temporal resolution properties, are widely used in the field of neural engineering. The main users of this BCI system are the people with motor disabilites or peripheral nerve damage as it enables people to interact with their surroundings involving minimal muscle movement [1], [6]. Different types of EEG control signals are used for different BCI designs and among them the most popular are: visually evoked potentials (VEP) and event-related potentials (ERP) [1], [6], [7]. These control signals (VEP and ERP) are higly researched approaches due to their feasibility in clinical practices, high information transfer rate (ITR), and

less training time [6], [7], [8].

BCI system consists of five stages: signal acquisition, preprocessing, feature extraction, classification, and control interface [9]. Figure 1.1 shows the general design for BCI systems. In the signal acquisition stage, the input from the brain can be either electrophysiological activity or hemodynamic activity. The information relayed from one neuron to another by the electro-chemical transmitters is known as the electrophysiological activity [6]. On the other hand, the hemodynamic activity exploits the property where glucose is delivered to active neurons at higher rate than the glucose delivered to inactive neurons, and as a result, the oxyhaemoglobin ratio in both areas of the neurons is easily distinguished [10]. In our study, the focus is on EEG-based BCIs where the signal is acquired from the scalp and neuronal activity from different brain areas is recorded using the electrodes placed on the distinct scalp sites.

In the preprocessing stage, the recorded signals are filtered and structured in a specific form which can be used in the next stage. During the feature extraction stage, the distinguishing information is identified from raw EEG. This stage is the most challenging one as the acquired EEG contains considerable distortion due to the noise from electromyography (EMG) and electrooculography (EOG) [1]. In the classification stage, the recorded signal is classified based on the feature space which has been specified in the feature extraction stage. Finally, in the control interface stage, the command, based on the classified response, is generated to interact with the surroundings such as in controlling computer cursor, commanding a wheel chair or spelling out a word [1], [6], [11].

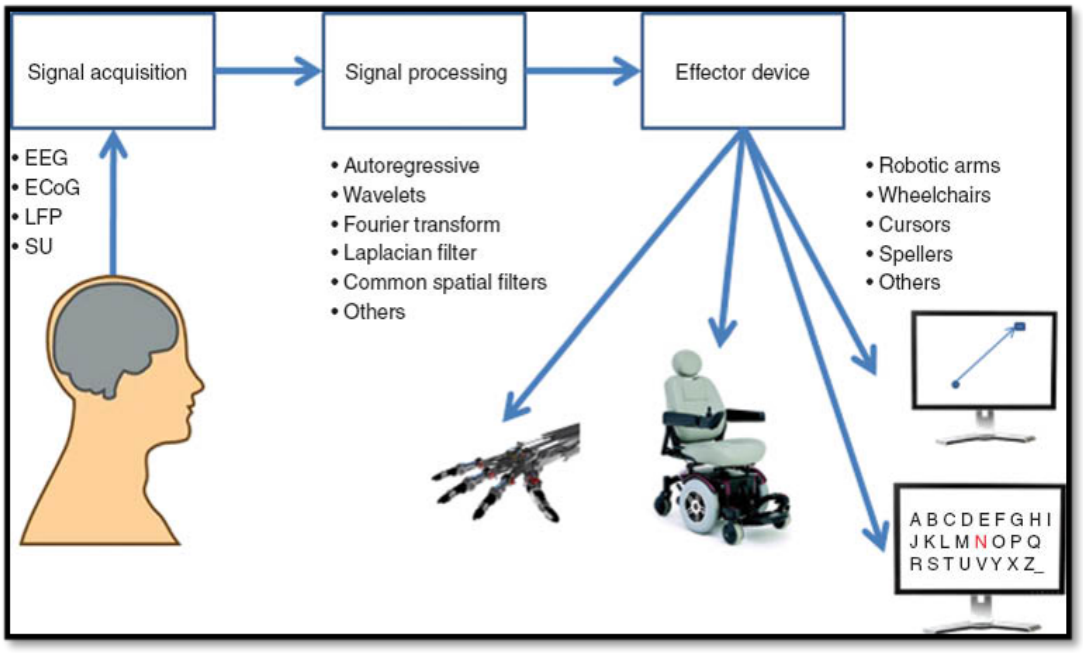


Figure 1.1: Simple flow diagram of any BCI design

1.1.2 Electroencephalography (EEG)

EEG is an electrophysiological monitoring method which is used for recording electrical brain activity caused by the currents (impulses) flowing through the synaptic connections between neurons. EEG signal is a widely used neuroimaging modality because it is non-invasive, portable, low-cost, with a high temporal resolution. EEG can easily be recorded by placing the electrodes on the scalp which can be either active or passive, and dry or wet. Its recording system comprises amplifiers, A/D converter, and a recording device (PC). After the electrodes sense the signal from the scalp, the amplifiers increase the EEG signals' amplitude in order for the A/D converter to digitize them with higher accuracy. The recording device then stores and displays the acquired data.

The EEG recording setup includes three electrodes: active, reference, and ground. Where the recorded EEG is the potential difference between active and

reference electrodes, the ground electrode is used to measure the differential voltages and to reduce the noise of the other two electrodes. Moreover, the electrodes used in this setup are usually manufactured using small tip-like pins of silver chloride (AgCl) [12]. The EEG signal amplitude is in the order of microvolts so a high gain, typically x100000, is required to achieve a good digitized output. As the EEG signal experiences background noise from different muscle activities, the contact impedance between the scalp and electrode should be kept below 10KΩ [13] in order to obtain a high quality EEG signal. For a good connectivity between the electrode and the scalp, conductive gel is used as a bridge, making this type of setup a wet electrode setup. Although this setup requires a continuous maintenance compared to a dry electrode setup (in which no conductive gel is used), the quality of EEG signal obtained is better with minimal background noise. Dry electrodes can also be used to record the EEG signal but a completely dry system might be subject to several challenges involving not just a very high impedance between electrode and skin, but amplification circuit with high input impedance as well [14]. Regardless of the electrodes used during the whole process, the user should stay stationary and certain precautions must be taken to minimize the electronic noise generated by both external sources (such as powerlines) and internal sources (such as thermal, shot, flicker, and burst noises) [15]. The EEG signal can be recorded using unipolar or bipolar measurements. In unipolar setup, the potential difference between active electrode and a fixed reference electrode is measured whereas in the bipolar setup the potential difference between two electrode sites is measured.

Furthermore, EEG signals are categorized in five sub-classes based on the frequency range of the signal: delta (δ), theta (θ), alpha (α), beta (β) and gamma (γ). These different frequency bands of EEG signals correspond to different activities or states of mind. The delta band consists of signals lying in the range of 0-4 Hz. Delta rythm is observed in adults only in the deep sleep state and it is not easy to monitor it because the low frequency muscle movements of neck or jaw can easily distort the EEG [16]. The theta band consists of signals lying in the range of 4-7 Hz. Theta rythm corresponds to concentration and cognitive activities such as mental calculations [17]. Alpha waves (8-12 Hz), on the other

hand, are observed in the occipital area of the brain, indicating visual processing, and their amplitude reduces with increasing mental effort [18]. The alpha rhythm is strongly observed in relaxing and eyes-closed state. Beta rhythms (12-20 Hz) are detected in the frontal and central lobe and are linked to motor activities; they are suppressed when motor action is performed [19]. Finally, the gamma rhythms (30-100 Hz) are associated with specific motor functions and perceptions [20].

In our experimental setup, electrode sensors are used to record EEG from scalp and the placement of electrodes is based on the International 10-20 system [21]. The system uses two reference points from the head to define the relative positions of electrodes. One of the reference point is nasion which is located in the middle of the forehead on top of the nose, and the other reference point is inion which is located under the bump on the back of the skull. The electrode locations are marked at intervals of 10% and 20% at the transverse and median planes [21] as shown in Figure 1.2. The letters in each location correspond to specific brain regions in such a way that A represents the ear lobe, C-the central region, Pg-the nasopharyngeal, P-the parietal, F-the frontal, Fp-the frontal polar, and O-the occipital area.

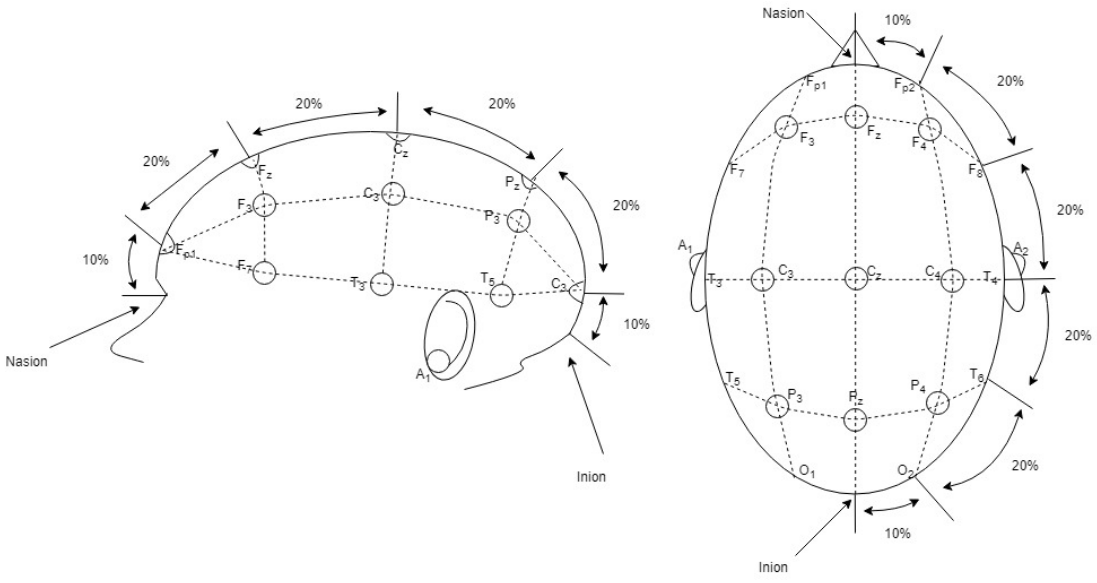


Figure 1.2: 10-20 International Electrode Placement System

1.1.3 Visual Evoked Potential (VEP)

VEP is the electrophysiological signal recorded from the position on the scalp corresponding to visual cortex [22]. VEP response is elicited in primary visual cortex when the visual field is stimulated, and the signal can be observed after averaging the recorded EEG. Evoked potentials are classified as transient evoked potentials (TVEPs) and steady-state visually evoked potentials (SSVEPs). If the visual stimulus is periodic with frequency higher than 6 Hz, the resulting VEP is called SSVEP which is typically analyzed in the frequency domain. On the contrary, TVEP response occurs when the visual stimulus frequency is less than 6 Hz [6]. Different visual stimulus patterns generate distinct VEP responses. Moreover, the input visual patterns are designed with the aim of achieving orthogonal VEP responses for different stimulus patterns in order to get reliable identification of the targets. A variety of light sources can be used as visual simulators and their flash pattern determines the type of VEP response generated. SSVEP BCIs are categorized in three groups: time-modulated VEPs (t-VEPs), frequency-modulated-VEPs (f-VEPs), and code-modulated VEPs (c-VEPs). The t-VEPs are the responses evoked in visual cortex when the input visual stimulus pattern for the targets are orthogonal in time [6]. On the other hand, the f-VEPs are the periodic responses generated when the input visual stimulus is a single frequency flash and the fundamental frequency of the response corresponds to the flash frequency [6], [8]. Finally, the c-VEPs are the evoked responses when the visual stimulus varies depending on the binary pseudorandom sequence, where bit '0' means OFF and bit '1' means ON.

1.1.4 Code-Modulated Visual Evoked Potential (c-VEP)

In recent years, the c-VEP paradigm has gained immense popularity among the BCI community due to its potential to achieve high information transfer rate (ITR) [8], [23], [24]. ITR is a standard performance metric used by BCI community to quantify the overall speed of the BCI system. In c-VEP paradigm, a pseudorandom binary coding sequence and its time-lagged version are assigned

to targets which flicker according to their own sequence. The c-VEP is observed in the EEG recorded from the occipital lobe, depending on the target that the subject fixates his gaze on [8], [23], [24]. This pseudorandom binary sequence comprises combinations of zeros and ones which correspond to the visual intensity of the stimulus. Typically, m-sequence is used as a pseudorandom binary sequence. It is nearly orthogonal to its time-lagged version and the autocorrelation function of the m-sequence is similar to the unit impulse function, making it more suitable for distinguishing between different targets in BCI application [25]. Variable (m) in the m-sequence defines the size of the linear-feedback shift registers required to generate the sequence, and that sequence's length is determined by the formula $2^m - 1$ [25].

For c-VEP BCI application, mostly monitor screens are used as a visual stimulus source with different targets (letters/numbers) presented as matrix on screen. Figure 1.3 shows the arrangement of symbols on screen. Each target on the screen modulates with the assigned pseudorandom binary sequence. The first target (A) modulates with the original m-sequence while the rest of the targets modulate with the time-lagged version of the original m-sequence [8], [23], [24]. C-VEP sequence is limited by the refresh rate of the monitor where each bit is displayed for the duration of one frame. In our project, we have used 127-bit length of m-sequence to modulate the visual stimuli. The background of the letters, presented on the screen, change the color according to their pseudorandom binary sequence. In our study, the background color flickers between green if the bit is '1', and blue if the bit is '0'. Letter A flickers with the original m-sequence; letter B flickers with 3-bit circular-shift of original sequence; letter C flickers with 6-bit circular-shift of original sequence and so on. The design in use can be seen from Figure 1.3, where all the letters' background have colors according to their pseudorandom binary sequence. For processing the c-VEP responses, the synchronizing trigger channel is required to mark the starting of the stimulation cycle [8]. Moreover, all the targets flicker with orthogonal pseudorandom binary sequence in time domain and the evoked responses are assumed to be near-orthogonal and distinguishable.

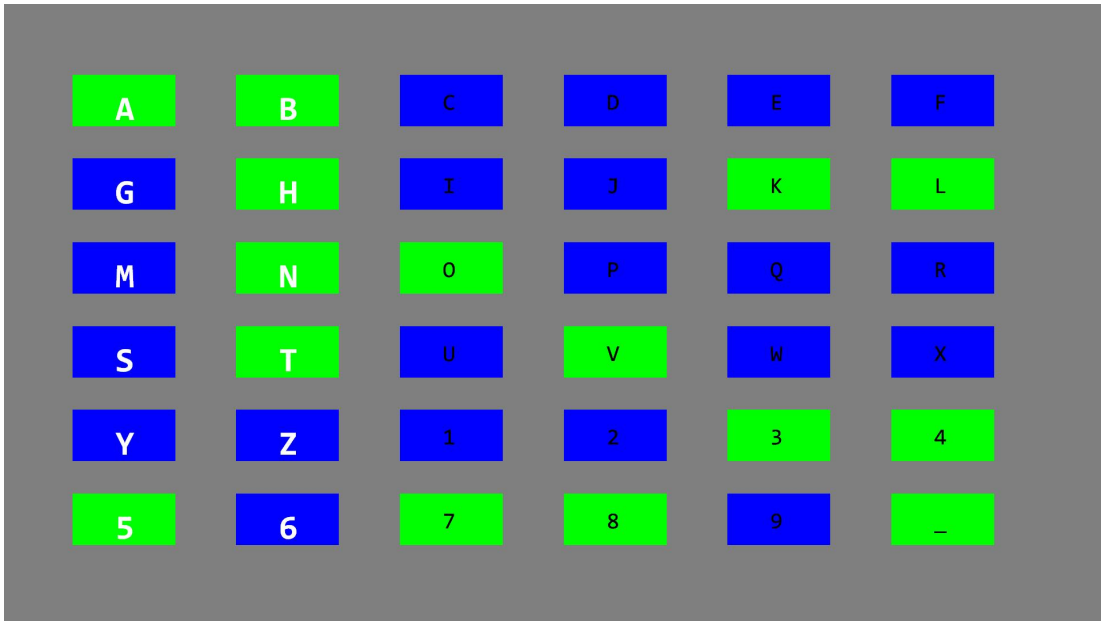


Figure 1.3: The figure shows one of the frames of our hybrid BCI speller design where 36 targets (letters/numbers) are presented as 6x6 matrix on screen with each target flickering with its different pseudorandom binary sequence. The first two columns are flashed simultaneously. If the subject performs a mental task at the onset of the flash, it generates the P300-evoked potential.

