Improving Brain-Computer Interface Performance By Using Dynamic Methods Based on Analysis of Cognitive State

by

Abdulrahman W. Aref

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biomedical Engineering)
in The University of Michigan
2018

Doctoral Committee:

Assistant Professor Cynthia Chestek, Co-Chair
Research Assistant Professor Jane E. Huggins, Co-Chair
Professor Douglas C. Noll
Assistant Professor William Stacey

Abdulrahman W. Aref

awaref@umich.edu

OCRID iD: 0000-0002-5210-6056

To my parents, who have been my greatest influences in my education and character.

To my family, those still here and those who have passed on.

Acknowledgments

I would like to express the deepest gratitude to the following that without their help, this research and dissertation would not be possible.

My parents, Walid Aref and Faika Said, who have provided me with every resource I ever needed to excel, without hesitation. I cannot thank you enough for all the unconditional love and support you give me when I need it the most.

My sister, Safaa Aref, and my two brothers, Mohammad Aref and Jalaleldeen Aref, have been by my side through the best and the worst and have all been a huge influence on the person I am today. I know we have not been able to spend a lot of time together over the past couple years, but you are always on my mind.

My research advisor, Dr. Jane Huggins, deserves the highest praise, as she has been my greatest mentor over the last several years. She has instilled in me the curiosity of a researcher and led by example. Thank you very much for being patient with me. I will forever remember you as one of my greatest teachers.

My friends, who have provided me with a steady amount of support, comfort, and inspiration at times of hardship, I would not have been able to complete this dissertation without you.

Preface

With the exception of the first and last chapters, this dissertation is comprised of manuscripts in various stages of preparation and submission. The overall story of my research has been molded out of these manuscripts, and I have attempted to minimize recurrent and overlapping text. However, some degree of recurrent text was unavoidable.

To report studies in a more organized manner, I split up one of the manuscripts into 3 different chapters. Chapters 3, 4, and 5 together comprise one manuscript to be submitted. Therefore, any repetition of topics in these chapters would be removed upon submission. Organizing my dissertation in this way, although not chronological, presents the research in a coherent manner that is easier to follow.

Table of Contents

List of Tables

List of Figures

Abstract

Communication for individuals with severe motor and speech impairments can be very difficult and they find the need for the assistance of augmentative and alternative communication (AAC) systems. Common commercialized AAC systems require some amount of voluntary control and are unusable by individuals with severe disabilities. Non-invasive brain-computer interfaces (BCIs) are an emerging means of communication for people with severe motor and speech impairments. BCIs allow the user to make selections on the computer just using their brain signals, electroencephalogram (EEG). However, although they are revolutionary for individuals that cannot control other available AAC systems, BCIs have several limitations. Two major limitations of BCIs are: 1) BCIs are static/synchronous in nature; 2) BCIs are susceptible to changes in user attention. Since people in the populations that need BCI technology the most (e.g. amyotrophic lateral sclerosis ;ALS) may experience attention impairments, incorporating attention-monitoring features into the BCI would improve BCI performance by reducing errors in these populations. This research presents two dynamic methods developed to help the BCI become more user-aware and allow users to control the BCI at their own pace. Using a well-established negative correlation between alpha band power in the EEG and attention, the first method used alpha band analysis to detect losses in user attention and abstained selections that were unattended to reduce potential errors. The second method, called P300-Certainty, abstained selections that do not reach a specified confidence level. To test both methods, off-line analysis was performed on recorded EEG from 30 subjects using the BCI for spelling. Subjects selected 9 sentences and at least 23 characters per sentence with additional corrections. Alpha band analysis and P300-Certainty were tested off-line, separately and together, on

this dataset to determine their efficacy at increasing BCI accuracy by abstaining potential errors. In addition, P300-Certainty was implemented in a BCI-facilitated cognitive assessment to reduce potential errors, as well as, only choosing selections when they reach a specified confidence level. The on-line performance of P300-Certainty was calculated from this data. Alpha band analysis was performed off-line on this on-line data to determine its efficacy at increasing P300-Certainty on-line BCI accuracy.

Alpha band power was shown to be significant between correct and incorrect character selections with a significance of $p = 0.01004$. Using this significance, alpha band analysis was used to classify selections as correct or incorrect based on the EEG, however it was only improved accuracy for a subset of subjects (subjects exhibiting high alpha variance). Off-line analysis of P300-Certainty was shown to increase accuracy from 82.01±12.59% to 88.82±8.85% by abstaining potential errors, with a statistical significance of $p = 0.038$. Furthermore, P300-Certainty and alpha band analysis used together, improved BCI accuracy, over all subjects, more than either method did alone. The increase was statistically significant ($p = 0.008$) when compared to the raw BCI accuracy. The on-line accuracy of P300-Certainty was 83.62 ± 9.14%.

Alpha band analysis and P300-Certainty abstain potential errors using different, yet orthogonal, methods of measuring attention. Each method abstains potential errors that the other may have not detected. In conclusion, this research has introduced two methods that quantify attention in orthogonal ways that increase BCI accuracy by abstaining potential errors more than either method alone. Using these methods together allows the BCI to be more user-aware and allows the user to type at their own pace.

Chapter 1
Motivation

Introduction

Communication for individuals with severe motor and speech impairments can be very difficult and requires the assistance of an augmentative and alternative communication (AAC) systems. Common commercialized AAC systems require some amount of voluntary control and are unusable by individuals who have the most severe disabilities (e.g. completely unable to move). Non-invasive brain-computer interfaces (BCIs) are an emerging means of communication for people with severe motor and speech impairments. BCIs allow the user to make selections on the computer just using their brain signals. One common BCI that is utilized for communication is the P300 BCI, which is controlled by the user's P300 response (found in user's brain signals). However, although they are revolutionary for individuals that cannot control other available AAC systems, P300 BCIs have several limitations. Two major limitations of BCIs are:

1. P300 BCIs are static/synchronous in nature
2. P300 BCIs are susceptible to changes in user attention

This dissertation includes two dynamic methods that help the BCI become more user-aware and allow users to control the BCI at their own pace.

Background

BCI

Wolpaw defined a BCI as a “non-muscular channel for sending messages and commands to the external world” [1]. Non-invasive brain-computer interfaces (BCIs) are an emerging means of communication for people with severe motor and speech impairments, since they do not require voluntary muscle control. There have been many studies proposing BCIs controlled by different aspects of a user’s brain signals, including visually evoked potentials [2,3], motor imagery [3-5], slow cortical potentials [6], and several others. One signal that is commonly used to control BCIs is the P300 response, which is a signal that appears in the electroencephalogram (EEG) about 300ms after a relevant, yet unpredictable event [7]. This dissertation focuses on improving performance of non-invasive P300-based BCIs (P300 BCI) by further analyzing the user’s cognitive state (attention).

P300 BCI

First introduced by Farwell and Donchin in [7], the P300 BCI is one of the easiest to learn and use. P300 BCIs have been used extensively in research to aid with communication in populations with severe motor and speech impairments [3, 8-10].

The P300 response, also called the oddball response, was first described in 1965 in [11], and is evident in the user’s EEG around 300ms after a rare and desired stimulus is presented among many others. Figure 1 shows an example of the P300 response found in the EEG that is used to control P300 BCIs.

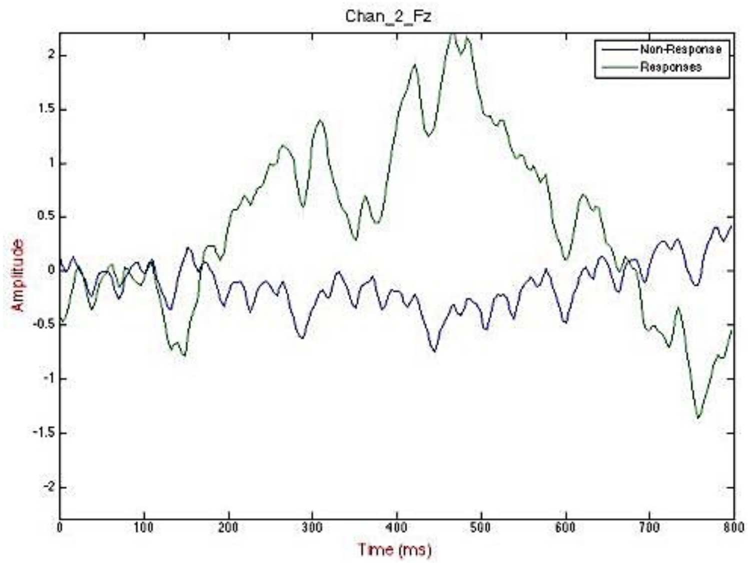

Figure 1: This graph shows an example of a P300 signal in response to a target (green line), while the blue line (centered around 0) represents the response (no P300 present) to a non-target. This graph was generated by averaging 30 responses to targets and non-targets. The amplitude is in uV.

The P300 Speller

The P300 Speller is a common P300 BCI design that allows spelling utilizing the user's EEG (P300 response). There are many variations of P300 BCIs, however the one being presented is the most basic, original design. The user is presented with a grid of selections (can be letters, numbers, icons, words, etc.), as shown in the example in Figure 2. The selections flash, in rows or columns in this application, for a user-customized number of times per character selection. The user is asked to concentrate on a selection and count how many times it flashes (whether it was a row flash or a column flash). The number of total flashes presented to the user is determined from the calibration data he/she completes before controlling the BCI. After all available selections flash for the user-specific number of times, the P300 Speller then determines which row and column elicited a reaction (P300 response) from the user. The selection at the intersection of the chosen row and column is the selection that is chosen by the BCI and is printed on the screen. The user then starts the process over again for the next character.

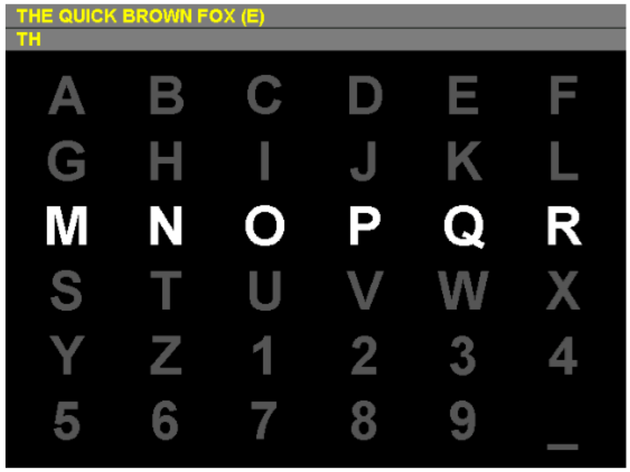

Figure 2: This is an example of the P300 Speller display. The 3rd row is intensified to show how the flashes look to the user.

All of the research reported in this dissertation was performed using a P300 BCI, with most using the P300 Speller.

P300 BCI Terminology

When talking about P300 BCIs, there are several terms to be familiar with that will help understand its mechanism of operation.

Flashes: This is defined as each individual stimulus presented to the user. Selections can flash in groups (e.g. rows and columns) or independently.

Sequences: Once all of the possible selections have flashed one time and only one time, this group of flashes is considered a sequence. Multiple sequences may be needed for the BCI to make a selection. This is due to the low signal-to-noise ratio of EEG.

Classifier Value: Before a user can control the BCI, they need to complete a calibration step. Step-Wise Linear Discriminant Analysis (SWLDA), other types of LDA, Least Squares (LS), or Support Vector Machines (SVMs) use the labeled responses from the calibration data to generate weights, as well as a user-specific number of sequences needed for accurate BCI selection. These weights are used to classify the user responses when controlling the BCI. When controlling the BCI, once all of the user-specific number of sequences are presented, the EEG responses within a certain time period (e.g. 0ms-800ms) after each flash are analyzed for the presence of a P300 response. The weights are applied to the EEG during that time period to help classify the responses. SWLDA, LS, or SVM are used to classify the presence or absence of a P300 response. After each sequence, all possible selections are assigned a score, called a **classifier value**, calculated by SWLDA to evaluate the presence of a P300 response. The higher the classifier value assigned to a selection, the more likely that a true P300 response was detected.

After all sequences have been flashed, the classifier values assigned to each row and column (in a row-column flash pattern) are combined over all sequences. The P300 Speller makes a selection based on which row-column have the highest combined classifier value. The intersection of the selected row and column is the selection that is made.

Attention and P300 BCI

The P300 BCI design is vulnerable to variations in attention due to the rapid nature of the flashes (31.25ms – 125ms per flash) and the length of time it takes to make a single selection (13.3s - 22.5s per selection depending on subject-specific performance). When controlling the P300 BCI, subjects use multiple sets of flashes (predetermined from a subject's calibration data) for the BCI to make a single selection. All of these flashes are taken into consideration by the BCI when making a selection. Large numbers of flashes tend to be more monotonous due to the length of time that user needs to attend to the BCI. If a user's attention wanders halfway through the number of flashes, then the selection the BCI makes (classification) has a larger probability to be incorrect or not be the one desired by the user. Users with a larger number of flashes need to pay attention to the BCI display for a longer period of time. On the other hand, some users need to have a larger number of flashes for optimal BCI performance. The ability to detect moments of decreased attention in the EEG of a BCI user would allow the BCI to not choose the corresponding selection, which would reduce total errors. In addition, this can also allow the BCI to only make selections when the user is paying attention.

Populations that need BCI technology the most (e.g., people with amyotrophic lateral sclerosis (ALS)) experience attention impairments [12].

Incorporating an attention-monitoring feature into the BCI would improve BCI performance by reducing errors for populations that need it the most.

Abstention of potentially erroneous selections

Traditional P300 BCIs tend to have a static nature. The BCI presents the user with a predetermined number of stimuli before making a decision. A decision is always made regardless of whether the user was attending to the BCI or not. This ends up with unintended or unattended selections being chosen by the BCI, most of which are errors. Since ideal BCI performance requires attention to the BCI, being able to detect losses of attention during use can help the BCI to reject potentially erroneous selections that are unintended or unattended. Rejecting selections that are potential errors is known as abstention.

Selection Confidence Methods

To be able to achieve a BCI that abstains potential errors, the BCI needs to be able to calculate a measure of confidence for each selection based on the user's responses to the BCI stimuli (flashes). If selections do not reach an appropriate confidence level, they are abstained. Methods that allow this functionality in a BCI are called selection confidence methods. In a sense, these selection confidence methods are an indirect measure of the user's attention to the BCI. Low confidence for a selection indicates that the user did not exhibit a P300 response when that selection was presented. Thus, the user was not attending to that selection.

