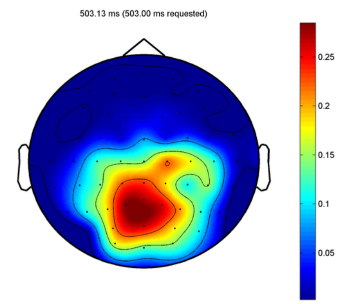
CHALMERS
UNIVERSITY OF TECHNOLOGY



Brain-Computer Interfaces in a car: Science fiction or a realistic concept?

An investigation regarding technical pre-conditions and potential

Master's thesis in the Systems, Control and Mechatronics and Biomedical Engineering MSc programmes

Adrian Edgren
Martin Kaalhus

Department of Signals & Systems
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2015

Brain-Computer Interfaces in a car: Science fiction or a realistic concept? An investigation regarding technical pre-conditions and potential

Adrian Edgren, 900421, adriane@student.chalmers.se
Martin Kaalhus, 920116, kaalhus@student.chalmers.se

Supervisor: Tomas McKelvey, Chalmers
Jan Nilsson, Semcon
Examiner: Tomas McKelvey, Chalmers

Master's Thesis 2015 Spring Semester (EX016/2015)
Department of Signals & Systems
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: An Emotiv Epoc headset, a Volvo V70 and a topography plot made in Matlab. The topography displays the potentials from the brain recorded from the scalp of a user by electrodes.

Keywords: Brain-Computer Interface, Machine Learning, Strategic Controller, Human-Machine Interface, Classification

Typeset in LaTeX
Gothenburg, Sweden 2015

Brain-Computer Interfaces in a car: Science fiction or a realistic concept?
An investigation regarding technical pre-conditions and potential
Adrian Edgren & Martin Kaalhus
Department of Signals & Systems
Chalmers University of Technology

Abstract

This study aimed to evaluate the feasibility of using a Brain-Computer Interface (BCI) based upon the P300 Event-related Potential (ERP) in an user-friendly setting in a vehicle. Compared to earlier studies within the field, this thesis evaluates the use of BCI systems in a new setting, taking steps towards a commercial product. The specific end-application in mind is a strategic controller in an autonomous car, using the system to issue commands such as overtake or changing lanes. This puts requirements on response-time and usability.

Using a commercially available EEG headset, the Emotiv EPOC headset, a number of tests have been performed to identify a suitable set of components to use in the specific setting with the selected hardware. Different signal processing and machine learning tools were applied and through evaluation of the performance it could be seen that using Independent Component Analysis (ICA) as a spatial filter combined with a Support Vector Machine (SVM) classifier for a 3x3 matrix of choices was the ideal choice of components for the current setup.

Using this system architecture, while performing tests in a moving vehicle, showed that the inherent disturbances of the vehicle environment did not have an significant impact on the performance. However, the overall response-time of the system and interface is longer than required in the specific setting. The hardware used during the thesis work is not ideal for the application, where repeatability of measurements as well as usability issues needs to be addressed.

Keywords: Brain-Computer Interface, Machine Learning, Strategic Controller, Human-Machine Interface, Classification

Acknowledgements

The authors would first of all like to thank Jan Nilsson at Semcon, where this thesis has been performed, for his continuous and energetic support, as well as interesting discussions about future implications of the results. Apart from Jan, we would also like to thank Anders Björklund as well as Karin Eklund for taking the thesis proposal under consideration and connecting us with Jan.

We would also like to thank the other thesis workers for meaningful discussions as well as giving a feeling of a small thesis work community. Without you, the work during the thesis would not have been as inspiring and rewarding.

For insightful technical feedback as well as reviewing our thesis, we thank our examiner Tomas McKelvey, at Chalmers University of Technology. For their continuous investments into future technology as well and more specifically into autonomous cars and new HMI concepts, we would like to thank Volvo Cars and in particular Claes Edgren, whose effort and partnership with Semcon in this research project made the thesis possible. In conjunction, with that we would like to thank Victoria ICT for interest in the project as well as reviewing the potential of the results.

The authors would also like to give thanks to A. L. Hodgkin as well as A. F. Huxley for their work developing an theory of action potentials, while dissecting frogs and squids. This work created a foundation for the field of electrophysiology, which in time led to further research regarding electroencephalography which forms a basis for this thesis.

Adrian Edgren Martin Kaalhus, Gothenburg, June 2015

Contents

1 Introduction

The power and versatility of the human mind has fascinated human kind throughout history and also made us the dominant species on Earth. The source of the mind is our brain, arguably the most complex object in the universe that we know of. For the majority of history, mankind have been limited to express our minds intentions through our muscles and skeletons to issue actions including speech. Using increasingly sophisticated mechanical and electrical tools we have empowered the mind to store moments to share with others through pictures and video, sending messages instantly across the globe, as well as control cars, planes and other vehicles with great force and precision.

Mankind is now standing before a new breakthrough where our limbs and their limitations can be bypassed by giving the control directly to the mind. This will open new opportunities and possibilities to new applications and empower people even more as well as give disabled a possibility to use today's technology in the same way as others. Apart from this development the tools we develop are becoming more and more intelligent and need less direct input from humans, since they rely their operation on their own minds and sensors. Combining this we are in the future able to create an autonomous car where the needed input is given solely by the power of our minds.

This development also enables people that earlier could not use systems because of their disabilities to use them in the same way and the same level as everyone else. The development of a Brain-Computer Interface for strategic use in a car as well as in other applications gives disabled the power to take control over their own lives and enables them to live a more active and full life. Since this system only uses brain waves for input even severe disabilities such as ALS no longer need to be a problem. The goal in the future is to be able to use different systems while only being limited to the power of our minds.

1.1 Brain-Computer Interface

The history of brain-computer interfaces starts with Hans Berger and his discovery regarding the electrical activity of the human brain, and with it the development of a way to measure it with electroencephalography (EEG) [1]. While using EEG a number of electrodes are placed on the scalp of the user. A standardised system of electrode positions is the 10-20 system, as shown in Figure 1. This picture shows one version of the system, which can be expanded with additional positions if necessary.

There has been rapid development in BCIs since the mid-1990s. Complex brain motor cortex signals has been recorded from neural ensembles and used to control external devices [3]. This have resulted in widely use of BCIs for several types of communications- and control systems in different areas of applications [4, 5]. Utilising this development this thesis aims to investigate the possibility of using today's technology as an HMI for controlling systems in a car.

1.2 Autonomous Cars

The development of driver-less cars started in the late 1970's [6]. Since then large steps to completely autonomous vehicles have been taken, with increasingly rapid development. This includes advances in a large number of fields including sensor technology, processing power, algorithm development to name a few. Although to this point in time most vehicles are still under test with different degrees of autonomy,

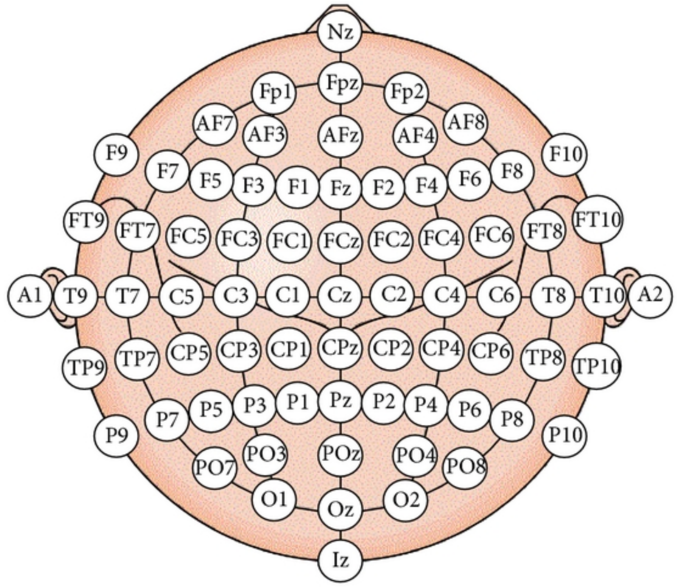


Figure 1: The international standard of electrode positions for EEG [2].

both because of technical and juridical reasons. This results in a large number of semi-autonomous cars with technology such as self-parking, adaptive cruise control, lane assist and dead-angle detection. Examples of car companies with advanced technology in this field are Mercedes, Audi, Tesla and Volvo where the first three are soon offering complete autonomy in highway situations [7]. Volvo are at the present testing autonomously driven cars as well as caravans of connected trucks in an urban setting in Gothenburg.

Several different techniques has been developed for issuing the needed input to the semi-autonomous cars. This includes among other things eye tracking [8] as well as remote control using smartphones [9]. With today's technology approaching complete autonomy of the cars, less and less input is needed from the driver, resulting in a need of a strategic controller to enable the user to navigate the car and issue commands to it. Research regarding using a BCI as a strategic controller has been performed with promising results [6]. Using motor imaging, which will be presented in 2.2, this study made it possible for the driver to choose between four options: left, right, forward and backwards. However, motor imaging requires effort from the user and is limited to a smaller number of possible user choices, compared to a P300-based system which will be discussed further.

1.3 General Requirements on Vehicle HMI

The HMIs that are controlling systems in the vehicle need to be clear, user-friendly, fast and easily used. A recent research project called AIDE categorised the functions in an vehicle-HMI into two parts; ADAS (Advanced Driver Assistance System) which function is to support the driver in the main task of driving the vehicle and IVIS (In-Vehicle Information Systems) which main functions does not concern the driving but other aspects of comfort in the vehicle, such as climate control, music and communication [10].

During the AIDE project several investigations into the requirements of the performance of HMI systems were performed. This included all aspects of driving and put the user's limitations and needs in focus [10]. The most important factors that the investigations resulted in are:

- Adaptive/compensatory behaviour
- Distraction
- Mental Workload

Although, due to the fact that the BCI is to be integrated in an autonomous vehicle the established rules and guidelines for designing vehicle HMIs will be changed and be comparable to developing a HMI in general, for example on a standard computer monitor. Some differences still remain though and new issues arise.

P. Gruhn discusses HMI design and their elemnts and summarise a good design into four points:

- Contrast
- Repetition
- Alignment
- Proximity

In short, this means that things that are different should look very different, to reuse graphical elements as much as possible, to align elements with some a visual connection to another as well as placing elements that belong together close to each other [11].

Designing an interface for an autonomous vehicle is new field, since autonomous vehicles have not existed earlier. It is hard to transfer experience from other automation domains based on:

- Different environmental variables
- Different type of decisions
- Different time frame
- Different available information

Important issues that arises include achieving and maintaining the trust of the user, dealing with complacency as well as having the mental workload of the user in mind [12]. Based on the novelty of the application area for HMIs in autonomous vehicles it is hard to evaluate performance and user needs, though an efficient and trust-worthy HMI is one of the key pieces to design a fully functional and user-friendly autonomous vehicle.

1.4 Objectives of this study

This project takes a leap in to a new, unique, area of this type of BCIs. Even though similar solutions have been investigated through other methods there is to the authors knowledge no other research regarding P300 signals in a car environment. The research performed within this study will capture the performance of the BCI, its internal and external variables, and discuss it from the following three perspectives:

- Accuracy
- Response time
- User experience

The reason to why these points have been chosen are many but in general accuracy and the time from choice to confirmed decision is of huge importance when it comes to car manufactures guidelines. However, since autonomous cars are not yet available for consumers there are no defined requirements for the response time. Nevertheless, a reasonable demand is that an user will be able to make a choice to turn of the freeway within the time from when the car passes the sign to when the exit appears.

The goal in this first leap towards a fully working system is to determine a number of parameters regarding the actual design of the BCI, rather than evaluating performance in detail. Not only will the physical appearance of the system be investigated but also two types of classifiers in combination with three types of spatial filters will be compared. Hence, the contribution from this study to further research projects lies in their process of decision regarding the basic outline in design of their system and most likely to what type of classifiers and initial processing of data to choose or not. The particular questions to be answered during this study is; ‘’What system components and configurations are important for a P300 based BCI to perform in a moving automotive vehicle?” and ”Is it possible to design a working and user-friendly system based on the P300 ERP, with the technology available today?”

1.5 Limitations

As mentioned in the previous part of this chapter the objective in this project is to investigate a set of system components and configurations in a BCI system, operating in a car. However this investigation will be performed off-line, i.e. the data will be collected in real time but the analysis and validation are executed off-line. The three points stated in Section 1.4 are only validated towards the theoretical possibilities of the optimal system set of variables. No real-time system with feedback for the user and internal classification will be investigated and validated within this project, although the outcome of the project gives suggestions in designing this system.

2 BCI – An overview

As introduced in Section 1.1 EEG has become a well-known and useful system for various form of medical research. In later days commercial products where EEG is included in BCIs have become more common. In this Section a more extensive background will be given to the BCI as an overall system but also to the three most common methods used in advanced BCI systems today. Only one of these are used in this study, the event related potential known as the P300 [13], but all three are presented to display that several alternatives exist.

Brain-Computer Interfaces are, as mentioned, systems that interprets the brain activity of a subject. This interpretation can be used to control an action and therefore a user is, through a BCI, granted a new channel for communication comparable to nerves and muscles that are the normal channels a body can control [14]. Hence, it is possible to enable a huge number of services and applications to individuals with different sort of disabilities that earlier was impossible for them.

On a research level EEG is an important tool since it givs an insight to the complex system that is the human brain. A brain contains a huge amount of neurons, 100 billions, and billions of connections between them [15]. The challenge for the BCI research is to interpret the activity in these connections with a limited resolution from the electrodes. Even small physical movements, e.g. using a finger, results in millions of neurons firing impulses in an exact order. If a BCI were able to detect this order, it is possible to use this interpretation to simulate the specific movement in hardware and the intended movement could be duplicated exactly. However, due to numerous reasons, limitations in the hardware within the BCIs and gaps in the knowledge of how the processes in the brain actually works, this is not possible with the technology available today.

The architecture of typical BCIs can be summarised into four blocks; an acquisition system (i.e. a headset), a data collector (i.e. an Analogue to Digital converter (A/D-converter)), a data processor (i.e. digital filters of various types) and a classifier. In addition to this is of course a user and in the end an action of some sort. An overview of a BCI can be seen in Figure 2. The acquisition part of the system can be of two types, invasive or non-invasive. An invasive BCI makes use of implanted electrodes attached directly on the surface of the brain [16]. This method has the advantage of a better signal to noise ratio (SNR) but the obvious drawback is the need for surgical procedure to attach the implantation. Non-invasive BCI uses electrodes on the surface of the scalp to acquire the signals that originally comes from the brain. This naturally gives lower potential since the electrodes are further from the source but are much more user-friendly since the BCI does not need any or at least very little preparation compared to the invasive method. In this project a non-invasive method will be used and all further discussions where BCIs are mentioned refer to the non-invasive case.

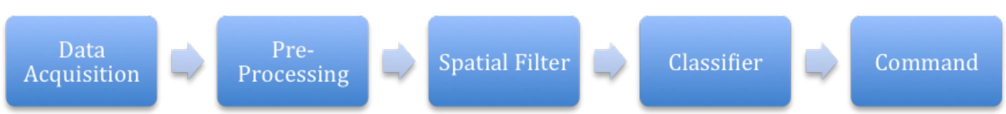


Figure 2: A overview of the components included in a BCI. The data processing part has been separated into two parts, initial signal processing, e.g. digital filters, and spatial filters, further discussed in Section 4.2

In a BCI the EEG data must be pre-processed before anything useful can be retrieved from it, as can be seen in Figure 2 there are several types of pre-processing. The one marked with pre-processing refer to digital filters to obtain the most relevant frequency spectrum of the signal and to reduce the occurrence of unnecessary frequencies. However, the spatial filter is also a type of pre-processing with the objective to reduce the set of channels by establishing the most important parts, further discussed in Section 4.2. This rigorous refinement of data is needed because of the low signal-to-noise ratio (SNR) due to the fact that the electrodes on the surface are recording the combined electrical activity originating from many different neurons in several lobes of the brain. This mix of different signals is then passed through the bone and human tissue of the skull that tend to act as a low-pass filter of the signal [3].

Within the signal, informative parts can be found both in the frequency- and time-domain. In Table 1 some interesting frequency spectrums or wave forms are listed. The energy level in the respective spectrum of the signal type Table 1 can be increased by an user on purpose. The energy level in the frequency spectrum Beta is an good example, it can be manipulated by concentration and is a common input to BCIs in game applications [4]. This is one possible method for manipulating a BCI but this requires a tremendous amount of willpower, training and time. Even if such an effort is performed, the resulting system is often slow and needs recalibration and is not discussed further in this paper [5].

In the time domain, several pattern recognition algorithms for motor imaging can be used [17, 18]. Three types of signals and there possible interpretation by BCIs are further introduced in the following sections, since they are the most promising methods for fast and safe decision-making in BCIs of today [19, 20].

Wave type	**Frequency**	**Action to enhance**
Gamma	31 Hz -	Reflect the consciousness . One common description is that Gamma and Beta waves are connected to attention, cognition and perception.
Beta	13 Hz – 30 Hz	Are associated to focus of the mind, e.g. solve mathematical equations, reading etc. Most common in the central and frontal lobes in the brain.
Alpha	7.5 Hz – 12 Hz	Are connected to relaxation. When an individual closes his or her eyes and think peaceful thoughts are situations when alpha waves typically occurs.
Theta	3.5 Hz – 7.5 Hz	Are connected to daydreaming and inefficacy. The lowest frequency levels of theta waves are present in a state between awareness and sleep. This type of waves are unusual in adult subjects.
Delta	0.5 Hz – 3.5 Hz	Are present in healthy person only during periods of sleeping.

Table 1: Basic wave shapes and their frequencies [21].

2.1 P300 Event Related Potential

In more than five decades the P300 event related potential (ERP) have been known to researchers in the field of neuroscience [13]. An P300 evoked ERP is associated with a protocol known as the

Oddball Paradigm. This protocol can be explained as a number of altering visual stimuli presented to an user. The user is instructed to search for a certain one, this special stimuli appears rare compared to other in the protocol but when it happens an ERP is triggered in the brain of the user and can be detected by a BCI [22]. A popular Oddball Paradigm application is the P300-Speller, this method will be further discussed in Section 2.1.1. There are other techniques to invoke an ERP response, these are not in the scope of this report and will not be discussed further but the authors recommend interested users to read further about them [23].

