KU LEUVEN

Spatiotemporal beamforming for decoding motion-onset Visual Evoked Potentials: a BCI study.

Arno Libert

Thesis submitted for the degree of
Master of Science in Artificial Intelligence,
option Engineering and Computer Science

Thesis supervisor:
Prof. dr. ir. M. Van Hulle

Assessors:
Prof. L. Geurts
B. Wittevrongel

Mentors:
B. Wittevrongel

Academic year 2017-2018

Preface

I would like to thank my professors from the previous year for giving me the opportunity to learn so much in so little time. I also want to thank my promotor and daily supervisor for assisting me in the creation of this work. Thanks go out to the subjects of my experiments without whom this work would be a lot shorter. Special thanks go to my parents for sacrificing so much for the education of me and my brothers. Finally, thanks to my friends for putting up with my asocial behaviour during the completion of this work.

Arno Libert

"Turns out there is such a thing as a crystal ball. It is simply shaped as a computer." – unknown.

Contents

Abstract

Patients suffering from amyotrophic lateral sclerosis (ALS), other muscular and neurodegenerative diseases or spinal cord injuries have difficulty communicating their intention to people. The recent developments in the field of technology and artificial intelligence give the patients a method to express themselves using brain-computer interfacing (BCI). BCI is aimed at establishing a direct communication channel between the brain and an external device (often a computer) bypassing the need for muscular control, including speech and gestures. One of the less common electroencephalography (EEG)-based paradigms relies on the motion-onset visual evoked potentials (mVEP), a lateralized potential over the parieto-occipital scalp area evoked in sync with the onset of a briefly moving stimulus to which the subject pays attention. In this work, the effect of the stimulus direction, the combination of the moving stimulus with the P300 paradigm and the feasibility for the use of mVEP in BCI is investigated in terms of both scalp activations and decoding performance. For the first time, a novel spatiotemporal beamforming algorithm is utilized, and its performance is compared to that of existing mVEP decoders. While a clear effect of stimulus direction on scalp activation is observed, this does not seem to be the case for BCI performance for any of the considered decoders. The combination of the mVEP and P300 paradigm does not show any influence on the scalp activation. This, combined with the increase in fatigue for the subject by focusing on flashing stimuli, makes the combination of mVEP and P300 unadvisable. Spatiotemporal beamforming has the advantage of being significantly less computationally expensive, and in this way more suitable for online BCI applications. Finally, a mVEP spelling console BCI was developed and tested. Test results show a large sensitivity to peripheral stimulation for mVEP. Thus, making the designed layout of the spelling interface a key factor in the creation of mVEP based BCI.

List of figures and tables

List of figures

List of tables

List of abbreviations and symbols

List of abbreviations

ALS: amyotrophic lateral sclerosis
BCI: brain-computer interface/interfacing
EEG: electroencephalography
mVEP: motion-onset visual evoked potential
ITR: information transfer rate
ERP: event-related potentials
CPU: central processing unit
AP: action potentials
ATP: adenosine triphosphate
Hz: Hertz
VEP: visual evoked potential
t-VEP: temporal VEP
f-VEP: frequential VEP
c-VEP: code-modulated VEP
SVM: support vector machines
MDF: modified finite newton
LDA: linear discriminant analysis
LDF: linear discriminant function
SWLDA: stepwise LDA
PP: peak picking
LCMV: linearly constrained minimum variance
stBF: spatiotemporal beamformer
PCA: principal component analysis
SNR: signal-to-noise ratio
ECoG: electrocorticography
PSD: power spectral density
ERD: event-related desynchronization
ERS: event-related synchronization
CAR: common average reference
MEG: magnetoencephalography
fMRI: functional magnetic resonance imaging
BOLD: blood oxygenation level dependent
NIRS: near-infrared spectroscopy
LT: left translation
RT: right translation

List of symbols

1.1.4. Transport of the action potential
V: membrane potential
V_{Na} : sodium equilibrium potential
V_K: potassium equilibrium potential

1.3.1. Support vector machines
ξ: penalty
x_i: training sample i
t_i: training label of training sample i
y: classifier of the form $y = w * x + b$

C: regularization parameter
β: function of (w, b), newton point, proposed solution
m: number of channels
n: number of datapoints per channel
r: number of classes

1.3.2. Stepwise linear discriminant analysis
$N_r(\mu_i, \Sigma_i)$: multivariate Gaussian with mean vector μ_i and covariance matrix Σ_i
x: observation
π_i: probability of x belonging to class i
U: Fisher's linear discriminant function
m: number of channels
n: number of datapoints per channel
r: number of classes

1.3.3. Spatiotemporal beamformer
w_{sp}: spatial filter
S: EEG signal
A: activation pattern
a_{sp}: spatial activation pattern
m: number of channels
n: number of datapoints per channel
r: number of classes
$X \epsilon \mathbb{R}^{(mn)xr}$: column wise flattened version of EEG trials S
$\Sigma \epsilon \mathbb{R}^{(mn)x(mn)}$: covariance matrix of X
$a \epsilon \mathbb{R}^{(mn)x1}$: column wise flattened version of A

1.4.3. Information transfer rate
M: number of possible selections per minute
N: number of available targets/classes
P: accuracy

1. Introduction

Patients suffering from amyotrophic lateral sclerosis (ALS), other muscular and neurodegenerative diseases or spinal cord injuries have difficulty communicating their intention to people. The recent developments in the field of technology and artificial intelligence give the patients a method to express themselves using brain-computer interfacing (BCI).

Brain-computer interfacing has been witnessing growing interest in recent years as the applications are broad: from a communication tool [1] for the disabled to an augmented control facility in gaming [2], [3]. The most popular way to gauge brain activity is with electroencephalography (EEG) and for this non-invasive type of BCI several paradigms have been introduced [3]–[7]. A lesser known paradigm is based on the motion-onset visual evoked potential (mVEP), evoked by a suddenly and briefly moving target. It has been used to classify stimulus responses into target/non-target ones and in this way decode the user's intention [8]–[11]. Utilizing a spatiotemporal beamformer, an advanced classification technique previously used for the classification of other paradigms [12]–[16], an on-line spelling console was created. An advantage to the use of mVEP lies in the fact that, unlike other paradigms, the stimuli required to elicit mVEP is not as taxing on the eyes allowing the BCIs to be used for prolonged periods of time.

In previous works at other research groups, such a spelling console has already been developed [10], [11], [17], [18]. The goal of this study is to create a similar BCI and compare the information transfer rate (ITR) to that of its predecessors.

In the following pages, the development steps of the on-line spelling console BCI using mVEP and spatiotemporal beamforming is described. First, the technical background and terms used are explained in detail. Second, the materials and methods for data collection and signal processing are described. Third, the results of the experiments are analyzed. Last, these results are discussed and compared to other similar works.

In this section, a detailed explanation of all aspects utilized in this work will be given. The first section attempts to give a basic overview of the workings of the brain. The second section will focus on the variations of event-related potentials (ERP) and the stimuli required to generate the response. The third section explains the classifiers used. The fourth section clarifies the definition of brain computer interfaces and the current day standards.

1.1. Neural Background

In computer science, a central processing unit (CPU) is the most vital part of any working system. The CPU controls workflow, ordering and ensures any circuit in the system receives the correct input and manages that circuits output. A human can be described as an incredibly complex system, which in turn requires an absurdly complex CPU: the brain.

The brain is unarguably the most important organ in any complex organism. The organ controls the needs of the organism, creates a response for sensorial input, memory, emotions, … Researchers have only scratched the surface of the workings of the

complex organ and most functions and interconnections happening in the brain remain a scientific mystery.

In this section, the setup of the brain will be explained starting from the workings of a neuron to the generation of action potentials responsible for ERP. All information is taken from [19].

1.1.1. Neurons

If we stick to our analogy of the brain as a CPU, we could simplify a nerve cell or neuron as being one of the many transistors of the CPU. Neurons receive, conduct and transmit signals. Each neuron has a specific role in the nervous system, receiving and generating electrical signals from a variation of sources. The meaning and significance of a signal is dependent on the role of the cell. Several different neuron types and their signals are:

- motor neuron: signals are commands for muscle contraction;
- sensory neuron: signals are information of a stimulus type (e.g. light, force, heat, …) at a certain location;
- interneuron: signals are part of the computation in the brain, generally a combination of sensory information and the generation of motor commands.

In the follow section a basic explanation of neural functions is given. Starting with the layout of the cells and ending with the means used to communicate between each other.

1.1.2. Sketching a nerve cell

A neuron has three determining features: the cell body or soma, dendrites and an axon. The cell body is the biosynthetic center of the cell, containing the nucleus and almost all ribosomes. The dendrites and the axon are nerve fiber cells. The dendrites are branching, tubular cells which can be compared to antennae to create a large surface area for the reception of signal from other cells. The axon, which is generally single and longer than the dendrites, conducts action potentials (AP) away from the soma to distant targets. When the axon gets closer to its destination, it splits up into many axon terminals to distribute the signal to many destinations simultaneously.

In Figure 1 an image of a neuron is shown. On the image the most key features of a nerve cell are displayed. Several other components shown are the myelin sheaths formed by glial cells (supportive cells) to form an insulator around the axon. The nodes of Ranvier are gaps in the myelin sheaths where the axon can be influenced by surrounding cell fluids. Myelination has two distinct advantages:

1. action potentials travel faster along the axon and
2. energy conservation (only need to be generated in the nodes of Ranvier).

Conduction speed alongside an axon is correlated with the thickness of the axon. The signal being transported alongside the axon “jumps” from node to node.

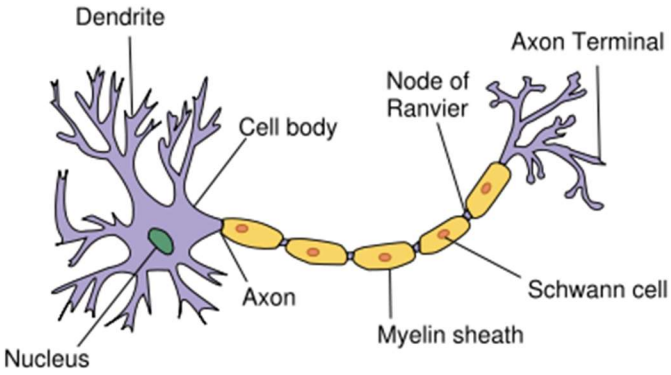

Figure 1: typical neuron with features indicated [20].