Unlike f-VEP, c-VEP requires a training stage where the user is instructed to fixate his gaze onto the reference letter (A) and EEG is recorded for the N number of repetitions of 127-bit code sequence. The reference template for the letter A is then obtained by averaging the EEG over N trials. Another benefit of c-VEP includes less training time because the templates for the rest of the targets can be obtained by circularly shifting the template of reference letter by the corresponding circular-shift introduced in the original m-sequence for each target [8], [24], [23]. During the test stage, code sequence for each symbol is repeated only once and the response is correlated with the templates obtained for each symbol. The symbol with the maximum correlation is selected as the target subject looked at.

1.1.5 P300 Evoked Potential

The P300 is a large positive signal generated in the parietal and central lobe areas of the brain. It peaks at about 300 ms in response to a mental task carried out when an infrequent stimulus occurs in an 'oddball' paradigm [11], [26], [27], [28], [29]. A typical P300 response has four peaks: N100, P200, N300, P300. The first letter of the peaks represents the positive or negative and the following number is the average latency of the peaks in millisecond [30]. In P300-based BCIs, the visual stimuli are presented as letters or symbol commands on the screen and all the rows and columns are flashed randomly while EEG is monitored [11], [26], [28], [29]. The subject fixates his gaze onto the target and depending on the flash of the target, the subject is instructed to keep the flash count in memory. As a result of this cognitive processing, occurring only at the instants when the target is flashed, P300 response is generated in the brain for the corresponding target [1], [28]. Figure 1.3 shows the first two columns flashed randomly and if the subject performs the mental task in response to the visual stimulus, it generates the P300 response.

The P300 potentials can be recorded without any training but the variations reported from one subject to another can not generalize the system for all the users [28]. Hence, training is performed on every subject to calibrate the spatiotemporal filters which helps the system adapt to the subject-specific P300 brain responses [28]. In a 36 letter/number matrix on screen, there are 6 rows and 6 columns, denoting that one single trial consists of 12 flashes of row/column. In the test stage, a single trial of stimulus is not sufficient to provide a reliable detection of the target. Although, large number of trials during the test increase the accuracy of the system, they cause the ITR to decrease significantly [6]. The advantage of P300 BCI is that it depends on the cognitive processing unlike VEP responses which primarily depend on the subject's gaze and vision [6].

1.1.6 Previous Hybrid BCIs

Recently in BCI community, researchers have focused on hybrid BCIs which combine two paradigms, making the system more accurate, fast, and reliable in use [31], [32], [33], [34]. Different physiological signals can be used to develop a hybrid BCI such as EEG, electrocardiogram (ECG), electrooculogram (EOG) or any other hemodynamic signal. Moreover, in the initial stages of the development of hybrid BCIs, researchers have put more focus on P300, SSVEP, event-related synchronization (ERS), event-related desynchronization (ERD) and sensory motor rhythms. The decision-making process involves the combination of two or more physiological measures mentioned above so that the patient's intent could be communicated in a more accurate manner. Moreover, if the patient responds adequately to both the combinations of these physiological measures, the information throughout will be improved [32].

A large number of studies have been carried out concerning hybrid BCIs, including the one by Luth et al. [35] which proposes a hybrid system incorporating P300 and SSVEP for a rehabilitation robotic system control. Moreover, in 2010, Pfurtscheller et al. [36] investigated the performance and feasibility of the hybrid BCI based on ERS and SSVEP. It was aimed at creating an efficient brain-switch (detection of a single brain pattern) with minimal or zero false activation rate [36]. Allison et al. [37] introduced a hybrid BCI based on ERD and SSVEP for a 2-D cursor control. Furthermore, the same application of 2-D cursor control based on motor imagery and P300 was introduced by Long et al. [34]. In addition to all these important developments in the study of hybrid BCIs, in 2013, Yin et al. [38] proposed a hybrid BCI speller combining P300 and SSVEP paradigms which involved the letter/symbol matrix. Significant improvement in accuracy was observed with this combination as online classification accuracy of P300 was measured as 68.65%; the hybrid's was 93.85%. The ITR of the P300 was measured as 33.8 bits/min and of the hybrid was 56.44 bits/min [38]. Yin et al. [39] also researched on different designs of hybrid BCI speller and reported the maximum average accuracy of ITR being 95.18% and 50.41 bits/min.

1.1.7 Model Based Paradigms for BCI Design

Researchers have also focused on modeling the brain responses which will allow us to study the transient and steady-state responses of the brain for a different visual stimuli [40], [41], [42], [43]. Robinson et al. [40] proposed a physiological based corticothalamic model aimed at reproducing different features of EEG which includes ERPs and discrete spectral peaks in all the frequency bands of EEG. Roberts et al. [41] uses Robinson's model to explain the non-linear behavior of cortical activity when periodic visual stimulus is presented to a subject. Robinson's model consists of large number of second order filters and non-linear sigmoidal blocks. High non-linearity due to feedback loops and second order filters make it difficult to optimize the parameters in such neural models, therefore, linear models are proposed. Zhang et al. [42] proposed second order linear system to model the transient response of SSVEP. Nagel et al. [44] proposed a moving average model which predicts the brain responses to arbitrary visual stimulus patterns. In the same year (2018), Safi et al. [43] proposed a Box-Jenkins model for the SSVEP recognition.

In online BCI applications, target classification is performed during test stage. Nagel and Safi have used model based approach for classification of targets and reported improvements in accuracy of BCI systems [44], [43]. Model parameters are optimized during the training stage and responses for all the targets are simulated using the model. In test stage, the response for a specific target is correlated with all the simulated responses and by looking at the maximum correlation, target letter/symbol is chosen. This model based approach for classification of targets has an advantage of reduced training time and increased number of targets. In c-VEP BCI explained before we mentioned the use of m-sequence and the reference templates are obtained by circularly shifting the template of letter 'A' which reduces the training time. However, if we use arbitrary binary sequences for all the letter/numbers we would need to perform training for all the symbols separately. In that case, following the model based approach which optimizes a model on arbitrary visual input sequences, we can reduce the training to only one letter. As a result, we can also increase the number of targets because no additional training

would be required. Furthermore, these models can be used to study different variations of visual stimuli and identify the optimal input sequences that brain responds to significantly. Although, there has not been much study on the modeled responses of c-VEP, we can use the linear and non-linear models proposed for SSVEP to explain the c-VEP responses observed in our experiments.

1.2 Objective and Scope

Inspired from studies detailed in the previous section, we propose a novel hybrid BCI speller with 36 targets (letters/numbers) presented as a 6x6 matrix on screen. The proposed hybrid BCI speller uses P300 and c-VEP paradigms, and this thesis investigates its feasibility and performance. Predominantly, c-VEP paradigm is considered as the most efficient BCI paradigm since it requires less training, provides high ITR, and allows for application with a large number of targets [8], [23], [24]. It is already evident from previous studies that the ITR of the system increases at higher stimulus presentation rates such as 120 Hz and 240 Hz. Although the ITR increases, high accuracy might not be achieved with a single trial. Thus, for high accuracy the number of trials need to be increased which in turn reduces the ITR significantly. Therefore, the goal of this study is primarily to improve the accuracy of the BCI speller while maintaining a high ITR.

For this purpose, experiments have been carried out in order to test the feasibility of the P300 and c-VEP hybrid system. Refresh rate of the monitor is set to 120 Hz which allows us to achieve high ITR and improved accuracy [24]. During the whole procedure, accuracy of the BCI systems and the ITR is used as the main performance metric. The experimental data is acquired using the hybrid paradigm after which observations are made in order to identify the behaviours of P300 signal and c-VEP. Moreover, accuracy and ITR of both c-VEP and hybrid design are compared during the offline analysis. In our study, different fusion techniques used for combining the P300 and c-VEP features, are also compared.

In the second part of this thesis, we have implemented a moving average model to simulate the c-VEP response which are further used in the online experiments for classification of targets. Forward model (input code sequence is used to predict EEG) for 60Hz and 120Hz monitor refresh rates are implemented and compared. Our goal is to fit a linear model which can improve the training time compared to the case where training is required for all the symbols.

1.3 Organization of Thesis

This thesis has three chapters: materials and methods, results, and discussion/conclusion.

In chapter 2, (Materials and Methods) the experimental setup, procedure, and design is discussed in detail. It discusses the methods used for the preparation of subjects and the data acquisition. Furthermore, the processing of data: preprocessing, P300 data processing, c-VEP data processing, and different fusion techniques are explained in this section. Next, we explain the implementation of the forward and backward models.

In chapter 3, (Results) the results obtained from the experiments are analyzed and compared with the results of different c-VEP and hybrid paradigms. The quality, pattern of the signals and the results obtained using different fusion techniques are discussed in this section. Furthermore, comparison is drawn between classification results of linear models and classification results of online experiments.

In chapter 4, (Discussion and Conclusion) the general discussion is given and results are summarized.

Chapter 2

Materials and Methods

2.1 Hybrid Stimulus Design

The hybrid speller consisted of 36 targets (letter/numbers) presented as a 6x6 matrix on screen. Random flashes of row and column, as well as flickering of the background of each target according to its pseudorandom binary sequence, elicit the P300 and c-VEP responses in specific areas of brain respectively. The default background screen of the speller was gray and each target was assigned a rectangular cell of 90 x 180 pixels.

Figure 2.1a shows the c-VEP stimulus design where each target is displayed on a green or blue background. The backgrounds of targets were arranged according to the corresponding bit of each target's binary sequence, where bit value '1' corresponds to green background and bit value '0' to blue background. In c-VEP studies, usually the code length of 63-bit is used, allowing the stimulation of maximum 32 targets with a 2-bit lag. In order to stimulate 36 targets on screen, a 127-bit long m-sequence is required. The m-sequence assigned to reference letter (A) is as follows: 10100011110010001011001110101001111101000011100010010011011010 11011110110001101001011101110011001010101111111000000100000110000. The original m-sequence was shifted circularly, one after another, by 3-bits and the

resulting 36 orthogonal sequences were assigned to the letters/numbers in alphanumeric order. Since there has not been enough research performed on the optimum time lag of the pseudorandom sequence, arbitrarily 3-bit lagged versions were chosen. Refresh rate of the monitor, in our experiments, was set to 120Hz so a single trial of c-VEP took $\frac{127}{120} = 1.0583s$.

In P300-based speller BCIs, each of the 6 rows and 6 columns of the stimulus were randomly flashed once during a single trial, with the total number of flashes being 12 in one trial. Figure 2.1b shows the P300 stimulus design where 2 rows were flashed simultaneously. For a single row/column flash, we had active and passive duration of a flash. Single flash had the active time where font color became white and the passive time was where the font color returned to default black. P300 flash time was mostly chosen between 120ms-150ms for active and passive times which made the single flash time of 240ms-300ms. In our experiments, we set the active time to 100ms and the passive time to 150ms. For 12 row and column flashes, the run time for a single trial of P300 became $0.25 * 12 = 3s$.

Since P300 would dominate the timing of the hybrid design, the P300 design was modified to reduce the single trial time by half. By grouping each row/column with the following row/column (i.e. first and second row, third and fourth row, fifth and sixth row), the total number of flashes became 6, consisting of 3 groups of rows and 3 groups of columns. For simplicity, we will refer to the groups of rows and columns as a single row and column. The run time for a single trial of P300 became $0.25 * 6 = 1.5s$. Although we can reduce the timing further by limiting the active and passive times, the user in that case would not be able respond to the stimulus correctly. If only P300 paradigm is considered, we had only 9 targets which are groups of symbols with each group containing 4 symbols. Figure 2.2 shows the grouping of rows, columns, and targets. P300 and c-VEP stimulus designs were superimposed to make a hybrid speller as shown in Figure 2.1c.

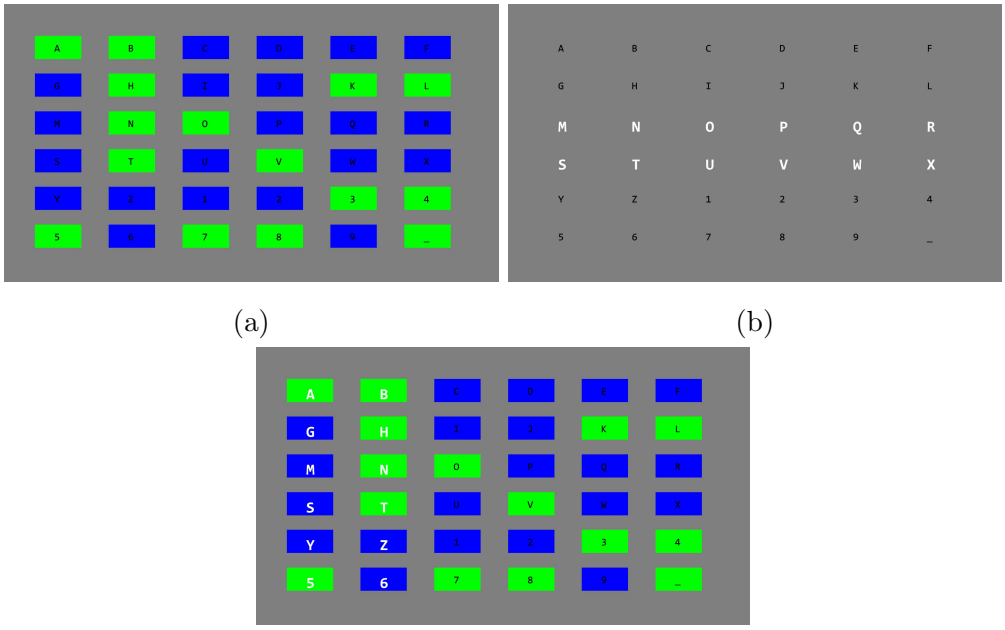


Figure 2.1: The figure explains the design and the superimposed hybrid frame. Note that all these frames were captured at different time instants during a test session just to give an idea about the design. The monitor refresh rate was set to 120 Hz and each frame was displayed on the screen after $\frac{1}{120} = 8.3ms$. (a) We turned off the P300 flashes and only the c-VEP stimulus was shown. Here the cell background of each target is green or blue depending on the corresponding bit sequence of each letter/symbol, green if the bit value is '1' and blue if the bit value is '0'. (b) This frame shows only the P300 stimulus and at that instance only, second and third rows were flashed. Single flash had the active time (font color becomes white) of 100 ms and the passive time (font color returns to default black) of 150 ms. (c) This frame shows the superimposed hybrid stimulus frame. P300 set the timing of the hybrid stimulation sequence; in the training stage, the seqeunce was repeated for 50 P300 trials and in the test stage, only a single trial was performed.

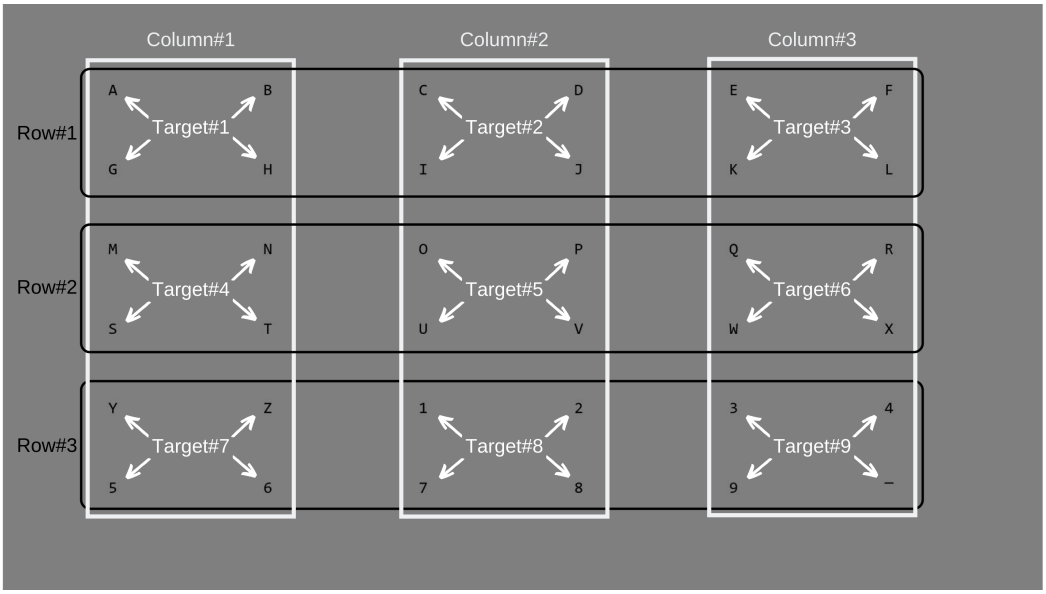


Figure 2.2: Default frame of the hybrid speller design showing the grouped, rows, columns, and targets for P300 Paradigm.