The neurophysiological process that underlies an P300 ERP response is unknown, yet there are a several known factors which are affecting it. In this Section, task related factors affecting the P300 will be discussed together with some theories regarding the neurophysiological origin.

The appearance of a P300 signal is a small electrical potential in the range of 6 to 20 μV with a frequency in the span 1-16Hz. It rises about 250-400 milliseconds, normally around 300 ms and therefore it is called P300, from the the time when the stimulus appear in front of the subject. The signal peak lasts for less than 100ms, see Figure 3. The most important areas to retrieve measurements are the parietal lobe and occipital lobe see Figure 4 [24]. However, even if measurements are recorded from the correct positions and in an accurate time span it is hard to distinguish each ERP signal from the noise included in the data. From previous studies it is obvious that it is possible to trigger an ERP of P300 type in a subject and to classify it correctly using several filtering techniques [25]. These methods are further discussed in Section 4. By using data from several occasions where a P300 is present, classification process can be made more accurate.

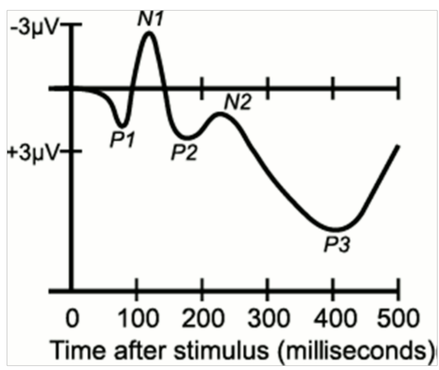


Figure 3: The characteristic of a P300 signal. In the Figure is the P300 peak marked as P3. P2 and P1 are peaks occurring 200 resp. 100 milliseconds after the visual stimuli appears to the user. Note the reversed scale of the voltage on the y-axis, from [25].

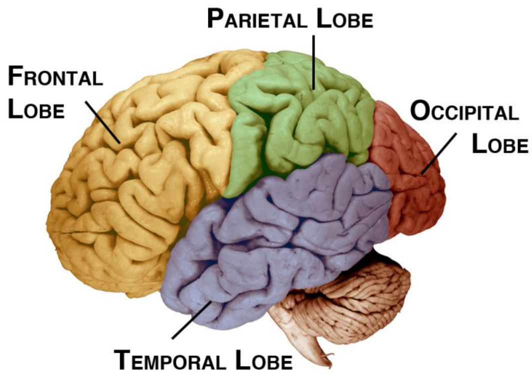


Figure 4: A overview of the lobes within the human brain, image from [26]

2.1.1 P300-Speller Application

In this type of application the P300 and the Oddball paradigm are exploited to determine a letter chosen by a subject to be typed. The application consists of a 6×6 array of characters in an alphabet as displayed in Figure 5.

If the user looks at a chosen letter and the particular row is illuminated, this will trigger an P300 signal. By also altering the columns to be illuminated a coordinate can be established by observing an ERP for both row and column separately [27]. The sequence in which the rows and columns flashes is chosen randomly and since there are five non-targets stimuli in each direction are it is rare to get target stimuli. A common protocol for this application is to instruct a subject to count the number of times a certain letter is flashing. By making the algorithm to flash every row and column a pre-defined number of times and to save the occurrences of the chosen letter it is possible to evaluate the data and to train a classifier in the task to distinguish target from the non-targets [24]. There are many types of classifiers and several of them will be discussed further in Section 4.

The main reason to why the system is set to flash each column and row several times is to use averaging in the quest of eliminating noise [28]. Hence, by making use of the average value for the signals across the flashes of a certain row it is possible to reduce the noise in the data concerning this particular row. This method also is used when an user actually tries to write by manipulate the application, the letter to be written has to be illuminated several times before a decision is made [24].

The amount of time a row or column is illuminated is called the ‘’stimulus interval” and the time between these flashes, when the matrix is blank, is called ‘’inter-stimulus interval” (ISI). The length of ISI varies from one study to another and there is no standard to follow. However, in one study a test was performed with a longer and a shorter duration, of ISI 350 ms and 175 ms respectively, where the shorter interval gave a better performance. Also the size of the matrix was altered, a 3×3-matrix was used, and this seemed to give improved results [29]. In one study the P300 speller is tested in a noisy environment and compared to a test group which used the system in a controlled environment and the result indicated that the system could perform acceptable also in the noisy environment. In fact the difference in accuracy were only 5%. The same study displayed statistical result that an altering between rows and columns compared to an altering of element in the speller matrix is beneficial with an ISI of 75 ms if the time for flash is 100ms [30].

Figure 5: The typical P300-speller alphabet [31]

2.2 Motor Imaging

Motor imaging (MI) is another possibility of interacting with a computer trough a BCI In such an application the subject is instructed to imagine a movement but not execute the actual physical response, e.g. think about moving the left arm but actually don't use the arm. This kind of MI makes several areas in the brain more active. Yet, these areas are not necessary the same areas that normally activates when the particular physical action is performed [32]. Nevertheless, since it is possible to activate new regions in the brain for several imagined movements it is possible to find an unique pattern for each MI in the signals if several channels are used. By interpreting the signals during a training protocol it is possible to use a classifier to distinguish the commands and to use this trained classifier in live application.

This method has been used successfully in several fields of medical research especially when it comes to rehabilitation of patients affected by stroke [33, 34, 35]. Also, the method is used to help individuals that have lost limbs or are paralyzed to manipulate robotic replacements and aids [36, 37].

Not only has MI been used in the field of medicine but also in commercial products and purposes. Since the limits for this system is small there a numerous of different studies on new applications for this method. The most spread use of MI in commercial purposes is manipulation of virtual worlds, e.g. games and virtual car simulations [38, 39]. However, MI have also been used in studies performed in co-operations between universities and car manufacturers. In Germany an example of these resulted in a team of researchers which have been able to drive a car using MI [6].

The drawback of MI large is the amount of training which is needed, days in some cases [40]. Also, it is required for the user to re-train or recalibrate the system every time it is used. Some studies shows that the voltage from the potential within the brain tend to get lower during longer session with BCIs and this has the affect that MI applications needs to recalibrate during use to maintain the same response time [6]. Users that are eager to learn this system can get a very stable and secure application with 100% accuracy of the decisions they make using MI.

2.3 Steady-state visual evoked potentials

A third promising candidate is known as steady-state visual evoked potentials (SSVEPs). In this method the stimuli is presented to the subject similar to the one used in the P300 case. Yet, the triggered signal formed in the brain is completely different. The process in SSVEP starts with a screen that displays several potential options to an user. By making every option to flash with certain frequencies it is possible to determine which of the options the subject is focusing on. It should be mentioned that the main difference from the set up of a P300-speller is that in this case all options are flashing simultaneously but with an unique frequency. The classification from the BCI is performed by the detecting the highest energy level in the particular frequency band [41]. The most important part of the brain where this phenomenon can be obtained is the occipital-parietal region [42], see Figure 4.

There are several advantages with this method in a BCI. One of the most important one is that no training is required since the increase of energy in the frequency spectrum is performed spontaneously. Also the response time is, using this method, relatively fast, $2-7$s [41, 43]. However of course are there some downsides connected to this method also. An initial recording is required every new session to establish the mean energy levels in the frequencies related to different choices. This type of calibration can be needed in the middle of the sessions as well. Many studies of SSVEP are only considering two types of choices and it is just recent that the method has been tested for extended menus [44, 45]. There is also an aspect related to the user experience of this system since several potential choices entails a screen with several flashing objects. It is most likely that such an application within a car can be perceived as disturbing. Nevertheless, studies on the SSVEP in BCIs have showed encouraging results regarding accuracy and this method seems very promising in the future of the field in mobile BCIs.

3 System Components

Within this project the design of a BCI system for use in the investigation is taking place. Naturally, this system is consisting of several components both of hardware and software types. In this section the hardware parts are presented together with the environments and tools where the various type of software is to be created.

3.1 Hardware Emotiv EPOC

In this project the EEG headset EPOC from Emotiv is used as an EEG hardware device. The EPOC device has been used in various research studies before and there are a great amount of information available regarding its basic performance in controlled environments. The device is a 14-channel wireless EEG headset. The electrodes are positioned at $AF_3, AF_4, F_7, F_3, F_4, F_8, FC_5, FC_6, T_7, T_8, P_7, P_8, O_1, O_2$ according to the international 10-20 location system which as mentioned in Section 1.1. The EPOC has also two reference electrodes positioned below both of the ears. The data are transmitted to a suitable device through a 2.4 GHz wireless USB dongle. The sample rate is 128Hz with 0.51μV resolution. However, the true sampling is performed at 2048Hz and the data is down-sampled to the mentioned rate. There is a single A/D-converter built in the headset used to convert the continuous EEG readings into discrete time data. The EPOC device has a digital 5th order Sinc filter that restricts the bandwidth of the obtained data to between 0.2 and 45 Hz. In addition to this, power line noise is filtered from the signal by two notch-filters at 50 and 60 Hz [46].

The Emotiv EPOC is not licensed as a medical grade system, instead it is labelled to be suited for game applications. Yet, studies have confirmed that it does capture real EEG signals and ERPs as well. In one study is the EPOC compared to a medical grade device, a 128 channels cap with Silver Chloride electrodes and a sampling rate of 2048Hz, for the determination of P300 signals. Both devices reached above 75% in classification certainty and there were only 8% difference between them, to the medical grade cap advantage of course [47].

As the EPOC seems to capture real EEG in a satisfying manner it has become quite popular in academic research. This is actually not the only reason to the growth in popularity, the mobility of the EPOC in combination with a relative low price opens up a new era for BCIs. In addition to this there is a growing interest for machine learning and data processing techniques, which have resulted in an explosion in the field of easy access BCIs and their applications [48, 6, 38]. The mentioned attribute regarding mobility is something that is most desirable in this project and is one of the main reasons to why this device is chosen. Yet, there were other candidates of hardware available and a market investigation was performed before the choice was made.

3.2 Software

During the course of this thesis several software products will be used in development in research. The most significant will be presented briefly in this section.

OpenViBE: OpenViBE is a program for easily developing system that rely on EEG signals as input and acts a software platform dedicated to designing, testing and using brain-computer interfaces. The package includes a Designer tool to create and run custom applications, along with several preconfigured and demo programs which are ready for use. The software is used for real-time neuroscience (that is, for real-time processing of brain signals). It can be used to acquire, filter, process, classify and visualise brain signals in real time [49].

Python:

Python is a powerful well-known platform-independent programming language and is a widely used general-purpose, high-level programming language. Its design philosophy emphasises code readability, and its syntax allows programmers to express concepts in fewer lines of code than would be possible in languages such as C++ or Java. The language provides constructs intended to enable clear programs on both a small and large scale.

Python supports multiple programming paradigms, including object-oriented, imperative and functional programming or procedural styles. It features a dynamic type system and automatic memory management and has a large and comprehensive standard library [50].

Emotiv software:

Emotiv is an emerging company developing and manufacturing EEG BCI devices with relevant software for all kinds of applications. Examples of these software that are covered within this project is:

- **TestBench:**
  TestBench is Emotiv's own software for developing EEG applications with several built-in functions. This software will be primarily be used in conjunction with OpenViBE to verify that the

EPOC system is working as it should, as well as gain an understanding and knowledge of the system [46].

- **Emotiv Control Panel:**
  Another tool for development provided from Emotiv are the Emotiv Control Panel, this have a lot in common with TestBench but with another set of basic functions. However, within this project are not the actual interface used other than for make sure that good contact is established before tests are executed. Instead are files which are included in the installation used to enable contact between the own written Python program and the hardware.

Matlab/Simulink:

MATLAB is a powerful calculation tool. This also includes a simulation and developing tool called Simulink. The name stands for matrix laboratory and is a multi-paradigm numerical computing environment and fourth-generation programming language. Developed by MathWorks, MATLAB allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other languages, including C, C++, Java, Fortran and Python. Although MATLAB is intended primarily for numerical computing, an additional package, Simulink, adds graphical multi-domain simulation and Model-Based Design for dynamic and embedded systems [51].

4 Theory

In this section further theory regarding the implemented system are displayed. The first part consist of a description of classifiers in general and also a presentation of the two types used in this study. These are linear discriminant analysis (LDA) and support vector machines (SVM). In the last section the concept of spatial filters is explained and the three types used in this project are presented, together with there respective mathematical theory.

4.1 Classifiers

Classifiers, the part of the BCI where data is classified and the BCI base its decision on this classification. Thus, the purpose of using a classifier in a BCI is to classify data retrieved from the EEG into groups of classes. The methods used in this project perform this task by dividing the data space into subspaces, each subspace holds an unique type of class. In the two dimensional case this can be seen as two types of dots divided by a vector which separates the two types apart from each other. This is visualised in Figure 6, the vector is however labelled as a hyperplane in this figure. The reason to this is the simple fact that the classifiers can be used in a multidimensional case as well as the two dimensional and then naturally a vector is equal to a plane. Nevertheless, in both cases implies a larger margin between the data points and this hyperplane a better classification of the data [52].

The method Linear Discriminant Analysis (LDA) can be used to find such a vector or hyperplane. As the name states this is a linear method and LDA can not only be used as a classifier but also as a dimensionality reducer, i.e. a spatial filter. Such filters are further discussed in Section 4.2 and they are applied before the actual classification [53]. However, another method for separating the classes is the Support Vector Machine (SVM). This method essentially solves the same problem as the LDA and is also linear. The difference between them is small and related to how this plane is found and how the angle of it is calculated [54]. However, both methods have the objective to maximise the distance to the found plane and each class [55].

Within this project is two classes enough to separate, if a data set holds a P300 wave or not. The classifiers therefor are giving an output in form of an integer of binary type, i.e. an one implies that a P300 is present and a zero implies the opposite [56]. The training data, as well as later unknown data, is a vector of samples from the pre-processed EEG measurements, denoted as $\mathbf{x}$. This vector correspond to a given output y where y is 1 or 0.

Both the LDA and the SVM are used in this project and to enable the classification of unknown data, training of the algorithms is necessary. As with almost all machine learning processes. This training is performed with a data set of the same type as in the real application but the outcome of the classification must be known. By feeding in n vectors of known data, $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, .., \mathbf{x}_n]^T$ corresponding to a vector of the correct classes, denoted as $\mathbf{y} = [y_1, y_1, ..., y_n]^T$, in the LDA and SVM algorithm in a software, MATLAB [57][58] or Python [52], the classifiers $\mathbf{W}_{LDA}$ and $\mathbf{W}_{SVM}$ are trained.

$$\mathbf{W}_{LDA} = LDA(\mathbf{X}, \mathbf{y})$$
$$\mathbf{W}_{SVM} = SVM(\mathbf{X}, \mathbf{y})$$

These can then be used to predict the class of an unknown vector $\mathbf{x}'$ provided this vector is of same nature as the vectors in $\mathbf{X}$.

Different types of these classifiers exists but the main difference from one to another is the function which decide the characteristic of the hyperplane, called kernel function, used in the training stage. In this project the linear type of kernel is used, this means that the hyperplanes are linear but there exist several other possibilities, e.g. exponential, polynomial etc.[54].

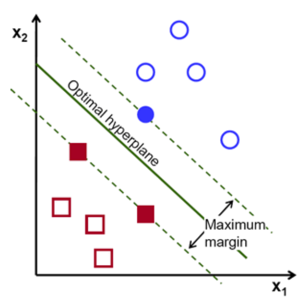


Figure 6: A visualisation of the classification with two types of classes present [59]. In this figure the support vectors in the SVM are present. However, the visualisation is valid for both the LDA and the SVM methods.

4.2 Spatial Filtering

A common component in constructions of BCIs is the usage of spatial filters. The signals in the brain can be compared to the mix of voices in a room full of people that are talking, at a party for instance. Under some presumptions in theory it is possible to use a spatial filter to sort out every unique voice from the mix of sounds. This can be performed quite well for the case when that there are an equal number of microphones as individuals in this room and the mixing matrix is known and is well conditioned. As mentioned are there many similarities to this party and the signals within the brain [60]. Even though all signals in the brain originating from a certain area. The signals are spread to many other parts of the brain. Thus, electrodes on the scalp are retrieving a mix of signals from different areas and it is hard to distinguish each signal from the noise originating from other positions. In other words there is a low SNR in the raw EEG signal in each channel.

There are several algorithms for constructing a spatial filter used in P300 potential based BCIs, xDAWN, Independent Component Analysis (ICA) and Principal Component Analysis (PCA) to mention a few. These algorithms have in common that they attempt to estimate a subspace of the channels or values from the data in each channel which contains the desired signal. In addition to this reducing of important attributes in the data, these algorithms also performs a maximisation of the Signal to Signal plus Noise Ratio (SSNR), i.e. tries to enhance the quality of the estimated signals. In test performed with and without spatial filters on EEG data which contained P300 signals gave those signals filtered with a spatial filter higher accuracy in the classification [61]. This is the reason why this theory is worth to discuss in this paper. Applying a spatial filter before one of the earlier discussed classifiers will result in a subspace of the relevant channels and give faster decisions in addition to lowering the computational cost, also will the resulting accuracy of the classification increase. In the remaining parts of this section a more formal description of ICA and PCA are presented since they have showed promising result in use with P300 ERPs [62, 63]. Also will the final section describe a simpler method similar to spatial filtering that make use of an average signal level.

4.2.1 ICA

In this section a description of independent component analysis is given.The purpose of this algorithm is as mentioned to sort out relevant channels and to minimize noise originating from other sources than the relevant one for P300 ERPs [64, 65]. Independent component analysis is more common in applications build with own code, e.g. in MATLAB and Python or C, however also tools like openVibe enable its use.

To define ICA we make use of a statistical model, ”latent variables”. By assuming that it exists a observation of linear mixtures $x_1, ..., x_n$ of n independent components

$$x_j = a_{j1}s_1 + a_{j2}s_2 + ... + a_{jn}s_n, \text{ for all } j \tag{1}$$

The time index t, which normally are used here, is dropped and in the ICA model it is assumed that each signal mixture x_j as well as each independent component s_k is a random variable instead of a time signal. If we refer to the problem with voices in a party, mentioned in Section 4.2, are the observed microphone signals, $x_j(t)$, then a sample of this random variable. An assumption is that both x_j and s_k have zero mean. However, if this would be false it is possible to centre the observable variables, x_i for $i = 1, 2..., n$, by subtracting the sample mean, which alter the model to be zero mean.