1.1.3. Signal propagation

When the action potential arrives at an axon terminal, more precisely the synapses of the axon terminal, it will cause a release of neurotransmitter molecules. The neurotransmitter is picked up by the dendrite of a nearby neuron and cause a new action potential, thus passing the signal on from cell to cell. In this section, a brief explanation of the transport of the action potential and chemical communication between cells will be given.

1.1.4. Transport of the action potential

The action potential enters a cell either through an external input (e.g. light-sensitive neurons with incident light) or through another neuron. The signal enters at the dendrite and must now travel alongside the membrane to the axon terminal to pass on the electrical pulse. How a neuron achieves the conduction requires a deeper insight into the cell membrane of a neuron.
The cell membrane is riddled with channels permeable to a certain inorganic ion. A change in charge distributions alongside this membrane will cause a voltage difference or membrane potential. An increase in membrane potential is called depolarization, while a decrease is called hyperpolarization. The neural signaling depends on channels of which permeability is regulated by gated channels. There exist two types of gated channels: voltage-gated and ligand-gated. The voltage gated channels are required for the passing of the action potential. The ligand-gated channels are needed for the synaptic communication between cells.

In figure 2 the driving force behind the signal propagation is depicted: the Na^+-K^+ pump. The Na^+-K^+ pump pumps Na^+ out and K^+ in the cell using metabolic energy (ATP). Which causes the concentration of K^+ to be much larger inside than outside the neuron and the concentration of Na^+ to be much larger outside the cell than inside. The difference in concentration causes a want for an influx of $Na^+$and an efflux of K^+. The state at which the net current flow of ions is termed the resting potential. A neuron at rest is kept at rest by the Na^+-K^+ pump since there exist leak channels through which ions escape the cell. The flow of ions is driven by a combination of a voltage gradient across a membrane and an ion concentration gradient. This electrochemical gradient is 0 if there is no net flow of ions through the channels and is determined by the Nernst equation which is:

- 50 … 65 mV in case of Na^+ equilibrium potential (V_{Na}),
- -70 … -100 mV in case of K^+ equilibrium potential (V_K).

The resulting net driving force, given a membrane potential V, is:

- $V - V_{Na}$, driving Na^+ into the cell and
- $V - V_K$, driving K^+ out of the cell.

Together with the conductance of the membrane channels this net driving force controls the actual flow of each ion through the membrane. The conductance of a membrane is the ease by which an ion can pass through the membrane, comparable to a resistance for electrical current.

Depolarization causes Na^+ channels to first open and afterwards become inactivated. The Na^+ channels are originally closed at resting potential. When the membrane potential rises, voltage gated Na^+channels open resulting in an influx of Na^+ ions. After approximately 0.5ms the channels shut again, even when the membrane is still depolarized. The Na^+channels are now in an inactivated stage and will remain so until either the membrane potential returns to resting state or a few milliseconds have passed. The influx of Na^+ ions causes the membrane to depolarize further, which in turn opens more voltage gated Na^+channels. The cycle continues until the Na^+ equilibrium is reached removing the drive for influx of Na^+. The voltage-gated K^+ channels are also opened when a membrane depolarization occurs, but open slower than their Na^+ counterparts. The K^+ channels cause an efflux of K^+ ions thus lowering the membrane potential and causing the Na^+channels to recover faster from their inactive state.

The action potential is caused by the depolarization of the membrane when a certain threshold of activity is crossed. If the threshold is not surpassed, no Na^+channels will open, and the charge of the membrane potential will passively spread over the membrane. The action potential provides rapid long-distance communication. The action potential is regenerated throughout its journey to the axon terminal. The entire process is depicted in Figure 3.

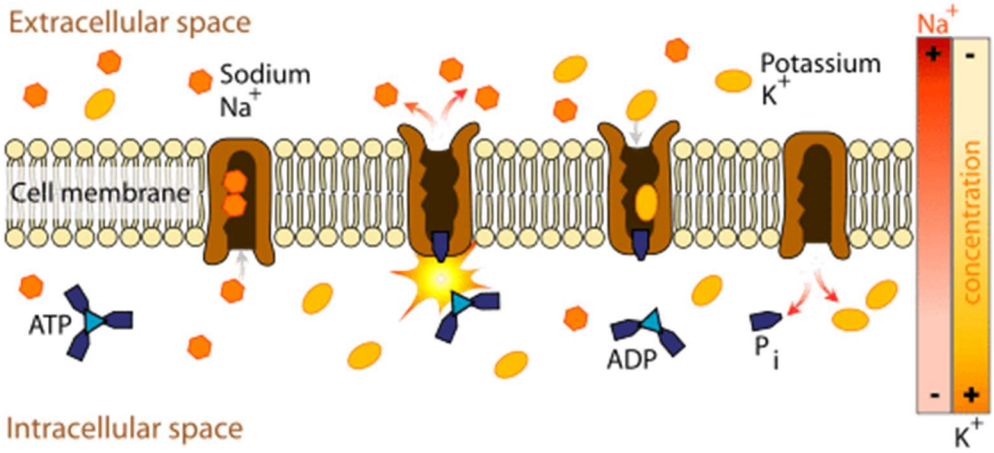

Figure 2: visual depiction of the Na^+-K^+ pump [21].

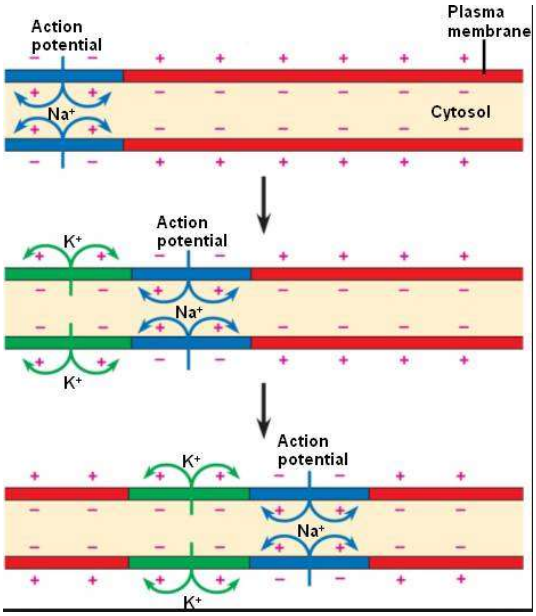

Figure 3: visual depiction of the propagation of an action potential alongside the cell membrane [22].

1.1.5. Extracellular

The action potential has reached the axon terminal and now needs to be passed on to the neighboring neuron. The neurons are not connected however and there is need of a new signaling system to pass on the information. The axon terminal has a different type of channel: Ca^+ voltage-gated channels. The Ca^+ channels differ from the Na^+channels in the following ways:

1. Permeable to Ca^+,
2. Not as easily inactivated.

The axon terminal also contains synaptic vesicles, containing neurotransmitter molecules.

The action potential depolarizes the axon terminal as it reached the end of the axon. The axon terminal contains Ca^+ voltage-gated channels which now open. The influx of Ca^+into the cell trigger the exocytotic release of a large amount of neurotransmitter molecules. The neurotransmitter is released in the synaptic cleft which is the space between an axon terminal of neuron A and a dendrite of neuron B. The dendrites of neuron B has ligand-gated channels which open when one of the neurotransmitter molecules binds to it. The ligand-gated channels allow a specific ion type (e.g. Na^+) to pass through. The depolarization of the cell membrane has now begun. This is depicted in Figure 4.

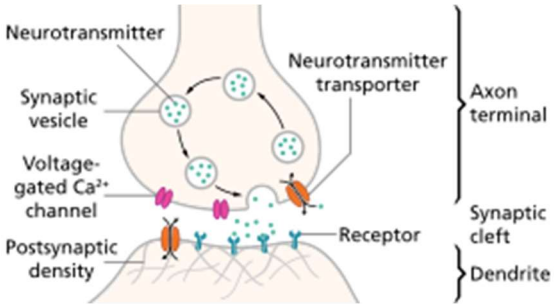

Figure 4: detailed view of the synapse at the end of an axon of a neuron, depicted is the transfer of an action potential to a second neuron [23].

The action potentials in massive quantities (millions of neurons in close proximity firing simultaneously) will be measured as a brain wave by EEG. To note is that this was a very simple and limited explanation of how this complex process works, the entire description with all details is outside the limits of this work. For more information, please consult [19].

1.1.6. Brain waves

The neurons firing action potentials create a brain wave. The brain waves, as other waves, are measured in cycles per seconds (Hertz (Hz)) or frequency of activity. Originally, researchers identified 4 types of brain waves and described their functions depicted in table 1 [24], [25]. The original waves were later divided into several other types [26]. An important fact to note here is that the classification of brain waves is more a convenience, brainwaves are not separate. All brain waves can be measured at any point, but the state of consciousness determines which one will be dominant. A list of supposed methods to alter the dominance of brain waves or “train” your brain to increase a certain type of brain wave is given in [27].