Several studies have used BCIs with abstention by calculating confidence values for selections in different ways [3, 4]. Methods that have been used to calculated selection confidence values are: Independent Component Analysis (ICA) [5] and Bayesian Linear Discriminant Analysis (BLDA) [3, 6]. These methods all calculated selection confidence values for selections and evaluated them against a predetermined confidence threshold to determine whether a selection was retained or abstained. The study using ICA showed an increase of accuracy from 90% to 92.1%, while the study using BLDA showed a non-significant decrease in accuracy from 79.44±29.98% to 74.40±27.16%.

Dynamic Stopping

Instead of having the BCI make selections after a fixed number of sequences, selection confidence methods can also be used to make selections as soon as their confidence level surpasses a predetermined confidence threshold, regardless of how many sequences have been presented. For example, a user may not always need the same number of sequences to make a decision. A user who appeared from the calibration data to need 10 sequences for the BCI to make an accurate selection, may actually require less sequences for some selections. If at 4 sequences, the selection confidence level of one of the selections surpasses the threshold, then that selection is chosen and the user moves on to type the next selection. The use of selection confidence methods in this manner is referred to as dynamic stopping.

In recent years, dynamic stopping has been explored as an aid to increase the asynchronous nature of the P300 Speller. Lenhardt et al. developed an algorithm to decrease the number of sequences depending on the user's performance [17]. The study showed a significant increase in the information transfer rate (ITR) when compared to typing using the standard P300 speller without dynamic stopping. More recently, several studies have used Bayesian Linear Discriminant Analysis (BLDA) to classify the selections, calculated selection confidence probabilities using Bayes rule, and evaluated them against a confidence threshold to achieve dynamic stopping. Once a selection's probability surpasses the confidence threshold, no more sequences are flashed and that selection is chosen. It was found that spelling time was decreases from 21 minutes to 10.47±5.69 minutes, while maintaining accuracy. Park and Kim developed a dynamic stopping model that uses a reward point system based on the P300 classifier values to give each row or column a point value. The row or column with the greatest point value would be considered the target after each flash (the target should ideally remain the same for each sequence) and would be selected once it reached a predetermined point value [18]. Park and Kim's model showed a significant increase in accuracy from 85.63±15.6% to 92.5±0.09% [18]. Very similar studies have been performed using different algorithms to achieve dynamic stopping [19, 20].

Selection confidence methods utilized to achieve dynamic stopping are very useful when using a P300 BCI by calculating a measure of attention and allowing the user to type at their own pace based on their attention to the BCI. The measure of attention provided by selection confidence methods is derived from the user's responses to the BCI flashes, which in a sense is an indirect measure of attention. A more direct measure of attention will be explored in the next section.

Alpha band analysis of EEG as a measure of attention

A correlation between alpha band power (EEG frequency band of 8-13 Hz) and attention is well established in Clinical Neurophysiology and Cognitive Neuroscience studies, with low alpha band power corresponding to high attention levels and high alpha band power corresponding to low attention levels [21-23]. Furthermore, studies (outside a BCI context) have demonstrated that pre-stimulus alpha band power also correlates with the amplitude of the P300 response elicited by different kinds of stimuli [21, 22].

Alpha band analysis of EEG in a BCI context

Several studies have also been conducted to analyze alpha band power during a visual-spatial attention task [24-27]. These studies each used a variation on an EEG BCI that presented left and right stimuli while the user's eyes were straight ahead (testing covert attention). Alpha band power calculated from the EEG recorded from the parieto-occipital cortex in the brain (area that processes visual input) was used to classify to which stimuli a user was attending. Although alpha differences between left and right were used to determine which side the user was attending to, the studies did not explore the correlation between alpha and accuracy of BCI selections.

The study detailed in Chapter 2 examines the benefits of using alpha band analysis of the EEG as a direct measure of attention to improve BCI performance by abstaining possible errors where the user exhibits low levels of attention (high levels of alpha).

Significance

Due to the static/synchronous nature of BCI and the susceptibility to changes in user attention, its use is limited to those populations without severe motor and speech impairments or those with impairments that still retain some voluntary muscular control. Furthermore, populations that need BCI technology the most (e.g., people with amyotrophic lateral sclerosis (ALS)) experience attention impairments [3]. Incorporating an attention-monitoring feature to the BCI would improve BCI performance by reducing errors for populations that need it the most.

A brief description of populations that would benefit from a dynamic attention-monitoring BCI are as follows:

Amyotrophic lateral sclerosis (ALS): With incidence of about 2 out of 100,000 people [28], and often known as Lou Gherig's Disease, ALS is characterized by the progressive loss of motor control with the retention of sensory and cognitive function. ALS is one of the main motivators for BCI research since some of those affected reach a level of motor and speech impairment where AAC systems are very difficult to use. The progression of impairment in people with ALS gets so severe, that their breathing muscles cease to work and require the use of a ventilator. At this point, they are given a choice whether to use a ventilator to keep them alive or not. Many more people with ALS may choose to reach this state of disability and use a ventilator if an effective BCI was available. Currently BCIs provide limited aid to these individuals. Dynamic attention-monitoring BCIs have great potential to help these individuals communicate.

Cerebral palsy (CP): CP is a congenital condition that includes a large variety of motor deficiencies. With incidence of 2.4 out of 1000 births [29], about 16% of people with CP exhibit impairments severe enough to warrant the use of a BCI [30].

Neuromuscular disease, spinal cord injury, brainstem stroke, and traumatic brain injury, are all conditions that may result in severe motor and speech impairments and warrant the use of a BCI for communication.

Some individual with the conditions listed above experience severe levels of impairment to the point where BCIs may be the only option for communication.

Note that there are many other conditions that cause motor and speech impairment that may benefit from the use of BCI, and the attraction to BCI technology to these populations may increase as BCI performance and effectiveness increases.

Outline of Chapters

The primary goal of this research was to develop methods to improve BCI performance by allowing them to be more aware of the user's cognitive state (attention). Two different approaches were taken to achieve this dynamic attention-monitoring BCI. The first approach, described in Chapter 2, was to use a well-studied attention-marker directly derived from the user's EEG (alpha band power) to improve BCI performance. This is achieved by reducing BCI spelling errors through abstaining selections during which users exhibit losses in attention.

The second approach, described in Chapter 3, was to develop a selection confidence method (the P300-Certainty algorithm) to improve BCI performance by abstaining selections that do not reach a specified confidence level.

Chapter 4 explores using the P300-Certainty algorithm to achieve dynamic stopping by allowing the BCI to make a selection once the confidence of a selection reaches a specified level. This permits the BCI to be much more dynamic and allows the user to control the BCI at their own pace.

To bring it all together, Chapter 5 combines the methods introduced in Chapter 2 (alpha band analysis) and Chapter 3 (P300-Certainty) to demonstrate the effectiveness of these two orthogonal methods in improving BCI performance more than either method used alone.

Finally, Chapter 6 includes a concluding discussion summarizing the contributions made by the work presented in this dissertation, including proposed future work based on these contributions.

Overall, the work presented in this dissertation describes methods that can be used to create a BCI that is resilient to wandering user attention and allows users to control the BCI at their own pace.

References

[1] Wolpaw, JR, Birbaumer, N, McFarland, DJ, Pfurtscheller, G., and Vaughan, TM, “Brain-computer interfaces for communication and control,” *Clinical Neurophysiology*, vol. 113, no. 6, pp. (2002): 767–791.

[2] Yoshimura, N, Itakura, N, “A transient VEP-based real-time brain-computer interface using non-direct gazed visual stimuli.” *Electromyogr.Clin.Neurophysiol.* 49 (2009):323-335.

[3] Nicolas-Alonso, LF, Gomez-Gil, J “Brain computer interfaces, a review.” *Sensors (Basel)* 12 (2012):1211-1279.

[4] An, X, Kuang, D, et al. “A Deep Learning Method for Classification of EEG Data Based on Motor Imagery.” (2014):203-210.

[5] Bai, O, Lin, P, et al. “A high performance sensorimotor beta rhythm-based brain- computer interface associated with human natural motor behavior.” *J.Neural Eng.* 5 (2008):24-35.

[6] Kubler, A, Neumann, N, et al. “Brain-computer communication: self-regulation of slow cortical potentials for verbal communication.” *Arch.Phys.Med.Rehabil.* 82(2001):1533- 1539.

[7] Farwell, LA, and Donchin, E, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials", *Electroenceph. Clin. Neurophysiol.*, vol. 70, no. 6, (1988): 510 - 523.

[8] Kubler, A, Furdea, A, et al. “A brain-computer interface controlled auditory event- related potential (p300) spelling system for locked-in patients.” *Ann.N.Y.Acad.Sci.* 1157 (2009):90-100.

[9] D'Avanzo C, Schiff S, et al. “A Bayesian method to estimate single-trial event- related potentials with application to the study of the P300 variability.” *J.Neurosci.Methods* 198 (2011):114-124.

[10] Zickler, C, Riccio, A, et al. “A brain-computer interface as input channel for a standard assistive technology software.” *Clin.EEG Neurosci.* 42 (2011):236-244.

[11] Sutton, S, Braren, M, Zubin, J, and John, ER, “Evoked-potential correlates of stimulus uncertainty,” *Science*, vol. 150, no. 700 (1965):1187–1188.

[12] Phukan, J, Pender, NP, and Hardiman, O. "Cognitive Impairment in Amyotrophic Lateral Sclerosis." *The Lancet Neurology* 6.11 (2007): 994-1003.

[13] Verschore, H, Kindermans P, Verstraeten, D, and Schrauwen, B "Dynamic Stopping Improves the Speed and Accuracy of a P300 Speller." *Artificial Neural Networks and Machine Learning – ICANN 2012* 7552 (2012): 661-68.

[14] Zhang, H, Guan, C, and Wang, C "Asynchronous P300-Based Brain--Computer Interfaces: A Computational Approach With Statistical Models." *IEEE Transactions on Biomedical Engineering* 55.6 (2008): 1754-763.

[15] Serby, H, Yom-Tov, E, & Inbar, G, “An Improved P300-Based Brain-Computer Interface,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 13 (2005): 89-98.

[16] Mainsah, BO, Collins, LM, Colwell, KA, Sellers, EW, Ryan, DB, Caves, K, & Throckmorton, CS, “Increasing BCI communication rates with dynamic stopping towards more practical use: An ALS study,” *Journal of Neural Engineering*, 12 (2015), 016013.

[17] Lenhardt, A, Kaper, M, and Ritter, H, "An Adaptive P300-Based Online Brain–Computer Interface." *IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING* 16.2 (2008): 121-30.

[18] Park, JY, and Kim, K, "A POMDP Approach to Optimizing P300 Speller BCI Paradigm." *IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING* 20.4 (2012): 584-94.

[19] Schreuder, M, Hohne, J, Treder, M, Blankertz, B, and Tangermann, M, "Performance Optimization of ERP-Based BCIs Using Dynamic Stopping." *ConfProc IEEE Eng Med Biol Soc.* 10 (2011): 4580-583

[20] Mainsah, B, Collins, LM, Colwell, K, Throckmorton CS, "Accelerating Dynamic Data Collection in P300 Spellers with Language Modeling." *EEG & Clinical Neuroscience Society.* (2012) Poster session

[21] Polich, J, "Updating P300: An Integrative Theory of P3a and P3b." *Clinical Neurophysiology* 118.10 (2007): 2128-148

[22] Hanslmayr, S, "Visual Discrimination Performance Is Related to Decreased Alpha Amplitude but Increased Phase Locking." *Neuroscience Letters* 375.1 (2005): 64-68.

[23] Gleiss, S, and Kayser, C "Acoustic Noise Improves Visual Perception and Modulates Occipital Oscillatory States." *Journal of Cognitive Neuroscience* 26.4 (2014): 699-711.

[24] Kelly, SP, "Increases in Alpha Oscillatory Power Reflect an Active Retinotopic Mechanism for Distracter Suppression During Sustained Visuospatial Attention." *Journal of Neurophysiology* 95.6 (2006): 3844-851.

[25] Kelly, SP, Lalor, EC, Reilly, RB, and Foxe, JJ, "Visual Spatial Attention Tracking Using High-Density SSVEP Data for Independent Brain–Computer Communication." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 13.2 (2005): 172-78.

[26] Kelly, SP, Lalor, EC, Reilly, RB, and Foxe, JJ, "Independent Brain Computer Interface Control Using Visual Spatial Attention-Dependent Modulations of Parieto-occipital Alpha." *Conference Proceedings. 2nd International IEEE EMBS Conference on Neural Engineering*, (2005): 667-670.

[27] Treder, MS, Bahramisharif, A, Schmidt, NM, Van Gerven, MA, and Blankertz, B, "Brain-computer Interfacing Using Modulations of Alpha Activity Induced by Covert Shifts of Attention." *Journal of NeuroEngineering and Rehabilitation* 8.1 (2011): 24.

[28] "CDC - Amyotrophic Lateral Sclerosis: FAQ." [Online]. Available: https://wwwn.cdc.gov/ALS/ALSFAQ.aspx#b.

[29] Miller, F, and SpringerLink (Online service), Cerebral palsy. New York: Springer, (2005).

[30] Himmelmann, K, Lindh, K, Cooley Hidecker, MJ, "Communication ability in cerebral palsy: A study from the CP register of western Sweden" *European Journal of Paediatric Neurology*. Vol. 17 Issue 6 (2013) 568-574.

[31] Donchin, E, Heffley, E, Hillyard, SA, Loveless, N., Maltzman, I., Ohman, A., Rösler, F., Ruchkin, D., Siddle, D., "Cognition and event-related potentials. II. The orienting reflex and P300" *Ann. NY Acad. Sci.*, 425 (1984): 39–57.

Chapter 2
Improving Accuracy of Individual BCI Selections Using Alpha Band Analysis

The text in this chapter is reformatted from a paper submitted to the Journal of Neural Engineering.