For convenience a matrix notation will be used. By denoting the mixtures $x_1, ..., x_n$ with $\mathbf{x}$ and likewise make $\mathbf{s}$ denote the vector with elements $s_1, ..., s_n$. In the same manner we denote by $\mathbf{A}$ the matrix with elements a_{ij}, in other word are the general notation, with matrices as bold uppercase and vectors denoted as bold lowercase, held. In the following description are all vectors considered as column vectors, i.e. a transpose of a vector $\mathbf{x}$ as $\mathbf{x}^T$ is a row vector. This new form of notation make it possible to write (1) as

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{2}$$

We denote the columns in $\mathbf{A}$ as $\mathbf{a}_i$ and equation 2 is equal to

$$\mathbf{x} = \sum_{i=1}^{n} \mathbf{a}_i \mathbf{s}_i \tag{3}$$

The statistical model in equation 2 is known as the ICA model. This model is a generative model, in other words it describes how the process of mixing the components $\mathbf{s}_i$ generates the observed data. The reason to why "latent variables" are mentioned in the beginning of this section is that the independent components are latent variables and, i.e. they cannot be observed directly. In fact, also the mixing matrix, $\mathbf{A}$, is assumed to be unknown. The only observation is the vector $\mathbf{x}$ and both $\mathbf{A}$ and $\mathbf{s}$ is required to be estimated to be used. This estimation must be performed under as general assumptions as possible.

A very simple assumption is the starting point for ICA, and this assumption is that the components $\mathbf{s}_i$ are statistically independent, for further description of this assumption see Section 3 in the work of A. Hyvärinen et. al [64]. Below it will be revealed that also must make an assumption that the independent components must have non-Gaussion distribution. Yet, in the basic model is no such assumption made, instead are the distributions considered unknown otherwise the problem would be considerably simplified. However, we make one simplification by assuming the unknown mixing matrix to be square although this assumption can sometimes be altered. Nevertheless, when the estimation of the matrix $\mathbf{A}$ is completed it is possible to compute its inverse, denoted as $\mathbf{W}$, and obtain the independent component as

$$\mathbf{s} = \mathbf{W}\mathbf{x} \tag{4}$$

There are several alternatives to determine the matrix $\mathbf{W}$. Regardless of method, ICA aims to maximize the non-Gausian attributes of the source, i.e. to minimize the mutual information. If the variables do not follow a Gaussian distribution, ICA cannot separate them [24].

In various applications a noise term is added to the model since it is more realistic to assume that the measurements include a great amount of noise. For simplicity the noise term is omitted in this description. Such an alteration might be included in the filter, even though many applications seems to be sufficient without it [64].

ICA is very closely related to a method known as *blind source separation* (BSS). In BSS the word "source" refer to the original signal, i.e. independent components, like the individual that speaks at the party with mixed voices mentioned in Section 4.2. The word "Blind" means that there is very little known about the mixing matrix, which containing all the voices. And this is something that correlates very well with the conditions in a BCI and ICA are one of the most widely used filter for BSS which makes it a good candidate to be used in the system described in this report.

4.2.2 PCA

A method similar to ICA is the earlier mentioned Principal Component Analysis (PCA). This is a technique used to transform a number of variables, with a potential correlation, into a smaller number of variables denoted as principal components. In a multichannel BCI it would be possible to analyse the data from a number of channels within a certain time span and use PCA to get a single vector containing the most important components. Then it would be possible to perform a classification of this vector, in this section the theory behind PCA is given [66].

The main idea of PCA could be stated in one single sentence: *Is it possible to obtain a new basis that is a linear combination of the original one which re-expresses the data in an optimal way?*. The refereed ”optimal way” in this sentence is when the elements in the original data, presented in the new basis, become sorted in order of importance. This can be stated in the following mathematical manners:

Assuming that the original data set is represented in terms of an $m \times n$ matrix, denoted as $\mathbf{X}$. In $\mathbf{X}$ the n columns are the observations, i.e. the samples, and the m rows are the variables, i.e. the channels. The objective in a PCA is to linearly transform the matrix $\mathbf{X}$ into a new matrix $\mathbf{Y}$ of the same dimensions and that fulfils the following criteria

$$\mathbf{Y} = \mathbf{PX} \tag{5}$$

for some matrix $\mathbf{P}$ with the dimension $m \times m$ and which sort the data in $\mathbf{Y}$ in the optimal way.

Equation 5 represent the change of basis. If the rows $\mathbf{P}$ are considered as row vectors, denoted as $\mathbf{p_1}, \mathbf{p_2}, \mathbf{p_3}, \ldots, \mathbf{p_m}$, and the columns of $\mathbf{X}$ to be column vectors denoted as $\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}, \ldots, \mathbf{x_n}$ the equation 5 can be represented in the following manner

$$\mathbf{PX} = (\mathbf{Px_1}, \mathbf{Px_2}, \mathbf{Px_3}, \cdots, \mathbf{Px_n}) = \begin{pmatrix} \mathbf{p_1} \bullet \mathbf{x_1} & \mathbf{p_1} \bullet \mathbf{x_2} & \cdots & \mathbf{p_1} \bullet \mathbf{x_n} \\ \mathbf{p_2} \bullet \mathbf{x_1} & \mathbf{p_2} \bullet \mathbf{x_2} & \cdots & \mathbf{p_2} \bullet \mathbf{x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{p_m} \bullet \mathbf{x_1} & \mathbf{p_m} \bullet \mathbf{x_2} & \cdots & \mathbf{p_m} \bullet \mathbf{x_n} \end{pmatrix} = \mathbf{Y}$$

In equation 4.2.2 are $\mathbf{p_i}, \mathbf{x_j} \in \Re^m$ and from this it is clear that $\mathbf{p_i} \bullet \mathbf{x_j}$ is the normal Euclidian inner product, i.e. a dot product. From this the conclusion can be drawn that the original data, matrix $\mathbf{X}$, is projected onto the columns of $\mathbf{P}$. Thus, the new basis for representing the columns of $\mathbf{X}$ are the rows of $\mathbf{P}$. These rows will also be the directions of the searched principal components.

As a next step the issue of defining the new basis must be addressed. Principal component analysis defines independence's between the components by considering the original variance, i.e. the variance in the original data $\mathbf{X}$. The technique strives to de-correlate the original data by maximizing the variance. This maximization is performed by altering the directions and the representation which gives maximum variance is then used to define the new basis.

However, this does not reveal how the PCA algorithm solve the original problem in equation 5 and find s the optimal new basis. The answer is that it uses the covariance matrix of $\mathbf{Y}$ denoted as $\mathbf{C_Y}$. Since $\mathbf{X}$ is consisting of m variables and n samples it is not difficult to think about $\mathbf{X}$ in terms of m row vectors of length n as

$$\mathbf{X} = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m,1} & x_{m,2} & \cdots & x_{m,n} \end{pmatrix} = \begin{pmatrix} \mathbf{x_1} \\ \mathbf{x_2} \\ \vdots \\ \mathbf{x_m} \end{pmatrix} \in \Re^{m \times n} \qquad , \mathbf{x_i^T} \in \Re^n$$

Since there exist a row vector for each variable which contains all samples for the particular variable the sample covariance matrix $\mathbf{C_X}$ can be defined as

$$\mathbf{C_X} = \frac{1}{n-1}\mathbf{X}\mathbf{X}^T = \frac{1}{n-1}\begin{pmatrix} \mathbf{x_1}\mathbf{x_1}^T & \mathbf{x_1}\mathbf{x_2}^T & \cdots & \mathbf{x_1}\mathbf{x_m}^T \\ \mathbf{x_2}\mathbf{x_1}^T & \mathbf{x_2}\mathbf{x_2}^T & \cdots & \mathbf{x_2}\mathbf{x_m}^T \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x_m}\mathbf{x_1}^T & \mathbf{x_m}\mathbf{x_2}^T & \cdots & \mathbf{x_m}\mathbf{x_m}^T \end{pmatrix} \in \Re^{m\times m}$$

Covariance can be considered as a measurement of the correlation of two variables [67]. The PCA method is built on assumption that the variables in $\mathbf{Y}$ should be as uncorrelated as possible. In the theory of covariance this is equivalent to say that variables in $\mathbf{C_Y}$ should be close to zero since covariance matrices are always positive definite or semi-definite. As mentioned before large variances are important and therefore two demands can be stated on the matrix $\mathbf{C_Y}$

- Maximize the variance, i.e. maximize the diagonal entries.
- Minimize the covariance between variables, i.e. minimize the off-diagonal entries.

Hence, it is clear that the sought matrix $\mathbf{C_Y}$ is diagonal since the minimum possible covariance is zero. In other word if we choose the transformation matrix $\mathbf{P}$ in such a manner that $\mathbf{C_Y}$ is diagonal will the main objective be achieved.

From this point another assumption is to be made, the vectors in the new basis, $\mathbf{p_1}, \mathbf{p_2}, \ldots, \mathbf{p_m}$, are orthogonal and in addition to this they are also orthonormal. With this in mind and by considering the formula for the covariance matrix, $\mathbf{C_Y}$ in equation 4.2.2, and the interpretation of $\mathbf{Y}$ as terms of $\mathbf{X}$ and $\mathbf{P}$ in equation 4.2.2 $\mathbf{C_Y}$ can be expressed as

$$\mathbf{C_Y} = \frac{1}{n-1}\mathbf{Y}\mathbf{Y}^T = \frac{1}{n-1}\mathbf{P}\mathbf{X}\mathbf{P}\mathbf{X}^T = \frac{1}{n-1}(\mathbf{P}\mathbf{X})(\mathbf{X}^T\mathbf{P}^T) = \frac{1}{n-1}\mathbf{P}(\mathbf{X}\mathbf{X}^T)\mathbf{P}^T \tag{6}$$

This can be expressed as

$$\mathbf{C_Y} = \frac{1}{n-1}\mathbf{P}\mathbf{S}\mathbf{P}^T \qquad \text{where} \qquad \mathbf{S} = \mathbf{X}\mathbf{X}^T \tag{7}$$

Here it is worth to distress the fact that $\mathbf{S}$ is a symmetric matrix with dimension $m \times m$. Since every square symmetric matrix is orthogonally diagonalisable. Thus, it is possible to write $\mathbf{S} = \mathbf{E}\mathbf{D}\mathbf{E}^T$ where $\mathbf{E}$ is an $m \times m$ orthonormal matrix with columns equal to the orthonormal eigenvectors of $\mathbf{S}$. The matrix $\mathbf{D}$ is a diagonal matrix and the diagonal entries are the eigenvalues of $\mathbf{S}$

Here is a crucial choice for the transformation matrix $\mathbf{P}$ made. By selecting the eigenvectors of $\mathbf{S}$ as rows in $\mathbf{S}$ the following equality holds, $\mathbf{P} = \mathbf{E}^T$. Thus can a new expression of the covariance matrix be determined as

$$\mathbf{C_Y} = \frac{1}{n-1}\mathbf{P}\mathbf{S}\mathbf{P}^T = \frac{1}{n-1}\mathbf{E}^T(\mathbf{E}\mathbf{D}\mathbf{E}^T)\mathbf{E}$$

Since $\mathbf{E}$ is an orthonormal matrix is $\mathbf{E}^T\mathbf{E} = \mathbf{I}$, where $\mathbf{I}$ is the $m \times m$ identity matrix. Hence will a special case for the choice of $\mathbf{P}$ result in

$$\mathbf{C_Y} = \frac{1}{n-1}\mathbf{D}$$

In the described method have an automated sorting process of the importance of the principal components based on variance been performed. The largest variance corresponds to the leading principal component and the second largest to the second principal component etc. This is used to sort the data in the diagonalisation stage, when the eigenvalues of $\mathbf{S} = \mathbf{X}\mathbf{X}^T$ are obtained they are sorted in descending order and placed in the diagonal of $\mathbf{D}$. The same sorting procedure is used to place the columns in $\mathbf{E}$.

Hence have the objective been fulfilled and the diagonalising of the covariance matrix of the transformed data are completed. The rows of $\mathbf{P}$ are the principal components and the eigenvectors of matrix $\mathbf{X}\mathbf{X}^T$. Since the rows are in order of importance they give the information of how "principal" each principal component is. Then the most important information can be passed to the classifier by selecting a subspace from the first part of the principal components.

The above given theory of PCA are based on eigenvector decomposition however there is another possibility of solving a PCA problem based on method known as singular value decomposing (SVD). In fact the two methods are so closely related that it is hard to separate the one from the other, the theory of SVD is not given here since the PCA algorithm used in this project is based in the eigenvector decomposition but the authors recommend interested readers to deepen there knowledge in the work of Jonathon Shlens, " A Tutorial on Principal Component Analysis" [68].

The method described in this section is mainly a series of linear algebra operations and this is useful in real time applications due to the shorter computational time compared to non-linear methods. The resulting vector of principal components can, as mentioned, be input to a classifier and therefore is this method a good candidate to be used for improved robustness in the decision-making process of a BCI. However, if the resulting vector is fed into a classifier directly nothing is achieved by using this method. Since the PCA method only sort the values according to relevance in the original vector, the classifier would not notice any difference since only the order of the values is changed. To solve this, only a subspace of the resulting vectors of each the values in each of the relevant channels is used.

4.2.3 Grand Averaging

In above sections two rather advanced methods for simplifying the data stream before classification have been described. Within this section a third method are presented, the use of a grand mean value. The method, called as grand averaging (AVG), is quite simple but has showed great potential in other projects [24].

The method can be described as removing the in all channels common ground level of the signals from each individual channel. In other words by calculating the mean value of the complete data set collected after one illumination of an element in the matrix of choices, an epoc, and removing this from each channel. This will have the affect that the signal level in each channel will fluctuate around the level of the mean value of the noise. As a result the impact on the signals of certain brain patterns that occur only in specific areas of the brain, such as a P300 ERP, is enhanced.

More formal can this be written in the following manners. If the matrix $\mathbf{X}$ of size $n \times m$ holds the m values from n channels from a pre-defined time span. If the rows in matrix $\mathbf{A}$ is denoted as $\mathbf{x}_j$ for $j = 1, 2, .., n$ the corresponding mean value of each row, $\bar{\mathbf{x}}_j$ for $j = 1, 2, .., n$, is calculated as

$$\bar{\mathbf{x}}_j = \frac{1}{m} \sum_{i=1}^{m} x_{ji}$$

where x_{ji}, for $i = [1, 2, ..., m]$, is each element in vector $\mathbf{x}_j$. The mean value of the vector $\bar{\mathbf{x}}_j$, a scalar denoted as M, is then given by

$$M = \frac{1}{mn} \sum_i \sum_j x_{i,j}$$

And by subtracting M from each element, x_{ji}, in $\mathbf{A}$ the mutual signal level is removed

$$\mathbf{D}_{i,j} = x_{1,j} - M$$

Where $\mathbf{D}$ is the matrix which holds the final data with the mean value from the complete data set removed from each element.

5 Method

The method in this project has two parts, the first is how to create and set up the BCI to be used in test and evaluations. The second part is the method to be used in the actual tests and how to secure and evaluate the results from these. This section will follow this division and first a description regarding the preparation and construction will be discussed and the later part will describe the test-protocol.

5.1 Scientific approach and Literature Study

Even though the main focus of this project is in the test of a BCI in a vehicle environment the design of the BCI itself is a requirement to reach this goal. Within this building part there were several components that needed to be considered and a first step was of course to perform a rigorous literature study in the subject. The entire Section 2 can be seen as a brief summary of this study.

This investigation is using a qualitative approach, where research questions are formulated along with hypothesises. These are then evaluated in a number of tests, where results are taken from and answers to the questions are given. Additionally, the study is exploratory and evaluates several concepts and approaches to find a suitable solution with today's technology.

5.2 Design

Since the hardware itself is provided by Emotiv as stated in Section 3.1 the design is mainly focused around the software and the algorithms used within it. Looking into the requirements on the software there are three parts that needs to be developed.

5.3 Software

As stated earlier in Section 5.2 the focus of this project are the development of algorithms and software for BCI control as well as an evaluation of the concepts feasibility as an strategic controller in a car.

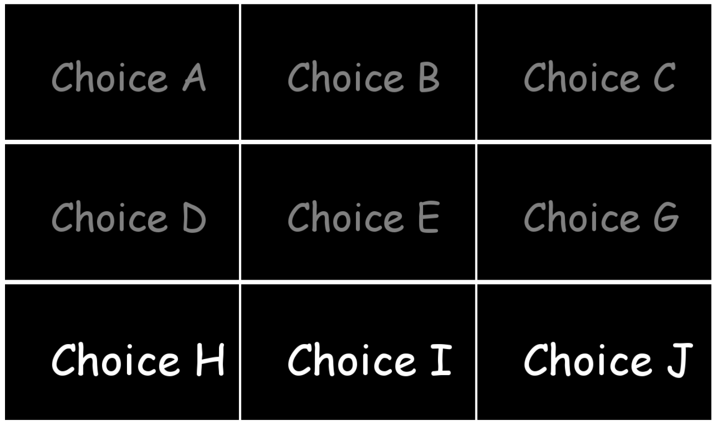


Figure 7: The graphical grid with choices, in this version is the size 3×3 and the bottom row is flashed.

The software is developed to be as simple, effective and scaled-down and should be able to run on an everyday laptop. Although some pre-processing of the signal is done already in the hardware it could prove beneficial to further improve the quality of the data that is fed into the classification algorithms. This is done with the best candidate of spatial filtering and additional digital filters.

From the literature study classification algorithms best suited for the application where chosen as candidates for a system in a car. These are then implemented to suit the specific requirements in an offline analysis program. The developed algorithms are then tested and verified on data from user tests. This validation gave an answer to which algorithm that have the potential to be integrated in a real-time system.

To enable testing of the concept the stimuli that will trigger the ERP is developed. With the P300 concept this includes a grid with blinking rows and columns, see Figure 7. Simultaneously are a data collecting part executed and values are saved and used in the analysis phase.

The software implements these concepts on screens of two different sizes that will be used in the system. This is to enhance the possibility that the screen size matter when it comes to performance of the system. Also it is possible to change the size of the grid, in Figure 7 are the size 3×3.