Table 1: different brain waves, the frequencies of the brain waves and the function of the brain waves

Brain Wave	Effects on Brain Function
Delta (0.5 50 4 Hz)	Are predominant during sleep Should be low while awake Repair the brain Serve as emotional radar Are responsible for intuition and unconscious thought In abundance, can interfere with emotional or cognitive processing
Theta (4 to 8 Hz)	Present during pre-sleep or trance state Promote insight and meditation In abundance, can create inattentiveness, distractibility, and lack of focus
Alpha (8 to 12 Hz)	Promote relaxation Serve as gateway for restorative sleep In abundance, can make you spacey, unmotivated, inattentive, and depressed
SMR (12 to 15 Hz)	Are related to calm external attention Regulate impulsivity and hyperactivity Promote body awareness Help control anxiety and anger Promote the inhibition of movement
Beta (15 to 20 Hz)	Are related to active external attention Enhance cognitive processing Improve concentration, attentiveness, focus
High Beta (20 to 36 Hz)	Are related to body tension Promote a high state of arousal Result in excitement, anxiety, stress Related to Post Traumatic Stress Disorder (PTSD)
Gamma (36 to 64 Hz)	Are linked to intellectual comprehension Are related to creativity Promote integrative thinking

1.2. Variations of event-related potentials

1.2.1. Introduction and definition

Two facts have been established in the previous section. First, the brain waves resulting from action potential firing of neurons can be measured by EEG recording equipment. Second, certain brain waves contain information about the mental state and environmental stimuli. These statements alone do not help us much in the search for an EEG measurement viable for BCI classification. The most interesting cases for BCI are the responses time-locked to an easily reproducible event. The responses were termed event-related potentials. The reproducible events can be cognitive control operations, affective operations or memory-related operation [28]–[30]. The systems are distributed in various parts of the brain and are activated in different time intervals. The ERPs are measured using the same equipment as EEG, their meaning however is fundamentally

different. Visually evoked ERPs are a completely non-invasive, objective, and inexpensive investigation, which provides information about early functional changes of the visual pathway and visual brain cortex (sometimes recognizable prior to detectable morphological changes observed by imaging techniques). The interest in research for clinical diagnostics has increased in recent years after dropping post introduction of the MRI [31].

As previously stated, the equipment for recording ERP and EEG is the same. Amplitudes of raw EEGs can be very large (mainly due to slow drifts), larger than those seen in ERP (as ERP's are filtered). The resulting signal-to-noise ratio is small for a single trial. To improve to SNR, EEG fragments or epochs are averaged over many trials. An epoch is part of the whole recorded EEG data consisting of a pre- and post-stimulus part. An ERP is obtained by averaging the epochs [29], [30]. ERP paradigms can be classified into large categories based on the system under study. The classes are:

- sensory tasks,
- motor tasks,
- attention tasks,
- emotion tasks,
- memory tasks,
- and cognitive control tasks.

ERPs are elicited by several methods often used in research: dichotic listening tasks (different auditory tones are played), passive oddball tasks ("target" stimulus appears in a series of non-target stimuli), presenting the stimuli at different intensities (louder auditory cue or larger visual intensity), two-click paradigm (two auditory click are played), …Generally speaking, ERPs are modulated by [29], [30], [32]–[39]:

- modality (e.g. visual, auditory, somatosensory),
- physical features within the modality (e.g. position, spatial frequency, color, orientation),
- stimulus category (e.g. living objects vs. nonliving objects, faces vs tools),
- rate of stimulus presentation (e.g. varying stimulus intervals),
- stimulus intensity (e.g. loudness of the acoustic stimulus, contrast).

Each ERP occurs more dominantly in the part of the brain consistent with its function. An overview of the (right) side of the brain is given in Figure 5. Any references to brain regions are referring to this image.

In this section, an overview of several visually-induced ERPs is given together with a brief explanation of their meaning and stimuli and any applications or diagnostic meaning they may have. The ERPs discussed are P300 and several variants of visually evoked potentials. Finally, it is important to note that all ERP variants are variable in either latency or magnitude for different subjects.

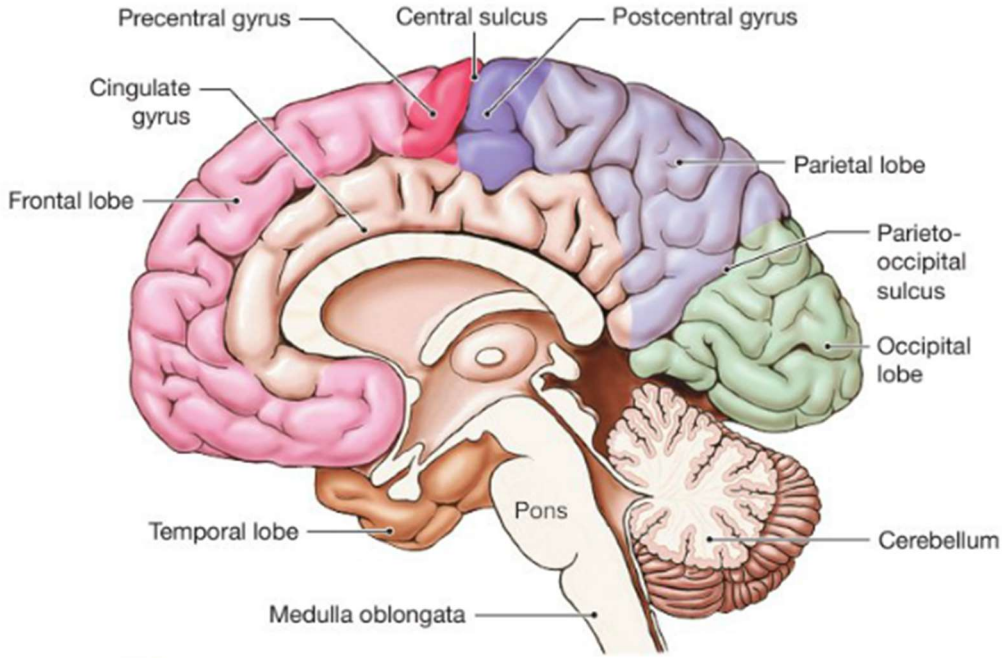

Figure 5: depiction of the right side of the brain with several locations indicated [40].

1.2.2. The P300

The P300, first named so in [41], is parietocentral positivity which occurs when a subject detects and informative task-relevant stimulus [41]–[43]. The name spawns from the occurrence of the peak post-stimulus, i.e. the ERP is a positive peak elicited 300 ms post-stimulus in a young adult subject. The P300 can be triggered by both visual and auditory stimuli [42], [44], [45]. The P300 is most commonly investigated with "oddball" paradigms. Target-related responses are elicited in the parietal cortex, novelty-related activations mainly in the inferior parietal and prefrontal regions. Stimulus modality-specific contributions come from the inferior temporal and superior parietal cortex for the visual and from the superior temporal cortex for the auditory modality [44]. The P300 wave may represent transfer of information to consciousness [42]. A visual representation of the P300 is given in Figure 6.

The amplitude of the P300 increases with lower probability and higher discriminability of targets. The latency increases when targets are harder to discriminate from standards but not when response times increase for other reasons [42], [44], [45]. As a direct consequence of this fact, the P300 has become a tool to separate the mental chronometry of stimulus evaluation from response selection and execution [30], [44], [46], [47]. The P300 amplitude decreases with age and is also lower in patients with decreased cognitive abilities [8], [42], [48].

In diagnostic research it has been attempted to link the P300 response to genetic diseases and psychological disorders [29], [42], [44], [48]–[51]. The P300 has been reported to change in patients diagnosed with epilepsy [29], [50], schizophrenia [29], [50], depression [48], [49] and Parkinson's disease [51]. Another use for the P300 is the

creation of brain computer interfaces for people suffering from motor and/or communication disabilities [4], [6], [52]–[56] and has recently been tested for gaming applications [2]. The P300 generation can be mentally draining on the subject utilizing the BCIs since the subject is required to focus on flashing stimuli which is an unfortunate side-effect.

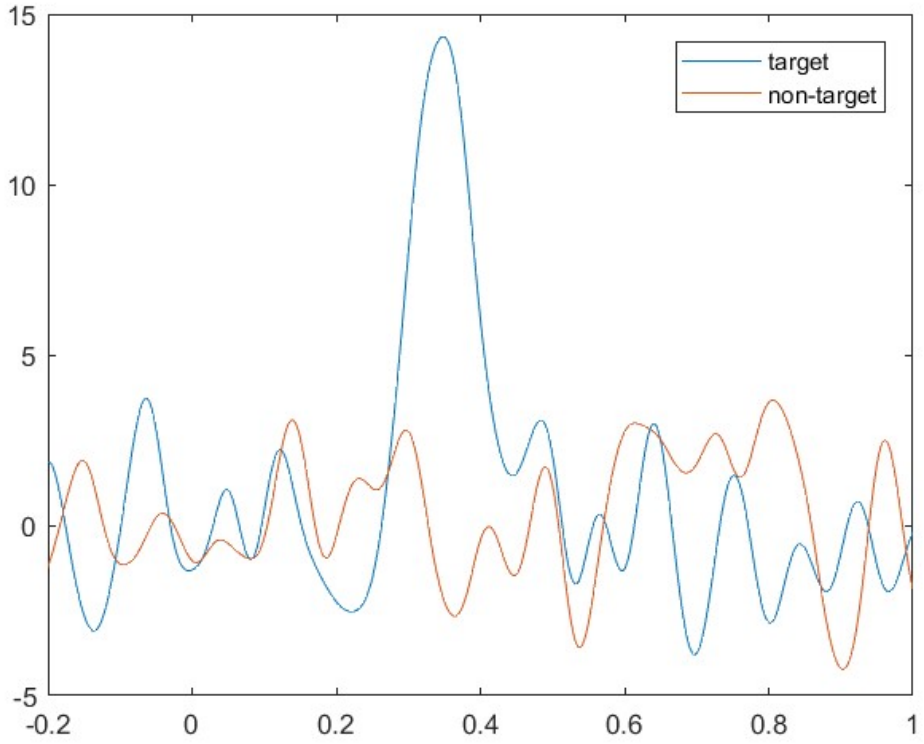

Figure 6: processed EEG recording of the target (blue) and the non-target (red) of an oddball paradigm clearly showing the P300. The x-axis displays the time given stimulus happens at 0, the y-axis displays the amplitude of the signal in microvolts.

1.2.3. The visual evoked potentials (VEP)

Several VEPs are elicited depending on the type of stimulus presented. The different types can be divided into temporal VEP (t-VEP) [9], [57], [58], frequential VEP (f-VEP) [5], [16], [59]–[62], code-modulated VEP (c-VEP) [13], [63]–[65] and motion VEP [2], [8], [9], [11], [17], [18], [31], [66]–[70]. These responses all share the trait that they are subject dependent. An example of each type will be briefly explained in terms of stimuli required to elicit the VEP and some of their applications.