Abstract

This study used off-line analysis of data from a P300-based Brain-Computer Interface (BCI). It has long been known that increased power in the EEG alpha frequency band (8-13Hz) is negatively correlated with the user's attention to the current task. Lack of user attention during BCI use is likely to cause errors in BCI selections. Recorded EEG from 30 subjects using the BCI was analyzed to determine the relationship between normalized alpha band power and accuracy of BCI selections. The alpha power for each character selection was compared to the correctness of each character, which was significantly different for correct versus incorrect characters with p-value of 0.01004, and thus indicating that it could be used to predict accuracy of the selection. Taking advantage of the significant difference between correct and incorrect characters, a machine-learning method to classify accuracy based on alpha for each subject was used to block erroneous selections in an off-line analysis. It was also found that the alpha variance exhibited by the user predicted the effectiveness of alpha-based classification in improving BCI accuracy. Specifically, users exhibiting high alpha variance, in the calibration data, showed improvement in BCI accuracy from alpha classification, while users exhibiting low alpha variance did not show improvement in BCI accuracy using alpha classification. For users exhibiting high variance, the mean BCI accuracy for

raw BCI performance and BCI performance with an alpha-based classification were 80.24±10.70% and 87.5±8.57%, respectively, which was statistically significant with a p-value = 0.041. Thus, alpha band power can be used to quantify attention in a BCI context and improve the BCI accuracy for subjects exhibiting high alpha variance. This suggests that lapses in attention cause some, but not all, BCI errors.

Introduction

Many individuals with severe motor impairments need assistive technologies to aid them in communication. Of those individuals, the ones that have little-to-no motor control or are locked-in cannot use common commercialized augmentative and alternative communication systems since most require at least some amount of motor control. Non-invasive brain-computer interfaces (BCIs) are an emerging means of communication for people with severe motor and speech impairments. One signal that is commonly used to control BCIs is the P300 response, which is a signal that appears in the electroencephalogram (EEG) about 300ms after a relevant, yet unpredictable event [1]. Farwell and Donchin developed the first P300 speller based on the P300 response [2]. The P300 speller presents the user with selections from a matrix by intensifying their color (flashing them from grey to white) and chooses the selection to which user's EEG exhibits P300 responses.

Background

The P300 Speller

The P300 Speller is a common P300 BCI design that allows spelling utilizing the user's EEG (P300 response). The user is presented with a grid of selections (can be letters, numbers, icons, words, etc.), an example of which is in Figure 2 in Chapter 1. The selections flash in groups (rows or columns) for a user-specific number of times per character selection. To make a selection, the user is asked to concentrate on a selection and count how many times it flashes, which is intended to help focus the user's attention. The total number of flashes presented to the user is determined from the calibration data he/she completes before being able to use the BCI. After all available selections flash for the user-specific number of times, the P300 Speller then determines which row and column elicited a reaction (P300 response) from the user. The selection that is chosen by the BCI and printed on the screen is the selection at the intersection of the chosen row and column. To type the next character, the user then starts the process over again.

Due to the rapid nature of flashes (31.25ms – 125ms per flash) and the length of time it takes to make selections (13.3s - 22.5s per selection depending on subject-specific performance), the P300 BCI design is vulnerable to variations in attention. When controlling the P300 BCI, subjects use multiple sets of flashes (predetermined from a subject's calibration data) for the BCI to make a single selection. When making a selection, all of the data from these flashes are taken into consideration by the BCI. Larger numbers of flashes tend to be more monotonous since the user needs to attend to the BCI for a longer length of time. If a user's attention wanders halfway through the number of flashes, then the selection the BCI makes (classification) has a larger probability to be an error or not intended by the user. Users with a larger number of flashes need to pay attention to the BCI display for a longer period of time. Having the ability to detect moments of decreased attention in the EEG of a BCI user would allow the BCI to abstain the corresponding selection, which would reduce total errors.

Many people with severe impairments (e.g., people with amyotrophic lateral sclerosis (ALS)) that need BCI technology the most, experience attention impairments [3]. Incorporating an attention-monitoring feature into the BCI could thus be an important part of moving BCIs out of the lab and into the clinic where these users can have access to and benefit from them.

Alpha band analysis to predict attention.

There is a well-established correlation between alpha band power (EEG frequency band of 8-13 Hz) and attention, where high alpha band power corresponds to low attention levels and low alpha band power corresponds to high attention levels (negative correlation) [4, 5, 6].

Alpha band analysis in a BCI context.

Several studies have been conducted to analyze alpha band power during a visual-spatial task, in which the users were using EEG to actively choose between two selections using covert attention [7-10]. These studies each used a variation on an EEG BCI that presented left and right stimuli while the user's eyes were straight ahead (testing covert attention). Although alpha differences between left and right were used to determine which side the user was attending to, the studies did not explore the correlation between alpha and accuracy of BCI selections.

Using the relationship between alpha band power and attention, the study reported here used BCI data to analyze the relationship between alpha and the accuracy of selected characters. The hypothesis for this study is that a higher alpha band power will be correlated with a higher probability for the corresponding selection to be incorrect.

Methods

BCI Setup

The P300 Speller used in this study is a modified version of a general-purpose brain-computer interface [11], called BCI2000 [12], developed by Schalk et al.

This study used recorded data while subjects performed BCI copy-spelling tasks [11]. This recorded data was acquired using EEG from a 16-electrode cap with electrodes at F3, Fz, F4, T7, C3, Cz, C4, T8, CP3, CP4, P3, Pz, P4, PO7, Oz, and PO8 sampled at 256 Hz. Data was from 30 subjects (14 males, 16 females, ages: 18yo-78yo, average age of 46.7± 20.0 years, including 11 with ALS). Subjects typed 9 sentences with at least 23 characters per sentence with additional corrections. All together, the subjects made a total number of 9723 selections [11]. Before using the BCI, the user completed a calibration step that allows the BCI to recognize their P300 response. This calibration data was used to calculate subject-specific parameters that allowed them to use the BCI effectively, details reported in [11].

The spelling accuracy for each sentence for each subject was calculated as the number of correct selections over the number of total selections. The subjects typed these sentences using copy-spelling with correction, therefore each subject's accuracy was calculated depending on the number of correct characters they selected compared to the number of total characters they selected while typing a sentence. The selections needed to make corrections were included in this accuracy calculation.

For all subjects, alpha band power was calculated for the following segments of data (visual presentation of intervals can be found in Figure 3):

(1) Pre-character (Pre-char) interval (3.5 seconds): The EEG data corresponding to the interval before a character is typed. Note that this data is (except for the first character) between 2 separate character selection processes (after the flashes stop for one character selection and before they start for the next character selection).

(2) Character (Char) interval (ranged from 13.13 seconds to 22.5 seconds): The duration varied depending on the subject-specific number of flashes determined from the calibration data. The EEG data corresponds to the interval during the character selection process where the selections are flashing in rows and columns. This is the time where the EEG quality directly affects the BCI function since this is the only part of the EEG that is analyzed by the BCI.

(3) Post-character (Post-char) interval (3.5 seconds): The EEG data immediately after the flashes stop for the character, corresponding to the interval after each character selection. Note that the Post-char interval for one character is also the Pre-char interval for the next character.

<table>
<tr><td>Pre-Char (n) Interval (3.5s)</td><td rowspan="2">Char (n) Interval (13.13-22.5s)</td><td>Post-Char (n) Interval (3.5s)</td></tr>
<tr><td>Post-Char (n-1) Interval (3.5s)</td><td>Pre-char (n+1) Interval (3.5s)</td></tr>
</table>

Figure 3: The 3 different intervals that were analyzed in the character-by-character alpha analysis. Note that current character is (n), and (n-1) and (n+1) denote overlaps from the previous and following character, respectively.

Data Analysis Methods

Artifact rejection was applied to electrodes with segments of data excluded from analysis if any electrode exceeded 500mV. To preserve resolution, the alpha band power was calculated for each electrode separately. For the alpha band analysis included in this study, all 16 electrodes were used to make sure no useful features were left out. The alpha band power, defined as the average power in the alpha frequency band (8-13Hz) found in the EEG signal (averaged over all electrodes), was calculated for each of the 3 intervals for each character selection.

Normalized alpha band power was calculated for the pre-char, char, and post-char intervals. To calculate the alpha band power, the frequency band power during time segments was determined by using fast Fourier transform (FFT) in MATLAB. To normalize alpha, the baseline alpha for each subject was defined as the average alpha band power in the first 3 character selections of the calibration data. Note, that the alpha in each of the 3 intervals was normalized separately using the corresponding baseline alpha calculated from the 3 intervals during the first 3 character selections of the calibration data. This portion of the data was selection for normalization to avoid character selections later in the calibration data that may have lower attention levels. Normalizations were performed separately for each electrode.

For each subject, to normalize alpha, the alpha band power was subtracted from the baseline alpha then scaled by dividing by the baseline alpha (resulting in normalized alpha). Finally, the normalized alpha for all electrodes were averaged to produce the mean normalized alpha for each interval during each character selection.

Note that varying levels of alpha between and within subjects' EEG data can cause the alpha power values to vary exponentially between and within subject data. To remove the exponential gaps of alpha power between and within subjects, the $\log_{10}$ was taken of alpha values, and the alpha variance was calculated using these $\log_{10}$ alpha values, as used by Polich in [4].

Based on the variance in alpha band power exhibited in the calibration data, the subjects in this study fell into two groups: subjects exhibiting high alpha variance (average variance of log alpha = $3.78 \pm 0.76 \log_{10}uV^4$), and subjects exhibiting low alpha variance (average variance of log alpha = $2.74 \pm 0.11 \log_{10}uV^4$). Further analyses were therefore performed and reported for the high and low alpha variance subjects separately and for all subjects together to discover if differences in alpha band variance exhibited in the BCI subject's calibration data affect the utility of alpha band analysis for BCI performance improvement.

The normalized alpha band power during the Pre-char intervals, the Char intervals, and the Post-char intervals were labeled depending on whether the corresponding character selected was correct or incorrect. Statistical analysis was performed to find the relationship of alpha band power (during each of the 3 intervals) and the accuracy of the corresponding characters selected (correct or incorrect). The significance of any differences in mean normalized alpha between correct and incorrect characters was tested with a two-tailed t-test for unequal variances.

Alpha analysis to improve BCI Accuracy

The potential benefit of using normalized alpha power as a gatekeeper to retain or abstain individual selections was then evaluated. For each subject, linear discriminant analysis (LDA) was used to classify selections as correct or incorrect based on normalized alpha power exhibited at each electrode. Alpha band power in the Char interval was used in this analysis, since this is the interval of time where the user is actively making BCI selections, and also showed a statistically significant difference between alpha during correctly and incorrectly selected characters. To test the accuracy of alpha to predict selections made by each subject, LDA was trained on the alpha exhibited at each electrode over all of the calibration data, then tested on the sentences that followed calibration.

BCI accuracy was recalculated, for each subject, using alpha to abstain character selections predicted to be errors. To determine whether a relationship exists between the change in BCI accuracy (using LDA) and the variance of the normalized alpha, the variance of the normalized alpha (over the calibration data) exhibited by each subject was compared to the results of the LDA.

Another metric that was used to evaluate performance is BCI Utility [13]. BCI Utility is defined as the ratio of the expected benefit per selection and the expected time per selection. Unlike other metrics, the BCI Utility metric not only quantifies the accuracy of selections and the rate of selections, but also takes error correction into consideration. The equation used to calculate BCI-Utility was as follows:

$$U = \frac{E(benefit/selection)}{E(time/selection)}$$

Results

To determine the relationship between normalized alpha and character accuracy, normalized alpha was evaluated for correct (10084 characters) and incorrect selections (2066 characters) from all subjects (Figure 4A). Using a two-tailed t-test for unequal variances, the statistical significance between the normalized alpha for correct and incorrect selections for each of the 3 intervals were $p=0.2327$, $p=0.01004$, and $p=0.0017$ for Pre-char, Char, and Post-char intervals, respectively. Figures 4B and 4C show the same breakdown while separating the high and low variance groups, respectively. The variances between the two groups were statistically significant with $p=2.38e\text{-}5$. For the high variance group (Figure 4B), the statistical significance for each of the 3 intervals were $p=0.034$, $p=0.0012$, and $p=0.0004$ for Pre-char, Char, and Post-char intervals, respectively. While for the low variance group (Figure 4C), only one interval was statistically significant (Post-char), with p-values of $p=0.332$, $p=0.164$, and $p=0.048$ for Pre-char, Char, and Post-char intervals, respectively.

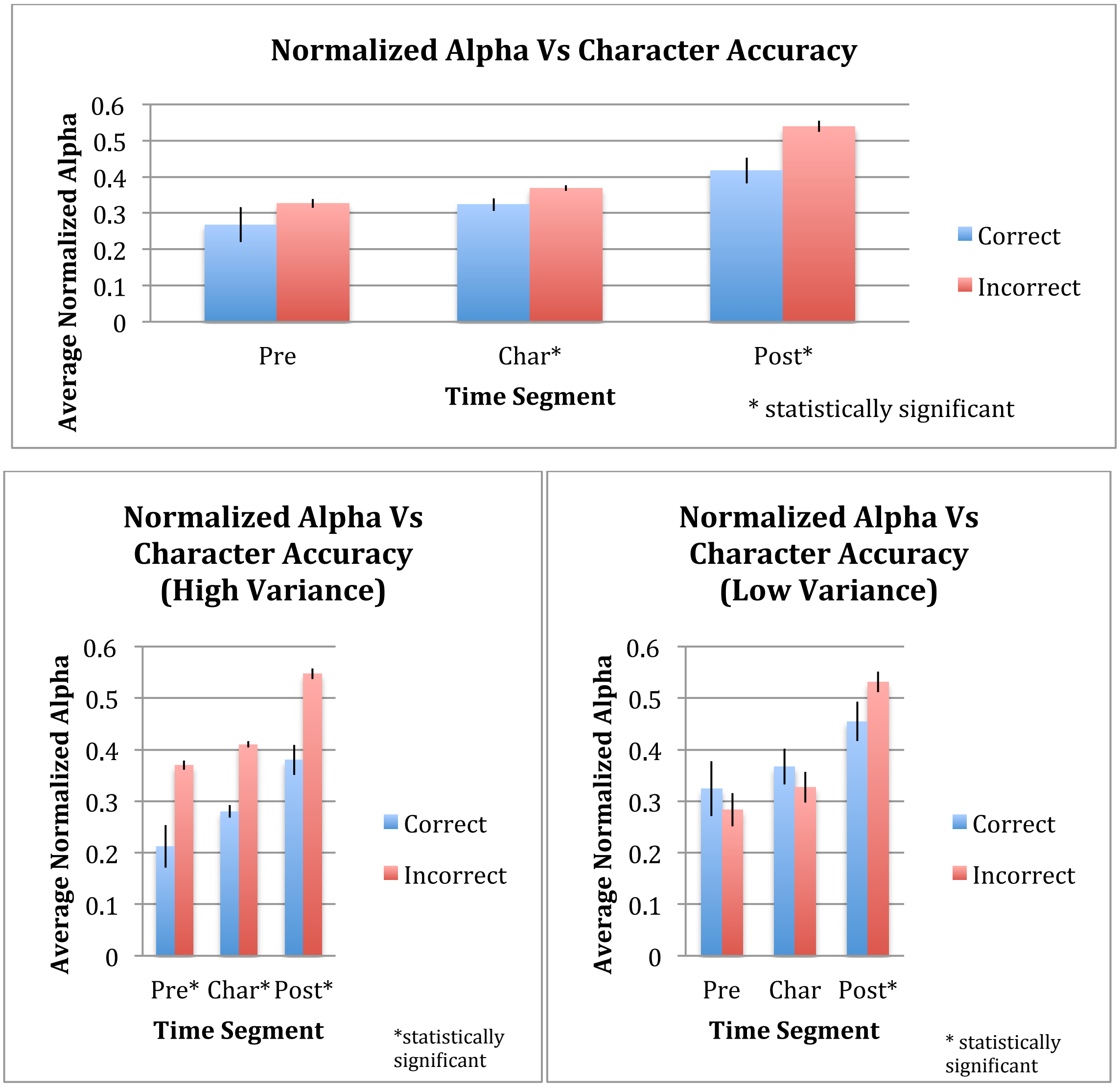

Figure 4: The mean of normalized alpha band power for each of the 3 intervals for correct and incorrect selected characters. The intervals that showed statistical significance are marked with an asterisk (*). **A (top):** All subjects**, B (bottom left):** Subjects exhibiting high alpha variance. **C (bottom right):** Subjects exhibiting low alpha variance.