5.4 Tests and evaluation criteria

The developed system are rigorously tested using a standardised test procedure. Because of the specific application safety is of out-most concern and together with the time to decision the top-priority. Since disturbances not are as crucial due to the nature of the autonomous driving, the safety aspect is reduced to the level that the system must be harmless to wear and use. Yet, the time to execute a command is important in autonomous cars as well as a car of ordinary type. Apart from these points the system must be reliable and user-friendly, which puts requirements on a small calibration time as well as the need for the user to learn the system.

5.4.1 Test Scheme

The tests are performed in two different sets or environments, office sets and car sets. The tests are performed for all of the parameters that can be chosen, manipulating only one at a time, in the test. This to achieve the correct causality in the results; the different parameters are listed below.

- 2 environments (office and car)
- 2 screen sizes (7" and 15")
- 2 size of matrices (2x2 and 3x3)
- 6 test persons
- 2 classification algorithms (LDA and SVM)
- 3 types of spatial filters (ICA, PCA & Grand Averaging)

The first 3 points in the list above are altered manually between the test persons, the last two are altered in the analysis of the data performed after the tests. In all test types the headset is turned 180 degrees from its normal heading, i.e. it is turned backwards. The reason to this is in the theory regarding the location of the origin of P300 ERPs, presented in Section 2.1. According to this the areas in the back of the head are of a higher importance in this type of applications and the used hardware will cover more of these areas when it used in this manner, the configuration of channels are presented in Section 3.1 and by relocating these position with 180 degrees rotation in Figure 1 it can be confirmed that several of the important positions from theory get covered in this new set-up.

Office tests

At first, tests are performed in a semi-controlled environment, without any feedback to the user, where every row and column blinks a fixed amount of times. The user is instructed to focus on one of the choices during the test. These tests can be seen as offline since the data is collected and analysed at a later time.

1. Introductory tests where one parameter varies per test.
2. Since no calibration is needed, tests can be separated into two sessions.
3. Approximately two weeks later the test person returns and performs two tests of one of the test types once more. These tests still do not give any feedback to the user, but will be analysed with a calibration based on data from the previous test of the same type.

In Section 2.1.1 it is mentioned that the duration of ISI in a P300 speller should not be too long and 175 ms seems to be a good value. This will also be used in these tests and therefore the test time can be approximated to last for 70.87 seconds for one choice in 3×3-matrix and a pre-defined number of flashes set to 45, $45 \cdot 0.175 \cdot 3 \cdot 3 = 70.87$. Only one parameter is changed between tests and there will be 4 individual test per user. However to give more statistical data every test are performed 2 times. From this will the total effective test time be closer to 4 hours for the entire test group.

The data from these tests are analysed in MATLAB and classification is performed with as little data as possible needed while maintaining a critical of certainty. As always, the certainty level will be chosen from the analysis using graphs and statistical methods to evaluate the relationship between data/time needed and the demand on certainty. These tests also give information regarding whether new calibration is necessary for each new use of the system or if user data can be saved over a period of time.

Car tests

The car tests will only take place in the target environment, the car, and is using the same set up as the office tests but only the 9.7-inch screen. The data of the implemented system is recorded and the analysis will be performed in the same manner as in the test for office and the results will be compared.

While performing the tests the user is placed in the back seat of the car, focusing on the screen while a driver puts the car into movement.

From the data recorded in these tests an evaluation of the optimal performance of the designed system in each of the four areas mentioned in Section 1.4 will be performed. This will lead to final conclusion if the system has the potential of achieving well, or if some parameters are in need of improvement before such a system is integrated in a vehicle and fully used.

5.4.2 Test Protocol

This section will describe the precise protocol used when performing the tests and their prerequisites.

The number of different test mentioned earlier will be listed and then performed in a rolling scheme for minimum bias. In other words test person 1 will start with test number 1, test person 2 will start with test number 2 and end with number 1.

Since the system is meant to be used in an ordinary environment setting, the environment parameters can not be controlled. Based on this it will suffice to keep the test in an office-like environment as well as in a car and compare the two, without closer measuring of light levels, sound, etc.

Test Group & Parameters:

There will be a total of 6 test subjects. The subjects will consist of a group of people, all being 20-30 years old, well educated and interested in the subject. Although, there will be several nationalities represented, as well as people from both genders. Having 6 different test subjects will result in a relatively general description of the population.

As mentioned earlier, the control of environment variables will be limited as it would be in a real user case. Apart from that, the variation in the variables will not have a large impact on the test results [30].

The testing procedure for the two environments has been as follows:

Office tests:

- Test subject, including relevant information and specific test is noted
- Headset is put on and adequate signal quality is achieved
- Data transmission is on and gives reasonable signals
- Show test program and instruct the test person accordingly what to do
- Start recording to a specific test data file, both data and blink file
- Perform the test, without feedback, note start and stop time of test
- Check that data transmission and signal quality is still reasonable
- Save the data and end the test session
- Ask questions about the user experience of the test and take notes

Car tests:

- Test subject, including relevant information and specific test is noted
- Headset is put on and adequate signal quality is achieved
- Data transmission is on and gives reasonable signals
- Show test program and instruct the test person accordingly what to do
- A driver set the car in motion and perform all task related to the the driving (test subject is seated in the backseat).
- Start recording to a specific test data file, both data and blink file
- Perform the test, without feedback, note start and stop time of test
- Check that data transmission and signal quality is still reasonable
- Save the data and end the test session
- Ask questions about the user experience of the test and take notes

If the test results are corrupt or the procedure was not followed as specified the test will be aborted and the results deleted.

From the results from these two sets of tests, theoretical performance and feasibility of the methods, based on this particular set up and hardware, will be assessed. The next step will be to implement a system accordingly to the result from the investigation regarding the methods and to evaluate how well such a system can perform.

6 Results and Analysis

From the analysis of the data retrieved in the tests described in the previous section (section 5.4) there are several results. These are presented later in this section in various forms, graphical, tables etc. and the authors want to emphasise the fact that they are only valid for this particular system set up. Hence, if a method is concerned to be the ”best” in some sense this is only valid for this unique system. As initial processing of the data are a digital Butterworth filter of second order applied to the signal with cutoff frequencies of 0.3Hz and 20Hz, this is since in theory are P300 ERP present below 16Hz, see Section 2.1. The mean value of each channel are also removed from the values, to make the signals to fluctuate around zero and the data are also divided into pieces, or batches, corresponding to 0.5 seconds of data starting when an illumination of an element occurred in the test. These batches are then possible to label with the corresponding illuminated element in the matrix and save to build a mean value of several batches. This initial phase of data processing are executed in all investigations within this chapter.

In Figure 8 all values from the first test, large screen in office environment and with a 3×3 matrix, are included for all users and the data stream are divided into the earlier mentioned batches of data, i.e. each piece represent values equal of 0.5 seconds. These 0.5 seconds starts when an element in the matrix of choices is illuminated, three sub-spaces of this data are created according to following types:

- Horizontal hits, i.e. data that represent a horizontal flash of target in focus.
- Vertical hits, i.e. data that represent a vertical flash of target in focus.
- Non-Hits, also denoted as false alarms and are data which represent flashes from all other targets than the one in focus of the user.

Each of the pieces of the same type are then added to each other and the values from one channel is chosen to be used. This results in a vector for each type and this is displayed in Figure 8 after the mean value from the vector is removed from each element. The chosen channel as the data stream is mounted at position P1 in Figure 1, channel AF3 in a headset used ”backwards” as described in Section 5.4.1. This is from the theory in Section 2.1 a good channel and it is also confirmed by the statistical methods for both the cases presented later in Section 6.1.4 to be the best channel to use in the grand average (AVG) method, described in Section 4.2.3, with 97% respectively 96% certainty depending on which of the attributes, presented in Section 6.1.4, that are in focus for the statistical measurement. The AVG method is similar to the method used to retrieve this result and therefore the same channel is used here for simplicity. The most interesting part in this result occurs between approximately 0.325 to 0.375 seconds, where the P300 ERP should influence the signal in theory, in Figure 8 it is clear that this is expressed as a higher voltage for the two types which represent an illumination of a choice in focus of the user. This can be seen as a validation that the system works properly.

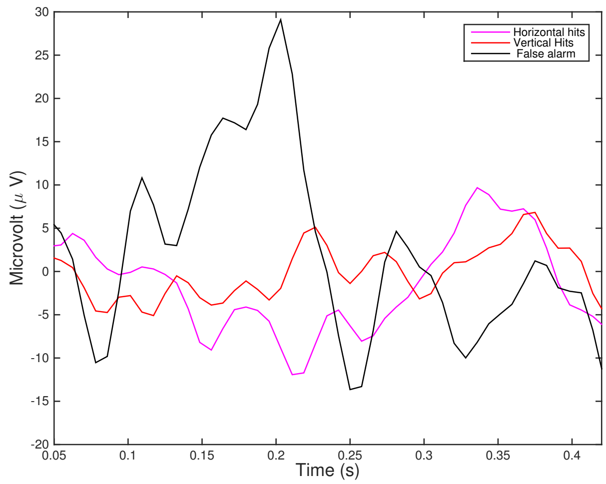


Figure 8: The complete data set from one test cut into pieces equal to 0.5 seconds after each type of flash (horizontal-, vertical- and non-hits). These pieces are then combined within the same type and a mean value is calculated. As can be seen the voltage level around 0.325 to 0.375 seconds is interesting since the false alarm are below both of the two types of hits.

6.1 Evaluation between test parameters

Within this section the first result from the investigation regarding the possible combination of algorithms are displayed. As the most crucial aspect of this system is to capture the occurrence of P300 ERPs, denoted as hits, and classify those in an efficient manner. However, the number of false classification or false alarms are actual something even more important in this application than to classify all occurrences of P300 correct since a false alarm would be devastating in a strategic controller. Thereby the result in this section will focus on the comparison between the performance, e.g. relation between correct and non-correct classifications, of both external (e.g. window size) and internal variables (e.g. spatial filtering) in combination with the question if the P300 ERPs are captured well or not. In the figures within this section the data from all test persons have been taken into account and the mean value from the resulting individual classifications from six users is used.

Within this investigation the most promising candidates of spatial filtering are combined with a classifier. To enable an investigation of the classifiers the resulting number of correct or non-correct (false-alarms) classifications from an uniquely trained classifier has been used for each test person and type of test. This classifier is trained with 85% of the data from each test and then are the validation based on the correct classification of the remaining ,for the classifier unknown, 15% of data. This procedure is used for all validation regarding the classification to ensure objective tests. Before the spatial filters, as mentioned in previous section, a digital filter of Butterworth type of the second order is applied with cutoff frequencies at 0.3Hz and 20Hz. The data are also divided in combination with the removal of the mean from each channel in the same manners as earlier described. In the grand average method (AVG) channel AF3 is used, position AF3 in the original channel configuration, corresponding to position P1 with the set up used in the tests. In the methods ICA and PCA a subspace of 4 channels are used, they correspond to the positions, P1, P2, PO3, P04. These have been confirmed by the statistically methods as the best to use in in the current configuration of the hardware.

6.1.1 Classification Procedure

The classification procedure can be described in several steps described below:

1. Data corresponding to 0.5 seconds, 64 values, is pre-processed according to earlier described filtering process, i.e. a Butterworth filter is used in combination with a removal of the mean value.
2. A selection of the channels is chosen for each method of spatial filter, e.g. in ICA and PCA a subspace corresponding to the positions, P1, P2, PO3, P04 is used.
3. The data from selected channels is processed in the spatial filter in the following manner:
   - ICA: The four channels are reduced to one, this is done automatically in the process. This vector is then the output from ICA.
   - PCA: The data in four channels are sorted in the process and the 16 most important values are chosen from each channels and put in one vector as output.
   - AVG: The data in channel AF3 is processed and is put in the output vector.
4. The output vector from the spatial filter is sent to the classifier and a output is given. This output is binary, a 1 if the classifier classify it as a P300 present and a 0 if a P300 is considered to not be present in the signal.

The above description can be performed for data set consisting of data from several batches as well. If several batches of data is included, the data corresponding to 0.5 seconds after several illuminations of a certain element in the decision matrix is added to each other and the the average is calculated. For example, if 5 batches of data are included step 1 is performed on each batch and the values in each channel are added to each other and divided by 5 before the newly created set is sent to step 2 and the process continues. A classification is then made with a new type of data set, based on several illuminations, which have the same size as if only one illumination is included.

Regardless if one or several batches is included, the classification has four possible outcomes:

1. Correct element and a classification output 1, i.e. a correct classification.
2. Correct element and a classification output 0, i.e. a missed classification.
3. Non-correct element and a classification output 1, i.e. a false classification (a false alarm).
4. Non-correct and a classification output 0, i.e. a a correct classification of a non correct choice.

6.1.2 Classification result for office environment

A natural start of this section is the result from the environment where the system will perform in best possible manner, the office. Even though the most interesting and important result will come from the test in the car environment the relevance of the results in the beginning of this chapter are high since they will display the maximum performance of the system in an uncontrolled environment.

To enable a good and visual result figures with histograms are used, Figures 9-32. In the figures the percentage of correct classifications are displayed as well as the percentage of false alarms for each and every method. To enhance the difference between correct and false classifications are, as mentioned in Section 4.2.3, several flashes used to build a mean-value, the number of flashes used are the values on the x-axis in the figures. Since one of the important aspects mentioned in the objective in Section 1.4 is the amount of time a decision take are the maximum number of flashes to build up an average value constrained to a maximum of 10. This gives a maximum theoretical time of 15.75 seconds to make a decision by using the system, regardless of the internal method of classification.

Since there could be a difference in the interpretation within the brain between a flash of a column compared to a row are the result also divided into in vertical and horizontal flashes. In all figures the green colours are used in the bars which represent the correct classifications and the red colours are of course the false alarms. The height (or the values on the y-axis in all figures) are as mentioned percentage in both cases.

In Figures 9-16 the large type of screen is used and this should provide the, in theory, best result of all test types. In Figures 9 and 10 are the LDA classification used, in Figure 11 and in Figure 12 are the SVM type of classifier used.

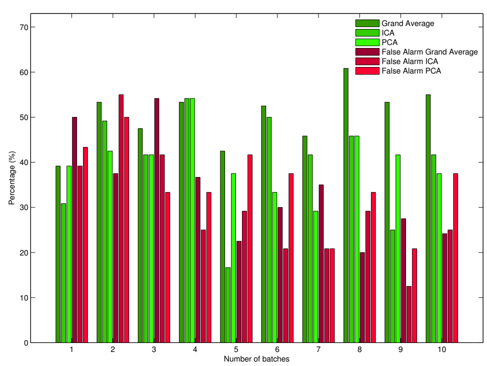


Figure 9: Horizontal result with LDA classifier with large screen and 3×3 possible choices in an office environment.

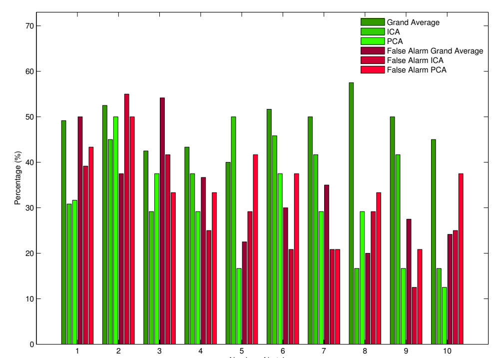


Figure 10: Vertical result with LDA classifier with large screen and 3×3 possible choices in an office environment.

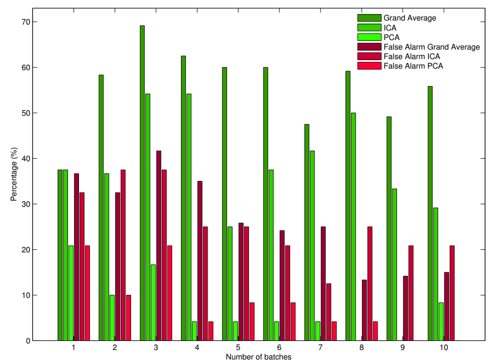


Figure 11: Horizontal result with SVM classifier with large screen and 3×3 possible choices in an office environment.

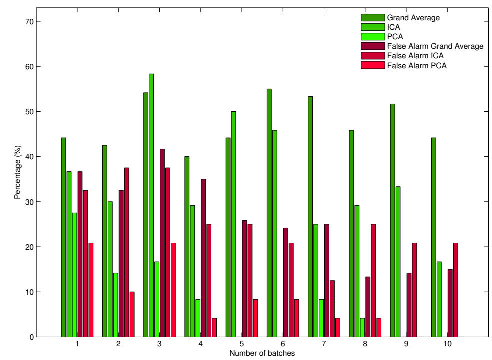


Figure 12: Vertical result with SVM classifier with large screen and 3×3 possible choices in an office environment.

In the result presented in Figures 13-16 are the same type of screen used as in the result from the previous histograms but the size of the matrix is changed to 2×2. All other aspects of the figures are in the same manners as earlier.

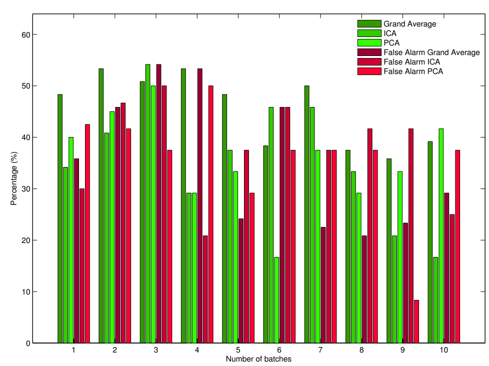


Figure 13: Horizontal Result with LDA classifier with large screen and 2×2 possible choices in an office environment.

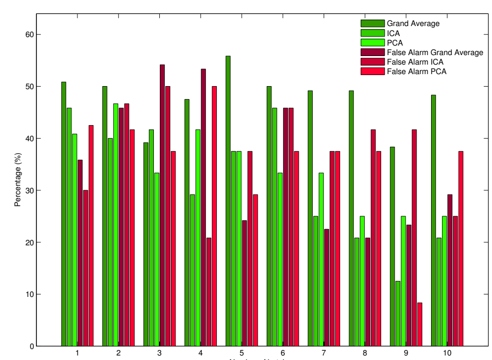


Figure 14: Vertical Result with LDA classifier with large screen and 2×2 possible choices in an office environment.