Flash VEPs, a type of t-VEP, are time-locked and phase-locked to flash onsets of gazed stimuli [58]. The response is elicited when a sequence of targets is flashed in a predetermined time pattern. The amplitude of the response is higher when the target is in the foveal visual field of the subject. Several targets, each flashing in a certain temporal pattern, can be flashed with the target being the response with the highest amplitude corresponding to the temporal pattern. An important thing to note is that stimuli must not overlap in order for this to work [9], [57], [58]. The phenomena has been used as a clinical

index to monitor anesthesia during surgical operation [71], [72], to indicate intracranial pressure [73], [74] and to alarm brain death [75], [76].

SSVEP, a type of f-VEP, are responses to a periodic visual stimulus recorded from the occipital area. Traditionally, the stimuli flicker between 6 and 30 Hz due to hardware restrictions (screen refresh rate and EEG bandwidth) [60], [61]. Previously, the targets must flicker at integer dividers of the screen refresh rate (usually 60 Hz) resulting in a limited number of possible targets and resulting information transfer rate [59]. However, several solutions have been offered to combat this limitation. First, a sampled sinusoidal stimulation profile where the stimuli are flickered at the same frequency but controlled by a phase shifted sinusoid. The method used a fuzzy logic classifier specifically designed for circular information [61]. Second, a profile with targets flickering at the same or varying frequency but at a different phase [7], [16], [77]–[79]. Applications are mainly BCI systems (e.g. spelling consoles) [5], [7], [16], [61], [77], [78] but several attempts have been made for integration within gaming [3]. Important to note is the need for exclusion of the alpha waves from the EEG signal as they interfere with the elicited response.

The c-VEP paradigm presents the subject with a sequence of high and low stimulus intensity with varying duty cycles, termed the code. The selectable targets are represented by a unique lagged version of this code [13], [63]–[65]. An often used code is the m-sequence used in [64] which is a pseudorandom binary sequence that has an autocorrelation function close to a unit impulse function and is nearly orthogonal to its time lagged version. Application of c-VEP exist in the creation of BCI systems [13].

Motion visually evoked potentials are a term used for any ERP generated by motion. There exist several documented variations of the mVEP [31]. The most prominent and promising variant is the motion-onset VEP, having the largest amplitudes and most consistent recording conditions [11], [31], [66], [68], [69]. The counterparts of the motion-onset VEP (mVEP refers to this from this point) is the motion-offset VEP which require a longer duration of motion which in turn causes adaptation [31], [80]. Motion-reversal VEPs represent responses to motion direction changes. The inter-subject variability of their shape makes them unusable for clear identification and classification [31], [52]. Steady-state motion-related VEPs using continuously moving stimuli also have a large inter-subject variability [31], [81]. mVEP is elicited when a translating line is presented in the field of focus. The movement must last between 120 and 200 ms followed by an inter stimulus pause lasting 60 ms. The stimulus consists of a negative peak around 200 ms post stimulus followed by a positive peak around 250-300 ms post-stimulus. Reports exists of a third positive peak 100 ms post stimulus [31] which is dominant when the subject becomes adapted to the stimulation sequence. To prevent the adaptation, the stimuli are shown in a random sequence. A visual representation is presented in figure 7. Applications of mVEP are the creation of brain computer interfaces for people suffering from motor and/or communication disabilities [9], [10], [18], [67], [70], [82] and has recently been tested for gaming applications [2], [83].

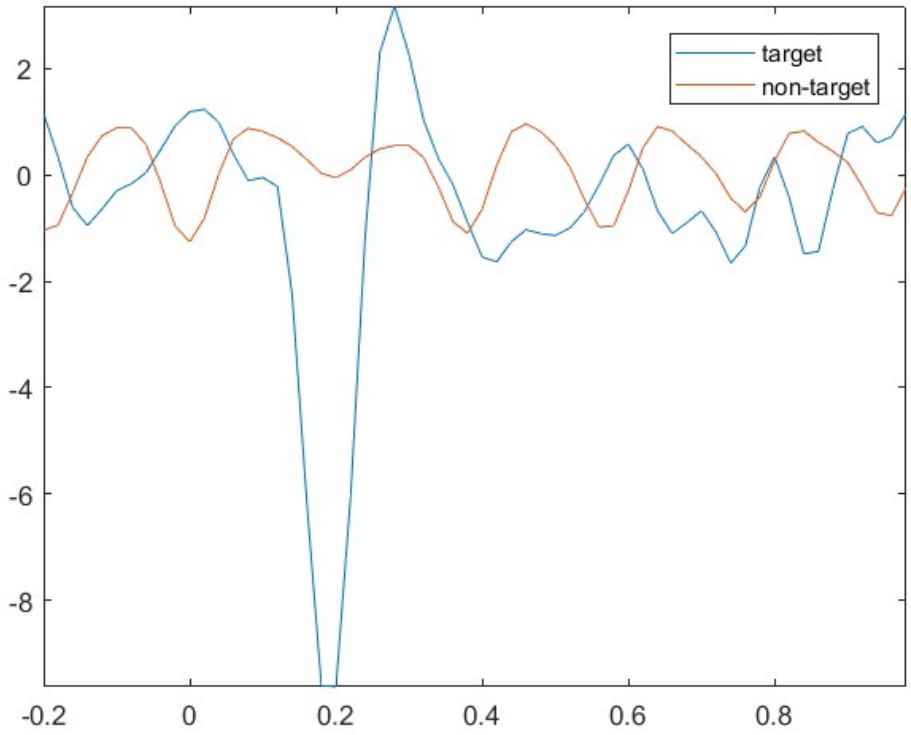

Figure 7: processed EEG recording of the target (blue) and non-target (red) response to motion-onset clearly showing the mVEP. The x-axis displays the time given stimulus happens at 0, the y-axis displays the amplitude of the signal in microvolts.

1.3. Classifiers

When a learning set L of multivariate observation are presented, supposing each observation is part of a predefined class k which has determining characteristics and a unique class label, the classes can be identified and distinguished using a classifier. A classifier requires two stages to be functional:

- **Training:** information is presented in a learning set of labeled observations to create a function (termed classifier) that separates the predefined classes as much as possible.
- **Classification:** new unlabeled observations are compared to the function to predict the class of the observation.

In machine learning, the training and classification operations are named supervised learning techniques. Together they form the task of class prediction.

The VEPs have dominant peaks elicited when the stimulus appears in the field of focus. The peaks can therefore be used to predict the target the subject is focusing on. The selection could be done manually, but the practical relevance of such a task would be limited and time-consuming. The automation of this process is done in the form of various classifiers. All visual BCIs, be it online or offline, utilize a classifier to predict the focus point of the subject. In this section several classifiers are compared for the classification

of mVEP: a support vector machine based on [84], stepwise linear discriminant analysis [85], peak picking and spatiotemporal beamformer [14].

1.3.1. Support vector machines

Linear support vector machines (SVMs) are tools for tackling large-scale data-mining tasks containing large number of examples and features coupled with a parse data matrix [84]. This and the accuracy of the SVM makes it a valid, and often used, classification method for ERP signals [9], [10], [17], [18]. A negative feature of the SVM classifiers is the training time required, making fast algorithms preferable. Linear SVMs attempt to maximize the margin between classes of the data. Any datapoint that breaks this margin is given a penalty (denoted by ξ) and is termed a slack. Two loss functions for imposing penalties on slacks are penalizing slacks linearly (penalty = ξ) named L_1-SVMs and penalizing slacks quadratically (penalty = $\xi^2/2$) named L_2-SVMs. Although L_1-SVMs are more popular since the method yields classifiers with less support vector leading to faster speeds. For linear SVMs the amount of support vectors is nullified when the final classifier is implemented using the weight vector in feature space [84]. The SVM described here is an L_2-SVM.

A binary classification problem with training samples $\{x_i, t_i\}_{i=1}^{m}$ where $x_i \in \mathbb{R}^n$ and $t_i \in \{1,0\}$. To obtain a classifier of the form $y = w * x + b$, an L_2-SVM solves the following problem:

$$\min_{(w,b)} \frac{1}{2} * (\|w\|^2 + b^2) + \frac{C}{2}\sum_i^m \xi_i^2 \text{ given } \forall_i: t_i(w * x_i + b) \geq 1 - \xi_i \quad (1)$$

Where C is the regularization parameter. $\frac{b^2}{2}$ was added so the standard regularized least squares algorithms can be utilized. To apply the least square algorithm, (1) must first be transformed to an equivalent formulation by eliminating ξ_i and dividing the objective function by C:

$$\min_{\beta} f(\beta) = \frac{\lambda}{2} * \|\beta\|^2 + \frac{1}{2} * \sum_{i \epsilon I(\beta)} d_i^2(\beta) \quad (2)$$

Where $\beta = (w, b),\ \lambda = \frac{1}{C}$, $d_i(\beta) = y_i(\beta) - t_i, y_i(\beta) = w * x_i + b, and\ I(\beta) = \{i: t_i y_i(\beta) < 1\}$. This SVM is further enhanced using a modified finite newton (MFN) algorithm [86]. MFN does iterations of the form:

$$\beta_{k+1} = \beta_k + \delta_k p_k \quad (3)$$

where p_k is based on a second order approximation of the objective function at β_k:

$$p_k = -H(\beta_k)^{-1} \nabla f(\beta_k). \quad (4)$$

The step size δ_k is chosen to satisfy an Armijo condition that ensures convergence, it is found by applying an exact line search allowing the direct application of convergence results from nonlinear optimization theory. Traditionally δ_k is found by applying a halving method of line search in the [0,1] interval. Since f is not twice differentiable at β where at least one of the d_i is zero, $H(\beta)$ is taken to be the generalized Hessian defined by $H(\beta) =$

$\lambda * J + C^T DC$ where J is the n x n identity matrix, C is a matrix whose rows are $(x_i^T, 1)$ and D is a diagonal matrix whose diagonal elements are given by:

$$D_{ii} = \begin{cases} 1 & if\ t_i y_i(\beta) < 1 \\ 0 & if\ t_i y_i(\beta) = 1 \\ 0 & if\ t_i y_i(\beta) > 1 \end{cases} \quad (5)$$

The third property contributes greatly to the overall efficiency of the method since the indices satisfying $t_i y_i(\beta) > 1$ do not affect $H(\beta)$ and p_k. The $t_i y_i(\beta) = 1$ cases are set to zero to keep the least squares nature of the problem, traditionally this would be set to a specific element of the interval [0,1]. The change allows the algorithm to calculate the Newton point ($\beta_k + p_k$), i.e. the solution of a regularized least squares problem, instead of the Newton direction (p_k).