MATLAB (The MathWorks, Inc., Natick, MA, USA) were used to design this BCI speller. Psychtoolbox-3 libraries were also used which provide the interface between computer hardware and MATLAB [45], [46], [47]. Moreover, the operating system Ubuntu 16.04 with a low-latency kernel was used allowing the psychtoolbox to reliably synchronize with the vertical screen retrace. Our setup consisted of two computers: stimulator station and recording station. All the code for the design mentioned above ran on the stimulator computer which, for synchronicity, also sent marker pulses simultaneously to the recording station.

2.2 Experimental Setup

2.2.1 Subjects

The experiments were performed on 7 healthy subjects (6 males, 1 female) with a mean age of 24.5 (range 19-26). All the participants had normal or corrected-to-normal vision. Three of the participants had previous experiences with the c-VEP based BCIs and the rest of them were novices. The study was authorized by the ethical committee of Bilkent University. Complete design of the BCI system and major details of the experiment were explained beforehand and the subjects were also asked to sign an informed consent form. In addition, subjects were asked to fill a questionnaire, which basically asked the users for any visual impairments, neurological diseases, epileptic history, or migraines.

2.2.2 Data Collection

For the EEG recording, actiCAP and V-Amp 16 channel EEG amplifier (2000 KHz sampling rate) by Brain Products (Brain Products, Gilching, Germany) were used. The actiCAP has 32 electrode sites and is designed according to the International 10-20 system. Moreover, 12 electrodes were applied in the experiment which spanned the occipital, parietal, central and frontal areas of the brain. EEG was recorded using the active wet electrodes from the electrode sites that corresponded to O1, Oz, O2, P3, Pz, P4, P7, P8, CP1, CP2, Cz, and Fz. All the signals were referenced to FCz electrode and the ground electrode was connected to the nasion on the forehead. Ten of the channels were used for P300 detection (O1, Oz, O2, P3, Pz, P4, CP1, CP2, Cz, and Fz) while for c-VEP detection, we used eight channels (O1, Oz, O2, P3, Pz, P4, P7, and P8). The ImpBox (Brain Products, Gilching, Germany) was used to monitor the electrode impedance which was kept below 20 kΩ for all the experiments. From the two computers used in our experiments, the recording computer used BCI2000 [48]

platform and FieldTrip buffer [49] to interface the stimulator computer and V-Amp.

2.2.3 Experimental Procedure

Since connectivity is crucial to the accurate recording of EEG signals, a small amount of conductive gel was applied onto the scalp before the electrodes were placed, in order to ensure better connectivity. For this experiment, EEG caps of different cuts and sizes could be used and in our lab, we worked with two different-sized caps: 56cm and 58cm. Each participant chose the EEG cap appropriate for their head size, which was then positioned carefully in a way that the O1, Oz, and O2 electrodes were placed exactly above the bony lump on the back of the head. The participant was seated approximately 80 cm away from a 25-inch LED monitor. In our experiment, Dell Alienware (AW2518HF) monitor was used which has a screen refresh rate of up to 240 Hz and a resolution of 1920 x 1080 pixels. While setting up the cap on subjects, we explained and displayed the stimulus design to familiarize them with the experiment.

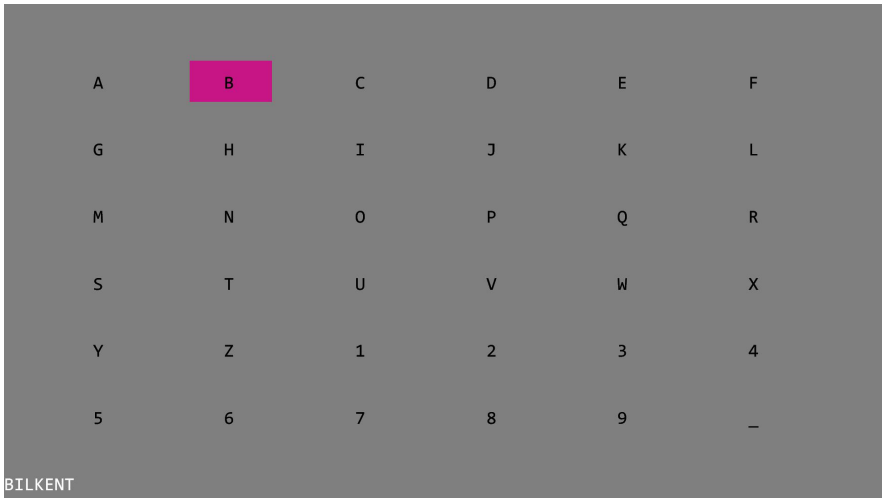


Figure 2.3: This frame shows the cue presented to the subject so that he should shift gaze and focus on the pink highlighted symbol. In our experiments, only 1 second was given for gaze-shifting. This frame was captured during the test stage which also provided the feedback (decided target) from the recording computer on the bottom left corner. However, in the training stage no such feedback was provided; only the user was given the cue to focus his gaze and the sequence started in a standard manner.

Firstly, training was performed on the reference letter 'A'. As soon as the experiment started, the subject focused his gaze on the cue presented on the screen as shown in Figure 2.3. After exactly one second, the visual stimulation sequence started, which is basically the flickering of the rectangular cells of each letter/number between green and blue color as can be seen in Figure 2.1c. In addition to the flickering, letters/numbers in each row and column were flashed 50 times as shown in Figure 2.1c. The subject was instructed to keep a mental count of random flashes of the focused letter which was flashed 100 times (e.g., one with each row flash and one with each column flash). Single trials of P300 and c-VEP took 1.5s and 1.06s respectively. In the training session, 50 P300 trials were performed which lasted for $1.5 * 50 = 75s$. Timing diagram in Figure 2.4 explains the stimulation sequence in more detail.

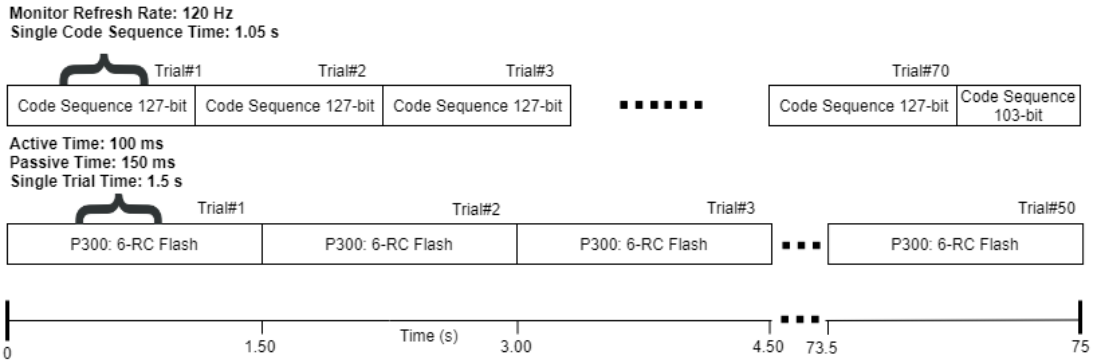


Figure 2.4: Single c-VEP code sequence took 1.06s and single P300 flashing sequence took 1.5s. In c-VEP, each bit was displayed for 8.3ms and single P300 flash was displayed for 250ms. In the duration of a single flash of P300, $\frac{250}{8.3} = 30$ c-VEP bits were displayed. The training session lasted for 75s and during that time, 50 P300 trials and 70 trials of 127-bit length c-VEP stimulation code were repeated. In addition to the 70 trials of c-VEP, we displayed starting 103 bits of the stimulation code to complete the 75s of the training session. However, the c-VEP responses from the last 103 bits were discarded in the signal processing stage.

In the test stage, the subject was supposed to spell the 20-letter sequence "BILKENTBCIEXPERIMENT". Each subject performed 5 test sessions with a break of 2 minutes after every session to avoid fatigue. The subject was instructed to follow the cue presented on the screen. After 1 second of the cue, the stimulation sequence started and the user focused on the target letter. In the test stage, only a single P300 trial was performed for each letter and this time, the subject kept a mental count of 2 for each letter. Timing diagram in Figure 2.5 explains the stimulation sequence for a single letter in more detail. The stimulus computer signalled the recording station as the stimulation for a single letter finished. All the signal processing was done on the recording computer and the response after classification was sent back to the stimulus computer as a feedback to the subject, which was printed on the left bottom corner of the screen as shown in Figure 2.3. As soon as the stimulator station received the feedback (decided target) from the recording computer, the next letter was highlighted and the subject had 1 second

to shift his/her gaze to the cued target. The stimulation, processing, classification, and feedback of a single letter lasted for 2.6 seconds and a single test session for 20 symbols took about 52 seconds.

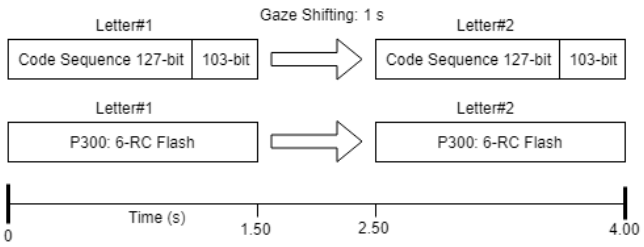


Figure 2.5: This figure shows the timing diagram of the test session for 2 letters. After 1 second of the cue presented on the screen, the code sequence started and the stimulation sequence continued for 1.5s. Again, for each letter, we repeated the starting 103 bits of c-VEP code sequence to complete the 1.5s test sequence for single letter. After the completion of the sequence, the stimulator computer raisec the TCP flag to notify the recording station. The recording computer performed the processing and sent the classified target letter back to the stimulator computer which was displayed on the left bottom corner of the screen. As soon as the feedback was received, stimulator computer highlighted the next symbol and the subject had to shift the gaze in 1 second after which the stimulation sequence for the second letter started. Note that during the processing stage, the responses from last 103 bits corresponding to c-VEP data were discarded for all the letters.

2.3 Signal Processing

As shown in Figure 2.6, the recorded EEG was processed simultaneously for the extraction of P300 and c-VEP features. The V-Amp provided the built-in high pass filter with cut-off frequency of 0.1Hz which eliminated the dc offset and other low frequency variations caused by the body movement. Moreover, marker channel on the stimulator computer sent a pulse which became high at the onset of every P300 row or column flash, and remained high for the duration of the active time (100ms). Details of data processing are explained in the following

sections.

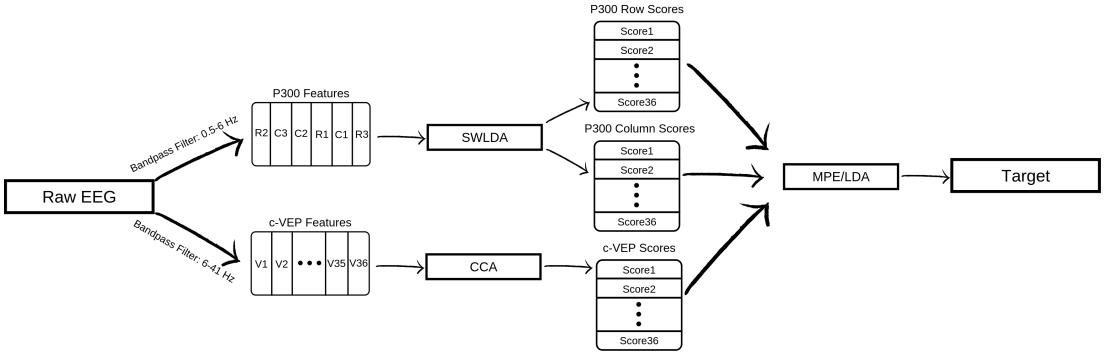


Figure 2.6: Schematic of the Hybrid Speller BCI incorporating P300 and c-VEP

2.3.1 P300 Detection

For P300 detection, signals from 10 electrode sites were considered: O1, Oz, O2, P3, Pz, P4, CP1, CP2, Cz, and Fz. P300 was a low frequency response so the recorded EEG was band-pass filtered with a cut-off frequency of 0.1Hz - 6Hz, thus eliminating the noise and other artifacts. As P300 had a maximum peak between 300ms and 500ms, 800ms-long epochs were taken for each row or column flash.

During the training, two P300 responses were elicited in the brain for a single trial. In this training stage, 50 P300 trials were performed which consisted of $50 * 6 = 300$ flashes. Among them, 100 were target responses (containing P300 signal) and 200 were non-target responses. Each epoch was taken to be 800ms long (1600 samples), starting from the onset of the corresponding marker of the flash. Two 3D arrays named $targetDataset$ and $nontargetDataset$ were formed and the sizes of the arrays were 100 x 1600 x 10 and 200 x 1600 x 10 respectively, where first dimension was the number of epochs, second dimension the epoch length, and third dimension the number of channels. EEG recordings were made at a sampling rate of 2KHz and to reduce the feature vector size, we down-sampled it to 20Hz by selecting every hundredth sample from the data. The feature vector was obtained by concatenating the data of all channels for each epoch. As a result, we obtained the new 2D arrays, $trainDataTarget$ and $trainDataNonTarget$,

with their sizes becoming 100 x 160 and 200 x 160 respectively.

Identifying P300 is a classification problem where one class is a target (P300 response) and the other class is a non-target. Different methods are used for P300 classification including linear discriminant analysis (LDA), step-wise linear discriminant analysis (SWLDA), support vector machine (SVM), matched filter, and independent component analysis (ICA). LDA is a widely-used method in BCI community as it requires less computation and provides good accuracy [50], [51]. Using the training data, LDA defines a hyper-plane which distinguishes the targets and non-targets based on their 160 features. It projects all data onto the hyper-plane which maximizes the distance between the two means of the projections. If there are C number of classes, LDA projects the data onto the hyper-plane with $C-1$ dimension. In our case there were two classes, target and non-target, and LDA projected the data in $\mathbb{R}^{160}$ space onto a line ($C-1=1$, where $C=2$) defined by the following equation:

$$y(x) = w^T x \tag{2.1}$$

where, x is the feature vector of size 1 x 160, and w is the weight vector. Note that y in equation 2.1 is a scalar not a binary label. Once the data is projected, we can decide on the threshold which separates the two classes. However, the P300 response was elicited for only one row/column, and we could distinguish the target directly by looking at which row/column had the maximum projected value [51].

$$target_{row} = \arg\max_{rows} [w^T x_{rows} + w_0] \tag{2.2}$$

$$target_{column} = \arg\max_{columns} [w^T x_{columns} + w_0] \tag{2.3}$$

Fisher's LDA is based on maximizing the ratio of the mean between the classes to the variance within the classes [52]. For two-class LDA, the objective function

is defined by the following equation:

$$J(w) = \frac{\mid \mu_1 - \mu_2 \mid}{\sigma_1^2 + \sigma_2^2} \tag{2.4}$$

where, μ_1, σ_1^2, μ_2, σ_2^2 are means and variances of the projections (y) of the data samples of two classes respectively. The solution for the weight vector, obtained by maximizing the objective function, is as follows:

$$\hat{w} = (\Sigma_1 + \Sigma_2)^{-1}(\mu_1 - \mu_2) \tag{2.5}$$

where, Σ_1, Σ_2 are the covariance matrices of two classes.

SWLDA is an extended version of LDA. It performs the feature selection (eliminates the less significant features) and achieves the best performance among all other techniques [53]. In this method, the features (predictors) were added one by one in the model and based on the p-values, the predictors were added or removed. Initially, SWLDA selected the most statistically significant feature with $p < 0.1$ and added it to the model. After that, new predictor was added and the p-value was computed again; if $p < 0.1$ for the new predictor, it was included in the model. Third step was to perform the backward step-wise analysis, which computed the p-value of each predictor by excluding only one predictor at a time and keeping the rest; the predictors with $p > 0.15$ were removed from the model. This iterative process was repeated until all the predictors satisfied the p-value criteria or the program completed the specified number of iterations [53].

In the test stage, a single trial P300 was performed for each symbol. Using the optimized weight vector obtained during the training session, the test data was projected onto a line. In MATLAB, for the 6 row/column flashes, we had the data arranged in a matrix called X with the size of 6 x 160. Equation 2.1 was used to project this test data and obtain 6 scalars. These 6 scalars were used as scores for the targets. We denoted these scalars as y_z where $z \in [1, 2, 3, 4, 5, 6]$.

For P300 detection, two type of scores were computed for each symbol: $score_i^{row}$ and $score_i^{column}$ where $i \in [1, 2, ..., 36]$. For $score_i^{row}$, all the symbols belonging to row1, row2 and row3 were assigned y_1, y_2, and y_3 respectively. On the other hand, for $score_i^{column}$, all the symbols belonging to column1, column2 and column3 were assigned y_4, y_5, and y_6 respectively. Further processing of these two types of scores is explained in section 2.3.3. Moreover, in MATLAB, 'stepwisefit' function was used to perform SWLDA.