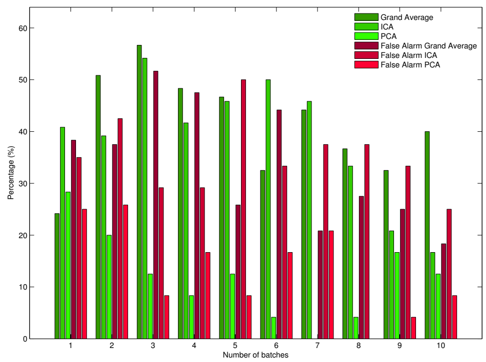


Figure 15: Horizontal Result with SVM classifier with large screen and 2×2 possible choices in an office environment.

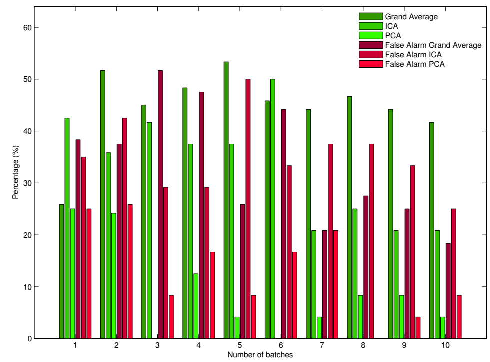


Figure 16: Vertical Result with SVM classifier with large screen and 2×2 possible choices in an office environment.

In the same manner as earlier the result for the tests are performed on the smaller screen displayed in Figures 17-24. As well in this case the result is divided into two types of sizes of the matrix, in types of classifiers and in horizontal and vertical flashes. The first four Figures, 17-20 present the result from test with a matrix size of 3×3 and in Figures 21-24 are the size 2×2 for the matrix used in the tests.

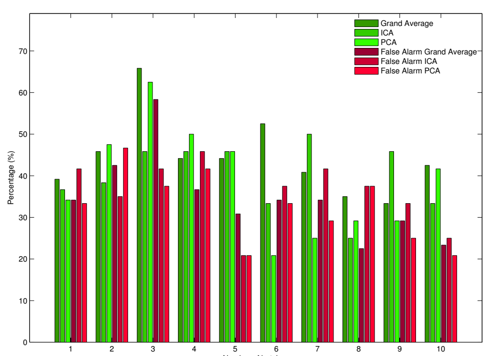


Figure 17: Horizontal Result with LDA classifier from test with small screen and 3×3 possible choices in an office environment.

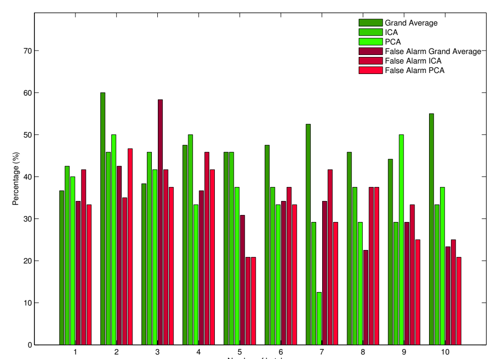


Figure 18: Vertical Result with LDA classifier with small screen and 3×3 possible choices in an office environment.

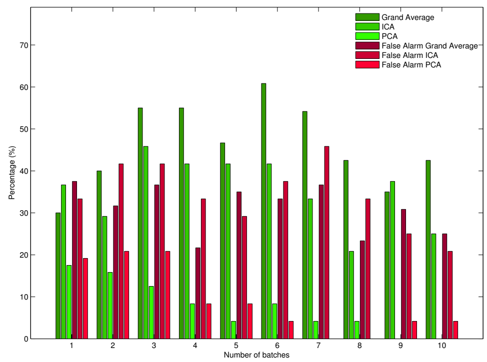


Figure 19: Horizontal Result with SVM classifier with small screen and 3×3 possible choices in an office environment.

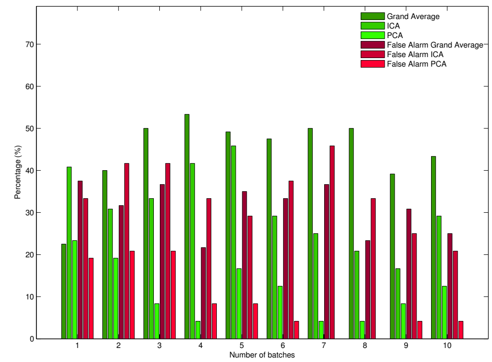


Figure 20: Vertical Result with SVM classifier with small screen and 3×3 possible choices in an office environment.

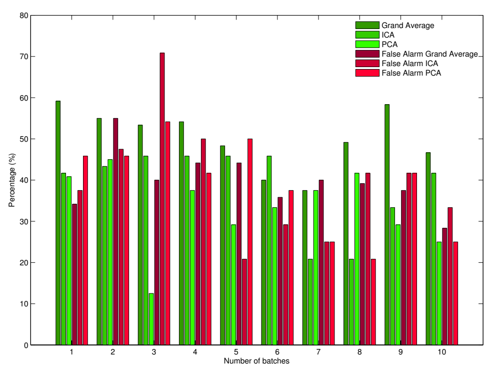


Figure 21: Horizontal Result with LDA classifier with small screen and 2×2 possible choices in an office environment.

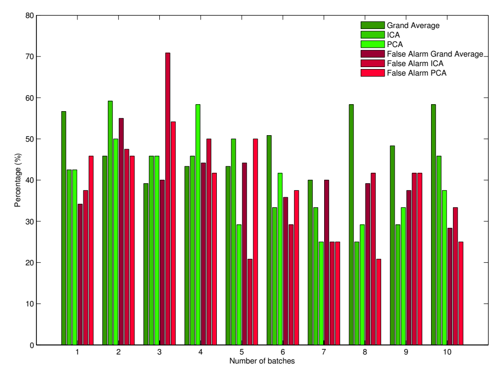


Figure 22: Vertical Result with LDA classifier with small screen and 2×2 possible choices in an office environment.

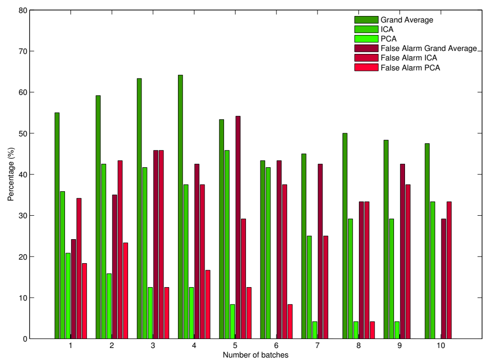


Figure 23: Horizontal Result with SVM classifier with small screen and 2 × 2 possible choices in an office environment.

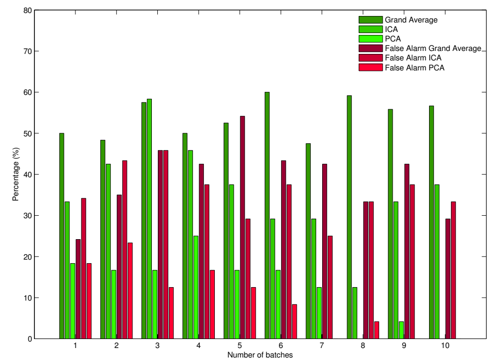


Figure 24: Vertical Result with SVM classifier with small screen and 2×2 possible choices in an office environment.

6.1.3 Classification result for vehicle environment

In this section is, from the main objective point of view, the most interesting result presented. As in the previous section will histograms be used to display the difference of performance between the investigated algorithms. Also, in this environment the horizontal and vertical flashes are separated for the choice in the grid where the user has his or her focus, furthermore the same approach is used regarding the presentation of the type of classifiers and matrix size. However, in this environment only one type of screen has been used. Since the system has constraints regarding physical size for use in a car only the smaller screen is considered and by this are also the result presented here fewer than in the previous section.

In Figure 25-28 the results are presented for the various algorithms, with all data from tests where a matrix of size 3 × 3 has been used in a car.

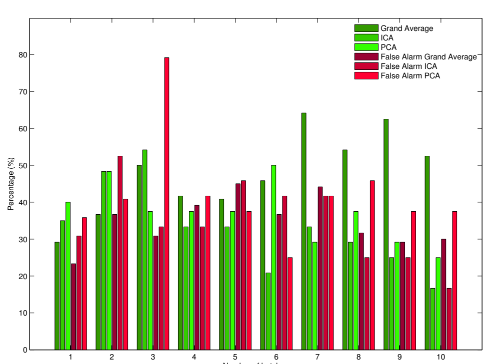


Figure 25: Horizontal Result with LDA from test with small screen and 3×3 possible choices in an car environment.

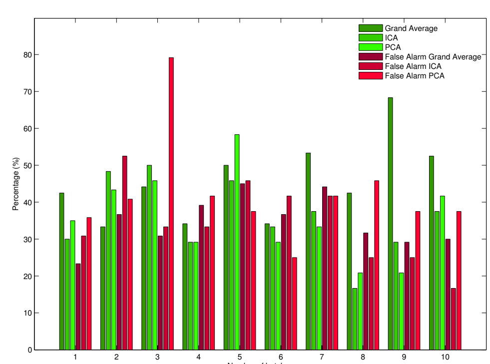


Figure 26: Vertical Result with LDA from test with small screen and 3×3 possible choices in an car environment.

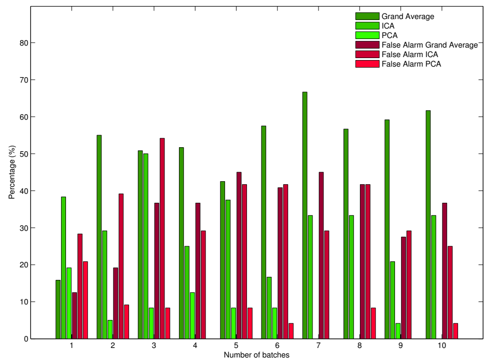


Figure 27: Horizontal Result with SVM from test with small screen and 3×3 possible choices in an car environment.

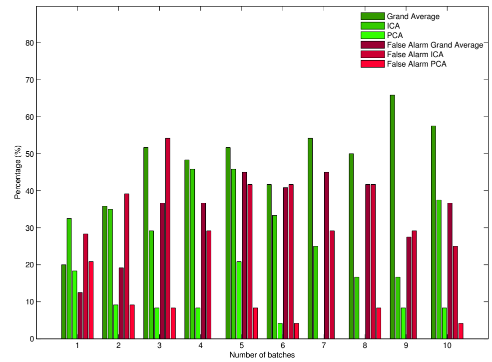


Figure 28: Vertical Result with SVM from test with small screen and 3×3 possible choices in an car environment.

In Figure 29-32 the results are presented for the various algorithms with all data from tests, where a matrix of size 2×2 has been used in a car.

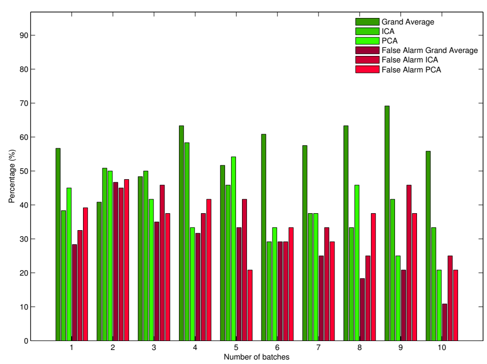


Figure 29: Horizontal Result with LDA from test with small screen and 2×2 possible choices in an car environment.

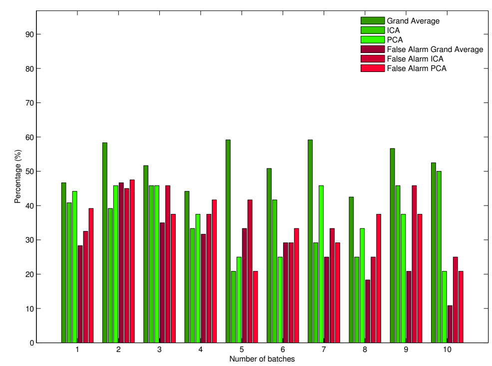


Figure 30: Vertical Result with LDA from test with small screen and 2×2 possible choices in an car environment.

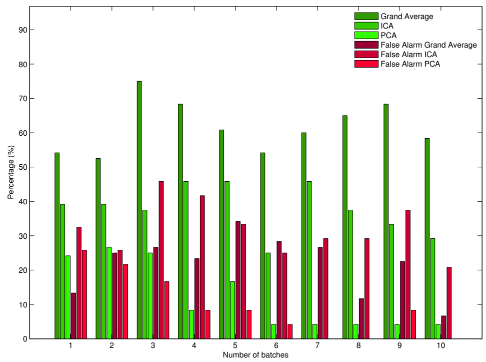


Figure 31: Horizontal Result with SVM from test with small screen and 2×2 possible choices in an car environment.

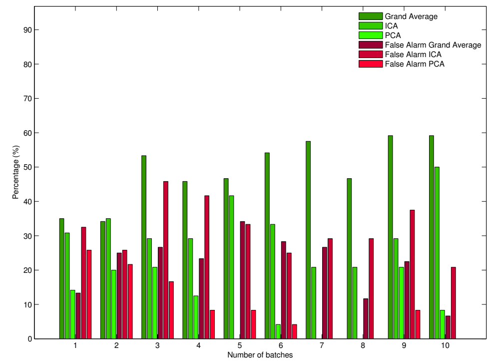


Figure 32: Vertical Result with SVM from test with small screen and 2×2 possible choices in an car environment.

6.1.4 Statistical Comparisons

As the result presented in the previous Section (section 6.1) mainly consists of graphical display they easily show trends etc. However, it may be hard to analyse the result in a proper manner from these figures. In this section the various test results are compared by the method known as a paired t-test or hypothesis evaluation [69]. This technique are used to compare two data sets to each other and will give a result of they are equal or of one has a higher mean. As mentioned in Section 6.1 false alarms are devastating for this system and thereby will the statistical comparisons of the variables for the system be centred around the number of false alarm that occurs.

However, for the scope and setting of this project a high difference between correct and non-correct classifications, which decreases the impact of false alarms, is a suitable parameter for validating the possible methods.

In this investigation the values from each algorithm are the ten mean values in the figures from the previous section. As mentioned earlier this is considered a good measurement for validation of performance and a larger difference will make it easier for a potential real-time system to take a decision, this is further investigated in Section 6.2. In the test in this section the initial hypothesis are two methods considered to be equal regarding the attribute tested and if this is disproved a new hypothesis are stated, this time are one of the data sets considered to have a more desirable form of the searched attribute, i.e. a higher difference between correct and non correct classifications.

In the test has the result from both of the investigated classification algorithms been compared for each filtering method and the best of these are later compared, best in the sense of most desirable of the two when the difference of false alarm and correct classification are considered. These comparisons are executed for both vertical and horizontal tests and the method with highest performance from each type are in the end compared to enable an evaluation if one is a significant better choice than the other, the result can be viewed in Tables 2 and 3 for horizontal and vertical case respectively.

Test type	Grand Average (AVG)	PCA	ICA	Best Method
Office Environment				
Big Screen 3 × 3	SVM	SVM	LDA	AVG (certainty≥99%)
Big Screen 2 × 2	LDA	SVM	SVM	AVG (certainty≥78%)
Small Screen 3 × 3	SVM	SVM	LDA	AVG (certainty≥92%)
Small Screen 2 × 2	SVM	SVM	SVM	AVG (certainty≥95%)
Car Environment				
Small Screen 3 × 3	SVM	SVM	SVM	AVG (certainty≥99%)
Small Screen 2 × 2	SVM	SVM	SVM	AVG (certainty≥99%)

Table 2: The result from the statistical comparison between the various type of filtering and tests in the horizontal case when a high difference between correct and non-correct classifications are the considered attribute. In each element is the best classifier for this attribute stated, i.e. LDA or SVM. In the column marked "Best Method" the filtering algorithm with highest difference are stated together with a certainty for this result.

Test type	Grand Average (AVG)	PCA	ICA	Best Method
Office Environment				
Big Screen 3 × 3	SVM	LDA	SVM	AVG (certainty≥97%)
Big Screen 2 × 2	SVM	SVM	LDA	AVG (certainty≥88%)
Small Screen 3 × 3	SVM	SVM	SVM	AVG (certainty≥98%)
Small Screen 2 × 2	SVM	SVM	SVM	AVG (certainty≥99%)
Car Environment				
Small Screen 3 × 3	SVM	LDA	SVM	AVG (certainty≥99%)
Small Screen 2 × 2	SVM	SVM	SVM	AVG (certainty≥99%)

Table 3: The result from the statistical comparison between the various type of filtering and tests in the vertical case when a high difference between correct and non-correct classifications are the considered attribute. In each element is the best classifier for this attribute stated, i.e. LDA or SVM.In the column marked "Best Method" the filtering algorithm with highest difference are stated together with a certainty for this result.

From the above Tables 2 and 3 the columns marked "Best Method" can be used for comparison between both vertical and horizontal types and also for different test types. The result from this comparison are then the, from the case when a high difference of correct classifications and false alarm are the searched attribute, the best candidate of the combination of the systems variables. In such a comparison the result is as follows:

1. Big Screen 3 × 3, Office environment
2. Big Screen 2 × 2, Office environment
3. Small Screen 3 × 3, Car environment
4. Small Screen 2 × 2, Car environment
5. Small Screen 3 × 3, Office environment
6. Small Screen 2 × 2, Office environment

In this list it should be noted that the certainty of the two top candidates compared to the others are above 98% but the certainty of the ordering of candidate 3-6 is equal or below 67%. The difference between vertical and horizontal case in each type are small enough to be seen as equal, i.e. there is a difference but the certainty of this difference are extremely low.

6.2 System test

All results in the previous section can be seen as result only regarding the methods and how well they perform when a lot of data from the same test are used for the training phase. However, in the actual system an initial training phase is executed where data is collected and used to calibrate the system to enable the following execution in real time. This scenario can be simulated by making use of the fact that the test persons have performed two tests of each type. The first of the two tests in every type is used as training data, then pre-processed data, in batches, from the the entire second test is fed in to the trained classifier. The number of classification on each possible choice is then observed and by counting a correct classification on a correct choices as one point in a scoreboard it is possible to evaluate the performance of the methods. Within the first section below the results from such a simulation are presented.

In the second section is a small investigation of the user friendliness presented. The focus is partly on the need for calibration and if it is possible use old data for calibration. As a second part is related to the user experience.