Thus, at one iteration, given a point β_k we set $I_k = I(\beta_k)$ and minimize f_{I_k} to obtain the Newton point, $\beta\ (denoted\ as\ \bar{\beta})$. Then a line search is performed in the interval $[\beta_k, \bar{\beta}]$ to yield the next Newton point. A step by step sequence is given below, proof for the convergence of the algorithm together with more details, are given in [84].

1. Choose a suitable starting β_0. Set k = 0, proceed to step 2
2. Check if β_k is the optimal solution of (2). If so, solution is β_k else proceed to step 3.
3. Let $I_k = I(\beta_k).\ Solve$

$$\min_{\beta} f_{I_k}(\beta),$$

   Let $\bar{\beta}$ denote the solution.
4. Perform a line search to decrease the objective function, f:

$$\min_{\beta \epsilon L} f(\beta),$$

   Where L = $\{\beta = \beta_k + \delta(\bar{\beta} - \beta_k) : \delta \geq 0\}$. Let δ^* denote the solution of this line search. Set $\beta_{k+1} = \delta^*(\beta - \beta_k)$ with k = k+1 and return to step 2.

The training of the SVM in this work is achieved by first transforming the labeled data shaped in an (mxn)xr matrix into an rx(mn) matrix, with m the number of recording electrodes, n the number of time samples per trial and r the number of trials in a recording. The training data is then normalized, labeled with a target/non-target label and loaded into the classifier. The output is an (mxn)x1 vector consisting of the solution of the classifier described above. To perform classification on the experiment data, the data is split up into the data corresponding to each possible class (i.e. each direction) and multiplied with the trained vector resulting in a score for each possible class. The winning class is the one with the highest score.

1.3.2. Stepwise linear discriminant analysis

Linear discriminant analysis (LDA) follows Fisher's assumption that both multivariate probability densities are multivariate Gaussian having arbitrary mean vectors and a common covariance matrix. i.e. take $f_1(\cdot)$ to be a $N_r(\mu_1, \Sigma_1)$ density and $f_2(\cdot)$ to be a

$N_r(\mu_2, \Sigma_2)$ density making the homogeneity assumption that $\Sigma_1 = \Sigma_2 = \Sigma_{XX}$. Information for this explanation is taken from [87]. The ratio of the two densities is now given by

$$\frac{f_1(x)}{f_2(x)} = \frac{\exp\{-\frac{1}{2}(x-\mu_1)^T\Sigma_{XX}^{-1}(x-\mu_1)\}}{\exp\{-\frac{1}{2}(x-\mu_2)^T\Sigma_{XX}^{-1}(x-\mu_2)\}}, \tag{6}$$

the logarithm of this function then becomes:

$$\log_e \frac{f_1(x)}{f_2(x)} = (\mu_1 - \mu_2)^T\Sigma_{XX}^{-1}x - \frac{1}{2}(\mu_1 - \mu_2)^T\Sigma_{XX}^{-1}(\mu_1 + \mu_2) \tag{7}$$

$$= (\mu_1 - \mu_2)^T\Sigma_{XX}^{-1}(x - \bar{\mu}), \tag{8}$$

where $\bar{\mu} = \frac{\mu_1+\mu_2}{2}$. Given that the second term of (5) can be rewritten as

$$\frac{1}{2}(\mu_1 - \mu_2)^T\Sigma_{XX}^{-1}(\mu_1 + \mu_2) = \mu_1^T\Sigma_{XX}^{-1}\mu_1 - \mu_2^T\Sigma_{XX}^{-1}\mu_2 \tag{9}$$

it follows that

$$L(x) = \log_e\{\frac{f_1(x)\pi_1}{f_2(x)\pi_2}\} = b_0 + b^Tx \tag{10}$$

Is a linear function of x, where

$$b = (\mu_1 - \mu_2)\Sigma_{XX}^{-1} \text{ and} \tag{11}$$

$$b_0 = -\frac{1}{2}\{\mu_1^T\Sigma_{XX}^{-1}\mu_1 - \mu_2^T\Sigma_{XX}^{-1}\mu_2\} + \log_e\frac{\pi_1}{\pi_2}. \tag{12}$$

Now x is assigned to $\Pi_1$if the logarithm of the ratio of the two posterior probabilities is greater than zero and otherwise assign it to Π_2, Π_1, Π_2 representing two different classes and π_1, π_2 representing the probability of x being a part of the respective class:

$$\begin{cases} if\ L(x) > 0 \rightarrow assign\ x\ to\ \Pi_1 \\ if\ L(x) < 0 \rightarrow assign\ x\ to\ \Pi_2 \end{cases}. \tag{13}$$

Important to note is the boundary $\{x\ \epsilon\ R^T | L(x) = 0\}$, R being the regions occupied by the class. The resulting equation is linear in x defining a hyperplane that divides the two classes. Rule (13) is generally referred to as Gaussian linear discriminant analysis and the part of the function L(x) in (8) that depends upon x,

$$U = b^Tx = (\mu_1 - \mu_2)^T\Sigma_{XX}^{-1}x, \tag{14}$$

is known as Fisher's linear discriminant function (LDF) [87]. The previous method works for only two classes, a method able to classify multiple target is required. The sample space must be partitioned into K nonoverlapping regions $R_1, R_2, \ldots, R_k$, such that observation x is assign to class Π_i if $x\ \epsilon\ R_i$. The partition is to be divided as to minimize the number of misclassified instances. Note that k two class linear discriminant analyses will not work since it produces regions not belonging to any of the k classes. Let

$$P(X\ \epsilon\ \Pi_i) = \pi_i\ with\ i\ \epsilon\ [1, K], \tag{15}$$

be the probabilities of a randomly selected observation X belonging to each of the different classes in the population and let

$$p(\Pi_i|x) = P(X \epsilon \Pi_i|X = x) = f_i(x)\ with\ i \epsilon [1, K], \tag{16}$$

be the multivariate probability density for each class. The resulting posterior probability that an observed x belongs to the ith class is now

$$p(\Pi_i|x) = P(X \epsilon \Pi_i|X = x) = \frac{f_i(x)\pi_i}{\sum_{k=1}^{K} f_k(x)\pi_k}. \tag{17}$$

Directly following from (17) is that the highest $f_i(x)\pi_i$ corresponds to the highest probable class. Assuming that for every class Π_i that $f_i(\cdot)$ is the $N_r(\mu_i, \Sigma_i)$ density, where μ_i is an r-vector and Σ_i is an (r x r) covariance matrix and that the covariance matrix for the K classes are identical, $\Sigma_1 = \cdots = \Sigma_k$, and equal to a common covariance matrix Σ_{XX}. The odds that x is assigned to Π_i instead of Π_j now becomes (similar to (8))

$$L_{ij}(x) = \log_e\{\frac{f_i(x)\pi_i}{f_j(x)\pi_j}\} = b_{0i} + b_{ij}^T x, \tag{18}$$

Where

$$b_{ij} = (\mu_i - \mu_j)\Sigma_{XX}^{-1} \text{ and} \tag{19}$$

$$b_0 = -\frac{1}{2}\{\mu_i^T \Sigma_{XX}^{-1} \mu_i - \mu_j^T \Sigma_{XX}^{-1} \mu_j\} + \log_e \frac{\pi_i}{\pi_j}. \tag{20}$$

The regression of a data set is the procedure where one attempts to model an equation to correspond with as many possible points of the given data set. There exists the backward elimination procedure which begins with the largest regression, using all variables, and subsequently reduces the number of variables in the equation until a decision is reached on the equation to use. There is also the forward selection procedure which attempt to achieve a similar conclusion working from the other direction, i.e. to insert variables in turn until the regression equation is satisfactory. The order of insertion is determined by using the partial correlation coefficient as a measure of the importance of variables not yet in the equation. The stepwise regression procedure, on which stepwise linear discriminant analysis (SWLDA) is based [85], is an improvement of the forward selection procedure. The improvements involve the reexamination of the incorporated variables at every stage. The variables added during an early stage may be superfluous at a later stage because of the relationship between it and several other variables now included in the regression.

To calculate the relationship between the variables, an F criterion for each variable in the regression is evaluated at every stage and compared with a preselected percentage point of the appropriate F distribution. This method provides a judgement on the contribution made by each variable as thought it was the most recent variable added into the regression. The variables that do not meet the predetermined F criterion are subsequently removed from the regression model before calculation continues.

A step by step explanation for the procedure used by the algorithm is given below. Note that initialization of the algorithm requires a data set, an entry criterion and an exit criterion [88].

1. The stepwise procedure starts with the simple correlation matrix and enters into the regression equation the X variable most highly correlated with the response, termed X_1.
2. Using the partial correlation coefficients as before, it selects the next X to enter regression by picking the variable whose partial correlation with the response is highest, termed X_2.
3. Given the newly formed regression equation $\hat{Y} = f(X_1, X_2)$, the method now evaluates the contribution X_1 would have had if X_2 were to be added first into the equation. If the F value is found statistically significant, X_1 is retained else it is removed.
4. The algorithm returns to step 2 until no variables remain the meet the entry criterion. Afterwards it terminates and returns the regression equation.

The training of the SWLDA in this work is achieved by first transforming the labeled data shaped in an (mxn)xr matrix into an rx(mn) matrix, with m the number of recording electrodes, n the number of time samples per trial and r the number of trials in a recording. The training data is then normalized, labeled with a target/non-target label and loaded into the classifier. The output is an (mxn)x1 vector consisting of the solution of the classifier described above. To perform classification, the unlabeled data is split up into the data corresponding to each possible label, averaged and multiplied with the trained vector resulting in a score for each possible class. The winning class is the one with the highest score.

1.3.3. Peak picking

Peak picking (PP) is a simple classification method which can be used when a class can be determined as the data containing the dominant negative or positive peak. The method was used in [67] to determine the focus point of the subject in a BCI by examining the most negative and positive peak within the expected temporal window. The algorithm implemented here is trained by calculating the temporal location of the minimum and maximum for the averaged test data and creating a search range with the respective locations as the center. Training is performed by calculating the minimum and maximum in the search range respectively naming the most negative and positive value as the peak corresponding to the mVEP. Classification is performed by searching within the trained range for the most negative and positive values, subtracting the minimum from the maximum and naming the highest value as the winner.