2.3.2 c-VEP Detection

For c-VEP detection, signals from 8 electrode sites were considered: O1, Oz, O2, P3, Pz, P4, P7, and P8. The recorded EEG was band-pass filtered with a cut-off frequency of 6Hz - 41Hz, thus eliminating the noise and other artifacts. As discussed earlier in section 2.2.3, the single trial of c-VEP took 1.058s which corresponded to 2117 samples at 2KHz sampling rate. The training was performed on letter 'A' and from Figure 2.4 it can be seen that there were 70 c-VEP responses. Each epoch was taken to be 1.058s long (2117 samples). It was observed that SSVEP brain responses had a latency of 100 to 150ms [54]. Excluding the initial 150ms of response improved the accuracy of c-VEP BCI speller [24]. Note that we needed to include the additional 150ms of response, at the end, in order to complete the epoch. Epochs were taken 150ms after starting from the onset of the corresponding marker of the flash. One 3D array, labelled T, was formed and the size of the array was 8 x 70 x 2117, where first dimension was the number of channels, second dimension the number of epochs, and third dimension the length of epoch. Identification in c-VEP was done by the template matching method, where template was computed for each symbol during the training stage. In the test stage, correlation of the response with all the templates was computed and assigned as scores for all the symbols.

Canonical correlation analysis (CCA) was performed on the c-VEP data and for that purpose, the data was arranged in two 2D arrays labeled X and Y. Y was formed by averaging the data over the 70 trials of c-VEP and then column-wise,

replicating the data 70 times. The size of Y became 8 x 148190. On the other hand, X was formed by concatenating the data of each trial and its size became 8 x 148190. We performed CCA on the averaged data (Y) and non-averaged data (X). The purpose of CCA was to find the basis vector which maximizes the correlation between the two variables when projected onto their corresponding basis vectors. The maximizing function for CCA is defined as follows:

$$\max_{W_x, W_y} \rho(X, Y) = \frac{E[W_x^T X Y^T W_y]}{\sqrt{E[W_x^T X X^T W_x] E[W_y^T Y Y^T W_y]}} \tag{2.6}$$

where W_x is the resulting weight vector(8 x 1) for variable X, and W_y is the weight vector (8 x 1) for variable Y. To summarize, the projected non-averaged (noisy) data $W_x^T X$ was correlated with the projected averaged data $W_y^T Y$. The next step was to calculate the reference templates for all the symbols. The reference template for letter A was obtained by projecting averaged EEG over 70 cycles onto the basis vector W_y, which had the size of 1 x 2117. For the rest of the symbols, we obtained the templates by successively introducing the circular shift of 3-bits in the reference template, obtained for 'A', where 3-bits corresponded to 50 samples. In MATLAB, we used 'canoncorr' function to perform CCA.

In the test stage, W_x coefficients were used for the spatial filtering which meant we projected the raw EEG data onto the basis vector W_x and then correlated it with all the templates to obtain c-VEP scores for each symbol $score_i^{cVEP}$ where $i \in [1, 2, ..., 36]$. Further processing is explained in the next section.

2.3.3 Fusion of P300 and c-VEP Scores

We implemented two fusion techniques, LDA and maximum probability estimation (MPE). In this step, we combined the P300 and c-VEP scores for all the symbols and decided which target the subject looked at.

2.3.3.1 Linear Discriminant Anaylysis (LDA)

LDA is a linear way of combining features and as described in section 2.3.1, it finds the hyperplane which maximizes the distance between the means of the classes while minimizing the variance within the class. It is a widely used method for combining the data from different modalities because it provides good accuracy and the computational cost is low [50], [51], [55]. For each symbol, we have 3 types of scores: $score_i^{column}$, $score_i^{row}$, and $score_i^{cVEP}$. LDA reduces this 3-dimensional feature space to a one-dimensional feature space where the decision is based on the magnitude of the projection. The hybrid scores are computed as follows:

$$Score_i^{hybrid} = w_1 score_i^{row} + w_2 score_i^{column} + w_3 score_i^{cVEP} \tag{2.7}$$

where $[w1\ w2\ w3]$ are the weights, defining the hyper-plane. Projections are defined by the equation 2.1, where X is a feature matrix (contains all the three types of scores for each symbol) of size 3 x 36, w is a weight vector of size 3 x 1, and y is the resulting projection vector. After this fusion, target was decided as the symbol with a maximum magnitude of the projection which can be represented as follows:

$$Target = \underset{i \in [1,2,...,36]}{\arg\max}\ [Score_i^{hybrid}] \tag{2.8}$$

2.3.3.2 Maximum Probability Estimation (MPE)

MPE algorithm combined the individual scores of P300 and c-VEP based on the Gaussian distribution for each paradigm obtained during the training session. This approach was proposed by Yin et al. [39] in which all types of scores were transformed into probabilities and then the symbol with maximum probability was classified as target. We assumed all the scores were Gaussian distributed. Each score type had target and non-target Gaussian distribution which were

denoted by T_d, N_d respectively, where $d \in [1, 2, 3]$. Note that $d = 1$ denotes the row type, $d = 2$ denotes the column type, and $d = 3$ denotes the c-VEP type. The mean and standard deviation for these distributions were obtained during training stage. The data for each type of score is arranged in the matrix as shown below.

$$\begin{bmatrix} score^d_{1,1} & score^d_{2,1} & ... & score^d_{i,1} & ... & score^d_{36,1} \\ score^d_{1,2} & score^d_{2,2} & ... & score^d_{i,2} & ... & score^d_{36,2} \\ \vdots & \vdots & & \ddots & & \vdots \\ score^d_{1,k} & score^d_{2,k} & ... & score^d_{i,k} & ... & score^d_{36,k} \end{bmatrix}$$

where, i^{th} column shows the scores for i^{th} symbol, k^{th} row shows the scores for k^{th} trial. We know from the design of the BCI speller that first score belongs to letter 'A' which was target in our training session and the rest of the symbols were non-target. For the target distribution (T_d) we computed the mean, μ_{T_d}, and standard deviation, σ_{T_d}, of the first column. For the non-target distribution (N_d) we computed the mean, μ_{N_d}, and standard deviation, σ_{N_d}, of the last 35 columns. The target distribution and non-target distributions were defined as $T_d \sim N(\mu_{T_d}, \sigma_{T_d})$ and $N_d \sim N(\mu_{N_d}, \sigma_{N_d})$ respectively.

In the test stage, we obtained 36 scores of each type and using the distributions obtained previously, we computed the likelihood of each score as defined by the following equations:

$$f(score^d_i|T_d) = \frac{1}{\sigma_{T_d}\sqrt{2\pi}} e^{-(1/2)((score^d_i - \mu_{T_d})/\sigma_{T_d})^2} \tag{2.9}$$

$$f(score^d_i|N_d) = \frac{1}{\sigma_{N_d}\sqrt{2\pi}} e^{-(1/2)((score^d_i - \mu_{N_d})/\sigma_{N_d})^2} \tag{2.10}$$

Next, we assumed that a symbol's 3 types of scores are conditionally independent from each other and also independent from the scores of different symbols. After computing the likelihood, we computed the total probability of each symbol being a target and non-target using the following equations:

$$p_i(Score_i|T_1, T_2, T_3) = \prod_{d=1}^{3} f(score^d_i|T_d) \tag{2.11}$$

$$p_i(Score_i|N_1, N_2, N_3) = \prod_{d=1}^{3} f(score_i^d|N_d) \tag{2.12}$$

where, $Score_i$ is the set of scores from all types $[score_i^1\ score_i^2\ score_i^3]$. We know that only one of the symbols is target and rest of them are non target so for each symbol we computed the likelihood of being a target given that all the other symbols were non-target. The equation below describes this step:

$$\begin{aligned} P_i(Score_i|T_1, T_2, T_3, Score_1...Score_{36}) = p_1(Score_1|N_1, N_2, N_3)\times \\ p_2(Score_2|N_1, N_2, N_3) \times ... \times p_i(Score_i|T_1, T_2, T_3) \\ \times p_{36}(Score_{36}|N_1, N_2, N_3) \end{aligned} \tag{2.13}$$

Computing probabilities using equation 2.13 gave us 36 probabilities for all the symbols and we chose the symbol with the maximum probability as the target symbol as described in the next equation:

$$Target = \underset{i\in[1,2,...,36]}{\arg\max}\ [P_i(Score_i|T_1, T_2, T_3, Score_1...Score_{36})] \tag{2.14}$$

2.3.4 Modeling of c-VEP Responses

Nagel et al. [44] proposed a moving average model, which was referred to as Ridge regression model, to study the VEP responses for arbitrary visual stimulus patterns. In this study, we implemented a moving average model for c-VEP paradigm. Data of 7 subjects was used for this study which was taken from BCI database of our lab. Data acquisition procedure and hardware used for this study was the same as explained in section 2.2. We performed offline analysis on data. Recorded EEG was band-pass filtered with a cut-off frequency of 6Hz - 41Hz, thus eliminating the noise and other artifacts.

Since data was acquired on 60Hz and 120Hz refresh rate of monitor, we modeled the responses for both refresh rates. The sequence for each target comprised 127 bits which took 2.117 seconds at 60 Hz refresh rate and the sampling frequency was set to 2000 Hz so there were 4233 samples for each sequence. In the case of 120 Hz refresh rate, the sequence took 1.058 seconds which corresponded to 2177 samples at 2000 Hz sampling rate. During the training, the EEG data was recorded 100 times for the sequence of target 'A'. Identification of the targets in model based c-VEP was done by correlating the EEG response with the simulated responses of all the targets and the symbol with the maximum correlation was classified as target.

CCA was performed on the c-VEP data where data was averaged over 100 trials. We performed CCA on the averaged data and non-averaged data, thereby increasing the signal-to-noise ratio (SNR). Note that the spatial filtering (CCA) was important to obtain a good fit model. Details of CCA are explained in section 2.3.2.

VEP responses usually last for 150ms-300ms so the input window of 250ms (500 samples) was chosen and the window length was denoted by M. The moving average model is defined as follows:

$$y_k = \beta_1 x_{k-M} + \beta_2 x_{k-M+1} + \ldots + \beta_M x_{k-1} \tag{2.15}$$

where y_k is the predicted sample of EEG at index k, x is the input code sequence and β coefficients are optimized parameters for the moving average model. As M input samples were used to predict the $(M+1)^{th}$ EEG sample, the first M EEG samples could not be predicted using this model. Beta coefficients were found using least squares method and the closed-form solution is given as follows:

$$\beta = ((X^T X + \lambda I)^{-1} X^T y)/\sigma(X) \tag{2.16}$$

where X is the n x M input matrix, y is the n x 1 response vector and λ is the ridge parameter set to 0.001. n is the number of windows used, for 60Hz we have $n = 433300 - M = 432800$ (where M=500) and for 120Hz we have $n = 211200$. We have the input vector of size 1 x 433300 and 1 x 211700 for 60Hz and 120Hz respectively. Matrix X was obtained by taking the n number of windows from input vector, which were sample wise shifted. Figure 2.7 explains the input and output relationship of the model.

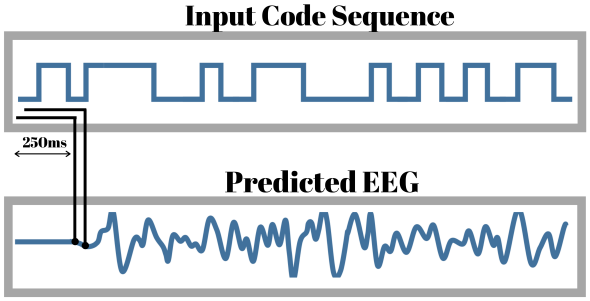


Figure 2.7: Model predicts the EEG using previous 250ms input code sequence. Note that this model does not predict the first 250ms of EEG and all the correlations and target identification is done without the first 250ms of EEG.

The model obtained during the training stage was used in the test stage to

simulate the responses of all the targets. In the test stage, experimental data was acquired for 20 letter word "BILKENTBCIEXPERIMENT". For 60Hz, single trial was used and for 120Hz two trials were performed. All the data was again spatially filtered and the response of each letter was correlated with the simulated responses. The symbol with maximum correlation was classified as a target letter.

2.4 Performance Evaluation

We used accuracy and ITR as measures to evaluate the performance of hybrid speller and model based c-VEP speller. As explained in section 2.2.3, during online test session we asked subjects to spell a 20-letter word. The accuracy was calculated as the number of correctly classified targets divided by the total number of targets. The ITR is a standard metric defined by the following equation [56]:

$$ITR = \frac{60}{T} \times (log_2 N + P log_2 P + (1-P) log_2 \frac{1-P}{N-1}) \tag{2.17}$$

where, N is the number of symbols in speller, P is the accuracy, and T is the time interval required for a single selection of target. In hybrid speller, T is 2.6s which is addition of time for single trial of P300 (1.5s), gaze shifting time (1s), and classification time (0.1s). In model based c-VEP, for 60Hz, T is 3.13s which is addition of time for single trial of c-VEP (2.12s), gaze shifting time (1s) and classification time (0.01s). Similarly, for 120Hz, T is 2.07s which is addition of time for single trial of c-VEP (1.06s), gaze shifting time (1s) and classification time (0.01s). Moreover, we used correlation between measured and predicted response to quantify the fit of the model.

Chapter 3

Results

3.1 Stimulus Measurements

Before performing experiments we test monitor response and delay in the marker pulse. For this purpose we use a simple photo-diode (BPW24R) circuit where the cathode is connected to a VCC, anode is connected to a resistor (470kΩ) and then resistor is grounded. As a stimulus source we use Dell Alienware (AW2518HF) monitor which has a screen refresh rate of up to 240 Hz and a resolution of 1920 x 1080 pixels. For the monitor response, we use the monitor screen background flickering between black and white at 60Hz, 120Hz, and 240Hz. While the background is flickering, we hold the photo-diode perpendicular to the screen and measurements are recorded using the digital oscilloscope. Figure 3.1 shows the obtained responses. We can clearly see in Figure 3.1 that as we go towards the higher frequency black to white response of each pixel worsens (the white light of the background takes some time to fully shine the screen). As explained earlier, ITR of the BCI system increases as we increase the stimulus presentation rate. We choose 120Hz refresh rate because it allows us to improve ITR while deteriorating the black to white response of each pixel minimally as shown in Figure 3.1b.

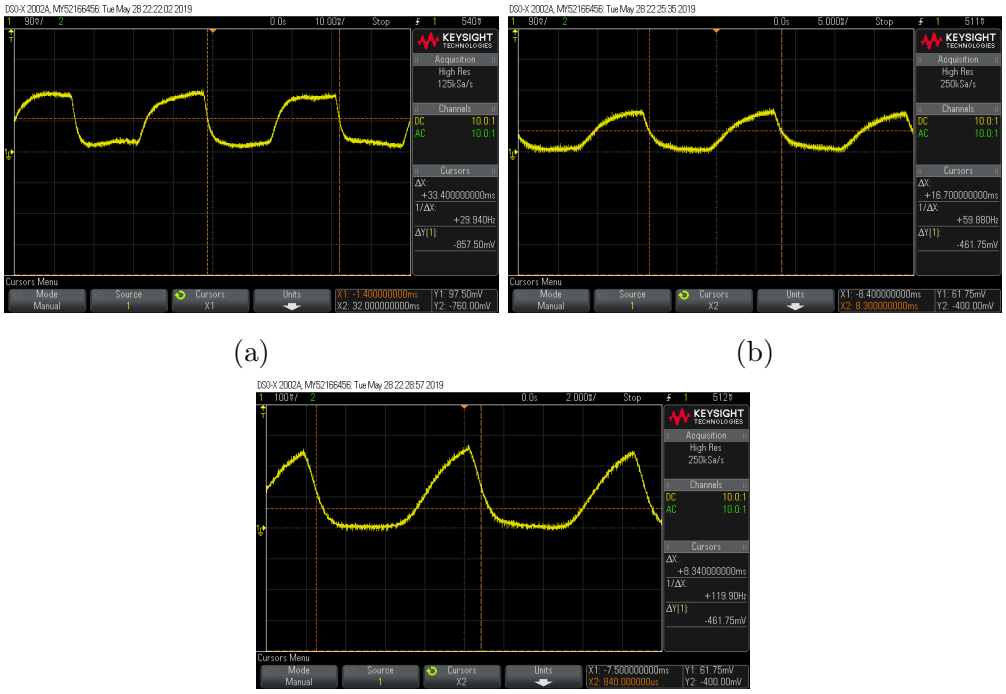
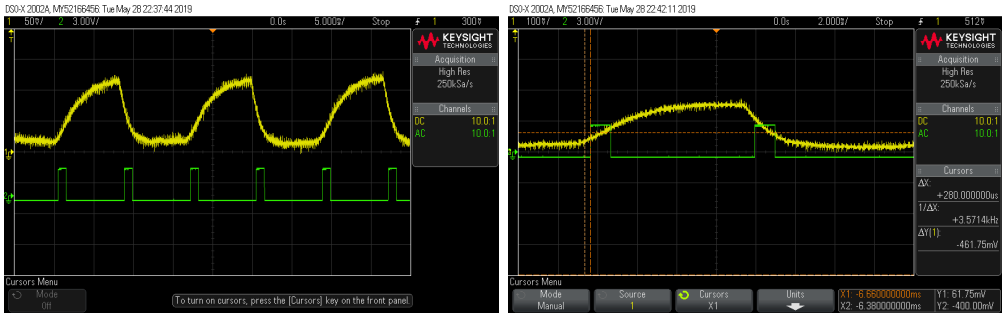


Figure 3.1: Photo-diode response for the screen flickering between black and white background at 60Hz, 120Hz, 240Hz. (a) This is response for 60Hz refresh rate where high corresponds to the white frame and low corresponds to the black frame. Cursors on screen show the period of one cycle of flicker where the frequency is measured as 29.94Hz. Note that one cycle of flicker has two frames (black and white) so the refresh rate of monitor is 2×29.94. (b) This is the response for 120 Hz. The frequency of one cycle of flicker is measured as 59.98 and refresh rate becomes 119.96Hz. (c) This is response for 240Hz. The frequency of one cycle of flicker is measured as 119.90Hz and refresh rate becomes 239.8Hz.