6.2.1 Classification test

As described in the introduction in this section is this investigation performed to simulate a potential behaviour with a initial calibration phase. To simulate such a process a good set up is essential and a description of this process is presented first.

An initial step is to divide the data into pieces (or batches), each batch holds values corresponding to 0.5 seconds after an illumination of an element in the matrix. Then the initial processing of data described in Section 6.1 is performed, i.e. a digital filtering of the data with a second order Butterworth-filter with cutoff frequencies of 0.3Hz and 20Hz etc. After this process all batches are sorted according to

- Hits, a P300 is expected to occur.
- Non-hits, no P300 is expected.

From this point each of the spatial filter methods are applied to an unique copy of the data and both types of classifiers are trained with further copies. This results in 6 uniquely trained classifiers, one for each method.

Next, the data from the second test is used as validation data. However, this time the batches are mapped to the corresponding illuminated element in the matrix. The earlier trained classifiers predict if a P300 occurs within each batch. Batches corresponding to the same element in the matrix are combined, the same type of mean value calculations as in Section 6.1 is used is this process, one by one. Data sets representing each possible choice are combined at most 10 times and classified, i.e. data corresponding to one specific choice is classified for $1, 2, .., 10$ combined batches. If 10 batches is reached the cycle starts over, i.e. the next batch corresponding to this element is classified alone. This is to simulate the possibility to wait for several illumination of one choice before a decision is made in the BCI.

By keeping track of the choice that is connected to each batch it is possible to keep a score of how many classifications of a P300 event each choice in the matrix get from a complete test. These points are then saved and can be represented in a scoreboard. This score can be used in a system to make a decision. It is possible to establish a basic requirement on the final scoreboard after a complete test. The points must be higher for the row and column where the user focuses than for the ones of each type where the user is not focusing. In all other cases a false decision will be made under the presumption of the same ratio, or at least almost the same ratio, between points holds over the entire test.

From the description of the test it is clear that the result will be a vector where each element is a type of score for each choice in the grid. A high score gives an implication that many classifications have predicted a P300 ERP to occur after this choice.

Data from all users are included in the results in each test and their combined score is divided by the number of users as a final result, i.e. if all user together get a score of 30 the results displayed is 5 ($\frac{30}{6} = 5$). The presentation is a combination of graphical display and tables were the result from the system analysis of the two test types in a car environment are displayed in Figures 33-36. The complete result for all test types and methods are presented in Tables 4-7. In the first two tables the results from tests performed in the office environment are presented and the results in the last two are retrieved from the data collected in the car environment. Note that a point is given for a correct classification regardless how many batches that are included in the process, i.e. a correct classification where 10 batches are used is of equal importance as if only 1 batch where included.

In Figure 33 and 34 the results for the tests with a 3×3 matrix are presented for LDA and SVM respectively. The height of the bars represent the number of classification scores and the bars with a red frame are false classifications. On the x-axis the labels are the matrix elements which represent the illuminations and the number starts in the top left corner of the matrix, i.e. ”Row 1” represent the top row and ”Column 1” represent the column far to the left etc. In the same manner the results for the data retrieved in the 2×2-matrix case are presented in Figure 35 and 36. The numerical values from these test can also be found in Tables 6 and 7 for the LDA and SVM classifiers respectively.

As seen only the SVM classifier in combination with AVG or ICA have a score on both the correct rows and columns that are above the non-correct choices of respectively type. Some results from a small investigation regarding how this relation between the scores holds through the test are displayed in Figures 37-40. In these figures it is possible to see how the correct choice ”take the lead” in the scoreboard around the centre of the test and this gives an optimal decision time of around 47 seconds (decision time is equal to flash time $\times$ number of choices $\times$ number of flashes, $0.175 \times 9 \times 30 = 47.25$ seconds).

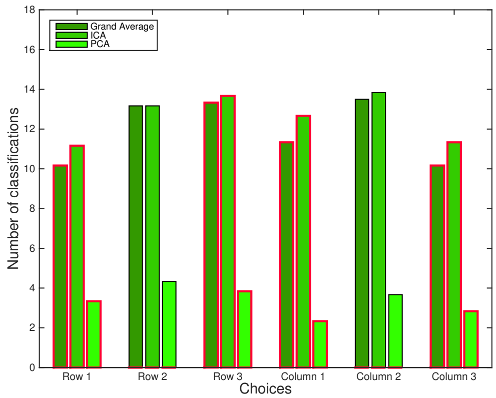


Figure 33: Results from the data retrieved in the test with a 3×3 matrix presented for the LDA classifier. The y-axis represent the number of classifications and the bars with a red coloured frame shows false classifications. The x-axis represent the corresponding flashing matrix elements, the numbering starts in the top left corner of the matrix.

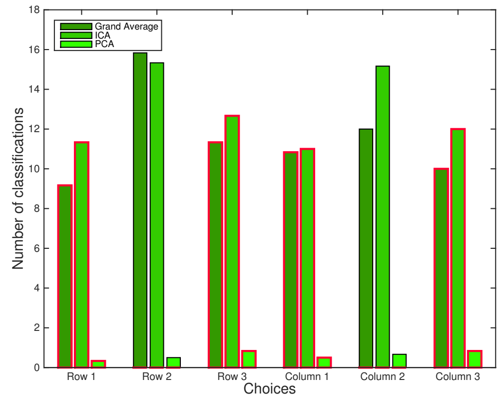


Figure 34: Results from the data retrieved in the test with a 3×3 matrix presented for the SVM classifier. The y-axis represent the number of classifications and the bars with a red coloured frame shows false classifications. The x-axis represent the corresponding flashing matrix elements, the numbering starts in the top left corner of the matrix.

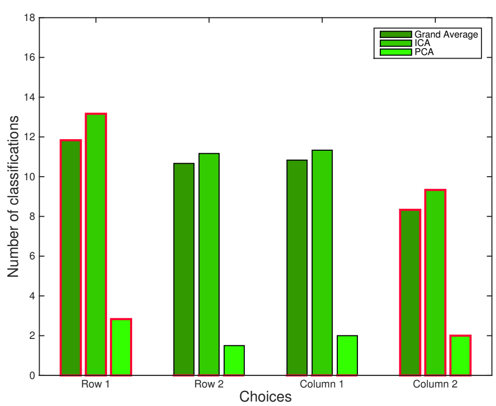


Figure 35: Results from the data retrieved in the test with a 2×2 matrix presented for the LDA classifier. The y-axis represent the number of classifications and the bars with a red coloured frame shows false classifications. The x-axis represent the corresponding flashing matrix elements, the numbering starts in the top left corner of the matrix.

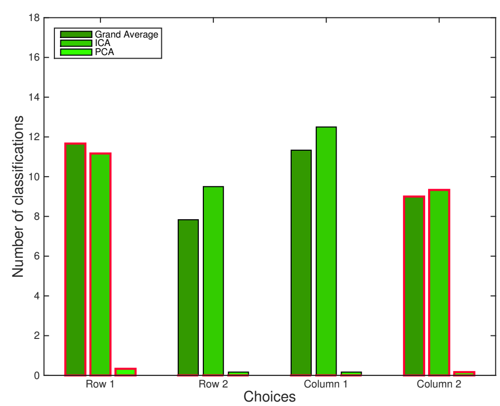


Figure 36: Results from the data retrieved in the test with a 2×2 matrix presented for the SVM classifier. The y-axis represent the number of classifications and the bars with a red coloured frame shows false classifications. The x-axis represent the corresponding flashing matrix elements, the numbering starts in the top left corner of the matrix.

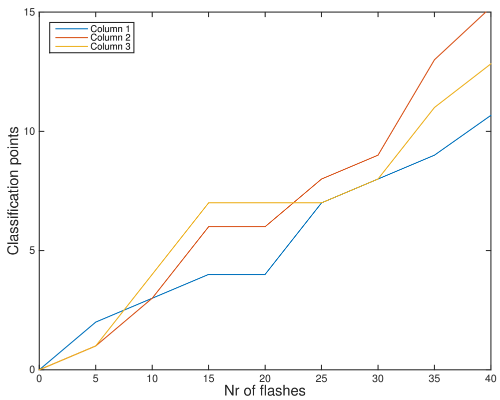


Figure 37: This figure displays the classification score of the possible columns in a system with a 3×3 matrix with a spatial filter of ICA type with a SVM classifier when a test are fed to the system in batches to be classified. On the x-axis are the number of times each column has been illuminated during the test.

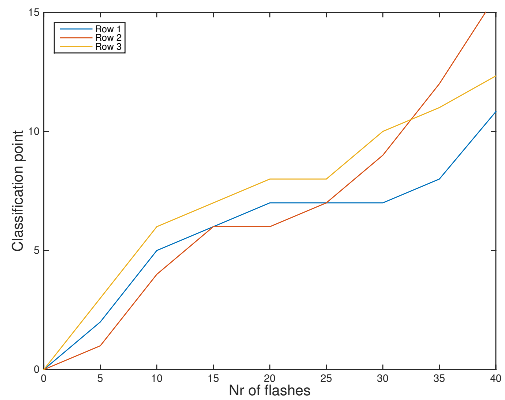


Figure 38: This figure displays the classification score of the possible rows in a system with a 3×3 matrix with a spatial filter of ICA type with a SVM classifier when a test are fed to the system in batches to be classified. On the x-axis are the number of times each row has been illuminated during the test.

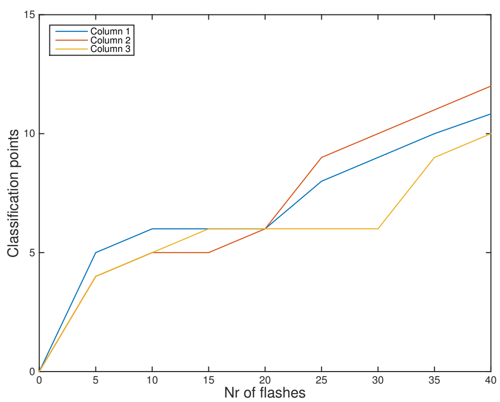


Figure 39: This figure displays the classification score of the possible columns in a system with a 3×3 matrix with AVG algorithm applied type with a SVM classifier when a test are fed to the system in batches to be classified. On the x-axis are the number of times each column has been illuminated during the test.

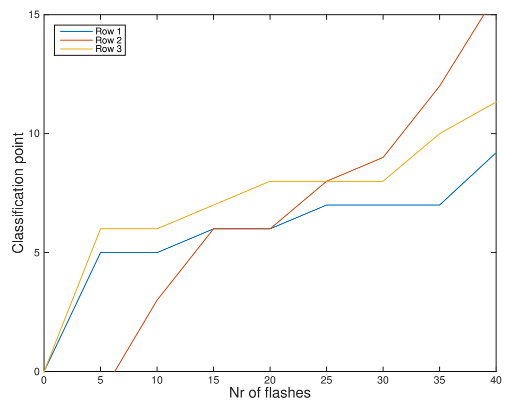


Figure 40: This figure displays the classification score of the possible rows in a system with a 3×3 matrix there the AVG data processing type with a SVM classifier are applied when a test are fed to the system in batches to be classified. On the x-axis are the number of times each row has been illuminated during the test.

Type of test and Matrix Element		**Classification Score**	
Large screen, matrix size 3×3	**Grand Average (AVG)**	**ICA**	**PCA**
Row 1	9.0	11.2	2.0
Row 2	**11.7**	**12.2**	**1.5**
Row 3	11.3	12.2	1.7
Column 1	10.8	11.7	1.5
Column 2	**10.5**	**12.5**	**2.7**
Column 3	10.7	12.3	2.2
Large screen, matrix size 2×2	**Grand Average (AVG)**	**ICA**	**PCA**
Row 1	13.2	12.0	2.5
Row 2	**10.8**	**11.3**	**2.0**
Column 1	13.3	14.2	1.5
Column 2	**11.8**	**12.5**	**2.2**
Small screen, matrix size 3×3	**Grand Average (AVG)**	**ICA**	**PCA**
Row 1	10.0	11.5	1.0
Row 2	**13.0**	**12.5**	**2.0**
Row 3	8.5	11.2	1.7
Column 1	10.7	12.7	2.0
Column 2	**13.7**	**12.8**	**1.0**
Column 3	9.8	12.2	2.0
Small screen, matrix size 2×2	**Grand Average (AVG)**	**ICA**	**PCA**
Row 1	11.2	9.7	1.0
Row 2	**10.7**	**12.3**	**1.5**
Column 1	**11.2**	**12.2**	**1.2**
Column 2	8.8	9.5	1.5

Table 4: The resulting score of classification for each method and test in office environment where LDA has been used as classifier. High score on Row 2, Column 2 and Row 2, Column 1 are correct classifications for matrix size 3×3 respectively 2×2, these are marked with bold font. High score on other elements in of the matrix are considered as false alarms.

Type of test and Matrix Element		Classification Score	
Large screen, matrix size 3×3	**Grand Average (AVG)**	**ICA**	**PCA**
Row 1	8.5	8.8	0.0
Row 2	**13.2**	**13.3**	**0.2**
Row 3	11.5	12.5	0.3
Column 1	7.5	8.7	0.0
Column 2	**12.8**	**15.5**	**0.0**
Column 3	11.0	11.3	0.3
Large screen, matrix size 2×2	**Grand Average (AVG)**	**ICA**	**PCA**
Row 1	12.5	13.8	0.2
Row 2	**9.2**	**10.0**	**0.3**
Column 1	**11.3**	**10.3**	**0.7**
Column 2	8.8	11.3	0.2
Small screen, matrix size 3×3	**Grand Average (AVG)**	**ICA**	**PCA**
Row 1	6.2	9.7	0.0
Row 2	**11.8**	**12.3**	**0.0**
Row 3	8.0	9.3	0.2
Column 1	9.2	10.2	0.0
Column 2	**14.0**	**14.0**	**0.0**
Column 3	9.0	11.3	0.0
Small screen, matrix size 2×2	**Grand Average (AVG)**	**ICA**	**PCA**
Row 1	13.0	13.3	0.3
Row 2	**7.3**	**10.2**	**0.0**
Column 1	**11.8**	**14.5**	**0.0**
Column 2	7.7	8.3	0.7

Table 5: The resulting score of classification for each method and test in office environment where SVM has been used as classifier. High score on Row 2, Column 2 and Row 2, Column 1 are correct classifications for matrix size 3×3 respectively 2×2, these are marked with bold font. High score on other elements in of the matrix are considered as false alarms.

Type of test and Matrix Element		Classification Score	
Small screen, matrix size 3×3	**Grand Average (AVG)**	**ICA**	**PCA**
Row 1	10.2	11.2	3.3
Row 2	**13.2**	**13.2**	**4.3**
Row 3	13.3	13.7	3.8
Column 1	11.3	12.7	2.3
Column 2	**13.5**	**13.8**	**3.7**
Column 3	10.2	11.3	2.8
Small screen, matrix size 2×2	**Grand Average (AVG)**	**ICA**	**PCA**
Row 1	11.8	13.2	2.8
Row 2	**10.7**	**11.0**	**1.5**
Column 1	**10.8**	**12.2**	**2.0**
Column 2	8.3	9.7	2.0

Table 6: The resulting score of classification for each method and test in car environment where LDA has been used as classifier. High score on Row 2, Column 2 and Row 2, Column 1 are correct classifications for matrix size 3×3 respectively 2×2, these are marked with bold font. High score on other elements in of the matrix are considered as false alarms.

Type of test and Matrix Element		**Classification Score**	
Small screen, matrix size 3×3	**Grand Average (AVG)**	**ICA**	**PCA**
Row 1	9.2	11.3	0.3
Row 2	**15.8**	**15.3**	**0.5**
Row 3	11.3	12.7	0.8
Column 1	10.8	11.0	0.5
Column 2	**12.0**	**15.2**	**0.7**
Column 3	10.0	12.0	0.8
Small screen, matrix size 2×2	**Grand Average (AVG)**	**ICA**	**PCA**
Row 1	11.7	12.0	0.3
Row 2	**7.8**	**10.2**	**0.2**
Column 1	**11.3**	**13.5**	**0.2**
Column 2	9.0	8.5	0.2

Table 7: The resulting score of classification for each method and test in car environment where SVM has been used as classifier. High score on Row 2, Column 2 and Row 2, Column 1 are correct classifications for matrix size 3×3 respectively 2×2, these are marked with bold font. High score on other elements in of the matrix are considered as false alarms.

By comparing the results related to the car environment presented in this section for the two types of classifiers it is possible to achieve two results:

1. The combinations of methods that perform best in a car.
2. The effect of matrix size on the results in a car.

These results are based on the highest points, values on the y-axis, of the possible choices regarding row and column. This since they, as mentioned in the introduction of this chapter, can be used to generate a decision in the system.

Only one type of matrix setup and classifier fulfils the earlier stated requirement on a higher score on the correct choice for both rows and columns, for the tests performed in a car. This is the 3×3 matrix size with a SVM classifier. However, the method of spatial filtering have two candidates that both fulfils this attribute, this is Grand Average and ICA. There is an advantage to the ICA method due to the higher difference in the column case.

6.2.2 User Friendliness

As an analysis regarding the user friendliness of a potential system has two types of investigation been performed and the description of the investigations as well as the results are presented here. The first of these studies is regarding the possibility of using old data as training for the system. The second one is an investigation which focus around how the users felt when they tested the system, both regarding the physical appearance as well as the issues of using the system. Both are in the subject of user friendliness but in two different aspects. In the first investigation the results are presented in values and figures, in the second more subjective observations presented as results.

Analysis with old calibration data

In Section 6.2.1 have, as mentioned, the data from the first test of the two performed by each user for each type of the tests been included in the calibration of the system. The second test has been used for the actual evaluation. The time span between the obtaining of data from the first and second test are very short and the headset where not removed or altered. Hence, the result in the previous sections do not give any information regarding if data retrieved from an earlier session can be used for calibration to a new test. There by are this investigated in the first analysis in this section and the reason to why it is written under the current title is that the result are telling if it is possible to shorten the set up time for ever new session with the system by using the data from previous sessions and make the system more attractive for users.