1.3.4. Spatiotemporal Beamformer

A system designed to receive spatially propagating signals often suffer from interference between the signals when the desired signal and interfering signals occupy the same temporal frequency band [89]. Using solely temporal filters, separating the signals would prove impossible. Given a different spatial source of each signal, a spatial filter may be used at the receiver to separate the signals. The design of early spatial filters allowed the formation of pencil beams in order to receive a signal radiating from a specific location and attenuate signals from other locations. This quality was termed “forming beams” [89]. A beamformer is a processor utilized with an array of sensors, collecting discrete spatial samples of propagating wave fields, and provides a versatile form of spatial filtering. Beamformers allow the separation of signals with frequential overlap

which originate from different spatial locations. Beamforming is applicable for the radiation and reception of energy. Spatial filtering utilizing beamformers has been implemented in radar, sonar, communication, imaging and several other fields [89].

Spatial filters can be designed for processing brain electrical activity pass information from a specified location while attenuating activity originating at other electrode locations. The power at the output of a spatial filter is an estimate of the neural power originating within the spatial passband of the filter. The calculation of multiple spatial filters, with a unique passband, allows for the creation of a map of neural power as a function of location. The output of the filter depicts a function of passband location. The spatial filtering method described in [90] is based on a linearly constrained minimum variance (LCMV) filtering.

Motivation for the use of multivariate techniques for EEG analysis lie in the separation of overlapping ERP components. A multivariate filter combines the EEG signals from multiple electrodes into one representative value. Two performance criteria are used as an estimation of the output. The sensitivity, which is a measurement of correlation with the actual amplitude of the ERP, and the specificity, which is a measure for the lack of correlation with structured interfering signals [14]. The multivariate filter used in this study was based on the spatiotemporal LCMV beamformer described in [14]. The spatiotemporal beamformer (stBF) is an extension of the LCMV beamformer to be a spatiotemporal filter for estimating the amplitude of ERP components in sensor space.

Beamformers take as input a spatiotemporal template of the component of interest. The LCMV beamformer was originally formulated a spatial filter $w_{sp} \epsilon \mathbb{R}^{mx1}$ with w representing a vector of the spatial filter and m the number of electrodes used for recording. When applied to the centered EEG signal S, it minimizes the variance of the result $w_{sp}^T S$:

$$w_{sp} = arg \min_{w_{sp}} w_{sp}^T S \left(w_{sp}^T S\right)^T = \arg\min_{w_{sp}} w_{sp}^T \sum_{sp} w_{sp} \tag{21}$$

where $\Sigma_{sp} \epsilon \mathbb{R}^{mxm}$ is the spatial covariance matrix of the signal S. To avoid trivial solutions of (21), a linear constraint on w_{sp} is implemented as follows:

$$a_{sp}^T w_{sp} = 1 \tag{22}$$

where $a_{sp} \epsilon \mathbb{R}^{mx1}$ is the spatial activation pattern. Utilizing the method of Lagrange multipliers, the solution of (21) becomes

$$w_{sp} = \frac{\sum_{sp}^{-1} a_{sp}}{a_{sp}^T \sum_{sp}^{-1} a_{sp}}. \tag{23}$$

To expand the formulation of the LCMV beamformer into a spatiotemporal beamformer, let $X \epsilon \mathbb{R}^{(mn)xr}$ be a matrix consisting of r columns x_i, which are the column wise flattened versions of the corresponding EEG trials S_i (i = 1, …, r), $\Sigma \epsilon \mathbb{R}^{(mn)x(mn)}$ be the covariance matrix of X, and $a \epsilon \mathbb{R}^{(mn)x1}$ be a vector containing the column wise flattened version of the shape of the spatiotemporal activation pattern A. The spatiotemporal beamformer $w \epsilon \mathbb{R}^{(mn)x1}$ is the result of the minimization of the variance of $w^T X$ constrained by $a^T w = 1$:

$$w = \frac{\Sigma^{-1}a}{a^T\Sigma^{-1}a}, \tag{24}$$

and applied to the data as a simple weighted sum:

$$y = sw, \tag{25}$$

where $s \in \mathbb{R}^{1x(mn)}$ indicates the concatenated rows of an epoch S and y represents the contribution of the activation pattern A in S.

The training of the spatiotemporal beamformer in this work is achieved by first transforming the labeled data shaped in an (mxn)xr matrix into an rx(mn) matrix, with m the number of recording electrodes, n the number of time samples per trial and r the number of trials in a recording. The training data is labeled with a target/non-target label and loaded into the classifier. The output are the weights of the beamformer. To perform classification, the unlabeled data is split up into the data corresponding to each possible class, averaged and multiplied with the weights of the beamformer resulting in a score for each possible class. The winning class is the one with the highest score.

1.4. Brain Computer Interfaces

It has been established in previous chapters that many components are required for a brain computer interface to function. An in-depth explanation for the term has not been given.

In this section, a definition for a brain-computer interface will be formulated together with the description of frequently used designs and a method to measure the information transfer rate will be given.

1.4.1. Definition

“Brain computer interfaces (brain machine interfaces, neuroprostheses) aim at creating a direct communication pathway between the brain and an external device, bypassing the need for an embodiment.”- Professor Van Hulle, BCI course at KULEUVEN.

BCI could provide a significant improvement of the quality of life of neurologically impaired patients suffering from amyotrophic lateral sclerosis, stroke, brain/spinal cord injury, muscular dystrophy, etc. In recent years developments have been made to implement BCI in gaming and entertainment. BCI can differ in paradigm, classifier and recording method. The first two have already been described in this work whilst the last will be described here.

1.4.2. Invasive and non-invasive BCI

The difference between invasive and non-invasive BCI lies in the location of the electrodes and the recording surface area. An overview between the differences in given in Figure 8. For invasive recording, electrodes are implanted intracranially resulting in the best signal quality but carries risks associated with the invasive surgical procedure.

Deep Invasive BCI can vary according to the depth of the implanted electrodes and mainly function on spike detection and spike sorting. Spike detection is performed by a variety of supervised or unsupervised method (e.g. window discriminator, power method,

principal component analysis (PCA), wavelets). Spike sorting, which is performed after spike detection, determines the number of captured neurons and assigns the different spikes to the neurons.

The change in field potential, which is a measurement of the reaction of several neurons, was classically achieved by implanting wire electrodes into the brain. A more modern approach is achieved by implanting micro-electrode arrays into the brain (e.g. UTAH array, Michigan Probes and NeuroProbes). Not all electrodes can be read out simultaneously. Hence, electrodes are selected according to preference, Signal-to-noise ratio (SNR) or expert scores [91].

ElectroCorticoGraphy (ECoG) is a partially invasive method where the electrodes are placed directly on the cortical surface. The methods utilized in EEG analysis can also be used on ECoG. Several examples not yet discussed are Power Spectral Density (PSD) or Event-Related Desynchronization/Synchronization (ERD/ERS) analysis. PSD is a measure of how power in a signal changes as a function of frequency and can be used to identify the subjects specific most reactive frequency bands by studying the difference between reference periods and active ones. ERD/ERS studies the relative power of groups of neurons which are (de)synchronously active or inactive. ERD occurs when band power decreases with relation to baseline condition and ERS when band power increases.

Non-invasive BCI most commonly use ElectroEncephaloGram recording. EEG is a measure of the brains voltage fluctuation as detected from scalp electrodes and approximates the cumulative electrical activity of neurons. EEG requires electrodes to be attached to the scalp and have the signal improved by applying conductive gel. The measured signal represents the difference between the voltages at two or more electrodes and hence requires a reference to identify significance (e.g. Single reference montage, Bipolar design, Common average reference (CAR)). The electrode placement is selected according to the regions on the scalp corresponding to the lobes of the brain. An overview of commonly used channels is given in Figure 9. EEG records mainly noise and is a 2-D representation of a 3-D reality which causes problems in the localization of the source of the electrical activity. In order to detect electrical activity of the brain from scalp recordings, it must be of sufficient strength and duration. The previously described action potentials when fired by neurons in close proximity, more precisely the post-synaptic potentials, fit this role perfectly since a small dipole is generated by the external post-synaptic potential. Important to note is that artifacts of other activities (e.g. blinking, muscle contraction) are also measured and must be rejected or corrected in the signal.

MagnetoEncephaloGraphy (MEG) measures the magnetic field surrounding the electrical dipole generated by the neurons in close proximity. The polarity is perpendicular to the current. If the current is running parallel to the scalp, the magnetic field exits the head from one side of the dipole and can be measured. If the current is perpendicular to the scalp, it cannot be measured. MEG is therefore more sensitive to the neurons parallel with the skull and makes source localization easier. The machine used to measure MEG is quiet, comfortable and doesn't require gel but is expensive and heavy, making large scale application difficult.

Functional Magnetic Resonance Imaging (fMRI) measures Blood Oxygenation Level Dependent (BOLD) signals which is an indirect measure of neural activity. When neural activity increases, so does the blood oxygen in this brain area which caused the fMRI

signal to increase. A region of interest is selected, and BOLD signal variance is studied following the presentation of a stimulus. An MRI is loud, heavy and expensive combined with a suboptimal temporal resolution (1 sample every few seconds) makes this method not feasible for large scale application.

Near-InfraRed Spectroscopy (NIRS) is based on the near-infrared range of light (between 700 and 1000 nm) which can pass through the skin, bone and other tissue relatively easily. Oxygenated and deoxygenated hemoglobin and the mitochondrial enzyme cytochrome oxidase have absorption bands within this range. NIRS measures differences in oxygenated hemoglobin, reduced hemoglobin, total hemoglobin and cytochrome oxidase concentration. If a tracer is utilized, measurement of cerebral blood flow and cerebral blood volume is possible. Advantages of NIRS is the measurement of blood oxygenation whilst the disadvantages lie in its difficulty in use, limited recording time, suboptimal temporal resolution and sensitivity to movement artefacts.