Next, we measure the time delay between marker pulse and monitor response. Figure 3.2 shows the marker pulse (blue) and the corresponding black to white response of pixel (yellow). We can see that monitor response is perfectly synchronized with the marker pulses. The delay between the marker and the flicker is around 280μs which can be ignored.



(a) (b)

Figure 3.2: Plots of the monitor response and the corresponding marker pulses. For the testing purpose we send the marker puslse after every frame to test the synchronicity of our system. We can see the pulse at every rising and falling edge of the monitor response.

3.2 Initial P300 Study

The conventional P300 with 6-row and 6-column flashes is implemented and superimposed with the c-VEP design. The experiments are performed on 3 different subjects to test the feasibility of the hybrid system. P300 signals are obtained during the training session, where 50 P300 trials are performed with 100ms active time and 150ms passive time. In 12 row/column flashes, the total time for a single sequence becomes $12 \times 0.25 = 3s$. Total number of flashes become $12 \times 50 = 600$, where 100 of them are target flashes and 500 are non-target flashes. Training is performed on letter 'A', where the subject keeps the mental count of the target flashes. The P300 signals for two different subjects are shown in the following figures. Different subjects are denoted with a capital letter S followed by a number (i.e. S1,...S7).

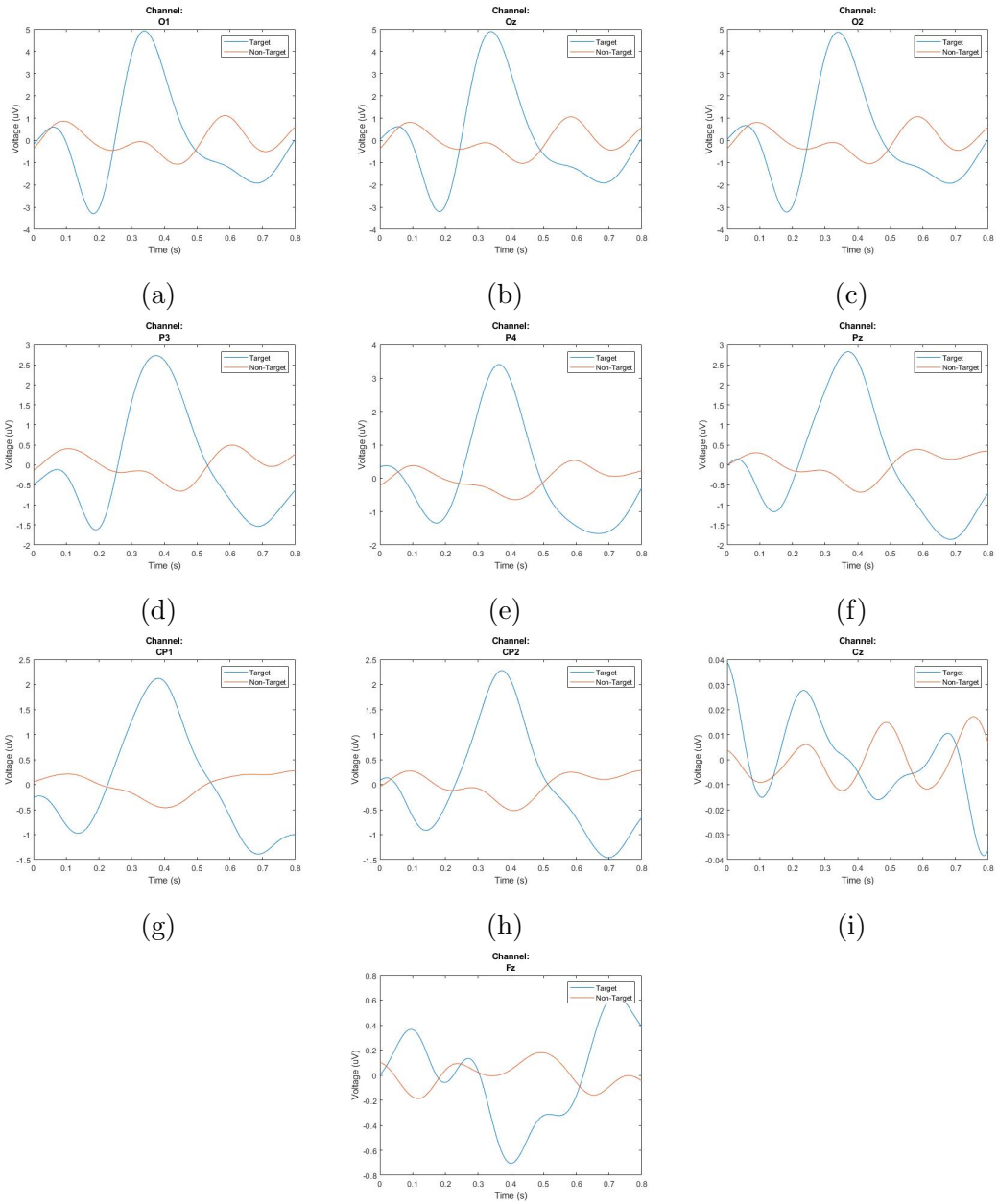


Figure 3.3: The non-target signal (orange) and target signal (blue) of S1 obtained from channels: O1, Oz, O2, P3, Pz, P4, CP1, CP2, Cz, and Fz respectively. Y-axis shows the voltage level in microvolts and x-axis shows the time in seconds. The graphs for P300 show the response for a single epoch length which is taken as $0.8s$.

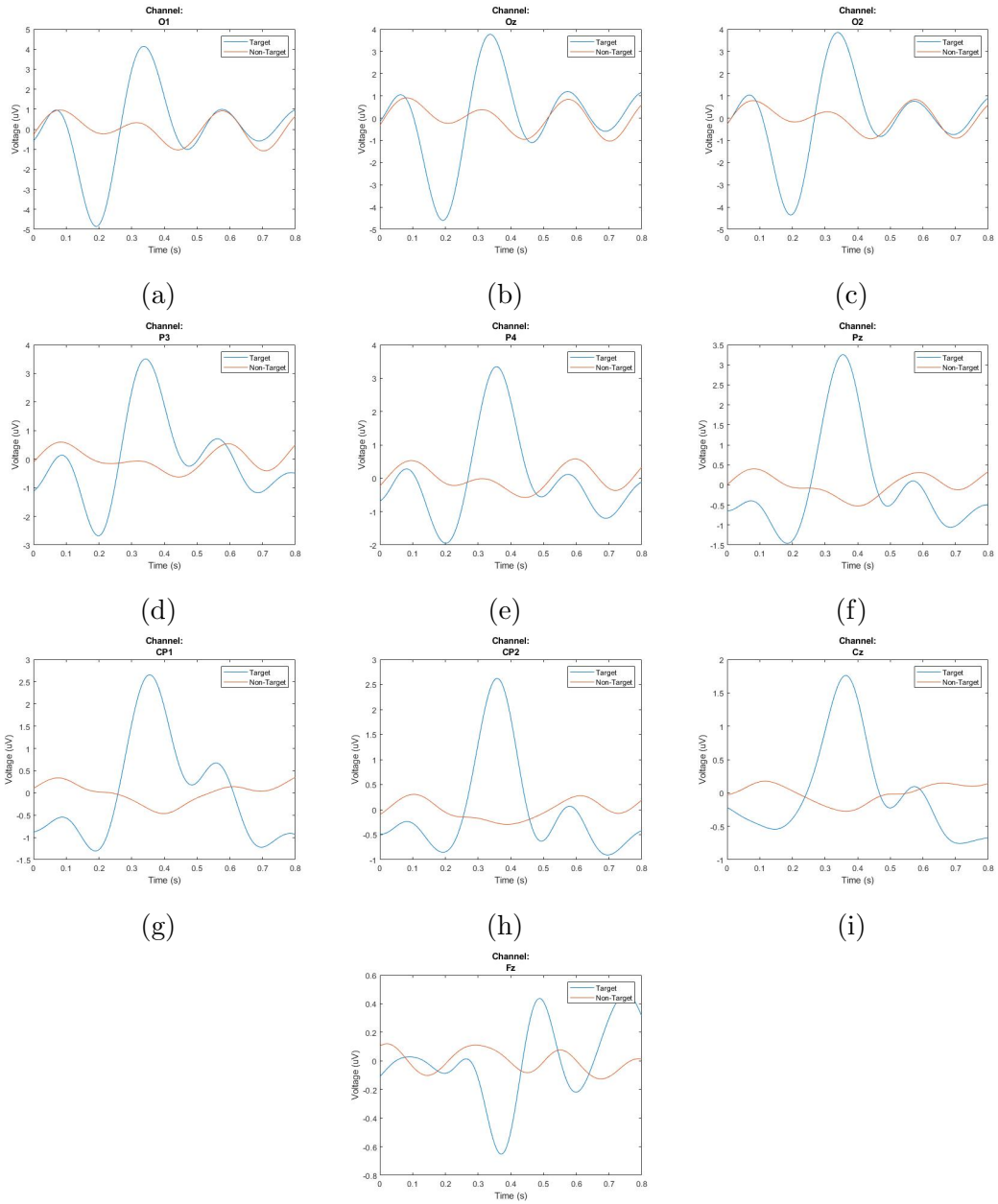


Figure 3.4: The non-target signal (orange) and target signal (blue) of S2 obtained from channels: O1, Oz, O2, P3, Pz, P4, CP1, CP2, Cz, and Fz respectively. Y-axis shows the voltage level in microvolts and x-axis shows the time in seconds. The graphs for P300 show the response for a single epoch length which is taken as $0.8s$.

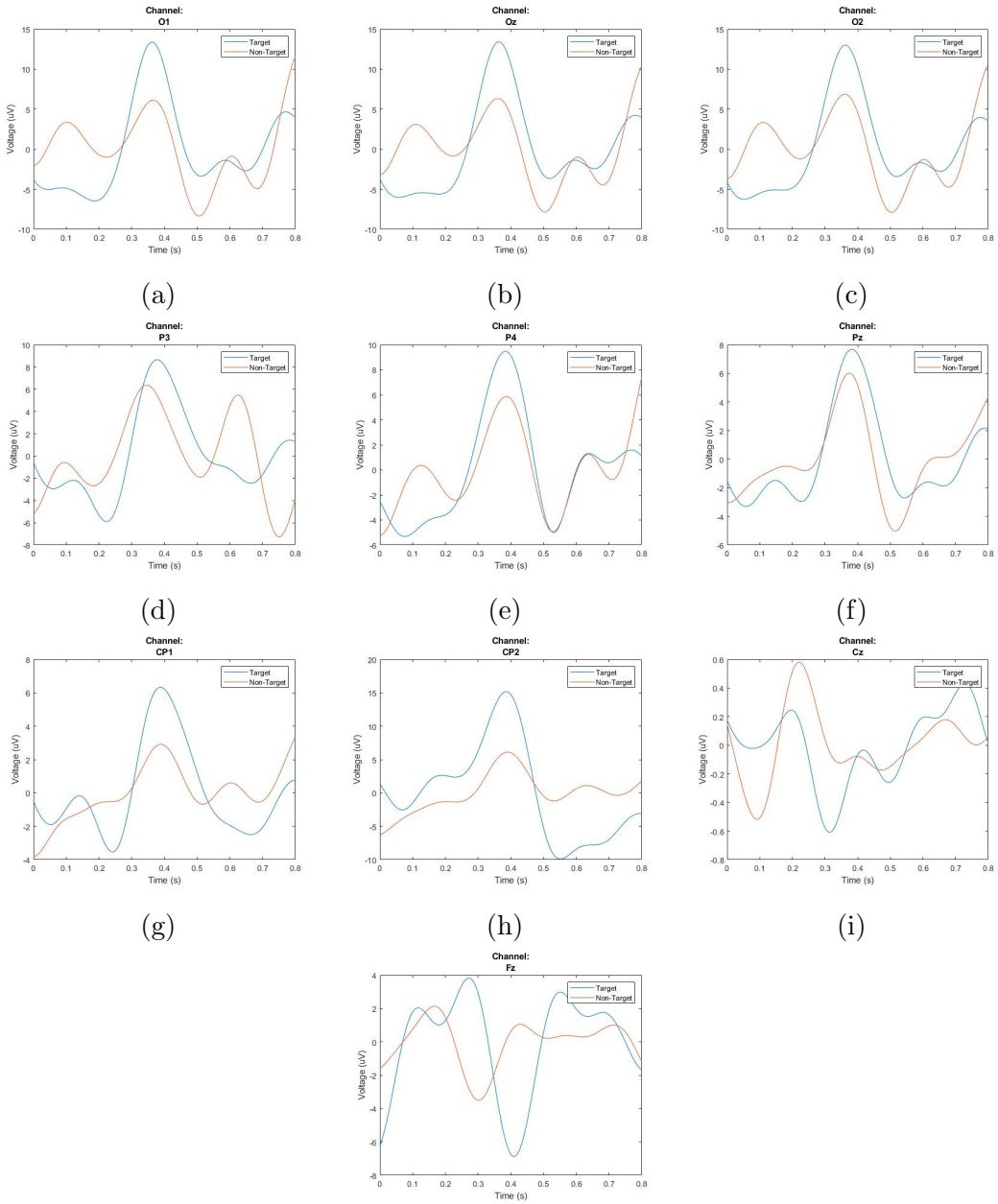


Figure 3.5: The non-target signal (orange) and target signal (blue) of S2 obtained from channels: O1, Oz, O2, P3, Pz, P4, CP1, CP2, Cz, and Fz respectively. Y-axis shows the voltage level in microvolts and x-axis shows the time in seconds. These are the target and non-target responses for a single trial.

In both cases (S1 and S2), as shown in Figure 3.3, and Figure 3.4, distinguishable peaks of P300 signals at around 300ms are observed. The target and non-target signals are obtained by averaging 100 target and 500 non-target responses respectively, from all the 10 channels. All the signals are referenced to the electrode site FCz which is on the top of the head. Moreover, one clear observation is that as we go away from the central lobe, the P300 signal becomes more prominent. Hence, P300 signal is not observed at Fz and Cz which are closest to the reference point. We can see the amplitude of the P300 peak increasing from $2.2\mu V$ (at channel: $CP2$) to $5.0\mu V$ (at channel: $O1$). Moreover, the N100 (the first negative peak) is also visible along with the P300 signal. The latency of P300 signals is between $300ms$ and $500ms$, where latency is defined as the time from the onset of the stimulus to the maximum of the positive peak. Although the amplitude varies slightly, the morphology of the signal is similar across the brain areas (e.g., parietal, central, occipital). Figure 3.5 shows the unaveraged (single trial) target and non-target responses at all the channels. Since we are using narrow band pass filter, most of the noise is filtered so we still obtain observable P300 peaks for the target. However, the target and non-target signals cannot be easily distinguished because of the baseline noise.