Within this investigation are the same data as in Sections 6.2.1 and **??** used for calibration and the system are in the same manners as in the previous sections feed a new test to give a classification score for each possible choice. However, this time have there been over 2 weeks since the users performed the test which is feed to the system in the calibration phase. As the type of test are the most promising one from Section 6.1.4, i.e. 3×3 matrix size on a large screen in office environment, chosen as candidate for this analysis. The points for a classification are given as in Section 6.2.1 and collected and displayed in the same manners to enable an easy comparison between the two studies.

The result are presented below in Figures 41, 42 and in Table 8, where row 2 and column is the correct choices. As can be seen the results from this study is compared to the one presented in Tables 4 and

5 worse in the sense that the score on the correct choices are lower compared to the non-correct ones. In fact if the best method in the study presented in the previous section, the ICA algorithm with the SVM classifier, are compared with its counterpart in this investigation the response from the system has changed from being possible of distinguish the correct choices regarding both row and column to not be able to do so. Hence, the result from this analysis therefor is that it seems to be impossible to use old data as calibration for the particular system set up. This result are also valid for the car environment since, as mentioned earlier, the best performing test among the different test types, i.e. large screen 3×3 matrix, are used in this investigation.

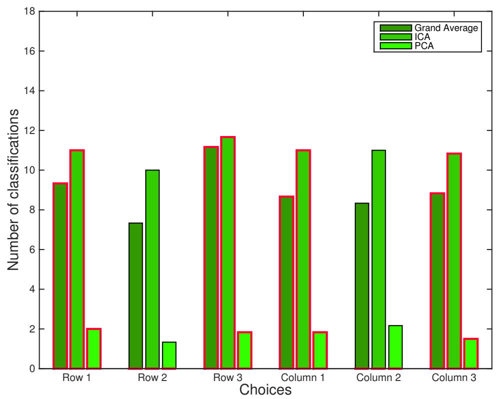


Figure 41: Results from the data retrieved in the test with a 3×3 matrix presented for the LDA classifier. The y-axis represent the number of classifications and the bars with a red coloured frame shows false classifications. The x-axis represent the corresponding flashing matrix elements, the numbering starts in the top left corner of the matrix.

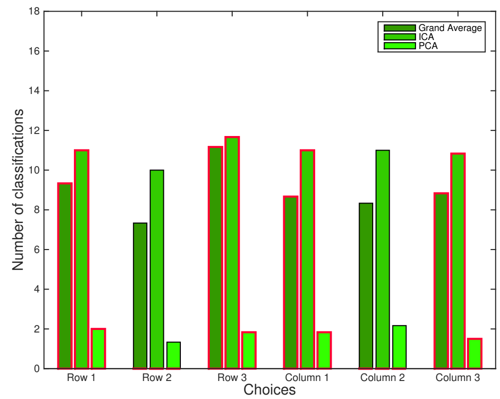


Figure 42: Results from the data retrieved in the test with a 3×3 matrix presented for the SVM classifier. The y-axis represent the number of classifications and the bars with a red coloured frame shows false classifications. The x-axis represent the corresponding flashing matrix elements, the numbering starts in the top left corner of the matrix.

Classifier type		Classification Score	
LDA	**Grand Average (AVG)**	**ICA**	**PCA**
Row 1	9.3	11.0	2.0
Row 2	**7.3**	**10.0**	**1.3**
Row 3	11.2	11.7	1.8
Column 1	8.7	11.0	1.8
Column 2	**8.3**	**11.0**	**2.2**
Column 3	8.8	10.8	1.5
SVM	**Grand Average (AVG)**	**ICA**	**PCA**
Row 1	8.5	9.8	0.2
Row 2	**6.0**	**9.0**	**0.2**
Row 3	9.8	11.2	0.3
Column 1	5.8	8.3	0.2
Column 2	**8.0**	**11.0**	**0.2**
Column 3	8.2	11.3	0.3

Table 8: The resulting score of classification for the two types of classifiers and filtering method for the first test in office environment, i.e. 3×3 matrix size with large screen. The number of classifications, or the score, on Row 2 and Column 2 are correct classifications, these are marked with bold font. Points on the other elements in of the matrix are considered as false alarms.

User experience

Even though it is a small area of the investigation performed within this project the user experience is something that should be considered during the test performed with user. All test persons have been asked some general questions regarding the feelings of the system both in office environment and in the vehicle.

The result from this small investigation are not hard numbers, instead the results and answers can be seen as the test subjects opinions regarding the system. If several subjects have got the same experience within this area are there most likely something that can be seen as a valid result in the addressed subject. With this said, the small questions regarding the system have resulted in following list of attributes.

- The system is not demanding to use in reasonable periods of time, however by keeping maximum focus for several minutes a feeling of exhaustion arises in several test subjects. In addition to this exhaustion the headset is uncomfortable to wear for long periods of time.
- The time between and the duration of each flash are considered quite fast and it is hard to keep full focus for each individual flash when a double enlightenment occur on the same choice (due to the randomly selected order of flashes). However, users does not find the system to be stressful.
- In the car environment where the glare from the light a factor to count with according to several users.
- When it come to the affect of the vehicle movement most of the users had trouble to focus during hard turns but not otherwise. In fact, the effect on the orientation of the screen from movements of the car seemed to be easy to deal with, accordingly to several users.
- None of the users testes experienced any dizziness or illness due to use of the system during movement.

7 Discussion

Since this project holds several pieces that are interesting to discuss this section is separated in several parts to make the discussion more easy to grasp for the reader. In this initial part a lighter discussion regarding the hardware, the physical appearance of the system, and general aspects of the tests will be held. Later the subjects will be more focused around the internal factors and the actual results from the analysis of the test data.

As a first point to be raised the choice has fallen on the purchased hardware. The EPOC headset from Emotiv is as mentioned in Section 3.1 a rather cheap equipment for use in the field of EEG and the benefits is a quite adaptable and easy-to-use system. This is something desirable since a manufacturer never would implement a very expensive medical grade EEG in a car due to various issues regarding usability and the disadvantage of using to advanced hardware for implementing a rather easy task. Another even more beneficial, from an economic point of view, approach should have been to use eye-tracking for instance. However, since the EPOC can compete economically with such system, and in theory also is faster and more accurate, it has been a good choice for this project. By looking at the result the hardware seems to be able to capture the potentials in a sufficient manner for the desired application, yet there is of course always room for improvements. However, attributes of equal or even more importance than the signal quality to improve are not regarding resolution of the measurement it is regarding the looks and attribute of user friendliness. By making use of dry electrodes instead of the type that requires saline solution, which currently is the only option in this model, would the set up process as well as the result be improved greatly.

The mentioned electrodes have been a issue during the set up phase in more than one occasion. One of the major concerns is the hair on the scalp, persons with long and/or thick hair have, naturally, a worse connection than a bald test person. On top of that the saline solution also seems to dry out faster when there is more hair. However, the hardest part is to establish the initial connections since the electrodes are a bit ”fuzzy”, due to the saline pads and it is hard to get them beneath the hair. Dry electrodes are most likely more solid in their structure and would be easier to apply correctly. Another benefit with such electrodes are the placement, most likely would dry electrodes require a harder pressure towards the scalp and in a headset with harder structure would the spread of the placement of electrodes be decreased compare to the rather soft plastic structure of this model. Yet, as used in a proof-of-concept system has the Emotiv headset been very useful and the transition between test persons have been as smooth as expected (even if taking out and insert electrodes is a little irritating).

Another concern regarding the set up of the tests is of course the integration of this system in the car interior. In the tests performed in this project test subjects has been seated in the backseat of the car with the screen attached to the chair in the front. However, in a final system the integration could possibly be on the centre stack in the car instead, although the position of the screen must be possible to tune. This is because the user must be able to be in a good angle to the screen and this also assumes that some kind of system determining the focus of the user is integrated with the BCI. This system should have the purpose of making sure that the user look at the screen and not have his or her focus on another thing when a decision in the BCI is made. In the tests during this project the users have been instructed to keep their focus on the screen during the whole test but this would of course not be the case in a real application. A suggestion to a solution for this problem is to make use of eye tracking. By using a system accurate enough to tell if the user looks at the screen or not as a condition for the use of the system would be enough, from a functional point of view, to implement and use it in a safe and secure fashion.

Also when it comes to possible improvement of the hardware directly used in the BCI, eye-tracking is a possibility. If the accuracy of the earlier suggested eye-tracking system is good enough the impact on the BCI could be extended further. By using the tracking of the focus of a user within the actual screen the illumination of elements in the grid of choices could be restricted to a smaller area where the user is focusing, this would improve the accuracy of the classification in the system even further. However, to the knowledge of the authors there are no eye-tracking device on the market which fulfils this demand on accuracy yet.

7.1 Result Comments

From Section 6 and especially from Section 6.1.4 it can be stated that a system built on some of the investigated methods based on the P300 ERP, seen from this empirical trials point of view, possible to be used in a vehicle setting. This statement is based on the assumption that is possible to develop an algorithm where the number of classifications where a P300 ERP is present for each element is used in the type of scoreboard introduced in the studies presented within Section 6.2.1 to end up in a decision. Such a method will reduce the impact of false classifications and have improved effect on the robustness of the system which is most desirable in a final system.

This kind of system will most likely work well under the assumption that it is possible to increase the difference between the number of false alarms and the correct classifications even more than what is achieved in Section 6.2.1. However, as can be seen in for example Figure 10 and 21 are there situations where the number of false alarms are equal or more than the correct classifications. These situations are seldom occurring after use of 5 batches of data to build a mean of. Thus, a suggestion is to make use of a weighting when distributing the points in the mentioned scoreboard, this weighting should be build on the number of data batches included in the classification.

According to the statistical analysis in Section 6.1.4, the result from the tests in the car outperforms its counterpart from the office in some parameters. This implies that although there are significant disturbances in the environment, other aspects of the system design can be more important, such as usability and focus of the user. However, even if the results from the tests in the car overall are encouraging and frankly quite good, the percentage of false alarm is also in this case relative high. This means that the constraints and demands on the algorithm for decision making increases. In order to achieve a perfectly safe system with the current setup a new type of investigation need to performed focusing on the origin of the false alarms.

The most obvious reason to false alarms are muscular artifacts. Even though the hardware have built in filters to deal such attributes are they most likely interfering with the signal. Hence should this area be the starting point in this further analysis and if a pattern in the signal can be found that tells about such artifact can they be dealt with by numerous types of pre-processing. Yet, if a fast and efficient algorithm for the decision making from the scoreboard approach reach the desired level of security would this described hunt for artifacts be unnecessary.

7.2 Experimental Setup

Since the aim of this project from the beginning has been to perform an evaluation of the feasibility of using a low-cost equipment for identifying, classifying and using P300 responses in a system suitable for the real world environment, surrounding parameters have not been controlled on purpose. There will be differences in a number of parameters such as temperature, mood and energy level on the test subjects, glare from light sources, movement of the users during tests and so forth. The intended result is that the system should be functional to use even with these uncertainties in the system. This also includes electrode position which will slightly differ from test to test, since the subjects differ in head geometry and the electrode mounts are flexible.

However, as stated earlier some parameters significant to the system performance have been controlled and kept track on. This includes, screen size, matrix size and most important environment change from car to office, which will differ greatly in the above mentioned external parameters. By using the system in both office and car environment it was possible to see if the car environment had a large impact on the performance of the system and give an indication if it was feasible at all to use the system in this setting.

When conducting the tests, Emotiv's software was used to determine the signal quality of the electrodes, which based on signal amplitude and difference towards reference signals gives a number to indicate the contact and signal quality of each channel. This is also shown graphical with a plot showing black, red, yellow or green depending on this quality for an easy overview. During the study we primarily used this to give a rough indicator that the system was properly initiated and the headset mounted, but also looked at the specific quality numbers. Since there, to the authors knowledge, does not exist any standards both regarding signal quality as well as overall performance measures of BCIs the measurements are hard to validate and is based on Emotiv's algorithms. The same applies to performance of the complete system which is based on the author's perspective of the important parameters in the system and its evaluation.

Another aspect of the experimental setup is the user base when performing the tests. This involved a total of 6 persons, consisting of thesis workers working in close proximity. The test subjects are all young and used to using modern technology, but differ in gender and nationality. To give a validated statistical evaluation of all aspects of the system a larger user base would be used, with a more thorough questionnaire. However, this is not the intention with the study and the results can be seen as indications for a concept based on the used hardware and software.

7.3 Design Choices

From the various results presented in Section 6 several conclusions can be drawn regarding the difference in performance of the investigated algorithms. In the graphical results it is easy to distinguish a trend, where the percentage of false alarms decreases with the number of batches used to build the mean value of the signal. This is something very positive since the classification seems to be better in at least one aspect. However, the increase of classifications is not as clear and can make the choice of number of batches to use harder.

Furthermore, for some cases different methods can achieve the best result in the same test, where averaging can give the best result for rows and ICA for columns, as in Figure 34. However, as seen in Section 6.2.1 it is relatively clear that the recommended combination of methods to use in a general term is using a 3x3 matrix, ICA and SVM classifier. This result is taken from the scores from the scoreboards used in both the mentioned sections.

Although, the results show that there are a lot of different components in the system which makes investigation complex, since there exists a numerous possibilities to combine these components. Apart from this, depending on the specific situation different components can give the best result. Due to this an optimal solution for all generic use cases can not be found.

For channel selection in this study, while using grand averaging, has analysed the data from the test group manually and from there chosen the optimal channel to use as a data stream while using the others for noise reduction. This has been the same for all test subjects, although in a final system this might not be the case. Hence, a system for evaluating and choosing the optimal channel or channels would be good to implement. While using ICA or PCA a subset of the optimal channels are used which have been picked manually and for this method it might also be beneficial to use an automatic channel selection.

When it comes to the searched possibility of include user profile and by this avoid the need for an obligatory calibration for every session are the results unfortunately very negative. This becomes very clear from the result in the Section 6.2.2 where data collected two weeks earlier are feed to the calibration phase of the system used for classification of newly obtained data. Even though the same user are performing the exact same test once more are the resulting score on classifications of the correct choices lower or on the same level as for the non-correct options. This property are also valid for all the investigated methods which implies that it does not matter which signal processing that are applied to the signals, the difference in the data are to great for the classifiers trained on the old data to handle. The reason to this lies most certainly in two factors;

1. Positions of the electrodes on the scalp of the user
2. The level of voltage measured

Since the headset not are as accurate as for example a cap in the manners electrodes are placed every set up could the bad performance displayed in Table 8 compared to the presented in Table 5 be due to that the positions of electrodes are altered between the tests. Thereby would the impact on the signals from a P300 ERP have a different appearance which would make the training of the classifiers useless, classification would be almost random. Another explanation is that the potential difference between some channels and the reference points are very different in level for each set up. This is most likely the case since the reference point are not exactly the same every time. Depending on how much the muscles in the tissue beneath the reference points are activated would a difference voltage level between the reference and every unique channel rise during the set up. It is quite a challenge to construct a system able of handle such a difference without a mandatory initial calibration every session. However, one possibility could be to make use of an enormous amount of data from several sessions and investigate if it is possible to train several sets of classifiers and depending on some parameter, which must be easy and fast to measure, choose a suitable subset to use for this particular session. Thereby would it be possible to have a small delay in the start up, instead of a full calibration, where the best possible of the classifiers are chosen.

7.4 Final Program & Implementation

A most natural first step to take from this project would to develop a final program, integrated in a car, to be seen as a proof-of-concept prototype and would show the intended use of the system. The performance must be balanced between speed and accuracy and the developed system evaluated in Section 6.2 would not be satisfying in its current state if it should be used in a car environment directly. However, it shows that the concept is viable in a car setting and with further development could create a functional system.

The implemented program for test in this project uses a 3x3 matrix of choices in a grid. However, this is not necessarily the best setup. In an interactive interface an alternative design could be to use the icons of the interface as elements in the BCI. In this case, the elements of the interface can be grouped in the same way as columns and rows group elements, but with the freedom of choosing the design to fit the interface design. One example is putting the elements in a circle, dividing it into sections left, right, top, middle, bottom. This would give 6 elements to choose from and is a simple variation of the interface, which of course can be further customised. This specific interface would then act as type of 3x2 matrix and the benefit would be that there could be an interface valid for both a system involving a BCI as well as without. For example could a tablet with touch based interaction that are implemented in the car be extended with a BCI to enable a new type of interaction without to change the interface layout.

There are of course other way than the above mentioned circle of new possible designs of the system. Another more BCI specific alternation is to divide the decision process into several steps, as a menu system. In that case the system would iteratively work towards a specific element in turns. In the specific case the system would first determine if a smaller part of the interface is the intended and then only blink elements in that area. From some of the result in Section 6.2 this type of system is actually something that have a possibility of increase the accuracy of the system. By making a system where the user first choose rows and then the system transpose the chosen row to only illuminate the options within it would only choices based on rows influence the decision making process, from the mentioned result would this have the potential to increase the performance of a final system.

7.4.1 Time to decision

While designing and evaluating a BCI system the time to decision is of vital importance. To be able to use the system in a real-time setting the decision time need to be minimised as much as possible, while maintaining a good level of accuracy. Since the aim of this study is to evaluate possibility of implementing a BCI in a autonomously driven car there will be a trade-off between speed and security in the system. The level of security needed in the system depends on the severity of an error. One possible implementation is using the system as a strategic controller in an autonomous car. In this case an error is severe since it can command the car to perform actions such as taking a highway exit without the intention of the driver. Although, it never interacts with the primary systems of the car such as steering and can not cause a safety issue.

Of course, with a higher level of security comes a longer decision time since more data is sampled to confirm the intention of the user. However, the security of the system is compromised if the time to decision is to long, e.g if the driver wants to change lane and the situation has changed during the time span elapsed between intention and execution. Other aspects of the time of decision is matrix size, blinking frequency and classification time inside the system. Even though the time for decision are at least estimated for the system validated in Section 33 to be around 47 seconds. This is a to long time span for decision in a car application but perhaps it is possible to also within the earlier discussed development of a decision algorithm to shorten this time.