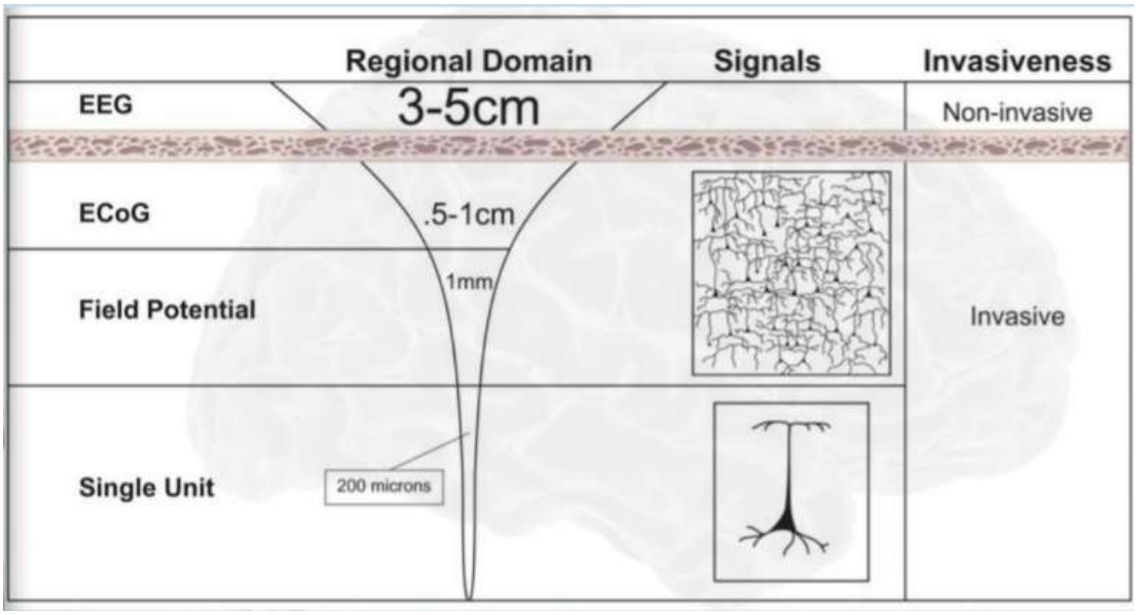

Figure 8: comparison between invasive and non-invasive recording methods showing depth, surface measured and signals [92].

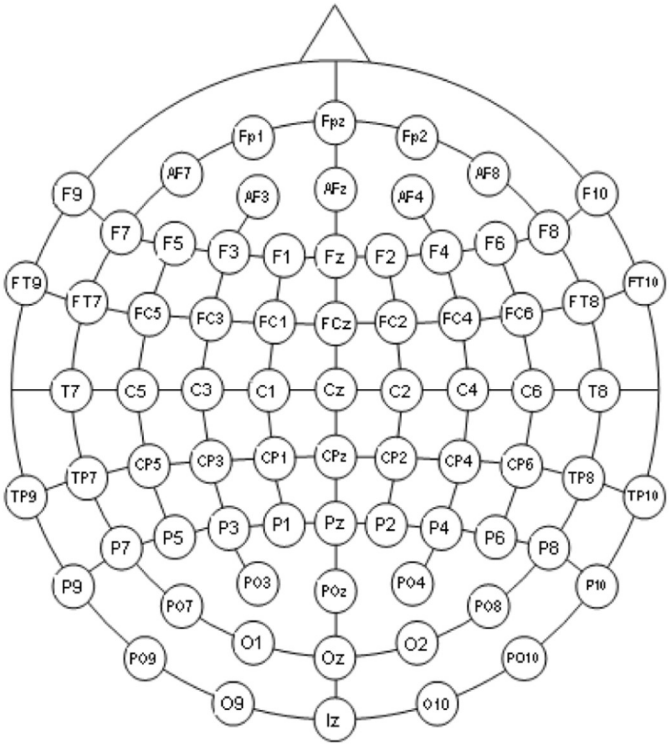

Figure 9: traditional layout of scalp recording sites for EEG [93].

1.4.3. Information transfer rate

A BCI transfers information from the brain to a computer. The Information Transfer Rate is a measurement of the number of accurate bits that can be transferred from brain to computer within a time interval. ITR is calculated using the formula [5], [58], [78]:

$$ITR = M * (\log_2 N + P * \log_2 P + (1-P) * \log_2 \frac{1-P}{N-1}),$$

with M the number of possible selections per minute, N the number of available targets and P the accuracy. The higher the value of ITR, the more information transferred per time interval. It is classically represented in bits/sec or bits/min.

1.5. Conclusion

In the past chapter, all required information to understand the technicalities of this work were described. First, a description of the generation of action potentials and the brain waves resulting from these action potentials. Second, various ERPs were described together with some examples of applications. Third, the classifiers utilized in this work were explained. Fourth, various possible BCIs were given.

Neurons can be exited both by other neurons and external stimuli. When a large number of neurons fire, a brain wave is generated which has a different meaning depending on the frequency of the wave. The post synaptic potential can be measured by several non-invasive measurements methods.

Many different types of VEP exists which are elicited following different stimuli. The stimuli are motion, code, frequency and time. The P300 is a different type of ERP caused by an oddball effect and is widely used in different BCI. The VEPs have different shapes and peaks which can be used for classification.

Many classifiers exist, the ones described in this work are stBF, SVM, SWLDA and PP. Classification is the procedure of studying labeled data to create a model and using this model to predict the label of unlabeled data.

BCI can vary in diverse ways: classification method, ERP, invasive and non-invasive. Invasive methods generate a better signal but have the risk of invasive surgery whilst non-invasive are easier to use but result in a noisy signal.

This work will focus on an EEG BCI comparing stBF, SVM, SWLDA and PP as classification methods for the mVEP and ITR will be used as a comparison with several other mVEP BCI.

2. Materials and methods

2.1. Introduction

The goal of this work was the creation of an on-line spelling console BCI based on mVEP utilizing the spatiotemporal beamformer classifier. To make this achievable a few questions had to be answered:

1. Can mVEP be classified by the spatiotemporal beamformer?
2. Does the movement direction influence the results?
3. Does adding other ERPs influence the results?

When an answer to these questions was formulated, the on-line spelling console was developed and tested. The materials and methods used to collect and analyze data are described in this chapter. Each section will answer the questions above respectively ending with the implementation of the spelling console.

2.2. Subjects

For the three experiments 12 healthy subjects (7 female, 5 male, average 25.33 ± 3.88) were recruited to participate in this study. 8 subjects had never participated in an EEG experiment before and the remaining 4 had no previous experience with visual BCI. Prior to the experiments, participants read, and when they agreed, signed a consent form previously approved by the ethical committee of the university hospital UZLeuven. All participants had normal or corrected-to-normal vision and were remunerated for their participation.

The testing of the BCI was performed on 2 healthy subjects (2 male) and underwent the experiment under similar conditions as listed above. A first proposal was tested on the 2 subjects resulting in suboptimal results, followed by a second proposal which was only tested on 1 of the 2 previous subjects to study any improvement.

2.3. Materials and statistics

The materials used for the experiments and the BCI were similar for all recording sessions. The equipment used remained identical and is described here. The interface was depicted on a 1920x1080 VIEWPixx monitor at a 120Hz refresh rate. Subjects were seated approximately 70 cm from the screen. All experimental interfaces were written in Matlab (v2017b) using the Psychtoolbox extension to ensure precise timing [94]. EEG data was collected from the subjects using 32 active Ag/AgCl electrodes evenly distributed over the scalp, as depicted by the grey nodes in Figure 10, using a SynAmps RT device operating at a sampling rate of 2000 Hz and using Curry 7 [95] to collect the data. The GND and REF electrodes were placed at AFz and FCz respectively.

For the three experiments, an EyeLink 1000 Plus device was used to track the gaze of the subjects to ensure correct behavior during the test. Important to note is that none of the data recorded by the eyetracker was included in the analysis of the classification performance.

A pairwise comparison between the accuracies of the four classifiers and between the accuracies of the two movement types was performed using a Paired Wilcoxon sign-

rank to determine significance between the classifiers and the movement types. The significance threshold was Bonferroni corrected and set to 0.0083 (=0.05/6) for the comparison of the different classifiers and set to 0.05 for the analysis between left translation (LT) and right translation (RT) and LT or RT and P300.

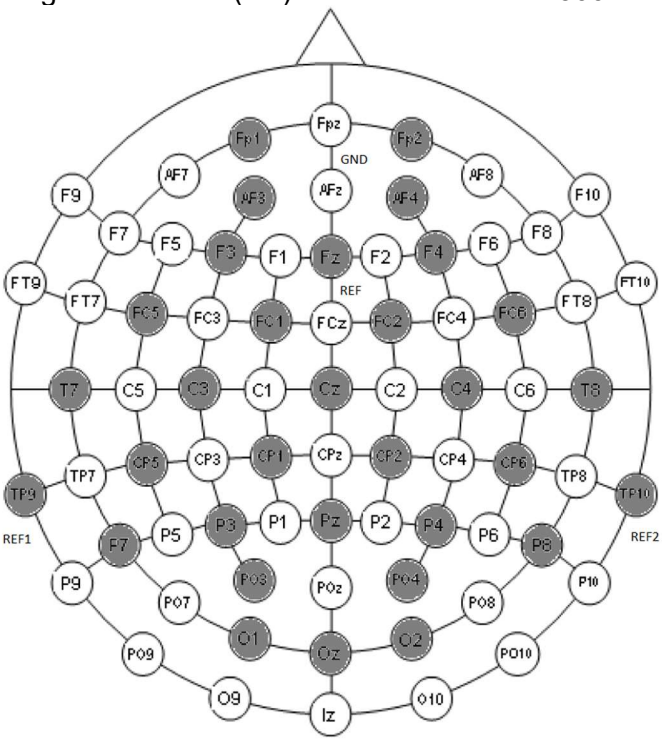

Figure 10: recording sites used for all experiments and the spelling console. Grey nodes depict a used site, REF is put at FCz and GND at AFz. Channels are referenced w.r.t. the mastoids at TP9 and TP10.