Since the character selection interval (Char, 13.13-22.5s) is much longer than the Pre-Char (3.5s) and Post-Char (3.5s) intervals, the significance of alpha during the Char interval may be diluted due to its length. When the Char interval is segmented into 3.5s intervals, the significance of alpha between correct and incorrect selections increases chronologically. Thus, the last 3.5s of the Char interval shows a higher significance of alpha between correct and incorrect selections ($p = 0.0021$) than the first 3.5s of the Char interval ($p = 0.087$).

Figure 5 shows the change in accuracy, for all subjects, achieved by abstaining characters during which subjects exhibited high alpha (low attention levels) during the character selection period. The subjects that showed improvement (increased accuracy) were those that exhibited much higher alpha variances (those represented in Figure 4B). Those exhibiting low alpha variances did not show improvement (those represented in Figure 4C) and thus using LDA for alpha classification was not helpful.

The mean BCI accuracy, calculated over all subjects, for raw BCI performance and BCI performance with alpha classification were 82.01±10.16% and 83.93±9.79%, respectively; however they were not statistically significant. The mean BCI-Utility, calculated over all subjects, for raw BCI performance and BCI performance with alpha classification was 2.65±1.61 and 2.78±1.60, respectively, however they were not significant. The mean BCI accuracy for subjects that exhibited high levels of alpha variance (15 out of 30 subjects; average variance of log alpha = 3.85±0.70 $\log_{10}uV^4$) for raw BCI performance and BCI performance with alpha classification were 80.24±10.70% and 87.5±8.57%, respectively, which was statistically significant with a p-value = 0.041. The mean BCI accuracy for subjects that exhibited low levels of alpha variance (15 out of 30 subjects; average variance of log alpha = 2.74±0.11 $\log_{10}uV^4$) for raw BCI performance and BCI performance with alpha classification were 84.02±14.5% and 79.35±13.92%, respectively, which was not statistically significant. Variance of log alpha (in the calibration data) was found to be practical in predicting whether or not alpha classification would help increase BCI accuracy.

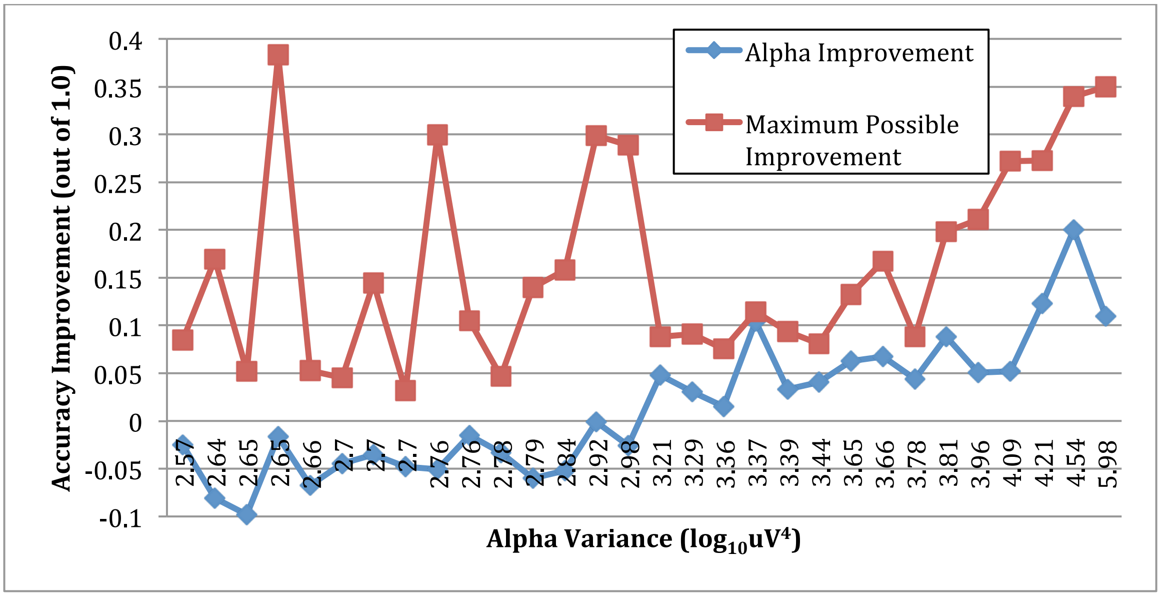

Figure 5: Graph showing relationship between alpha improvement in BCI accuracy and alpha variance. The accuracy improvement is presented in changes in accuracy on a 0 to 1 scale, where 1 is 100% accuracy. Alpha variance of subjects is sorted in order of increasing variance. The red line shows the maximum possible improvement to reach 100% accuracy for each subject as reference.

The high alpha variance group showed improved accuracy using alpha classification to abstain characters for the 15 of 30 subjects who exhibited high alpha levels (low attention levels). The high variance group included 6 ALS subjects, while the low variance group included 5 ALS subjects. The alpha variance exhibited by the ALS subjects did not show any statistically significant difference from the variances exhibited by controls, whether over all subjects or within the high and low variance groups.

The alpha variance exhibited by the low variance group caused the alpha classification to allow erroneous selections as well as reject correct selections in a somewhat random fashion. Their corresponding accuracies were decreased due to this phenomenon.

Discussion

A natural expectation would be that alpha band power during the Char interval would be most significant, since this interval is comprised of the EEG that is actually analyzed by the BCI. However, the predictive value of the alpha band, for all subjects, increases from insignificant in the pre-char interval ($p = 0.2327$) to maximum significance in the Post-Char interval ($p = 0.0017$), with the Char interval itself having a slightly lower significance ($p = 0.01004$). In Figure 4A, although the statistical significance of alpha between correct and incorrect characters increases from Pre-Char to Char to Post-Char, the magnitude of alpha increases for all selections (whether correct or incorrect). This suggests that starting a character (first segment of Char interval) may focus attention at first (lower alpha power), which then drifts away (higher alpha power) as the character selection progresses.

This implies that, practically, although the Post-Char interval showed slightly more significance than the Char interval, the latter portion of the Char interval can be used to predict the accuracy of the corresponding character. Using the Post-Char interval to predict selection accuracy may be useful for subjects with low alpha variance, however this will have a delay cost. While, for those that have high alpha variance, the prediction can be done simultaneously during the character selection interval (Char). In addition, for high variance subjects, since alpha in the Pre-Char interval is significant between correct and incorrect selections, the alpha during Pre-Char interval can be used to 'pause' the BCI when low levels of alpha are detected to prevent potential errors.

Study Limitations

The data analyzed in this study was for subjects who were supposed to be maintaining attention while copying text (copy-spelling) [11]. This is an unrealistic usage condition since they had no text composition tasks and few distractions. The only distraction inherent in the task was the need to problem-solve how to correct an error if it occurred, which is a limitation of the data analyzed in this study. In addition, users with low alpha variance were, in fact, paying attention. This motivates designing a study that thoroughly investigates the effect of distractors on

BCI usage. A study with a BCI designed to mimic real-world usage of a spelling BCI, including compositions tasks, strategic distractions, and periods where the user is not paying attention to the BCI display, would overcome this limitation and test alpha band power's attention-monitoring ability to its fullest.

Technical Application

The potential for an alpha-based classification to increase accuracy can be predicted by the alpha variance exhibited in the BCI user's calibration data. Users with $\log_{10}$ alpha variance of 3 $\log_{10}uV^4$ and above benefit from an alpha-based classification, and users with $\log_{10}$ alpha variance less than 3 $\log_{10}uV^4$ do not benefit from an alpha-based classification.

Conclusion

In conclusion, this study shows that alpha band power can be used to improve BCI performance by abstaining selections where losses of user attention are detected. However, this is only useful under conditions where the user is exhibiting high alpha variance. The average normalized alpha for correct character selections proved to be statistically significantly different than that for incorrect character selections. Using this significance, alpha classification using LDA was used to block any selections exhibiting low levels of attention (high alpha), which produced statistically significant increases in accuracy for subjects with high variance in normalized alpha. Using the user's attention level (measured by alpha band power), specific BCI settings (such as using alpha classification) may be added to improve BCI performance. The ability to detect attention in a BCI context results in a BCI that is more resilient to wandering user attention.

Acknowledgements

This work was supported by Grant #R21HD54697 from the National Institute of Child Health and Human Development (NICHD) US National Institutes of Health (NIH) and by Grant #H133G090005 from the National Institute on Disability and Rehabilitation Research (NIDRR). The content is solely the responsibility of the authors and does not necessarily represent the official views of NIDRR, the Department of Education, and NICHD of the NIH.

References

[1] Donchin, E, Heffley, E, Hillyard, SA, Loveless, N, Maltzman, I, Ohman, A, Rösler, F, Ruchkin, D, Siddle, D, "Cognition and event-related potentials. II. The orienting reflex and P300" *Ann. NY Acad. Sci*., 425 (1984): 39–57.

[2] Farwell, LA, and Donchin, E, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials", *Electroenceph. Clin. Neurophysiol.*, vol. 70, no. 6, (1988): 510 - 523.

[3] Phukan, J, Pender, NP, and Hardiman, O, "Cognitive Impairment in Amyotrophic Lateral Sclerosis." *The Lancet Neurology* 6.11 (2007): 994-1003.

[4] Polich, J, "Updating P300: An Integrative Theory of P3a and P3b." *Clinical Neurophysiology* 118.10 (2007): 2128-148.

[5] Hanslmayr, S, "Visual Discrimination Performance Is Related to Decreased Alpha Amplitude but Increased Phase Locking." *Neuroscience Letters* 375.1 (2005): 64-68.

[6] Gleiss, S, and Kayser, C, "Acoustic Noise Improves Visual Perception and Modulates Occipital Oscillatory States." *Journal of Cognitive Neuroscience* 26.4 (2014): 699-711.

[7] Kelly, SP, "Increases in Alpha Oscillatory Power Reflect an Active Retinotopic Mechanism for Distracter Suppression During Sustained Visuospatial Attention." *Journal of Neurophysiology* 95.6 (2006): 3844-851.

[8] Kelly, SP, Lalor, EC, Reilly, RB, and Foxe, JJ "Visual Spatial Attention Tracking Using High-Density SSVEP Data for Independent Brain–Computer Communication." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 13.2 (2005): 172-78.

[9] Kelly, SP, Lalor, EC, Reilly, RB, and Foxe, JJ, "Independent Brain Computer Interface Control Using Visual Spatial Attention-Dependent Modulations of Parieto-occipital Alpha." *Conference Proceedings. 2nd International IEEE EMBS Conference on Neural Engineering*, (2005): 667-670.

[10] Treder, MS, Bahramisharif, A, Schmidt, NM, Van Gerven, MA, and Blankertz, B, "Brain-computer Interfacing Using Modulations of Alpha Activity Induced by Covert Shifts of Attention." *Journal of NeuroEngineering and Rehabilitation* 8.1 (2011): 24.

[11] Thompson, DE, Gruis, KL, Huggins, JE, "A Plug-and-Play Brain-Computer Interface to Operate Commercial Assistive Technology" *Disability and Rehabilitation: Assistive Technology.* (2013).

[12] Schalk, G, McFarland, DJ, Hinterberger, T, Birbaumer, N, and Wolpaw, JR, "BCI2000: A General- purpose brain-computer interface (BCI) system," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, (2004):1034–1043.

[13] Thompson, DE, Blaine-Moraes, S, Huggins, JE' "Performance assessment in brain-computer interface-based augmentative and alternative communication", Journal of Neural Engineering. 11(3), (2014).

Chapter 3
The P300-Certainty Algorithm: Increasing BCI Accuracy by Abstention

Introduction

Individuals with severe motor and speech impairments have a lot of difficulty communicating and can benefit from the assistance of augmentative and alternative communication (AAC) systems. However, individuals with little-to-no motor control or locked-in syndrome are unable to use common commercialized AAC systems. Non-invasive Brain-Computer Interfaces (BCIs) have been an emerging means of communication for people with severe motor impairments. The P300 response is a signal that has been extensively used to control BCIs and appears in the user's electroencephalogram (EEG) about 300ms after a relevant, yet unpredictable event [1]. First introduced by Farwell and Donchin, the first P300 Speller was a BCI based on the P300 response [2]. By presenting the user with a matrix of selections, the P300 Speller allows users to choose selections using his/her brain signals (P300 responses).