8 Conclusion

From the result presented in Section 6 and the discussion held in the previous section one conclusion is very clear, the fact that the system investigated in this thesis is of a very complex nature. This statement does not only hold for the measurements but also for the actual validation of the system, since there is no existing standard definition, and it is hard to establish one, that determines the quality of how systems of this type should be validated. Nevertheless, from the different analysed attributes in this project the best performing set-up of the system in a car environment is a system with following parameter settings:

1. 3×3 grid of choices, as in Figure 7
2. An algorithm based on the independent component analysis as spatial filter
3. A support vector machine classifier

In Figure 43 these components are inserted in their positions with the used hardware and digital filter properties in the same manner as in the BCI overview presented in Figure 2 in Section 2. This can be seen as the final conclusions of which methods to be used within a potential final system. Yet, these conclusions are only valid for the specific system used in this project. In other words, if another hardware is included or if some other fundamental parameter is changed the results are compromised due to the complexity of the human brain and individual differences from test persons.

Even though there is a promising candidate for use in a system, suitable for the car environment, there are several steps left to be taken before a fully working product can be launched. The mentioned candidate definitely have potential to be used in a car, but further work have to be performed in the development of an optimal decision algorithm as well in suppressing the number of false alarms. In the the current set up the time to make a decision is around 47 seconds. This is unacceptable long in a system developed to be used for issuing commands, acting as a strategic controller in a vehicle.

Before an actual development of a working final system can begin it is crucial to alter the hardware for several reasons. Firstly, the electrode positions is not optimal to capture P300 ERPs, since the headset had to be turned 180 degrees. Secondly from an user perspective, it is uncomfortable to wear under long periods of time and the use of saline solution on the electrodes is definitely not optimal for long-term use.

The analysis performed also established that a calibration of the system is necessary for every new session. This makes it impossible to design a system based on user-profiles where calibration only is requested on rare occasions. This is something that decreases the user friendliness to a high degree.

Regardless of the earlier mentioned flaws in hardware and the unreachable attribute of user-profiles, there are very positive and promising results found within this thesis regarding if a BCI can be included in a car in the future. One is the fact that the results from the analysis performed on data retrieved in a car environment proved not to be the worst when comparing the attribute related to false alarms, in Section 6.1.4. This indicates that the car environment is a feasible application area for the system. Another promising result is the mentioned difference of scored points between a correct and an non-correct choice in the scoreboard in Section 6.2.

However, in combination with the earlier mentioned best performing candidate of algorithm and the physical appearance of the system the conclusion based on these results is clear. Hence, the main question of the project, which also is the title of this thesis, has got an answer; ”No, a Brain-Computer Interface in a car is not a realistic concept with the technology available today”.

Figure 43: A overview of a BCI with the components established in this investigation inserted. This can be related to Figure 2 where the spatial filter inserted is the ICA and the classifier is SVM, also the used headset and digital filter properties are stated.

8.1 Further Work

The results shown and discussed in this report is to be seen as a first and preliminary evaluation of the feasibility to use a P300 BCI system in a car setting. It concludes that there is potential of using a cheap and user-friendly system to be used as a control measure for strategic commands.

This initial study shows and discusses a set of usable tools as well as highlighting problems and difficulties in the system. A natural way forward is to further investigate the chosen methods and hardware with a larger user base and verify their performance in different settings. This would include developing an on-line system as a prototype to measure performance. From that study knowledge can be applied iteratively to the system to develop a final product.

References

[1] H. L., “Hans Berger (1873–1941), Richard Caton (1842–1926), and electroencephalography.” *Journal of Neurology, Neurosurgery, and Psychiatry*, no. 74, 2003.

[2] F. Sharbrough, G. Chatrian, R. Lesser, H. Lüders, M. Nuwer, and T. Picton, “American Electroencephalographic Society guidelines for standard electrode position nomenclature,” *J. clin. Neurophysiol*, vol. 8, no. 2, pp. 200–202, 1991.

[3] M. A. Lebedev and M. A. Nicolelis, “Brain–machine interfaces: past, present and future,” *TRENDS in Neurosciences*, vol. 29, no. 9, pp. 536–546, 2006.

[4] A. Nijholt, “BCI for games: A ‘state of the art’survey,” in *Entertainment Computing-ICEC 2008*. Springer, 2009, pp. 225–228.

[5] N. Birbaumer, “Breaking the silence: brain–computer interfaces (bci) for communication and motor control,” *Psychophysiology*, vol. 43, no. 6, pp. 517–532, 2006.

[6] D. Göhring, D. Latotzky, M. Wang, and R. Rojas, “Semi-autonomous car control using brain computer interfaces,” in *Intelligent Autonomous Systems 12*. Springer, 2013, pp. 393–408.

[7] C. Ziegler, “Tesla autopilot,” The Verge, 2014, http://www.theverge.com/2014/10/17/6982289/tesla-autopilot-is-a-non-revolution-for-self-driving-cars [2015-02-06].

[8] M. Wang and D. Latotzky, “Driving an autonomous car with eye tracking,” *Free University of Berlin*, March 2010, http://www.mi.fu-berlin.de/inf/publications/techreports/tr2010/B-10-03/B-10-03.pdf?1346662704.

[9] M. Wang, “Remote controlling an autonomouscar with an iphone,” *Free University of Berlin*, March 2010, http://edocs.fu-berlin.de/docs/servlets/MCRFileNodeServlet/FUDOCS$_{d}erivate_{0}00000000892/B-10-02.pdf$.

[10] e. a. Engström. J, Arfwidsson. J, “Meeting the challenges of future automotive hmi design: An overview of the aide integrated project,” *U.S. DEPARTMENT OF TRANSPORTATION*.

[11] P. Gruhn, “Human Machine Interface (HMI) Design: The Good, The Bad, and The Ugly (and what makes them so),” RockWell Automation, 2011, http://www.kirp.chtf.stuba.sk/ [2015-06-29].

[12] M. Nilsson, “Hav - human machine interaction (HMI) for autonomous vehicles,” Vinnova, 2014, http://www.vinnova.se/PageFiles/751326056/Presentation$_{f}fi_{a}utomation_{d}ay.pdf.pdf[2015-02-06]$.

[13] S. Sutton, M. Braren, J. Zubin, and E. R. John, “Evoked-Potential Correlates of Stimulus Uncertainty,” *Science*, vol. 150, no. 3700, pp. pp. 1187–1188, 1965. [Online]. Available: http://www.jstor.org/stable/1717778

[14] Wolpaw, J.R. and Birbaumer, N. and Heetderks, W.J. and McFarland, D.J. and Peckham, P.H. and Schalk, G. and Donchin, E. and Quatrano, L.A. and Robinson, C.J. and Vaughan, T.M., “Brain-computer interface technology: a review of the first international meeting,” *Rehabilitation Engineering, IEEE Transactions on*, vol. 8, no. 2, pp. 164–173, Jun 2000.

[15] H. Moravec, “When will computer hardware match the human brain,” *Journal of evolution and technology*, vol. 1, no. 1, p. 10, 1998.

[16] T. N. Lal, T. Hinterberger, G. Widman, M. Schröder, J. Hill, W. Rosenstiel, C. Elger, B. Schölkopf, and N. Birbaumer, “Methods towards invasive human brain computer interfaces,” 2005.

[17] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, “Brain–computer interfaces for communication and control,” *Clinical neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.

[18] F. Beverina, G. Palmas, S. Silvoni, F. Piccione, and S. Giove, “User adaptive BCIs: SSVEP and P300 based interfaces,” *PsychNology Journal*, vol. 1, no. 4, pp. 331–354, Jan. 2003, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.161.3731rep=rep1type=pdf.

[19] J. Mattout, M. Perrin, O. Bertrand, and E. Maby, “Improving BCI performance through co-adaptation: Applications to the P300-speller,” *Annals of Physical and Rehabilitation Medicine*, 2015.

[20] M. Hamedi, S.-H. Salleh, A. M. Noor, and I. Mohammad-Rezazadeh, “Neural network-based three-class motor imagery classification using time-domain features for BCI applications,” in *Region 10 Symposium, 2014 IEEE*. IEEE, 2014, pp. 204–207.

[21] E. A. Larsen, “Classification of EEG Signals in a Brain-Computer Interface System,” 2011, Norwegian University of Science and Technology.

[22] I. Berlad and H. Pratt, “P300 in response to the subject’s own name,” *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, vol. 96, no. 5, pp. 472–474, 1995.

[23] J. Polich, “Updating P300: an integrative theory of P3a and P3b,” *Clinical neurophysiology*, vol. 118, no. 10, pp. 2128–2148, 2007.

[24] E. Franz, “Development of a Fast and Efficient Algorithm for P300 Event Related Potential Detection,” Ph.D. dissertation, Temple University, 2014.

[25] H. Ekanayake, “P300 and Emotiv EPOC: Does Emotiv EPOC capture real EEG?” *Web publication http://neurofeedback. visaduma. info/emotivresearch. htm*, 2010.

[26] A. Benton, “Our brian is shrinking but our frontal lobe is growing,” Online, EvoAnth, The story of human evolution, 2014, http://www.evoanth.net/2014/03/13/our-brain-is-shrinking-but-our-frontal-lobe-is-growing/.

[27] M. Kaper, P. Meinicke, U. Grossekathoefer, T. Lingner, and H. Ritter, “BCI competition 2003-data set IIb: support vector machines for the P300 speller paradigm,” *Biomedical Engineering, IEEE Transactions on*, vol. 51, no. 6, pp. 1073–1076, 2004.

[28] B. Blankertz, K. Muller, G. Curio, T. M. Vaughan, G. Schalk, J. R. Wolpaw, A. Schlogl, C. Neuper, G. Pfurtscheller, T. Hinterberger *et al.*, “The BCI competition 2003: progress and perspectives in detection and discrimination of EEG single trials,” *Biomedical Engineering, IEEE Transactions on*, vol. 51, no. 6, pp. 1044–1051, 2004.

[29] E. W. Sellers, D. J. Krusienski, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, “A P300 event-related potential brain–computer interface (BCI): the effects of matrix size and inter stimulus interval on performance,” *Biological psychology*, vol. 73, no. 3, pp. 242–252, 2006.

[30] R. Ortner, R. Prückl, V. Putz, J. Scharinger, M. Bruckner, A. Schnuerer, and C. Guger, *Accuracy of a P300 Speller for different conditions: A comparison*. na, 2011.

[31] G. Schalk and J. Mellinger, *A Practical Guide to Brain–Computer Interfacing with BCI2000: General-Purpose Software for Brain-Computer Interface Research, Data Acquisition, Stimulus Presentation, and Brain Monitoring*. Springer Science & Business Media, 2010.

[32] J. Decety, “Do imagined and executed actions share the same neural substrate?” *Cognitive brain research*, vol. 3, no. 2, pp. 87–93, 1996.

[33] J. J. Daly, R. Cheng, J. Rogers, K. Litinas, K. Hrovat, and M. Dohring, "Feasibility of a new application of noninvasive brain computer interface (BCI): a case study of training for recovery of volitional motor control after stroke," *Journal of Neurologic Physical Therapy*, vol. 33, no. 4, pp. 203–211, 2009.

[34] E. Buch, C. Weber, L. G. Cohen, C. Braun, M. A. Dimyan, T. Ard, J. Mellinger, A. Caria, S. Soekadar, A. Fourkas *et al.*, "Think to move: a neuromagnetic brain-computer interface (BCI) system for chronic stroke," *Stroke*, vol. 39, no. 3, pp. 910–917, 2008.

[35] K. K. Ang, C. Guan, K. Sui Geok Chua, B. T. Ang, C. Kuah, C. Wangf, K. S. Phua, Z. Y. Chin, and H. Zhang, "A clinical study of motor imagery-based brain-computer interface for upper limb robotic rehabilitation," in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, Sept 2009, pp. 5981–5984.

[36] G. Onose, C. Grozea, A. Anghelescu, C. Daia, C. Sinescu, A. Ciurea, T. Spircu, A. Mirea, I. Andone, A. Spânu *et al.*, "On the feasibility of using motor imagery EEG-based brain–computer interface in chronic tetraplegics for assistive robotic arm control: a clinical test and long-term post-trial follow-up," *Spinal cord*, vol. 50, no. 8, pp. 599–608, 2012.

[37] J. P. Donoghue, A. Nurmikko, M. Black, and L. R. Hochberg, "Assistive technology and robotic control using motor cortex ensemble-based neural interface systems in humans with tetraplegia," *The Journal of Physiology*, vol. 579, no. 3, pp. 603–611, 2007. [Online]. Available: http://dx.doi.org/10.1113/jphysiol.2006.127209

[38] D. Coyle, J. Garcia, A. Satti, and T. McGinnity, "Eeg-based continuous control of a game using a 3 channel motor imagery bci: Bci game," in *Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), 2011 IEEE Symposium on*, April 2011, pp. 1–7.

[39] Q. Zhao, L. Zhang, and A. Cichocki, "Eeg-based asynchronous bci control of a car in 3d virtual reality environments," *Chinese Science Bulletin*, vol. 54, no. 1, pp. 78–87, 2009. [Online]. Available: http://dx.doi.org/10.1007/s11434-008-0547-3

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

[41] M. Cheng, X. Gao, S. Gao, and D. Xu, "Design and implementation of a brain-computer interface with high transfer rates," *Biomedical Engineering, IEEE Transactions on*, vol. 49, no. 10, pp. 1181–1186, Oct 2002.

[42] F. Beverina, G. Palmas, S. Silvoni, F. Piccione, and S. Giove, "User adaptive BCIs: SSVEP and P300 based interfaces." *PsychNology Journal*, vol. 1, no. 4, pp. 331–354, 2003.

[43] G. Muller-Putz and G. Pfurtscheller, "Control of an Electrical Prosthesis With an SSVEP-Based BCI," *Biomedical Engineering, IEEE Transactions on*, vol. 55, no. 1, pp. 361–364, Jan 2008.

[44] G. Calhoun and G. McMillan, "EEG-based control for human-computer interaction," in *Human Interaction with Complex Systems, 1996. HICS '96. Proceedings., Third Annual Symposium on*, Aug 1996, pp. 4–9.

[45] M. Middendorf, G. McMillan, G. Calhoun, K. S. Jones *et al.*, "Brain-computer interfaces based on the steady-state visual-evoked response," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 211–214, 2000.

[46] "Emotiv EPOC Brain Computer Interface and Scientific Contextual EEG," https://emotiv.com/product-specs/Emotiv%20EPOC%20Specifications%202014.pdf, accessed: 2015-01-28.

[47] M. Duvinage, T. Castermans, T. Dutoit, M. Petieau, T. Hoellinger, C. D. Saedeleer, K. Seetharaman, and G. Cheron, "A P300-based quantitative comparison between the Emotiv Epoc headset and a medical EEG device," *Biomedical Engineering*, vol. 765, pp. 2012–764, 2012.

[48] A. Campbell, T. Choudhury, S. Hu, H. Lu, M. K. Mukerjee, M. Rabbi, and R. D. Raizada, "NeuroPhone: brain-mobile phone interface using a wireless EEG headset," in *Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds.* ACM, 2010, pp. 3–8.

[49] I. Rennes, "Discover OpenViBE," OpenViBE, 2014, http://openvibe.inria.fr/discover/ [2015-02-06].

[50] Python, "Our Documentation - Python," Python, 2014, https://www.python.org/doc/ [2015-02-06].

[51] I. Rennes, "MATLAB - The Languague of technical computing," Mathworks, 2014, http://se.mathworks.com/products/matlab/ [2015-02-06].

[52] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[53] J. Ye, R. Janardan, and Q. Li, "Two-dimensional linear discriminant analysis," in *Advances in neural information processing systems*, 2004, pp. 1569–1576.

[54] N. Cristianini and J. Shawe-Taylor, *An introduction to support vector machines and other kernel-based learning methods.* Cambridge university press, 2000.

[55] A. M. Martinez and A. Kak, "PCA versus LDA," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 2, pp. 228–233, Feb 2001.

[56] J. Ye, R. Janardan, and Q. Li, "Two-dimensional linear discriminant analysis," in *Advances in neural information processing systems*, 2004, pp. 1569–1576.

[57] MATLAB, *version 8.2.0.701 (R2013b).* Natick, Massachusetts: The MathWorks Inc., 2013, http://se.mathworks.com/help/stats/discriminant-analysis.html?refresh=true.

[58] ——, *version 8.2.0.701 (R2013b).* Natick, Massachusetts: The MathWorks Inc., 2013, http://se.mathworks.com/help/stats/svmclassify.html.

[59] "OpenCV - Introduction to Support Vector Machines," http://docs.opencv.org/doc/tutorials/ml/introduction_to_svm/introduction_to_svm.html#introduction-to-support-vector-machines, accessed: 2015–02—03.

[60] E. Franz, "Development of a Fast and Efficient Algorithm for P300 Event Related Potential Detection," Ph.D. dissertation, Temple University, 2014.

[61] B. Rivet, H. Cecotti, R. Phlypo, O. Bertrand, E. Maby, and J. Mattout, "EEG sensor selection by sparse spatial filtering in P300 speller brain-computer interface," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, Aug 2010, pp. 5379–5382.

[62] R. Alkhater, "Real-time detection of P300 brain events: brain-computer interfaces for EEG-based communication aids," Ph.D. dissertation, Auckland University of Technology, 2012.

[63] C. Bugli and P. Lambert, "Comparison between Principal Component Analysis and Independent Component Analysis in Electroencephalograms Modelling," *Biometrical Journal*, vol. 49, no. 2, pp. 312–327, 2007. [Online]. Available: http://dx.doi.org/10.1002/bimj.200510285

[64] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Networks*, vol. 13, no. 4–5, pp. 411 – 430, 2000. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0893608000000265

[65] P. Comon, "Independent component analysis, A new concept?" *Signal Processing*, vol. 36, no. 3, pp. 287 – 314, 1994, higher Order Statistics. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0165168494900299

[66] M. Richardson, "Principal component analysis," *URL: http://people. maths. ox. ac. uk/richardsonm/SignalProcPCA. pdf (last access: 3.5. 2013). Aleš Hladnik Dr., Ass. Prof., Chair of Information and Graphic Arts Technology, Faculty of Natural Sciences and Engineering, University of Ljubljana, Slovenia ales. hladnik@ ntf. uni-lj. si*, 2009.

[67] D. C. Lay, *Linear Algebra and its Applications*, third edition ed. Addison-Wesley, 2006, p500-504.

[68] J. Shlens, "A tutorial on principal component analysis," *arXiv preprint arXiv:1404.1100*, 2014.

[69] D. D. Wackerly, W. M. III, and R. L. Scheaffer, *Mathematical Statistics with Applications*, sixth edition ed. Duxbury Advanced Series, 2002, p496-498.