3.3 Hybrid Experiment Results

3.3.1 P300 Results

This section discusses the results of the grouped P300. As known from section 2.3.1, after decimating, concatenating the data from all the 10 channels, and performing LDA, we obtain a feature vector of length 160. We have shown the feature vector obtained for S5 in Table 3.1. For simplicity, we arrange the features in a matrix of size 10 x 16 as shown in Table 3.1. Each column corresponds to a sample at a specific time and each row corresponds to a different channel. SWLDA selects the significant features statistically which are highlighted in Table 3.1. Out of 160, SWLDA selects only 15 features. This feature matrix also supports the

observations explained in section 3.2. Figure 3.6 shows averaged P300 data of S5 and marks the selected features. In S5, we can see N100 and P300 peaks in all the channels except Fz.

SWLDA selects distinct features and mitigates redundancy. Table 3.1 shows that 3 features are selected at $t = 300.5ms$ and are considered as three different variations of the response at $t = 300.5ms$. Figure 3.6b and Figure 3.6c show that signals are similar to each other, even the amplitude of the peaks are very close. Thus, adding the same features from these two signals will not increase the performance of our classifier. For instance, we can clearly see the P300 peak at $t = 300.5ms$ at O1, Oz, O2, P3 and P4 but SWLDA only selects the feature from channel O1 where the amplitude of the peak is $3.8\mu V$. Second feature corresponding to P300 peak at $t = 300.5ms$ is selected from channel Pz where the amplitude of the peak is $1.4\mu V$. The amplitude of the P300 peak at Pz is significantly different from the P300 peaks observed at O1, Oz, O2, P3 and P4 and that is why, SWLDA selects it as a feature. Third variation is observed in channel Cz, marked in Figure 3.6i, where the P300 response is delayed and significantly attenuated ($0.4\mu V$). Since this feature contributes to the overall fitting of data as a distinguishable feature, the algorithm selects it. However, considering the prior knowledge about the shape of the P300 signal, we can discard data from Cz and Fz. Moreover, according to the observation made in section 3.2, the P300 deteriorates significantly in the neighborhood locality of this channel which supports our choice of discarding it.

Table 3.1: Statistically significant features obtained after performing SWLDA. Rows correspond to different channels and columns correspond to different time instants. Note that '0' in the table refers to non-significant features.

Time (ms)	0.5	50.5	100.5	150.5	200.5	250.5	300.5	350.5	400.5	450.5	500.5	550.5	600.5	650.5	700.5	750.5
O1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1
Oz	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
O2	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0
P3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Pz	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
P4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
CP1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
CP2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cz	0	0	0	0	0	0	1	0	1	0	0	0	0	1	0	0
Fz	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

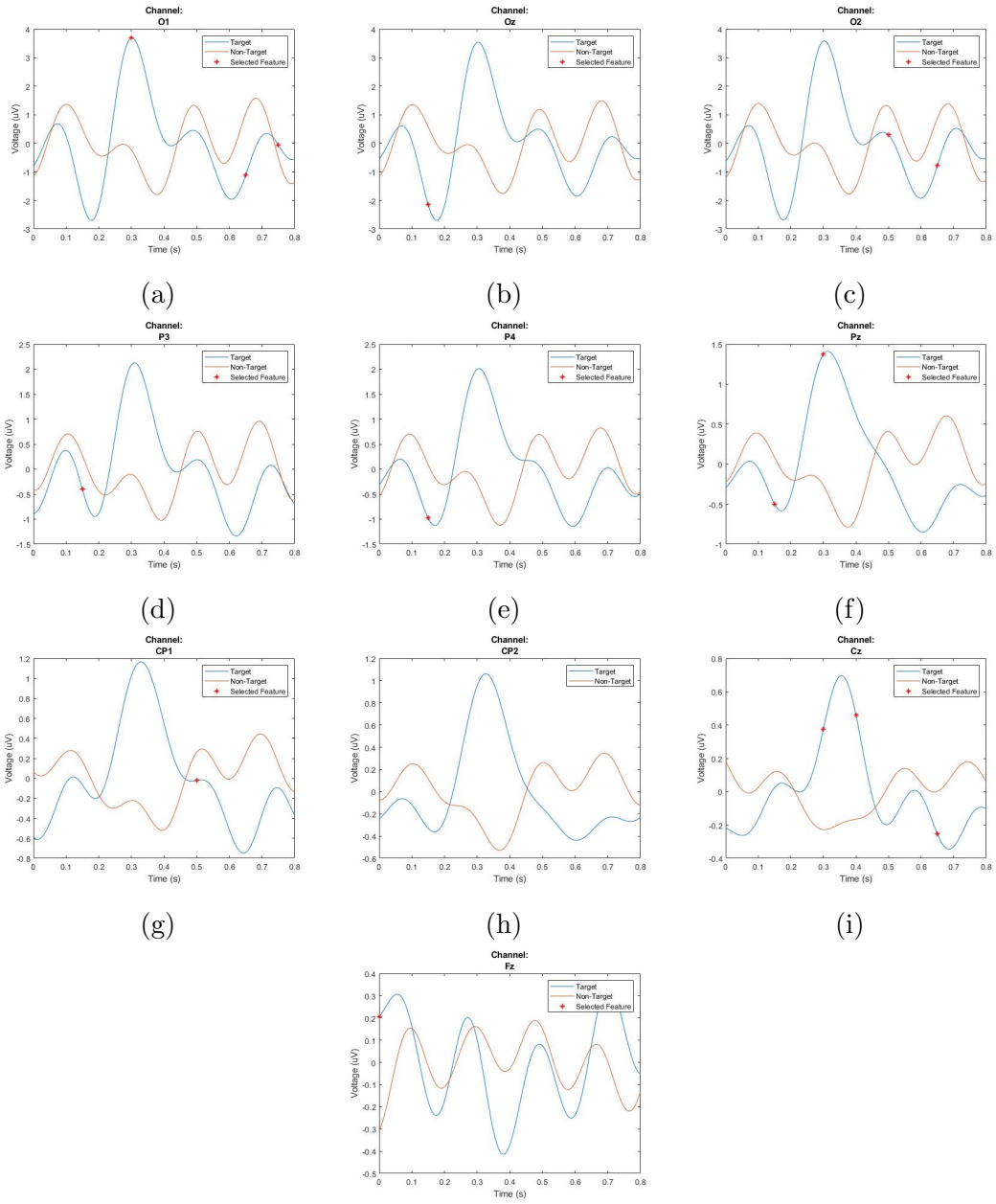


Figure 3.6: The non-target signal (orange) and target signal (blue) of S5 obtained from channels: O1, Oz, O2, P3, Pz, P4, CP1, CP2, Cz, and Fz respectively. Y-axis shows the voltage level in microvolts and x-axis shows the time in seconds. The selected features are marked in the plots with a red star.

3.3.2 c-VEP Results

We know from section 2.3.2 that reference templates are obtained by projecting the averaged EEG onto the basis vector W_y. Figure 3.7 shows the reference template and pseudorandom binary sequence for letter 'A'. We cannot see any relationship between input (visual stimulus) and output (reference template). It is an ongoing research topic where different techniques have been applied to model the brain responses [40], [41], [42], [43]. However, we use template matching method as explained in section 2.3.2 for identification of the target. Figure 3.8a shows the averaged response and single trial response from channel O1. The correlation coefficient of these two responses is 0.1016. On the other hand, Figure 3.8b shows the reference template and the single trial response projected onto the basis vector W_x (see section 2.3.2). The correlation coefficient in this case is 0.6002. If we compare Figure 3.8a and Figure 3.8b we can see the significance of spatial filtering.

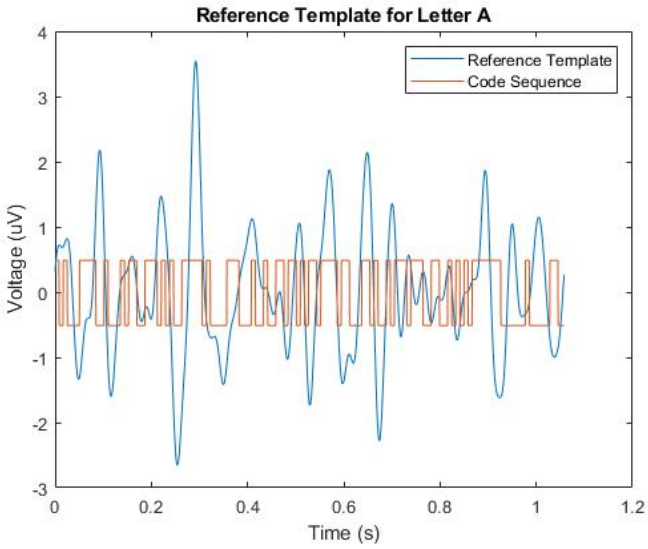
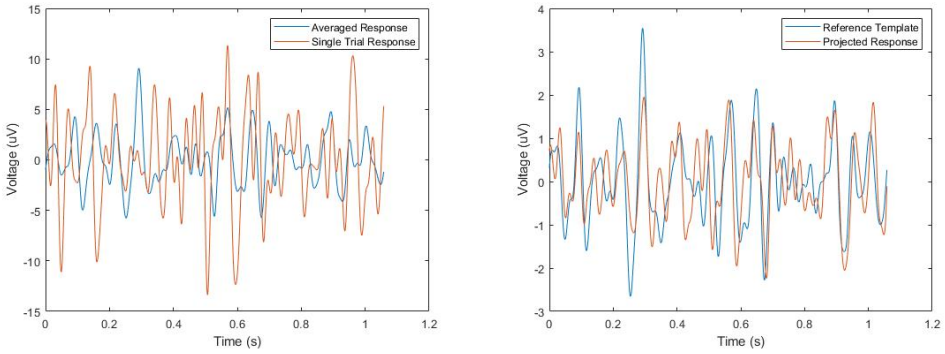


Figure 3.7: Reference template (blue) and m-sequence for letter 'A'. Y-axis shows the voltage level in microvolts and x-axis shows the time in seconds.



(a) (b)

Figure 3.8: Responses with and without spatial filtering. (a) averaged EEG and single trial response is plotted from channel O1. (b) reference template and single trial projected response.

We compute the reference templates for all the symbols. After projecting the single trial response onto the basis vector W_x, we correlate it with all 36 templates. Figure 3.9 shows the correlations of the single trial projected response of letter 'A' with all the templates. The response has maximum correlation (0.6002) with the first reference template, and the second maximum correlation (-0.3772) is with the reference template of letter 'B'.

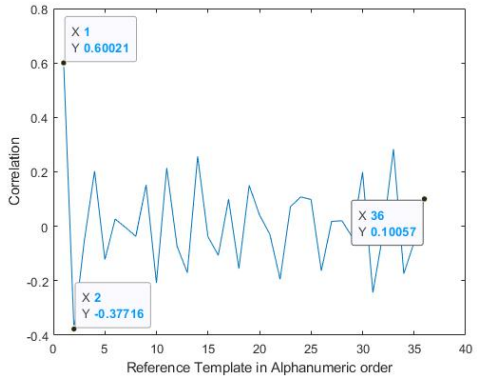


Figure 3.9: Correlation of a single trial projected response and all the reference templates. Y-axis shows the correlation coefficient and x-axis corresponds to the reference template of the symbols in alphanumeric order.

Table 3.2 shows the W_x basis vector obtained for each subject. Oz channel gives the maximum weights. From Table 3.2 we can see that channels with significant contribution to the projections are located at the occipital area (O1, Oz, O2).

Table 3.2: W_x coefficients obtained during training session for each subject.

Spatial Filter (Wx)								
Sub	**O1**	**Oz**	**O2**	**P3**	**Pz**	**P4**	**P7**	**P8**
1	0,0467	-0,1774	-0,1636	0,0402	0,0395	0,04	0,0462	0,1354
2	-0,1543	0,4207	0,007	0,0066	-0,0264	0,008	-0,0769	-0,1746
3	0,207	-0,3156	-0,0029	0,0479	-0,006	0,0189	-0,0123	0,0543
4	0,025	0,2583	0,0215	-0,0322	0,0156	-0,0824	-0,0747	-0,1099
5	0,1608	-0,517	-0,0351	0,0066	0,0322	-0,0035	0,1379	0,1953
6	0,0645	0,2623	-0,0132	-0,0068	-0,0247	-0,0726	-0,1502	-0,0585
7	-0,0094	0,5871	-0,1059	-0,0736	-0,0794	-0,0385	-1616	-0,1136

3.3.3 Accuracy and ITR of Hybrid BCI Speller

Table 3.3 and Table 3.4 show the accuracy and ITR values obtained for each subject. Note that the accuracy value reported here is the mean value of the five test sessions performed for each subject (see section 2.2.3). c-VEP ITR value, obtained using hybrid experiment data, is denoted by ITR_1 which uses T is 2.6s. On the other hand, for the individual c-VEP paradigm, we have computed ITR_2 which uses T as 2.16s. This difference arises because c-VEP single trial finishes early, thus in an individual c-VEP system, single trial would take about $1.06 + 1 + 0.1 = 2.16s$.

Average ITR and accuracy values for hybrid based c-VEP are 102.1 bits/min, and 92.5%; for individual c-VEP paradigm are 114.9 bits/min and 92.5%; for MPE based hybrid are 104.2 bits/min and 93.6% and for LDA based hybrid are 82.3 bits/min and 80.2% respectively. In comparison with the accuracy and ITR values of c-VEP alone, that is without simultaneously making use of the P300 data obtained during the hybrid experiments, MPE based hybrid has been successful in demonstrating 3-5% improvement in the accuracy and 2-11 bit/min improvement in the ITR, in 4 subjects. In the remaining participants, the accuracy either

decreased or remained unchanged. Furthermore, accuracy values for each subject are the mean of 5 accuracy values obtained for 5 test sessions for each subject. Standard deviation shows us that the accuracy variation within subject is around $\pm 5\%$, which corresponds to the misclassification of ± 1 letter.

Furthermore, we performed Wilcoxon signed-rank test for the accuracy values of c-VEP, MPE based hybrid, and LDA based hybrid. Wilcoxon's test is a non-parametric alternative of paired Student's t-test. Paired Student's t-test is performed under the assumption that data is normally distributed whereas Wilcoxon's test is more general and can be performed when data is not normally distributed. We know from equation 2.17 that ITR depends on accuracy values and the relationship is nonlinear. Non-linearity in the relationship of ITR and accuracy do not change the test results for ITR because Wilcoxon's test ranks the absolute difference values between pairwise data. The results obtained from Wilcoxon's test show us that the accuracy and ITR values of c-VEP alone and MPE based hybrid are not statistically different ($p = 0.293$) whereas the LDA based hybrid is statistically worse from c-VEP alone ($p = 0.016, \alpha = 0.05$).

Table 3.3: ITR and accuracy table for c-VEP and hybrid (MPE).

		c-VEP			Hybrid (MPE)	
Sub	**Impd** ($k\Omega$)	*Acc.* (%)	ITR_1**(bits/min)**	ITR_2**(bits/min)**	*Acc.* (%)	ITR **(bits/min)**
S1	20	90 ± 6.1	96.6	108.8	94.3 ± 4.6	105.3
S2	20	97 ± 3.7	111.3	125.2	98 ± 2.2	113.7
S3	20	96 ± 3.4	109	122.7	96 ± 3.4	109
S4	20	84 ± 5.3	85.7	96.5	87 ± 2.2	91.1
S5	50	96 ± 1.8	109	122.7	93 ± 2.2	102.6
S6	50	94 ± 4.5	104.6	117.8	91 ± 6.7	98.6
S7	20	91 ± 5.3	98.6	111.0	96 ± 1.8	109
Avg.	-	**92.5**	**102.1**	**114.9**	**93.6**	**104.2**

Table 3.4: ITR and accuracy table for hybrid (LDA). Accuracy and ITRs are the average of 5 experiments for each subject. Alongside the accuracy values, the standard deviations are also written.

	Hybrid (LDA)	
Sub	*Acc.* **(%)**	*ITR* **(bits/min)**
1	89.3 ± 3.9	95.3
2	82 ± 7.4	82.3
3	93 ± 5.5	102.6
4	82 ± 6.2	82.3
5	81 ± 5.3	80.6
6	81 ± 12.0	80.6
7	62 ± 4.7	52.2
Avg.	**80.2**	**82.3**

3.4 Results of Model Based BCI Paradigm

Model is successfully implemented and tested on offline data. To summarize, we have a m-sequence for letter 'A' and the training is repeated for 100 trials, using this data we obtain the weights for spatial filtering and train the moving average model. Next, we use this linear model to simulate the responses for all the targets. In the test stage, we correlate the acquired response with simulated responses and the symbol with maximum correlation is chosen as target.