Background

The P300 Speller detailed

The P300 Speller is a common P300 BCI design that presents users with selections from a matrix by intensifying their color (flashing them from grey to white) and chooses selections to which users' EEG exhibits P300 responses. Selections are presented to the user using a grid of selections. An example of this can be found in Figure 2 in Chapter 1. Before using the BCI, the user completes a calibration step; only one calibration step is needed for the P300 speller, such that the P300 speller learns how to recognize the user's P300 response. In the conventional version of the P300 speller, the selections are divided into 2 groups: the set of rows and the set of columns. The rows and columns flash at random, known as the row-column flash pattern. A sequence is defined as including one flash of every row and column. The P300 Speller checks for a response to each flash by applying a linear classifier to the EEG following the flash, producing a classifier value. In this study, we applied Step-Wise Linear Discriminant Analysis (SWLDA) to 0-800ms after each flash. A larger classifier value indicates the presence of a P300 while a smaller classifier value indicates the absence of a P300. For the P300 Speller to make an accurate decision, multiple sequences may be needed. The number of sequences a user needs depends on the signal-to-noise ratio of their EEG signal. After all the sequences are completed, the selection corresponding to the highest average P300 response is chosen. A user with a higher signal-to-noise ratio EEG signal would have a higher chance of being able to type in fewer flashes than one with a lower signal-to-noise ratio EEG signal. This is due to the fact that it is easier for the P300 Speller to detect and classify the P300 response when there is a higher signal-to-noise ratio in the EEG signal.

Overall view of the P300 Speller

An overall view of the P300 Speller can be found in Figure 6. First, the user focuses on the selection of interest on the BCI display. All of the selections flash for a user-specific number of sequences based on their calibration data. Once all of the sequences have been completed, the P300 Speller chooses the row and column that have the highest classifier values (most probable to have elicited a P300 response in the user). The intersection of the selected row and column is the selection that is chosen by the P300 Speller. The selection is then printed on the screen, then the next selection begins, the BCI display will start flashing for the user-specific number of sequences, after which the P300 Speller will make another decision, and so on.

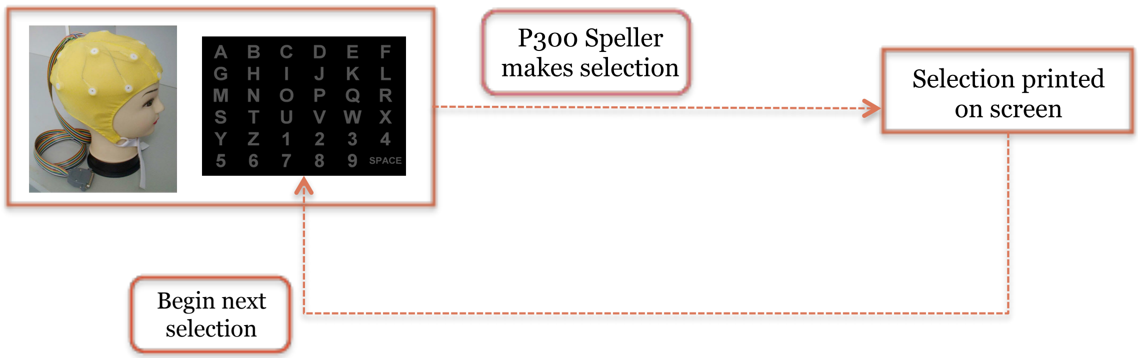

Figure 6: A block diagram summarizing the mechanism of action of the P300 Speller.

<u>Selection Confidence Methods</u>

Selection confidence methods are used with P300 BCI to reject potentially erroneous selections that do not reach a specified confidence level. Rejecting selections that are potential errors is known as abstention.

Several studies have used BCIs with abstention by calculating confidence values for selections in different ways [3, 4]. Methods that have been used to calculated selection confidence values are: Independent Component Analysis (ICA) [5] and Bayesian Linear Discriminant Analysis (BLDA) [3, 6]. These methods all calculated selection confidence values for selections and evaluated them against a predetermined confidence threshold to determine whether a selection was retained or abstained. The study using ICA showed an increase of accuracy from 90% to 92.1%, while the study using BLDA showed a non-significant decrease in accuracy from 79.44±29.98% to 74.40±27.16%.

This study introduces a new selection confidence method (the P300-Certainty algorithm) that generates selection confidence probabilities based on a secondary classifier applied to the classifier values from SWLDA. The secondary classifier uses a statistical test (U-test) since it has no parameters to adjust and is robust when it comes to outliers. This may provide a slight advantage over other existing selection confidence methods. Another advantage of P300-Certainty is that the sum of confidence probabilities of all selections is one, where each selection has an associated confidence probability called the certainty value.

Methods

BCI Setup

The P300 Speller used in this study is part of a general-purpose BCI, called BCI2000 [7], developed by Schalk et al.

The data used to test the P300-Certainty algorithm was recorded from subjects performing BCI copy-spelling tasks [8]. This data was acquired using EEG with 16 electrodes at F3, Fz, F4, T7, C3, Cz, C4, T8, CP3, CP4, P3, Pz, P4, PO7, Oz, and PO8 sampled at 256 Hz. The data was recorded from 30 subjects (14 males, 16 females, ages ranging from 18-78 years, average age of 46.7±20.0 years, including 11 with ALS). Each subject typed 9 sentences with at least 23 characters per sentence in addition to corrections [8]. The total number of selections made by all subjects was 12150 selections. Before using the BCI, the user completed a calibration step that allows the BCI to recognize their P300 response. This calibration data was used to calculate subject-specific parameters that allowed them to use the BCI effectively, details reported in [8].

The spelling accuracy was calculated for each sentence for each subject as the number of correct selections divided by the number of total selections. The sentences were typed using copy-spelling with correction, therefore each subject's accuracy was calculated depending on the number of correct selections they chose over the number of total selections they made while typing a sentence. This accuracy calculation also took into consideration the selections needed to make corrections.

The P300-Certainty algorithm

The P300-Certainty algorithm [9] was developed as a selection confidence method to increase BCI accuracy by reducing errors while using the P300 Speller. This is accomplished by calculating a probability of confidence for each selection, which we called the certainty value, and based on the users' corresponding EEG responses to the flashing of these selections. The certainty value of the selection chosen by the P300 Speller is evaluated against a user-specific threshold. The selection was typed if its certainty value surpassed the threshold, and abstained if it did not. Abstaining potentially erroneous selections would allow users to spell with a higher accuracy. Figure 7 updates the block diagram of Figure 6 to show the P300-Certainty used in the P300 Speller as a gatekeeper to make the decision to type a selection or abstain it.

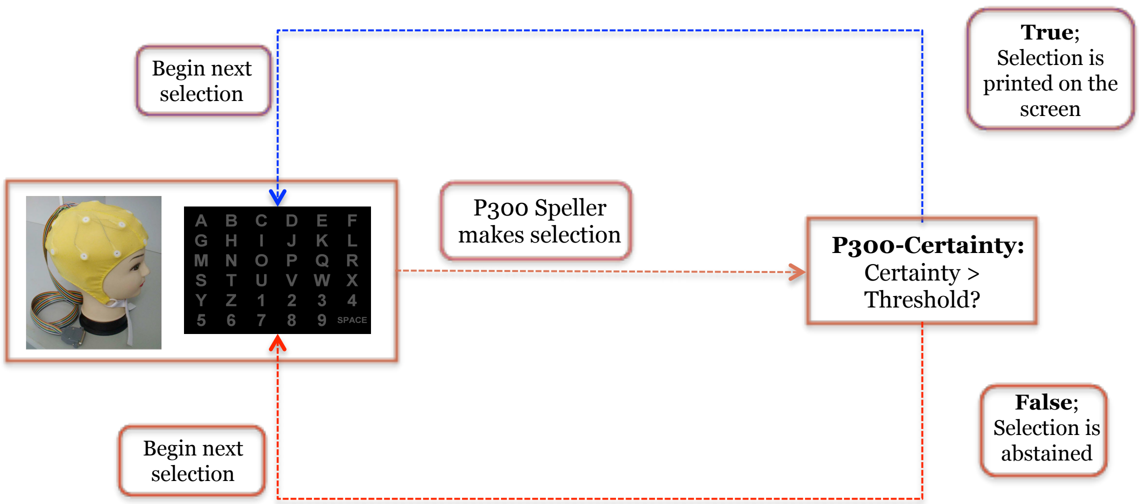

Figure 7: A block diagram showing how the P300-Certainty fits into the P300 Speller and demonstrating how it is used to decide whether to type a selection or abstain it.

Calculations behind the P300-Certainty Algorithm

Before getting into the details behind the P300-Certainty algorithm, there are a couple terms that need to be defined.

Target selection: This is the intended selection

Associated selections: These are the selections that flash together with the target selection. In a row-column flash pattern, these are selections along the same row or column as the target selection.

Unassociated selections: These are the selections that do not flash with the target selection. In a row-column flash pattern, these are selections that are not on the same row or column as the target selection.

At a conceptual level, the P300-Certainty algorithm has the following steps:

```
[For each response:
    - Record the response and note which stimulus it was elicited by.
    - If one response from each stimulus has not been recorded, wait
      until the next stimulus.
```

```
- For each possible selection, s:
   - Let f be the distributions of responses associated with s
   (flashed with s).
   - Let g be the distributions of responses not associated with s
   (did not flash with s).
         - Use hypothesis testing to test h_o: f=g, h_a: f>g,
   - Record the resulting p-value.
- For each possible selection, s:
   - s, belongs to a group of selections, S such that:
    - s is in S, all other selections in S are selections that
      flashed at the same time as s,
    - the number of targets in S is not more than S_max and not less
      than S_min,
- Let P_out be the product of all p-values of the selections not in S.
- Let P_in be the product of one minus the p-values of the selections
   in S.
- The certainty for s is log(P_out*P_in).
- If all stimuli have been intensified N times, forget all
   responses.]
```

For each possible selection, the distributions of the classifier values associated with the selection were compared statistically to the distributions of the classifier values unassociated with the selection. The statistical test used in this study to compare both groups was the U-test. The U-test was utilized to create one rank order within both sets of associated and unassociated selections. An attribute of the U-test that is attractive in this application is the fact that it has no parameters to adjust and is robust when it comes to outliers. The test measures how often a classifier value of each of the unassociated selections is smaller than that of the associated selections. The comparison made by the U-test results in a p-value for each selection in each set. For each selection, the products of one minus the p-value of the associated selections (let us call it P_{in}) was calculated, and the product of the p-values of all the unassociated values (let us call that P_{out}) was calculated. Finally, for each selection, the certainty value was calculated as the logarithm of (P_{in} X P_{out}). This results in the raw certainty value for each selection.

Furthermore, to make the raw certainty values easier to interpret as probabilities, they were normalized to be between zero and one, with the sum of all of the certainty values being one. This normalization was achieved using the softmax function [10]. The certainty value of a selection determined if the selection was to be typed or abstained depending on whether it was greater or less than the user-specific threshold, respectively.

The user-specific certainty threshold was calculated, based on the user's calibration data, for each subject using a Receiver Operating Characteristic (ROC) curve, where the true positive rate (sensitivity) and the false positive rate (100-specificity) are evaluated at many different thresholds. For each subject, the threshold that yielded the highest true positive rate and the lowest false positive was chosen as their user-specific threshold. There is a trade-off between the true positive rate and the false positive rate, and the threshold chosen was the one that optimized this trade-off for each subject.

To evaluate the performance of the P300-Certainty algorithm, the BCI spelling accuracy was calculated with and without P300-Certainty. Statistical significance was calculated using a paired t-test. Also, the numbers of total characters, correct characters, incorrect characters, correctly abstained characters, and incorrectly abstained characters were reported.

Another metric that was used to evaluate performance is BCI Utility [11]. BCI Utility is defined as the ratio of the expected benefit per selection and the expected time per selection. Unlike other metrics, the BCI Utility metric not only quantifies the accuracy of selections and the rate of selections, but also takes error correction into consideration. The equation used to calculate BCI-Utility was as follows:

$$U = \frac{E(benefit/selection)}{E(time/selection)}$$

Results

The improvement in BCI spelling accuracy using P300-Certainty is shown in Figure 8. The mean BCI accuracy calculated with and without P300-Certainty was 88.82±8.85% and 82.01±12.59%, respectively, and was statistically significant with a p-value = 0.038. The BCI-Utility calculated with and without P300-Certainty was 3.05±1.63 and 2.65±1.61, respectively, with a statistical significance of p = 0.024.

The mean of all user-specific certainty thresholds calculated using a ROC curve was 88.59±4.03%.

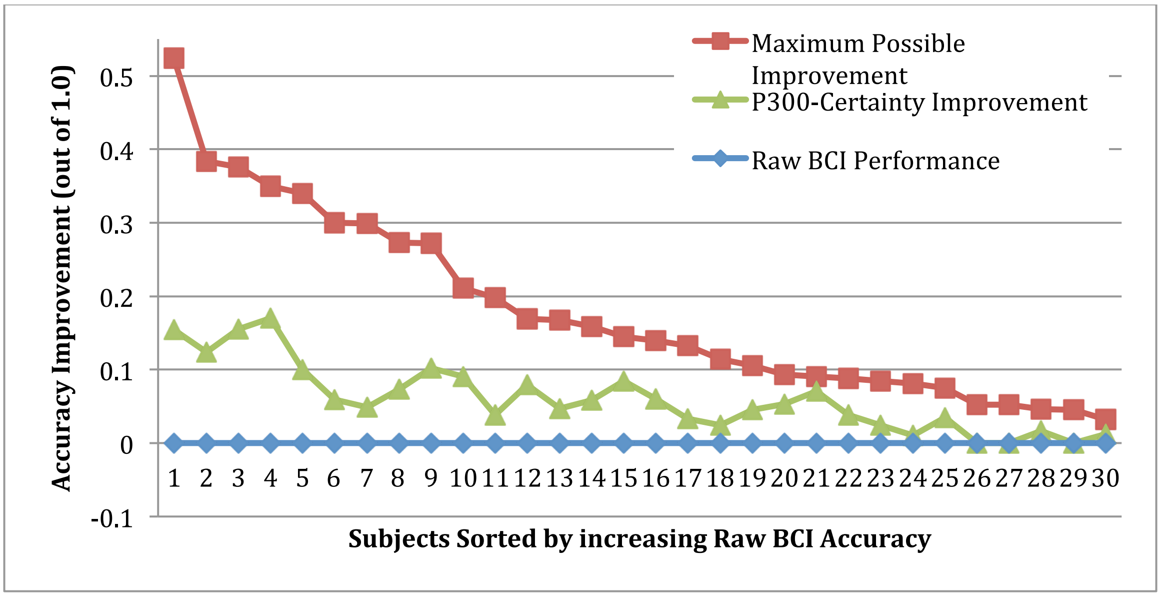

Figure 8: Graph showing P300-Certainty improvements in BCI accuracy. The accuracy improvement is presented in changes in accuracy on a 0 to 1 scale, where 1 is 100% accuracy. Subjects are sorted by increasing raw BCI accuracy. The red line shows the maximum possible improvement to reach 100% accuracy for each subject as reference. The blue line represents raw BCI accuracy as a reference.