After training, we simulate the responses for all 100 repeated trials. Correlation is calculated between measured response and simulated response to understand whether its a good fit or not. The average correlation is 0.357 and 0.396 for 60Hz and 120Hz respectively. Table 3.5 shows the correlation values between measured response and simulated response for all the subjects. Note that the subjects for this study are different than subjects who participated in hybrid experiments.

Table 3.5: Correlation between measured response and modeled response for 60Hz and 120Hz monitor refresh rate.

Sub	Correlation (60 Hz)	Correlation (120 Hz)
S1	0,409	0,394
S2	0,376	0,384
S3	0,400	0,354
S4	0,447	0,574
S5	0,271	0,342
S6	0,395	0,508
S7	0,204	0,214
Avg.	**0,357**	**0,396**

Figure 3.10 and Figure 3.11 show the single trial measured output and the corresponding model output of S4 for 60Hz and 120Hz respectively. Correlations are 0.447 and 0.574 for 60Hz and 120Hz respectively. We can see that model for both the refresh rates capture the major variations in measured response.

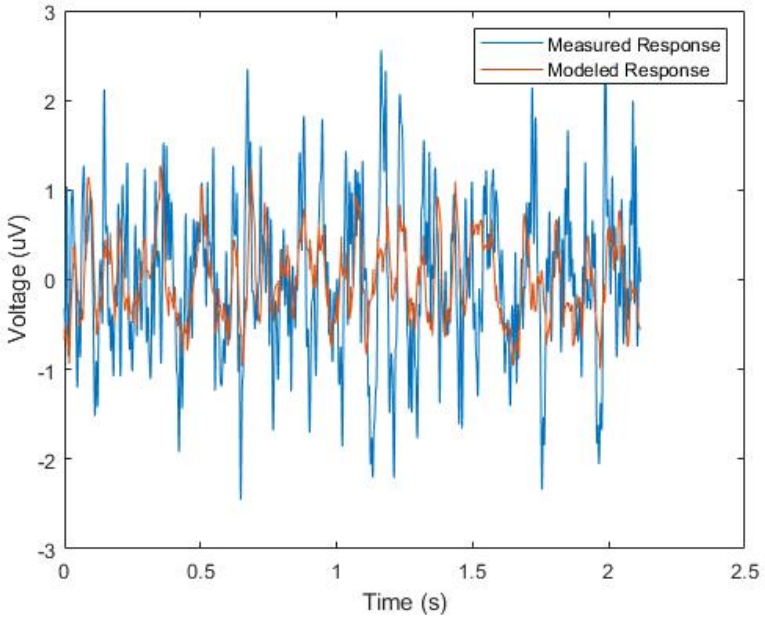


Figure 3.10: Measured response (blue) and modeled response (orange) of S4 for 60Hz experiments. Y-axis shows the voltage levels and x-axis shows time in seconds. The plot shows the data from the first trial of the training where the correlation coefficient is 0.447.

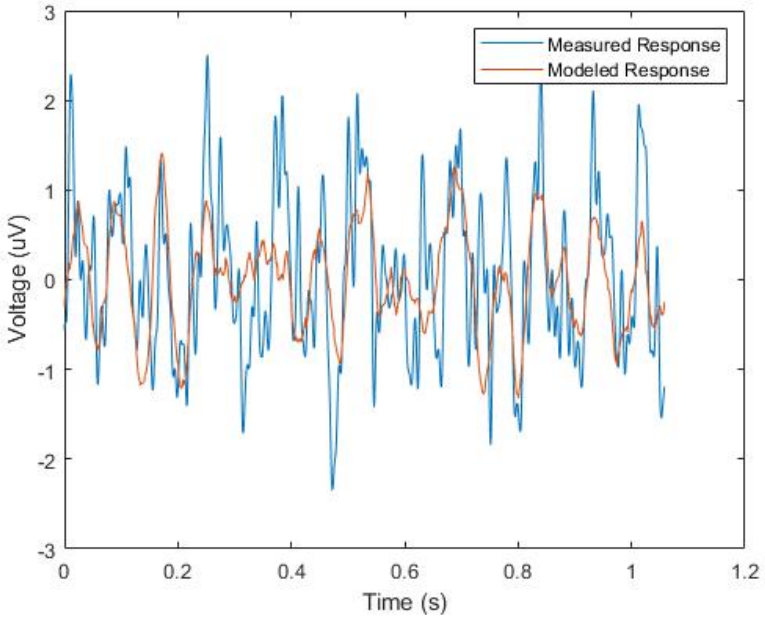


Figure 3.11: Measured response (blue) and modeled response (orange) of S4 for 120Hz experiments. Y-axis shows the voltage levels and x-axis shows time in seconds. The plot shows the data from the first trial of the training where the correlation coefficient is 0.574.

In the test stage, model is utilized and the accuracy and ITR values are computed. For 60Hz experiment, the ITR is computed using T as 3.13s. For 120Hz experiment, we have two accuracy and ITR values. Acc_1 and ITR_1 are the values for a single trial. Acc_2 and ITR_2 are the values for the averaged response over two trials. ITR_1 uses T as 2.07s whereas ITR_2 uses T as 3.13s. Table 3.5 shows the accuracy and ITR values for all the subjects. Average accuracy and ITR for 60Hz case are 87.1% and 76.4 bits/min. Average accuracy and ITR for 120 Hz case are 68.6% and 85.0 bits/min for single trial and 82.1% and 72.4 bits/min for two trials respectively. Note that the single trial accuracy values are low because in this data analysis we did not exclude initial 150ms of data as explained in section 2.3.2.

Table 3.6: Accuracy and ITR values obtained using the linear ridge regression model as a classifier.

	60 Hz		**120 Hz**			
Sub	*Acc.* (%)	ITR **(bits/min)**	Acc_1 (%)	ITR_1 **(bits/min)**	Acc_2 (%)	ITR_2 **(bits/min)**
S1	90	80,3	75	89,2	95	88,7
S2	95	88,7	70	79,7	90	80,3
S3	90	80,3	30	20,2	45	26
S4	90	80,3	100	149,9	100	99,1
S5	80	65,6	70	79,7	85	72,7
S6	90	80,3	100	149,9	100	99,1
S7	75	59.0	35	26,1	60	41,2
Avg.	**87,1**	**76,4**	**68,6**	**85,0**	**82,1**	**72,4**

Chapter 4

Discussion and Conclusion

In this study, a novel hybrid BCI speller with 36 selectable symbols is proposed, incorporating P300 and c-VEP paradigms, that achieves both high spelling accuracy and ITR. Fusion techniques have been applied in order to effectively combine the information of P300 and c-VEP at the score level. For this purpose, two different methods, MPE and LDA, have been implemented and compared. Experiments were performed on 7 participants to assess the feasibility and performance of the hybrid BCI speller's design. The accuracy and ITR value of MPE-based hybrid improved only by 1.1% and 2.1 bits/min, on average, respectively, compared to the c-VEP alone, that is without simultaneously making use of the P300 data obtained during the hybrid experiments. On the other hand, the average accuracy and ITR value of LDA-based hybrid worsened by 12.3% and 19.8 bits/min, respectively, compared to the c-VEP alone. Based on the mean of the average accuracy and ITR values of all the subjects, the performance can be ranked in the following order: MPE-based hybrid, c-VEP alone, and LDA-based hybrid. However, statistical tests indicated that the results of MPE-based hybrid and of c-VEP alone are not statistically different ($p = 0.293$). Moreover, the results of LDA-based hybrid and of c-VEP alone are statistically different ($p = 0.016, \alpha = 0.05$). Hence, based on mean accuracy, ITR results and statistical tests, we rule out the LDA-based hybrid.

There could be several reasons that the MPE-based hybrid did not show significant improvements compared to the c-VEP alone. One reason is the impedance problem in P300. Although the c-VEP paradigm is robust and can provide high levels of accuracy with impedance as high as 50kΩ, P300 signals become more difficult to observe at this level of impedance. Consequently, even in the experiments where the impedance is set to around 20kΩ, it becomes difficult to obtain good P300 responses. Furthermore, in c-VEP correlations are computed between reference templates and test signal. As known from section 2.3.2, CCA finds the weighting coefficients (W_x) and gives high weights to the relatively less noisy channels, making the system immune to noise. On the other hand, P300 uses SWLDA which is sensitive to noise, and even a small amount of noise can degrade the SWLDA classifier resulting into a wrong classification of targets. Thus, if the P300 and c-VEP paradigms are combined at high levels of impedance, the noisy P300 scores can deteriorate the c-VEP scores. This becomes evident from the trend observed in Table 3.3 where the accuracy and ITR go down when the impedance is 50kΩ.

In our initial P300 study, experiments were performed with different timings of P300 flash and the most feasible timing chosen was 250ms; 100ms active and 150ms passive time. As explained in section 2.3.1, the rows and columns are flashed randomly. However, randomizing the flashes can give rise to a problem where two flashes corresponding to the target symbol occur one after the other. In that case, it becomes difficult for the subject to follow up the mental count of the two flashes as the time difference between the two flashes is only 150ms (passive time). For instance, during the training when the first row and the first column corresponding to letter 'A' flash consecutively, it becomes difficult for the subject to register the mental count so fast with only 150ms in between for the cognitive processing. Compared to the conventional P300 where 12 rows/columns are flashed, our grouped P300 has a higher probability of flashing the target row and column successively.

From literature we know that P300 is generated in central and parietal lobes however, in our study we found out that P300 increases in its amplitude as we go towards occipital lobe and at Fz channel P300 is not observed. The reason

for this discrepancy is the selection of reference electrode position. For P300, positions corresponding to Cz, Cpz, right earlobe, and right mastoid are used as reference electrodes [53], [57]. Similarly for c-VEP, positions corresponding to Oz, earlobe, PO2, and POz are used as reference electrodes [8], [23], [58]. For hybrid experiments, positions corresponding to P8, T10, and right earlobe are chosen as reference electrodes [34], [38], [39]. However, we need to find the optimum reference electrode site for our proposed hybrid paradigm which may result in the improvement of quality of P300 and c-VEP signals.

Considering the results of the study, the proposed hybrid BCI speller would not be useful if the purpose is to increase the speed of the speller. This is because the independent c-VEP is fast enough to give an average ITR of 114.9 bits/min or higher, on its own. However, incorporating P300 and c-VEP is highly effective if the primary goal is to increase the accuracy, even if the ITR is compromised a little in the process. Although P300 does not play any significant role in increasing the ITR, it makes the BCI design more universal in a way that BCI system becomes useful for all kinds of people (e.g. a person with blurred vision would face problems working with c-VEP which depends on vision, whereas P300 stimulus, which depends on the cognitive ability, will assist that person in making the correct decision).

The present study requires further advancement in the performance of the hybrid paradigm. Thus, future work will help in further improving single-trial P300 in several ways. A sequence can be designed in which two target flashes do not occur one after the other. In addition, increasing the time in between the flashes will allow the user to perform the mental task conveniently. Furthermore, different techniques can be implemented for P300 which are less sensitive to noise e.g. CCA. We performed some preliminary work where CCA was implemented and the offline results showed some improvement compared to the results obtained using current technique. However, further research needs to be done in order to verify the new methods.

In the second part of this thesis, we implement a moving average model to simulate the c-VEP responses which is then used in test stage as a classifier.

Model based approach reduces the training time in a way that only one input sequence is used to obtain a model that fits to all the visual stimulus patterns. Data used in this modeling study is taken from BCI database of our lab. Average of the correlation between measured response and modeled response is 0.357 and 0.396 for 60Hz and 120Hz respectively. These correlations might not suggest that our model is very good, however, if we see the modeled and measured responses plotted together, it can be observed that the model captures major variations of the EEG correctly. Moreover, we use this model as a classifier and the accuracy and ITR obtained are 87.1% and 76.4 bits/min respectively for 60Hz. For 120Hz, we obtained two accuracy and ITR values. Acc_1 and ITR_1 are 68.6% and 85.0 bits/min respectively. Acc_2 and ITR_2 are 82.1% and 72.4 bits/min respectively, which are obtained using the averaged responses.

For the experiments, we train the subjects while setting up the hardware. The hybrid stimulus is displayed in working and we explain the subjects when to shift gaze and when to count. Even after taking all precautions, sometimes the obtained data turns out noisy as in the case of S7 in model based study. This subject may be an outlier, corrupting our data set, therefore specific criteria would be required for the screening of such subjects. Few subjects are unable to use BCI for communication purposes and this inability is referred to as BCI illiteracy. Suk et al. [59] proposed a method that predicts the BCI subject performance using spatio-temporal filters on 3 channels. Thus, a specific screening criterion in choosing the subjects, before carrying out the BCI experiments, could improve the data set considerably and help get rid of the outliers.

In c-VEP where m-sequence is used, letter 'A' is assigned m-sequence and the rest of the targets are assigned shifted versions of m-sequence which are orthogonal to each other. Therefore, the output responses of all the targets are expected to be uncorrelated (near orthogonal). However, from the recent work done by Basaklar [60], we understand that the output responses of m-sequence are not uncorrelated and in fact some of them are highly correlated which leads to misclassification of targets. To address this problem, the input stimulus sequence of all the letters can be optimized and in that case, the sequence may not be m-sequence which as a result would require the training for all the

letters. Nevertheless, we can use model based approach to improve the training time and use the optimized stimulus sequences simultaneously.

Future work may include the implementation of auto regressive moving average (ARMA) model and different non-linear models such as artificial neural network (ANN) for estimation of the EEG from any pseudorandom binary sequences. Robinson's corticothalamic model is physiological based model of brain which has linear, non-linear sigmoidal blocks and second order filters. Such a system can be put in the formalism of neural networks but the parameter identification was the problem. However, recent developments in neural networks give us the opportunity for fast parameter optimization which can be utilized for the task of modeling. Additionally, the inverse EEG to code model can also be implemented for estimation of the pseudorandom sequence from any given EEG and this EEG to code model can further be used for the task of target classification.

Bibliography

[1] J. R. Wolpaw, N. Birbaumer, D. J. Mcfarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.

[2] S. Gao, Y. Wang, X. Gao, and B. Hong, "Visual and auditory brain–computer interfaces," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 5, pp. 1436–1447, 2014.

[3] C.-T. Lin, C.-J. Chang, B.-S. Lin, S.-H. Hung, C.-F. Chao, and I.-J. Wang, "A real-time wireless brain–computer interface system for drowsiness detection," *IEEE transactions on biomedical circuits and systems*, vol. 4, no. 4, pp. 214–222, 2010.

[4] J. N. Mak and J. R. Wolpaw, "Clinical applications of brain-computer interfaces: current state and future prospects," *IEEE reviews in biomedical engineering*, vol. 2, pp. 187–199, 2009.

[5] M. V. R. Blondet, A. Badarinath, C. Khanna, and Z. Jin, "A wearable real-time bci system based on mobile cloud computing," in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 739–742, IEEE, 2013.

[6] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review," *sensors*, vol. 12, no. 2, pp. 1211–1279, 2012.

[7] S. Moghimi, A. Kushki, A. Marie Guerguerian, and T. Chau, "A review of eeg-based brain-computer interfaces as access pathways for individuals with severe disabilities," *Assistive Technology*, vol. 25, no. 2, pp. 99–110, 2013.

[8] G. Bin, X. Gao, Y. Wang, B. Hong, and S. Gao, "Vep-based brain-computer interfaces: time, frequency, and code modulations [research frontier]," *IEEE Computational Intelligence Magazine*, vol. 4, no. 4, pp. 22–26, 2009.

[9] M. B. Khalid, N. I. Rao, I. Rizwan-I-Haque, S. Munir, and F. Tahir, "Towards a brain computer interface using wavelet transform with averaged and time segmented adapted wavelets," *2009 2nd International Conference on Computer, Control and Communication*, 2009.

[10] S. Laureys, M. Boly, and G. Tononi, "Functional neuroimaging," *The Neurology of Consciousness*, pp. 31–42, 2009.

[11] L. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, 1988.

[12] C. M. Sinclair, M. C. Gasper, and A. S. Blum, "Basic electronics in clinical neurophysiology," in *The Clinical Neurophysiology Primer*, pp. 3–18, Springer, 2007.

[13] A. B. Usakli, "Improvement of eeg signal acquisition: An electrical aspect for state of the art of front end," *Computational Intelligence and Neuroscience*, vol. 2010, pp. 1–7, 2010.