Table 1: The breakdown of number of selections with and without the P300-Certainty algorithm.

	Total Selections Allowed	**Correct Selections**	**Incorrect Selections**	**Abstentions (Abstained Correctly)**	**Abstentions (Abstained incorrectly)**
Without P300-Certainty	12150	10084	2066	NA	NA
With P300-Certainty	11403	9937	1466	600	147

Table 1 shows the breakdown of the numbers of total characters, correct characters, incorrect characters, correctly abstained characters, and incorrectly abstained characters used in this study.

Discussion

The P300-Certainty algorithm improved accuracy for 27 out of the 30 subjects, without decreasing the accuracy for any subjects. This showed that P300-Certainty is successful as a selection confidence method that abstains selection based on a user-specific certainty threshold.

The data analyzed in this study was collected using a row-column flash pattern, however, P300-Certainty can be used in various other set-ups. Many other flash patterns have been developed that proved superior to row-column flashing, for example checkerboard flash pattern [12]. P300-Certainty can be used as a selection confidence method with these flash patterns by adjusting the associated and unassociated groups of selections. Where, for each selection, the associated group would be the selections that flash together with that selection, while the unassociated group would be the selections that do not flash together with it.

The next chapter discusses how the P300-Certainty algorithm was successfully adapted to a BCI set-up that uses single selection flashes in a 4-target set-up. Even though the core mechanisms of P300-Certainty were the same, a different statistical test (other than the U-test), was used to generate the certainty values due to the minimal number of stimuli in that set-up.

From the results presented in this study, P300-Certainty can be used to achieve dynamic stopping, by only making selections when a selection surpasses the certainty threshold. Since the data used in this study was recorded data, it is difficult to evaluate the impact of dynamic stopping on spelling accuracy. However, dynamic stopping can be achieved by incorporating P300-Certainty into the code of a P300 BCI and testing it on subjects in real-time. In the next chapter, P300-Certainty was used in real-time to achieve dynamic stopping in a BCI adapted for cognitive testing.

References

[1] Polich, A, “Brain-Computer Interface: New Directions” *Scientific Report 2006: Molecular and Integrative Neurosciences*. (2006)

[2] Farwell, LA, and Donchin, E, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials", *Electroenceph. Clin. Neurophysiol*., vol. 70, no. 6, (1988): 510 - 523.

[3] Verschore, H, Kindermans P, Verstraeten, D, and Schrauwen, B "Dynamic Stopping Improves the Speed and Accuracy of a P300 Speller." *Artificial Neural Networks and Machine Learning – ICANN 2012* 7552 (2012): 661-68.

[4] Zhang, H, Guan, C, and Wang, C "Asynchronous P300-Based Brain--Computer Interfaces: A Computational Approach With Statistical Models." *IEEE Transactions on Biomedical Engineering* 55.6 (2008): 1754-763.

[5] Serby, H, Yom-Tov, E, & Inbar, G, “An Improved P300-Based Brain-Computer Interface,” IEEE Transactions on Neural Systems and Rehabilitation Engineering, 13 (2005): 89-98.

[6] Mainsah, BO, Collins, LM, Colwell, KA, Sellers, EW, Ryan, DB, Caves, K, & Throckmorton, CS, “Increasing BCI communication rates with dynamic stopping towards more practical use: An ALS study,” *Journal of Neural Engineering*, 12 (2015), 016013.

[7] Schalk, G, McFarland, DJ, Hinterberger, T, Birbaumer, N, and Wolpaw, JR, “BCI2000: A General- purpose brain-computer interface (BCI) system,” *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, (2004):1034–1043.

[8] Thompson, DE, Gruis, KL, Huggins, JE, “A Plug-and-Play Brain-Computer Interface to Operate Commercial Assistive Technology” *Disability and Rehabilitation: Assistive Technology*. (2013).

[9] Aref, AW, Huggins, JE, “The P300-Certainty Algorithm: Improving accuracy by withholding erroneous selections”; *EEG & Clinical Neuroscience Society, EEG & Clin Soc*, (2012).

[10] Dunne, RA, Campbell, NA, “On the pairing of the softmax activation and cross-entropy penalty function and the derivation of the softmax activation function”, *Aust. Conf. on Neural Networks*, (1997): 181-185.

[11] Thompson, DE, Blaine-Moraes, S, Huggins, JE’ “Performance assessment in brain-computer interface-based augmentative and alternative communication”, Journal of Neural Engineering. 11(3), (2014).

[12] Townsend, G, et al., “A novel P300-based brain-computer interface stimulus presentation paradigm: Moving beyond rows and columns,” *Clinical Neurophysiology*, vol. 121, (2010): 1109–1120.

Chapter 4
On-line Performance of the P300-Certainty Algorithm

Introduction

Typically, clinical cognitive assessments require some kind of voluntary motor or speech responses. Therefore, many people with severe motor and speech impairments are unable to complete such assessments [1]. To mitigate this problem, different assistive technologies (e.g. touch screens, switches, and eye-trackers) have been used to allow those with disabilities access to these standardized cognitive assessments [2]. However, many of these assistive technologies still require small amounts of movement or speech for control, thus people with severe levels of motor or speech impairments still cannot access cognitive tests [2].

Promising efforts have been made by researchers to improve access to cognitive assessments for people with disabilities by using brain activity to make responses. Studies have used BCIs to enable access to cognitive tests for individuals with impairments [3,4]. Building upon this proof-of-concept, the study presented in this experiment uses a P300 BCI that allows users to take a cognitive assessment without the use of motor or speech responses and at their own pace. This BCI-facilitated cognitive assessment was tested with both controls and children with cerebral palsy (CP). To evaluate the performance of this BCI-facilitated assessment, the results of the cognitive test were compared to the results of the same users taking a traditionally administered assessment.

The P300-Certainty algorithm, first introduced in Chapter 3, was integrated into the BCI-facilitated cognitive test to allow users to make accurate selections at their own pace.

Methods

BCI Setup

To evaluate real-world effectiveness of the P300-Certainty algorithm, it was implemented into the code of BCI2000. The P300 BCI used in this study is built upon a general-purpose BCI, called BCI2000 [5], and is customized to administer a cognitive assessment, described in detail in [6].

This study used a BCI-facilitated version of a cognitive assessment known as the Peabody Picture Vocabulary Test – version 4 (PPVT-IV). The PPVT-IV measures verbal ability in standard American English vocabulary [7]. The standard method of administering the PPVT-IV involves the administrator of the PPVT-IV presenting the test-taker with 4 different pictures and speaking a word. The test-taker must then respond by either pointing at or verbally responding with the number of the picture that is best described by the spoken word.

Participants, ages 8 and up with the ability (can point or answer verbally) to complete the standard version of the PPVT-IV, were recruited from the University of Michigan Health System. The University of Michigan Institutional Review Board approved protocols and recruitment. Out of 30 participants that were recruited, only 21 participants completed the study. This experiment analyzes data from the 21 participants (10 with CP and 11 typically developing) that completed the study. The mean overall age of subjects was 16.75±5.54 years with ages ranging from 9-27.

This data was analyzed using 16 channels on an EEG cap with electrode locations at F3, Fz, F4, T7, C3, Cz, C4, T8, CP3, CP4, P3, Pz, P4, PO7, Oz, and PO8 sampled at 600 Hz.

In this setup, the BCI presented the user with four different pictures and a word was played through the speakers. Each of the four pictures had a numerical label (1-4) in the outside corner. These labels intensified (flashed) independent of each other. The user was instructed to focus their attention on the numerical label corresponding to the picture that they thought best described the word they were presented. To calibrate the BCI, the user was instructed to attend to a known answer for 30 different questions, with each selection flashing 10 times per question.

The P300-Certainty algorithm was used to achieve dynamic stopping (described in detail in the Introduction of this dissertation). No selections were made by the BCI until the certainty value (described in the Methods section of Chapter 3 of one of the four selections was at least 0.9 (or 90% certainty). In addition, a sliding window of five sequences was used, where only the EEG responses to the last five sequences were used to calculate certainty.

This allowed the BCI to wait for the user to attend to one of the targets before making any selections. After a selection was made (the certainty value of one of the four selections reached at least 0.9), the BCI entered a confirmation phase. This allowed the user to cancel a selection if it was not the one they intended or verify it if it was. This confirmation step was accomplished using the Hold/Release algorithm developed by Alcaide, described in [8].

The utilization of both dynamic stopping (P300-Certainty) and the confirmation step, allowed the BCI to operate in a very dynamic nature and users to control the BCI at their own pace.

The P300-Certainty algorithm

The P300-Certainty algorithm was adapted to this application by making 3 adjustments to accommodate the differing conditions. Due to the limited number of selections (4 selections instead of 36 presented in the P300 Speller in Experiment #1) instead of using the U-test to calculate the significance between selections, a t-test with unequal variance was used. The U-test needs at least 5 different selections to be able to determine statistical significance. The second difference was that the target that was statistically different was required to have a positive classifier value. This was due to the fact that by using a t-test, targets that had a very large negative classifier value came out to be significant.

The third difference is that dynamic stopping was achieved using P300-Certianty by allowing the BCI to only make selections when the certainty of a selection reached 0.9. In addition, instead of taking all sequences into consideration, only the last 5 sequences (3.75 seconds) were used to calculate P300-Certainty. This allows the BCI to adjust quickly after periods in which the user was not attending.

Data Analysis

To evaluate the performance of P300-Certainty on-line, the accuracy of the selections made by the user was calculated by comparing every time P300-Certainty made a selection (a selection reached a certainty of at least 0.9) to the final selection they chose for any question. For example, for a particular question, if a user wanted to choose selection 2, but the BCI chose selection 3 using P300-Certainty, the BCI would then proceed to the confirmation step where the selection can either be cancelled or verified. If the selection was cancelled, then the question was repeated and the user had the chance to make a selection again. If the user chooses selection number 2 the second time around and then verified it, then the first time when selection 3 was chosen would be considered an error since the last answer chosen by the user was selection 2. The accuracy of P300-Certainty was calculated in that manner over all subjects.

There are two exceptions to this. The first was if the BCI verified a selection that the user did not intend. Users were given the opportunity to alert the person conducting the study if any unintended verifications were made. These errors were taken into consideration when calculating the accuracy.

Second, before beginning the assessment, the users were advised, if they did not hear the word played by the speakers, to intentionally select any of the 4 selections and then cancel that selection just to cause the system to repeat the word. These instances were not logged in the study. So, some of the "errors" were actually intentional changes and the accuracy is underestimated.

Another analysis designed to provide additional insight into the future applications of P300-Certainty was that of the size of the sliding window used to achieve dynamic stopping. The window size used online by all study participants was five sequences long. Using the data collected from this study, off-line analysis was performed to determine which window size would yield the best performance for all subjects.

For each subject, the accuracy of selections was calculated using sliding window sizes of 1, 2, 3, 4, and 5 (Note that the sliding window size was fixed at 5 in the on-line study). This would offer information on ideal window size, including whether there is one that is uniform amongst all subjects. It is to be noted that analysis using sliding window sizes larger than 5 sequences was not possible due to the fact that this analysis was performed off-line and additional sequences could not be added.

Results

The mean of the accuracy of P300-Certainty on-line in a 4-target task used in this study over all subjects was 83.61±9.14%.

The impact of using varying sliding window sizes was analyzed off-line to determine what the ideal window size to use for P300-Certainty, results shown in Figure 9.

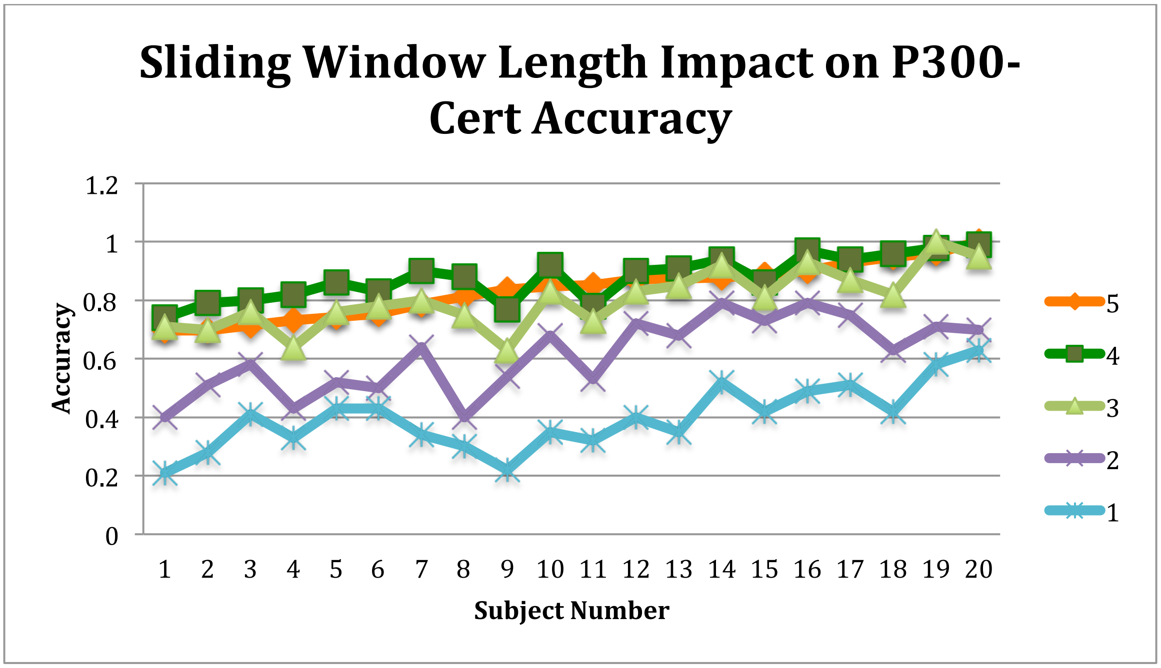

Figure 9: Off-line analysis of the impact of varying window length size on P300-Certainty accuracy.

In the on-line application, the minimum number of sequences in which a decision could be made was 5 sequences. In off-line analysis, the mean P300-Certainty accuracies, keeping the certainty threshold at 0.9, calculated for 5, 4, 3, 2, and 1 sequence(s) were 83.62 ± 9.14%, 87.7 ± 7.61%, 80.35 ± 9.92%, 61.15 ± 12.63% and 39.7 ± 11.12%, respectively. The statistical significance of using a window of 4 and 3 sequence(s) compared to a window of 5 sequences were all non-significant ($p > 0.05$). While using a window of 2 or 1 sequence(s) were statistically significantly different from a window of size 5 ($p < 0.05$), they decreased performance.