[14] G. Gargiulo, R. A. Calvo, P. Bifulco, M. Cesarelli, C. Jin, A. Mohamed, and A. van Schaik, "A new eeg recording system for passive dry electrodes," *Clinical Neurophysiology*, vol. 121, no. 5, pp. 686–693, 2010.

[15] W. M. Leach, "Fundamentals of low-noise analog circuit design," *Proceedings of the IEEE*, vol. 82, no. 10, pp. 1515–1538, 1994.

[16] A. Kübler, B. Kotchoubey, J. Kaiser, J. R. Wolpaw, and N. Birbaumer, "Brain-computer communication: Unlocking the locked in.," *Psychological Bulletin*, vol. 127, no. 3, pp. 358–375, 2001.

[17] B. Anand, G. Chhina, and B. Singh, "Some aspects of electroencephalographic studies in yogis," *Electroencephalography and Clinical Neurophysiology*, vol. 13, no. 3, pp. 452–456, 1961.

[18] A. Black, "The operant conditioning of central nervous system electrical activity," *Psychology of Learning and Motivation*, pp. 47–95, 1972.

[19] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," *Proceedings of the IEEE*, vol. 89, no. 7, pp. 1123–1134, 2001.

[20] K.-H. Lee, L. M. Williams, M. Breakspear, and E. Gordon, "Synchronous gamma activity: a review and contribution to an integrative neuroscience model of schizophrenia," *Brain Research Reviews*, vol. 41, no. 1, pp. 57–78, 2003.

[21] J. H. H., "The ten-twenty electrode system of the international federation," *Electroencephalography and Clinical Neurophysiology*, vol. 10, pp. 371–375, 1958.

[22] J. V. Odom, M. Bach, C. Barber, M. Brigell, M. F. Marmor, A. P. Tormene, and G. E. Holder, "Visual evoked potentials standard (2004)," *Documenta ophthalmologica*, vol. 108, no. 2, pp. 115–123, 2004.

[23] M. Spüler, W. Rosenstiel, and M. Bogdan, "Online adaptation of a c–vep brain–computer interface(bci) based on error–related potentials and unsupervised learning," *PLoS ONE*, vol. 7, Jul 2012.

[24] B. Wittevrongel, E. V. Wolputte, and M. M. V. Hulle, "Code–modulated visual evoked potentials using fast stimulus presentation and spatiotemporal beamformer decoding," *Scientific Reports*, vol. 7, Aug 2017.

[25] S. W. Golomb, *Shift register sequences*. Aegean Park Press., 1982.

[26] S. Sutton, M. Braren, J. Zubin, and E. John, "Evoked-potential correlates of stimulus uncertainty," *Science*, vol. 150, no. 3700, pp. 1187–1188, 1965.

[27] H. Cecotti and A. J. Ries, "Best practice for single-trial detection of event-related potentials: Application to brain-computer interfaces," *International Journal of Psychophysiology*, vol. 111, pp. 156–169, 2017.

[28] J. Mak, Y. Arbel, J. W. Minett, L. M. McCane, B. Yuksel, D. Ryan, D. Thompson, L. Bianchi, and D. Erdogmus, "Optimizing the p300-based brain–computer interface: current status, limitations and future directions," *Journal of neural engineering*, vol. 8, no. 2, p. 025003, 2011.

[29] E. Donchin and D. B. Smith, "The contingent negative variation and the late positive wave of the average evoked potential," *Electroencephalography and clinical Neurophysiology*, vol. 29, no. 2, pp. 201–203, 1970.

[30] M. Kaur, P. Ahmed, A. Soni, and M. Q. Rafiq, "Wavelet transform use for p300 signal clustering by self-organizing map," *International Journal of Scientific and Engineering Research*, vol. 4, no. 11, pp. 1631–1638, 2013.

[31] B. Z. Allison, D. J. McFarland, G. Schalk, S. D. Zheng, M. M. Jackson, and J. R. Wolpaw, "Towards an independent brain–computer interface using steady state visual evoked potentials," *Clinical neurophysiology*, vol. 119, no. 2, pp. 399–408, 2008.

[32] B. Z. Allison, C. Brunner, V. Kaiser, G. R. Müller-Putz, C. Neuper, and G. Pfurtscheller, "Toward a hybrid brain–computer interface based on imagined movement and visual attention," *Journal of neural engineering*, vol. 7, no. 2, p. 026007, 2010.

[33] C. Brunner, B. Z. Allison, D. J. Krusienski, V. Kaiser, G. R. Müller-Putz, G. Pfurtscheller, and C. Neuper, "Improved signal processing approaches in an offline simulation of a hybrid brain–computer interface," *Journal of neuroscience methods*, vol. 188, no. 1, pp. 165–173, 2010.

[34] J. Long, Y. Li, T. Yu, and Z. Gu, "Target selection with hybrid feature for bci-based 2-d cursor control," *IEEE Transactions on biomedical engineering*, vol. 59, no. 1, pp. 132–140, 2012.

[35] T. Luth, D. Ojdanic, O. Friman, O. Prenzel, and A. Graser, "Low level control in a semi-autonomous rehabilitation robotic system via a brain-computer interface," in *2007 IEEE 10th International Conference on Rehabilitation Robotics*, pp. 721–728, IEEE, 2007.

[36] G. Pfurtscheller, T. Solis-Escalante, R. Ortner, P. Linortner, and G. R. Muller-Putz, "Self-paced operation of an ssvep-based orthosis with and without an imagery-based "brain switch:" a feasibility study towards a hybrid bci," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 18, no. 4, pp. 409–414, 2010.

[37] B. Z. Allison, C. Brunner, C. Altstätter, I. C. Wagner, S. Grissmann, and C. Neuper, "A hybrid erd/ssvep bci for continuous simultaneous two dimensional cursor control," *Journal of neuroscience methods*, vol. 209, no. 2, pp. 299–307, 2012.

[38] E. Yin, Z. Zhou, J. Jiang, F. Chen, Y. Liu, and D. Hu, "A novel hybrid bci speller based on the incorporation of ssvep into the p300 paradigm," *Journal of neural engineering*, vol. 10, no. 2, p. 026012, 2013.

[39] E. Yin, T. Zeyl, R. Saab, T. Chau, D. Hu, and Z. Zhou, "A hybrid brain–computer interface based on the fusion of p300 and ssvep scores," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 4, pp. 693–701, 2015.

[40] P. A. Robinson, C. J. Rennie, and D. L. Rowe, "Dynamics of large–scale brain activity in normal arousal states and epileptic seizures," *Physical Review E*, vol. 65, Nov 2002.

[41] J. Roberts and P. Robinson, "Quantitative theory of driven nonlinear brain dynamics," *NeuroImage*, vol. 62, no. 3, pp. 1947–1955, 2012.

[42] S. Zhang, X. Han, X. Chen, Y. Wang, S. Gao, and X. Gao, "A study on dynamic model of steady-state visual evoked potentials," *Journal of neural engineering*, vol. 15, no. 4, p. 046010, 2018.

[43] S. M. M. Safi, M. Pooyan, and A. M. Nasrabadi, "Ssvep recognition by modeling brain activity using system identification based on box-jenkins model," *Computers in biology and medicine*, vol. 101, pp. 82–89, 2018.

[44] S. Nagel and M. Spüler, "Modelling the brain response to arbitrary visual stimulation patterns for a flexible high-speed brain-computer interface," *PloS one*, vol. 13, no. 10, p. e0206107, 2018.

[45] D. H. Brainard, "The psychophysics toolbox," *Spatial Vision*, vol. 10, pp. 433–436, Jan 1997.

[46] D. G. Pelli, "The videotoolbox software for visual psychophysics: transforming numbers into movies," *Spatial Vision*, vol. 10, pp. 437–442, Jan 1997.

[47] M. Kleiner, D. Brainard, D. Pelli, A. Ingling, R. Murray, and C. Broussard, "What's new in psychtoolbox–3," *Perception*, vol. 36, no. 14, pp. 1–16, 2007.

[48] G. Schalk, D. Mcfarland, T. Hinterberger, N. Birbaumer, and J. Wolpaw, "Bci2000: A general–purpose brain–computer interface (bci) system," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1034–1043, 2004.

[49] R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen, "Fieldtrip: Open source software for advanced analysis of meg, eeg, and invasive electrophysiological data," *Computational Intelligence and Neuroscience*, vol. 2011, pp. 1–9, 2011.

[50] V. Bostanov, "Bci competition 2003-data sets ib and iib: feature extraction from event-related brain potentials with the continuous wavelet transform and the t-value scalogram," *IEEE Transactions on Biomedical engineering*, vol. 51, no. 6, pp. 1057–1061, 2004.

[51] D. J. Krusienski, E. W. Sellers, F. Cabestaing, S. Bayoudh, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "A comparison of classification techniques for the p300 speller," *Journal of neural engineering*, vol. 3, no. 4, p. 299, 2006.

[52] R. A. Fisher *et al.*, "138: The use of multiple measurements in taxonomic problems.," 1936.

[53] D. J. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "Toward enhanced p300 speller performance," *Journal of neuroscience methods*, vol. 167, no. 1, pp. 15–21, 2008.

[54] F. Di Russo and D. Spinelli, "Electrophysiological evidence for an early attentional mechanism in visual processing in humans," *Vision research*, vol. 39, no. 18, pp. 2975–2985, 1999.

[55] S. Shahid, G. Prasad, and R. K. Sinha, "On fusion of heart and brain signals for hybrid bci," in *2011 5th International IEEE/EMBS Conference on Neural Engineering*, pp. 48–52, IEEE, 2011.

[56] J. R. Wolpaw, H. Ramoser, D. J. McFarland, and G. Pfurtscheller, "Eeg-based communication: improved accuracy by response verification," *IEEE transactions on Rehabilitation Engineering*, vol. 6, no. 3, pp. 326–333, 1998.

[57] H. Serby, E. Yom-Tov, and G. F. Inbar, "An improved p300-based brain-computer interface," *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 13, no. 1, pp. 89–98, 2005.

[58] Y. Wang, R. Wang, X. Gao, B. Hong, and S. Gao, "A practical vep-based brain-computer interface," *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 14, no. 2, pp. 234–240, 2006.

[59] H.-I. Suk, S. Fazli, J. Mehnert, K.-R. Müller, and S.-W. Lee, "Predicting bci subject performance using probabilistic spatio-temporal filters," *PloS one*, vol. 9, no. 2, p. e87056, 2014.

[60] T. Başaklar, Y. Tuncel, and Y. Z. Ider, "Effects of high stimulus presentation rate on eeg template characteristics and performance of c-vep based bcis," *Biomedical Physics & Engineering Express*, vol. 5, no. 3, p. 035023, 2019.

Appendix A

Data Acquired During the Experiments

All the experimental data is uploaded on the BCI hard drive of our lab and saved in folder named "Hybird P300 & c-VEP". We have data of eight initial experiments in the folders named "denemeX" where X is the experiment number. These experiments were done to test the simulator for different variations of setup (e.g., different active and passive times, lights on and light off inside the room, different zooming effects of letters, different cognitive tasks for P300, and different gaze shifting time). Then the folders that are named as "Date Experiment" where Date is the date when we conducted the experiment. This experimental data was used for 'Initial P300 Study' and final 'Hybrid Experiments'. Each folder has a log file, which mentions the name of subject, type of experiments, and impedance of subject. The experimental data used for the 'Hybrid Experiment Results' is in the following folders: 03.05.2019 Experiment, 03.07.2019 Experiment, 03.08.2019 Experiment, 03.09.2019 Experiment, 03.11.2019 Experiment, 03.25.2019 Experiment, and 04.04.2019 Experiment. We have also compiled a excel sheet (Hybrid BCI Experiments) which has all the accuracy and ITR values of all the subjects. Note that the experimental data is also included in the CD submitted along with this thesis.

Data for modeling study is saved on hard drive of our lab in folder named "CVEP". This folder has the data from the study of high stimulus presentation rates on c-VEP based BCIs spellers, done by Toygun Başaklar. Each folder has a log file, which mentions the name of subject, and type of experiment.

Appendix B

Software Used in Experiments and in Post-processing

We have codes running on 2 computers: stimulator computer and recording computer. All the codes are added in CD, submitted along with the thesis. "BCI GUI 2.zip" file has the code for recording computer and "Hybrid Stimulator.zip" has the code for stimulator computer.

1. Run BCI2000 software on recording computer, set the configuration and start the 'fieldtrip buffer'.

2. From the "BCI GUI 2" folder run the "GUI.m" script on recording computer and the GUI will pop up. This GUI can be used for individual P300, individual c-VEP, and hybrid (c-VEP and P300) experiments. On the third column, we have hybrid experiment options. There are 5 tabs for hybrid experiment column: parameters, record hybrid training data, generating classifier, online testing, and offline testing.

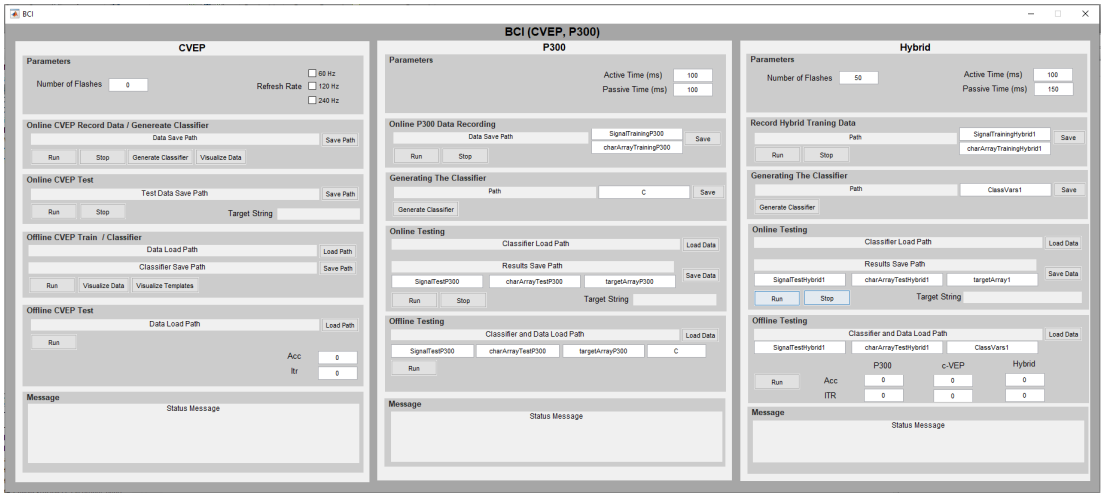


Figure B.1: GUI running on recording computer

3. For training session, enter number of flashes, active time, and passive time. In the "Record Hybrid Training Data", select the save directory and press run.

4. From the "Hybrid Stimulator" folder run the "GUI.m" script on stimulator computer and the GUI will pop up. This GUI has also 3 different stimulators combined: individual P300, individual c-VEP, and hybrid (c-VEP and P300). Different effects for symbols can be selected and different size of the fonts can be chosen.

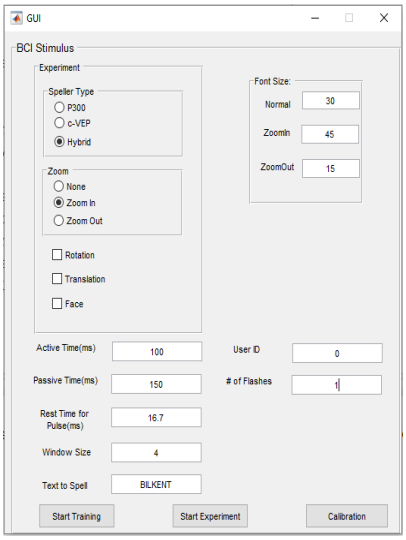


Figure B.2: GUI running on stimulator computer

5. Since on the recording computer training script is running, select hybrid option and then enter 'A' in "Text to Spell" field. Also, for P300 flashes select zoom in option. Enter number of flashes, active time, passive time and press "Start Training" button. Now the training will start!

6. On recording computer, select the directory in "Generating the classifier" tab and press "generate classifier" button. Once the classifier is generated, the message box will display a message.

7. For test session, load the classifier first (that we generated in previous step) from the "Online Testing" tab. Next, select the directory to save the experimental data and press run.

8. On stimulator computer, select hybrid option and then enter 'BILKENT-BCIEXPERIMENT' in "Text to Spell" field. Also, for P300 flashes select zoom in option. Enter number of flashes, active time, passive time and press "Start Experiment" button. Now the test session will start!

9. During the test session, the results of online classification will be shown in the "Target String" text field on recording computer.

10. On the recording computer, GUI provides offline analysis option under the heading "Offline Testing".