Discussion

This study shows that the P300-Certainty algorithm can be used effectively to achieve dynamic stopping, thus allowing users to make accurate selections at their own pace.

Using a window of 4 or 3 sequences was not statistically significant, meaning that similar accuracies may be achieved using either size window when compared to a window of 5 sequences. Due to using a smaller window of 4 or 3 sequences, P300-Certainty could make a BCI selection faster than using a window of 5 sequences.

On the other hand, using a window size of 2 or 1 sequence(s) may allow P300-Certainty to make selections at a faster rate, but the accuracy of selections would decrease significantly.

Conclusion

In conclusion, P300-Certainty was successfully used on-line to achieve dynamic stopping and allow users to type at their own pace. The results also suggest that using a smaller window size will not affect accuracy negatively and allow the BCI to make selections at a faster rate. In the next chapter, the impact of using alpha band analysis (detailed in Chapter 2) and P300-Certainty together is explored.

References

[1] Yin Foo R, Guppy M, et al., "Intelligence assessments for children with cerebral palsy: a systematic review." *Dev.Med.Child Neurol.* 55(10) (2013):911-918.

[2] Meyers AR, Andresen EM "Enabling our instruments: accommodation, universal design, and access to participation in research." *Arch.Phys.Med.Rehabil.* 81(12 Suppl 2) (2000):S5-9.

[3] Byrne JM, Dywan CA, et al., "An innovative method to assess the receptive vocabulary of children with cerebral palsy using event-related brain potentials." *J.Clin.Exp.Neuropsychol*. 17(1) (1995):9-19.

[4] Perego P, Turconi AC, et al., "Cognitive ability assessment by Brain-Computer Interface validation of a new assessment method for cognitive abilities." *J.Neurosci.Methods* 201(1) (2011):239-250.

[5] Schalk, G, McFarland, DJ, Hinterberger, T, Birbaumer, N, and Wolpaw, JR, "BCI2000: A General- purpose brain-computer interface (BCI) system," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, (2004):1034–1043.

[6] Huggins JE, Alcaide RE, "Brain-Computer Interfaces to Assess the Cognitive Capabilities of Locked-In Children." (2013)

[7] Dunn LM, Dunn DM, Peabody Picture Vocabulary Test – Fourth Edition. (2007).

[8] Alcaide-Aguirre RE, Huggins JE, "Novel hold-release functionality in a P300 brain-computer interface." *J.Neural Eng*. 11 (2014):066010.

Chapter 5
Combining Methods of Analyzing Users' Cognitive State to Improve BCI Performance

Introduction

For BCIs to make the move from the lab to a clinical setting, they need to be effective in real-world conditions. Many studies using BCIs are performed in a lab research setting, where unplanned distractions are sparse and possibly undocumented when present. A BCI that has the ability to detect losses in user attention and allows them to control the BCI at their own pace would be more resilient to wandering user attention and more applicable in real-world applications.

The addition of an attention-monitoring method would allow the BCI to be more user-aware. The effectiveness of alpha band power in the EEG as a measure of attention to abstain BCI selections where the user experiences low levels of attention, has been shown to improve BCI accuracy in users that exhibit high alpha variance. This was discussed in detail in Chapter 2 [1].

The P300-Certainty algorithm, presented in Chapters 3 and 4, is another method used to evaluate user attention. P300-Certainty is a selection confidence method that has been shown to improve BCI accuracy by abstaining BCI selections that do not reach a certain selection-confidence level.

Alpha band analysis and P300-Certainty abstain potential errors using different, orthogonal, methods of measuring attention. Each method abstains potential errors that the other may have missed.

Building on the results of the studies discussed in Chapters 2, 3, and 4, this study takes it one step further by combining alpha band analysis and P300-Certainty to evaluate the change in BCI performance of these two methods together versus each method alone.

Methods

BCI Setup

This study contains two different analyses combining alpha classification (Chapter 2) and P300-Certainty (Chapter 3). Although both of these analyses evaluate the effect of using both alpha and P300-Certainty on BCI performance, they use different datasets. The first analysis, which was performed off-line, tested the effect of both alpha classification and P300-Certainty (combining the methods reported in Chapters 2 and 3). The second analysis is off-line analysis of the effect of alpha classification on P300-Certainty data collected on-line (combining the methods reported in Chapters 2 and 4). The descriptions of these two analyses have been separated into different subsections below, to reduce confusion.

Off-line P300-Certainty and Alpha Analysis

Using the results from Chapters 2 and 3, an off-line analysis was performed to evaluate how alpha band analysis and P300-Certainty perform together as orthogonal methods to abstain potential errors versus either of them alone. This study used the same 30-subject dataset and BCI setup described in Chapters 2 and 3 for analysis.

Alpha band analysis and P300-Certainty were used just as described in the Methods sections of Chapters 2 and 3, respectively.

The outputs of both methods were combined using AND logic. For a selection to be typed, both methods must classify it as correct. Figure 10 shows an overview of how both alpha and P300-Certainty act as gatekeepers for BCI spelling.

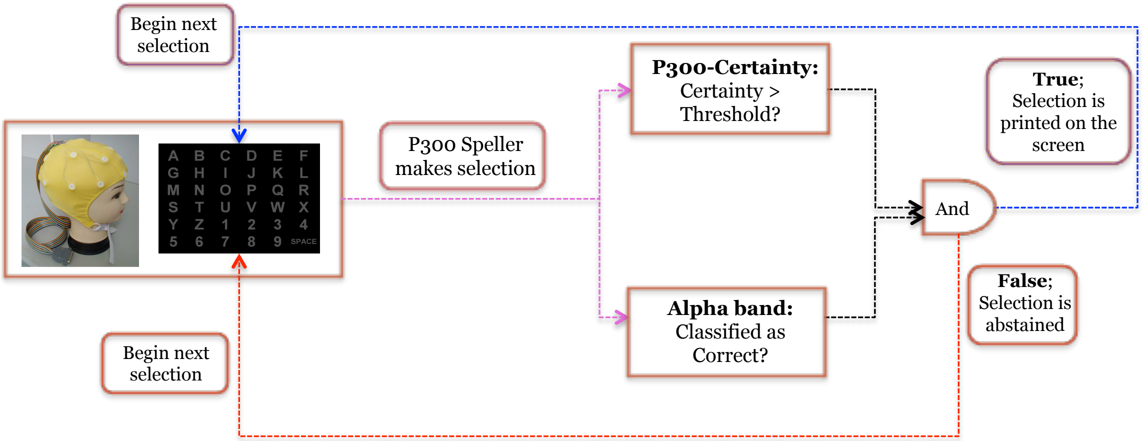

Figure 10: Overview of how the P300-Certainty and Alpha band classification act as gatekeepers to make BCI selections. Note that in addition to the selection made by the P300 Speller, P300-Certainty takes the classifier values as inputs, while Alpha takes the EEG data as input.

Based on the variance in alpha band power exhibited in the calibration data, the subjects in this study fell into two groups: subjects exhibiting high alpha variance (average variance of log alpha = 3.78±0.76 $\log_{10}uV^4$), and subjects exhibiting low alpha variance (average variance of log alpha = 2.74±0.11 $\log_{10}uV^4$). Further analyses were therefore performed and reported for the high and low alpha variance subjects separately and for all subjects together to discover if differences in alpha band variance exhibited in the BCI subject's calibration data affect the effectiveness of alpha band analysis for BCI performance improvement.

To evaluate the combination of alpha and P300-Certainty, the accuracy and BCI Utility (first presented in Chapter 3 [2]) of both together were calculated, for each subject, and statistically compared to the raw BCI accuracy and the performance of each method alone.

Off-line alpha analysis on On-line P300-Certainty data

Another analysis performed in this study was to evaluate, off-line, the impact of alpha analysis on P300-Certainty (on-line) accuracy. The data and BCI setup used in this study is the same 21-subject dataset and BCI setup described in Chapter 4. Alpha band analysis was performed on this dataset in the same way described in Chapter 2.

As reported in Chapter 2, it should be noted that varying levels of alpha between and within subjects' EEG data could cause the alpha power values to vary exponentially between and within subject data. To remove the exponential gaps of alpha power between and within subjects, the $\log_{10}$ was taken of alpha values, and the alpha variance was calculated using these $\log_{10}$ alpha values, as used by Polich in [2].

Based on the variance in alpha band power exhibited in the calibration data, the subjects in the study reported in Chapter 4, fell into two groups: 13 subjects exhibiting high alpha variance (average variance of log alpha = $3.82 \pm 0.37 \log_{10} uV^4$), and 8 subjects exhibiting low alpha variance (average variance of log alpha = $2.04 \pm 0.63 \log_{10} uV^4$). Further analyses were therefore performed and reported for the high and low alpha variance subjects separately and for all subjects together to discover if differences in alpha band variance exhibited in the BCI subject's calibration data affect the utility of alpha band analysis on P300-Certainty on-line data for BCI performance improvement.

To assess the performance of alpha analysis on P300-Certainty (on-line) data, the accuracy of alpha analysis with P300-Certainty was calculated and compared to the accuracy of P300-Certainty alone. The statistical significance between the aforementioned accuracy was done using a two-tailed paired t-test for unequal variances. To determine whether a relationship exists between the change in P300-Certainty on-line BCI accuracy (using alpha classification) and the variance of alpha, the variance of the alpha (over the calibration data) exhibited by each subject was compared to the results of alpha classification, as performed in Chapter 2.

Results

Off-line analysis results

The results of the off-line analysis of both methods are presented in Figure 11. The mean accuracy for raw BCI performance, P300-Certainty alone, only alpha classification, and both P300-Certainty and alpha classification were 82.01±12.59%, 88.82±8.85%, 83.93±9.79%, and 89.81±7.05%, respectively. P300-Certainty alone and the combination of P300-Cert and alpha classification produced statistically significant improvements over raw BCI accuracy with p-values of $p = 0.038$ and $p = 0.016$, respectively. Alpha classification did not produce statistically significant improvements ($p > 0.05$).

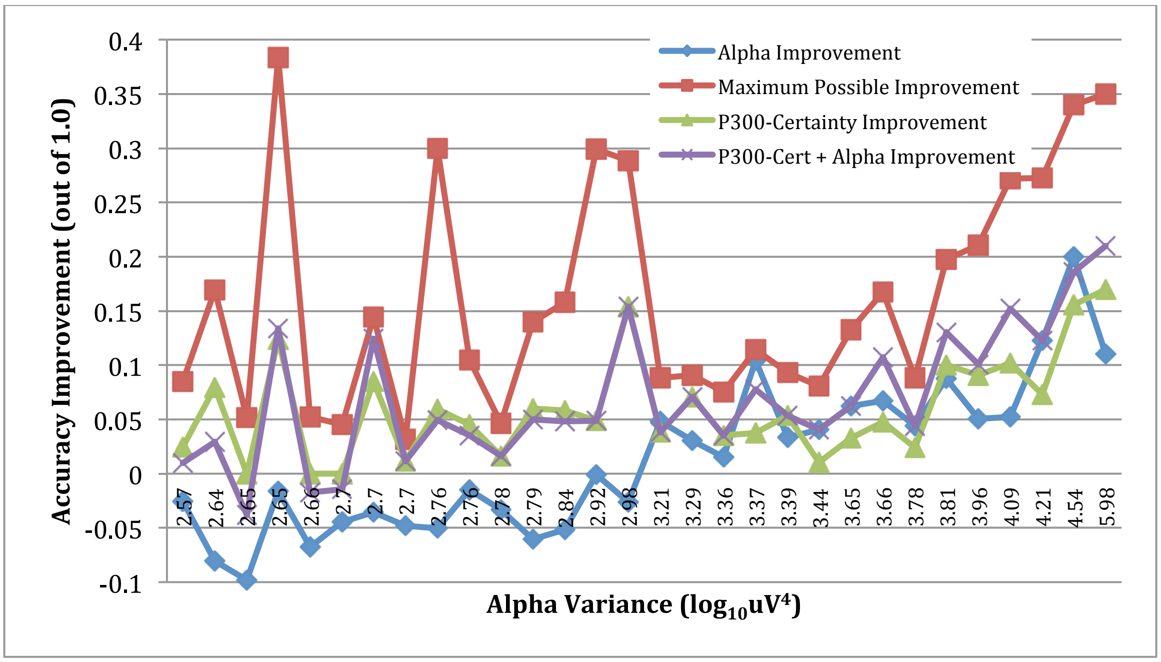

Figure 11: Graph showing alpha variance versus off-line BCI accuracy improvement using alpha, P300-Certainty, and P300-Cert + Alpha. The accuracy improvement is presented in changes in accuracy on a 0 to 1 scale, where 1 is 100% accuracy. Alpha variance of subjects is sorted in order of increasing variance. The red line shows the maximum possible improvement to reach 100% accuracy for each subject as reference.

Table 2: BCI-Utility and BCI accuracy calculated for Raw BCI, Alpha only, P300-Certainty only, and P300-Certainty and Alpha together.

	Raw BCI performance	BCI with P300-Certainty	BCI with Alpha analysis	BCI with P300-Certainty + Alpha analysis
BCI-Utility	2.65±1.61	3.05±1.63 (p = 0.024)	2.78±1.60 (p > 0.05)	3.13±1.67 (p = 0.018)
BCI Accuracy	82.01±12.59%	88.82±8.85% (p = 0.038)	83.93±9.79% (p > 0.05)	89.81±7.05% (p = 0.016)

Table 3: The breakdown of number of selections for Raw BCI, Alpha only, P300-Certainty only, and P300-Certainty and Alpha together.

	Total Selections Allowed	Correct Selections	Incorrect Selections	Abstentions (Abstained Correctly)	Abstentions (Abstained incorrectly)
Raw BCI Performance	12150	10084	2066	NA	NA
BCI with Alpha analysis	11281	9541	1740	326	543
BCI with P300-Certainty	11403	9937	1466	600	147
BCI with P300-Certainty + Alpha analysis	10741	9483	1258	808	601

Table 4: The breakdown of average accuracy and significance for the high variance group, low variance group, and both groups together.

	Raw BCI Accuracy	BCI with Alpha	BCI Accuracy with P300-Certainty + Alpha
High Variance Group	80.92±10.72%	87.5±8.57% (p = 0.041)	90.47±5.85% (p = 0.0063)
Low Variance Group	83.09±14.52%	79.35±13.92% (p > 0.05)	87.36±10.48% (p > 0.05)
All Subjects Together	82.01±12.59%	83.93±9.79% (p >0.05)	89.81±7.05% (p = 0.016)

Off-line alpha analysis on On-line P300-Certainty data results

The mean accuracy for on-line P300-Certainty and both on-line P300-Certainty and off-line alpha classification together were 83.62±9.14% and 86.81±6.98%, respectively; however they were not statistically significant. Figure 12 shows the impact of alpha classification on P300-Certainty on-line BCI accuracy.

The mean BCI accuracy for subjects that exhibited high levels of alpha variance (13 out of 21 subjects; average variance of log alpha = 3.82±0.37 $\log_{10}uV^4$)

for P300-Certainty on-line BCI performance and P300-Certainty on-line BCI performance with alpha classification were 76.35±5.88% and 85.69±6.98%, respectively, which was statistically significant with a p-value = 0.0279. The mean BCI accuracy for subjects that exhibited low levels of alpha variance (8 out of 30 subjects; average variance of log alpha = 2.04±0.63 $\log_{10}uV^4$) for P300-Certainty on-line BCI performance and P300-Certainty on-line performance with alpha classification were 91.14±5.31% and 86.86±5.43%, respectively, which was not statistically significant with a p-value = 0.1612. The variances between the two groups were statistically significant with p=2.18e-7. Variance of log alpha (in the calibration data) was found to be practical in predicting whether or not alpha classification would help increase P300-Certainty on-line BCI accuracy.

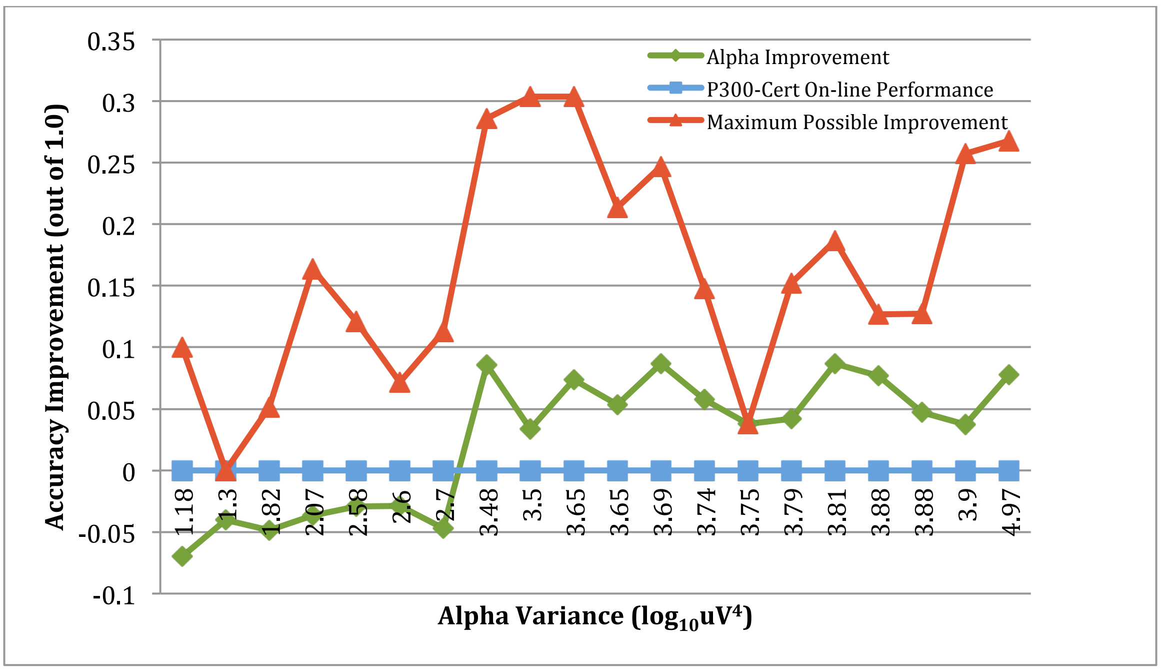

Figure 12: Graph showing relationship between alpha improvements in P300-Certainty on-line BCI accuracy and alpha variance. The accuracy improvement is presented in changes in accuracy on a 0 to 1 scale, where 1 is 100% accuracy. Alpha variance of subjects is sorted in order of increasing variance. The red line shows the maximum possible improvement to reach 100% accuracy for each subject as reference. The blue line represents raw BCI accuracy as a reference.

The high alpha variance group showed improved accuracy using alpha classification to abstain characters for the 13 of 21 subjects who exhibited high alpha levels (low attention levels). The high variance group included 7 CP subjects, while the low variance group included 3 CP subjects. The alpha variance exhibited by the CP subjects did not show any statistically significant difference from the variances exhibited by controls, whether over all subjects or within the high and low variance groups. There was no significant difference between accuracy improvements of CP subjects and those of typically developing subjects.

The alpha variance exhibited by the low variance group caused the alpha classification to allow erroneous selections as well as reject correct selections in a somewhat random fashion. Their corresponding accuracies were decreased due to this phenomenon.

Discussion

Both P300-Certainty and alpha classification increase accuracy by abstaining erroneous selections. Although alpha classification alone increases accuracy in subjects exhibiting high alpha variance, using it together with P300-Certainty increases accuracy in 27 out of 30 subjects, shown in Figure 11. Improvements in accuracy for all subjects show a statistically significant difference ($p = 0.016$) between raw BCI accuracy and BCI accuracy using P300-Certainty and alpha classification. While these subjects showed decreased accuracy when using alpha classification alone, using both methods improves accuracy for 12 out of the 15 subjects exhibiting low alpha variance. This is due to the fact that P300-Certainty and alpha classification are orthogonal methods used to abstain errors using different measures of attention. Each method abstains potential errors that the other may have missed.

For example, if a user were attending to something outside the BCI display, alpha band analysis would indicate that they were attending, however P300-Certainty would abstain the selection due to the lack of the user's EEG responses to the BCI stimuli (flashes). On the other hand, if the user were looking at the BCI display however not attending completely to the flashes, alpha band analysis would indicate that they were not attending and the selection would be abstained, even if P300-Certainty classified the selection as valid.

Using AND logic to combine the decision of P300-Certainty and alpha band analysis biases the BCI to abstain potential errors. Since the goal of this study was to reduce errors and increase overall BCI performance, BCI-Utility was used to determine how these abstentions affected BCI performance by analyzing the benefit gained. The abstentions by P300-Certainty and alpha have been shown to increase BCI accuracy, and BCI-Utility was used to see if these abstentions affected the overall benefit including rate.

BCI-Utility has shown to increase from that of raw BCI performance for P300-Certainty alone, alpha band analysis alone, and P300-Certainty and alpha band analysis together (presented in Table 2). For P300-Certainty alone and P300-

Certainty + Alpha, BCI-Utility was statistically significant from raw BCI, however for alpha alone it wasn't significant. This indicates that the abstentions generated by all methods increased overall benefit and that error correction did not affect the utility of the BCI.

Applying off-line alpha band analysis to the native P300-Certainty data showed a similar relationship as shown in [1]. This emphasizes that alpha variance in users' calibration data is indicative of whether or not alpha band analysis is useful to users' BCI performance. Users with high alpha variance in their calibration data experienced an increase in accuracy, while users with low alpha variance experienced a decrease in accuracy. However, if both P300-Certainty and alpha band analysis were used together, as shown in Figure 11, the combination allows for better performance than either method alone.

Study Limitations

P300-Certainty was used to make the BCI-facilitated PPVT-IV dynamic and more user-aware. However, the study did not include tasks using the BCI without P300-Certainty. To better study alpha classification and P300-Certainty, they need to be tested together on-line with experiments tailored to test performance with and without these methods.

Technical Application

The potential for an alpha-based classification to increase P300-Certainty on-line BCI accuracy can be predicted by the alpha variance exhibited in the BCI user's calibration data. Similar to the findings reported in Chapter 2, users with $\log_{10}$ alpha variance of 3 $\log_{10}uV^4$ and above benefit from an alpha-based classification, and users with $\log_{10}$ alpha variance less than 3 $\log_{10}uV^4$ do not benefit from an alpha-based classification.

Conclusion

In conclusion, combining both methods improves accuracy more than using either method alone. P300-Certainty and alpha classification are orthogonal methods used to abstain potential errors and make the BCI more user-aware by quantifying attention from different angles. Alpha band analysis directly quantifies attention using a studied measure of attention (alpha band power), while P300-Certainty indirectly quantifies attention by evaluating the statistical distribution of the user's EEG responses to the BCI stimuli.

References

[1] Aref, AW, Huggins, JE, “Improving Accuracy of Individual BCI Selections Using Alpha Band Analysis”. *Journal of Neural Engineering*. (2017): In Submission.

[2] Thompson, DE, Blaine-Moraes, S, Huggins, JE’ “Performance assessment in brain-computer interface-based augmentative and alternative communication”, Journal of Neural Engineering. 11(3), (2014).

[3] Polich, J, “Updating P300: An Integrative Theory of P3a and P3b.” *Clinical Neurophysiology* 118.10 (2007): 2128-148.

.

Chapter 6
Conclusion

Brain-Computer Interfaces (BCIs) have great potential to allow individuals with severe motor and speech impairments to communicate. However, traditional BCIs are static and are susceptible to losses in user attention. A BCI that has the ability to detect users' variations in attention would allow the BCI to abstain selections that are unattended. This dissertation introduces two methods (alpha band analysis and P300-Certainty) that measure attention in orthogonal ways. Alpha band analysis is used as a measure of attention directly derived from the user's EEG. Alpha band analysis increased accuracy in users exhibiting high alpha variance. P300-Certainty generates a confidence value for each selection and only types selections that reach a specific confidence level. This increased BCI accuracy by abstaining potential errors. This dissertation shows that using two orthogonal methods to detect and accommodate for losses in user attention can create a BCI that is more user-aware. With this, users are able to control the BCI at their own pace.

Contributions

Alpha band analysis as a measure of attention in P300 BCI

An attention-monitoring method using alpha band analysis was developed and presented. Alpha was shown to allow the BCI to monitor user attention levels and abstain predicted errors, however its usefulness was limited to users exhibiting high alpha variance. A statistically significant increase in accuracy was shown ($p < 0.05$) for individuals exhibiting high alpha variance.

P300-Certainty (Selection Confidence Method)

The P300-Certainty algorithm is a selection confidence method that was developed to improve BCI performance by abstaining selections that do not reach a specified confidence level (certainty threshold). P300-Certainty was shown to be a powerful method that allows the BCI to abstain predicted errors and increase BCI performance. A statistically significant increase in BCI spelling accuracy ($p < 0.05$) was observed.

In addition, P300-Certainty was implemented to achieve dynamic stopping and tested on-line. Dynamic stopping allowed the BCI to make selections only when they reached a specific certainty threshold. This allowed users to control the BCI at their own pace.

Using two orthogonal methods to quantify attention

Using both alpha band analysis and P300-Certainty was shown to increase BCI performance more than either method used alone. Since both methods were used to quantify attention in different ways, each method abstained errors that the other method may have not detected.

Study Limitations

The data analyzed in this dissertation was for subjects who were supposed to be maintaining attention while copying text (copy-spelling) [1] or taking a BCI-facilitated cognitive assessment [2]. This is an unrealistic usage condition since they had no text composition tasks and few distractions. The only distraction inherent in the task was the need to problem-solve how to correct an error if it occurred (basically recognizing the need to backspace), which is a limitation of the data analyzed in this study. In addition, users with low alpha variance were, in fact, paying attention. This motivates designing a study that thoroughly investigates the effect of distractors on BCI usage. A study with a BCI designed to mimic real-world usage of a BCI, including compositions tasks, strategic distractions, and periods where the user is not paying attention to the BCI display, would overcome this limitation and test the attention-monitoring ability of alpha band analysis and P300-Certainty to their fullest.

Future Work

Alpha band analysis to be tested on-line

Using recorded data, alpha band analysis increased BCI accuracy for users with high alpha variance. However, to test alpha band power's attention-monitoring ability to its fullest, a study is needed that is designed to mimic real-world usage of a spelling BCI, including composition tasks and strategic distractions.

It was shown in this dissertation that alpha band analysis improves BCI accuracy in users exhibiting high alpha variance. However, for users exhibiting low alpha variance, alpha band analysis decreases accuracy. Using this difference between users with high and low alpha variance, alpha band analysis can be used to monitor alpha variance and only be activated in periods with high alpha variance. This way alpha classification can be used on all subjects, and be deactivated during periods of low alpha variance so that user accuracy does not decrease.

Testing different machine-learning methods for alpha classification

Alpha classification using Linear Discriminant Analysis (LDA) increased BCI accuracy for users with high alpha variance. However, different machine-learning methods (e.g. Support Vector Machines or a clustering algorithm) may prove to be more effective at alpha classification.

Alpha band analysis also has the potential to improve P300 BCI calibration

Before being able to control a P300 BCI, the BCI needs to be trained to identify the user's P300 response. Attention to the BCI is crucial during calibration for the ideal BCI performance. Alpha band analysis can be used during calibration to identify and omit parts of the calibration data where the user lost attention to the BCI. This will allow the BCI to be trained on attended data, which will ultimately allow the user to control the BCI without as many errors.

P300-Certainty and alpha band analysis to be tested on-line together

To test the full potential of both P300-Certainty and alpha classification as methods to abstain potential errors by identifying periods of low user attention, both methods need to be evaluated on-line simultaneously. This study would investigate the effect of distractors on BCI usage. A study with a BCI designed to mimic real-world usage of a BCI, including compositions tasks, strategic distractions, and periods where the user is not paying attention to the BCI display, would test the attention-monitoring ability of alpha classification and P300-Certainty to their fullest. It would be ideal if people with severe motor and speech impairments were recruited for this study so that the effectiveness of both methods could be tested with end-users.

Overall, this dissertation presents research that can help create dynamic attention-monitoring P300 BCIs that are resilient to wandering user attention and allow users to control the BCI at their own pace.

References

[1] Thompson, DE, Gruis, KL, Huggins, JE, "A Plug-and-Play Brain-Computer Interface to Operate Commercial Assistive Technology" *Disability and Rehabilitation: Assistive Technology.* (2013).

[2] Huggins JE, Alcaide RE, "Brain-Computer Interfaces to Assess the Cognitive Capabilities of Locked-In Children." (2013).