2.4. mVEP classification for left and right movement

The experiments described in this section were designed to both answer the first and second question stated in the introduction of this chapter. The visual interface consisted of nine rectangular boxes (2.5 x 1.25cm) arranged in a 3 x 3 matrix, with a 6cm inter-target distance. The experiment consisted of 10 blocks. At the beginning of each block, a fixation point (+) was shown in the center of one of the targets (cued target) and gazed at by the subject. The stimulation sequence consisted of a vertical line segment that started to traverse a pseudorandomly selected target for 140 ms; 60 ms after that the line segment of another target started to move (thus, 60 ms inter-stimulus interval). Each target was stimulated 10 times in a pseudorandom order during one block. Subjects were asked to mentally count the number of times the line segment of the cued target moved. Subsequent blocks were separated by a 600 ms break during which the fixation point was displayed on another target, which was then immediately gazed at by the subject. A visual representation of one set of blocks is given in Figure 11. In total, all targets were cued four times in pseudorandom order yielding a set of nine blocks. All subjects repeated the experiment twice, once with the stimulus (i.e. the traversing line) moving from left to right (RT), and once with the stimulus moving from right to left (LT). Both experiments were counterbalanced across subjects to nullify any fatigue. A visual representation of the LT and RT tests are given in Figure 12 and Figure 13 respectively.

Data processing for both experiments were performed offline. First the raw EEG signal was re-referenced to the average of the mastoid signal (TP9 & TP10) and bandpass filtered between 1 and 10 Hz using a 4^{th} order Butterworth filter. The signal was then cut into 0.95-s epochs from 200 ms pre-stimulus onset for baseline correction to 750 ms post-stimulus onset, baseline corrected, downsampled to 50Hz and stored for further analysis. The procedure resulted in 3240 epochs of which 360 for each target direction, 40 for the intended target and 320 for non-targets. Prior to classification, the 200 ms baseline was removed from the epochs as this interval did not contain any stimulus-evoked activity. Important to note is that recording was performed on a separate machine than the one running the experiment to nullify interference.

Training of the classifiers was performed by loading the epochs, which should contain the mVEP located between 0 and 500 ms post stimulus, into the classifier. The accuracy of the classifier was tested utilizing a stratified 4-fold cross-validation where one set of blocks was kept out of each validation run. The trials for 3 sets of blocks were labelled as target/non-target and loaded into the classifier for training after transforming the data in the required shape described for each classifier in chapter 2. The fourth, unlabeled, block was then classified (after transformation). Classification was performed by first averaging the epochs to each of the 9 targets, the actual number of averaged epochs used is a parameter in judging the decoder's accuracy. The highest score resulting from classification was indicated as the target. The given label was afterwards compared to the actual target labels to determine the accuracy.

Note that channel selection has shown to improve decoder accuracy, also for the beamformer [12], the mVEP is predominantly present over the occipital pole and contralateral for the LT and RT, hence, as a result, the following channels were selected, LT and RT respectively:

{Oz, O1, PO3, P7, P3, Pz, CP1, CP5},
{Oz, O2, PO4, P4, P8, Pz, CP2, CP6}

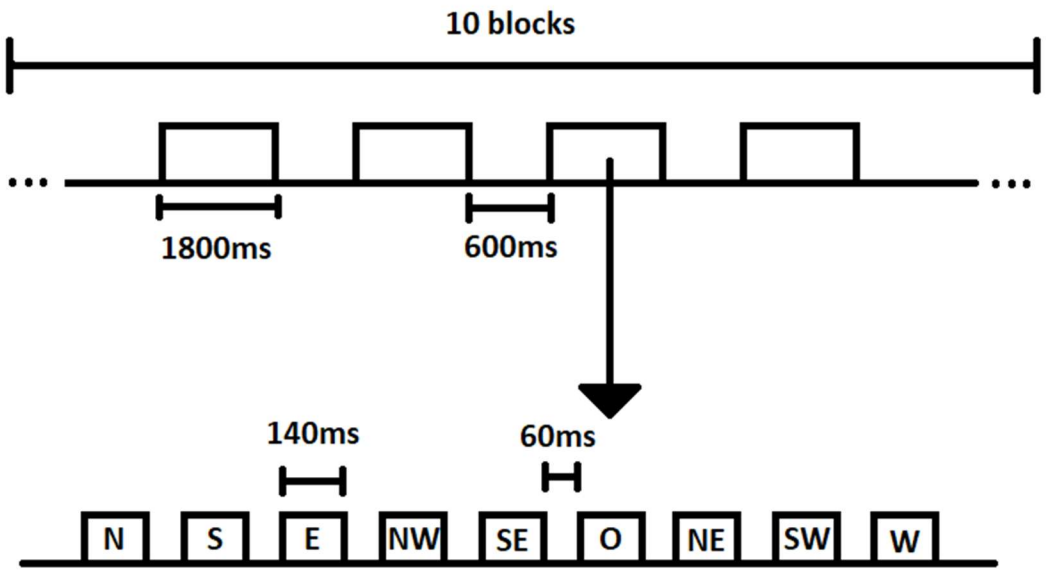

Figure 11: visual representation of the timing schedule of one block, the experiments consisted of 10 blocks.

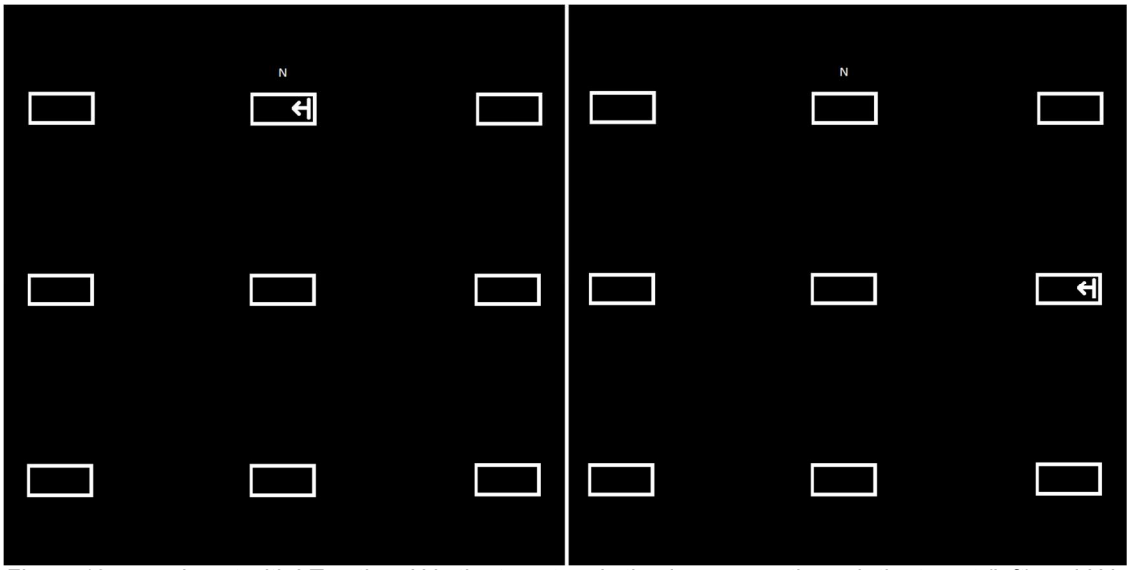

Figure 12: experiment with LT option. N is the target and stimulus passes through the target (left) and N is target and stimulus does not pass through the target (right).

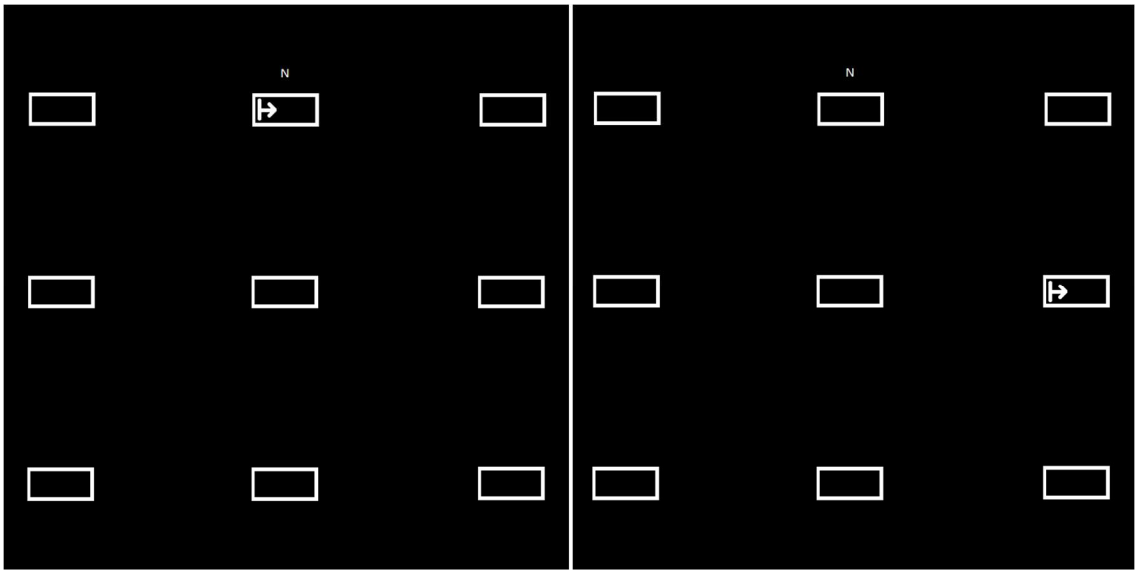

Figure 13: experiment with RT option. N is the target and stimulus passes through the target (left) and N is target and stimulus does not pass through the target (right).

2.5. mVEP and P300

The data of the previous experiments was analyzed offline and based on the shape and consistency of the mVEP either LT or RT was selected as a comparative. The P300 experiment was similar to the experiment of the respective movement described above but showed a sudden change in color together with the moving stimulus. A visual representation of the experiment is given in Figure 14. 6 of the aforementioned subjects were selected to undergo the rightwards movement variant of the P300 experiment for comparison and 6 were selected to undergo the leftwards movement.

The experimental data was analyzed in an analogous way as described above resulting in an equilibrium in epoch size and count for each target. The training of the classifiers was performed in an analogous way as described and since the P300 channels where the P300 is predominantly measured are already included in the selected channels, these were left unaltered. Accuracy is calculated in an analogous way.

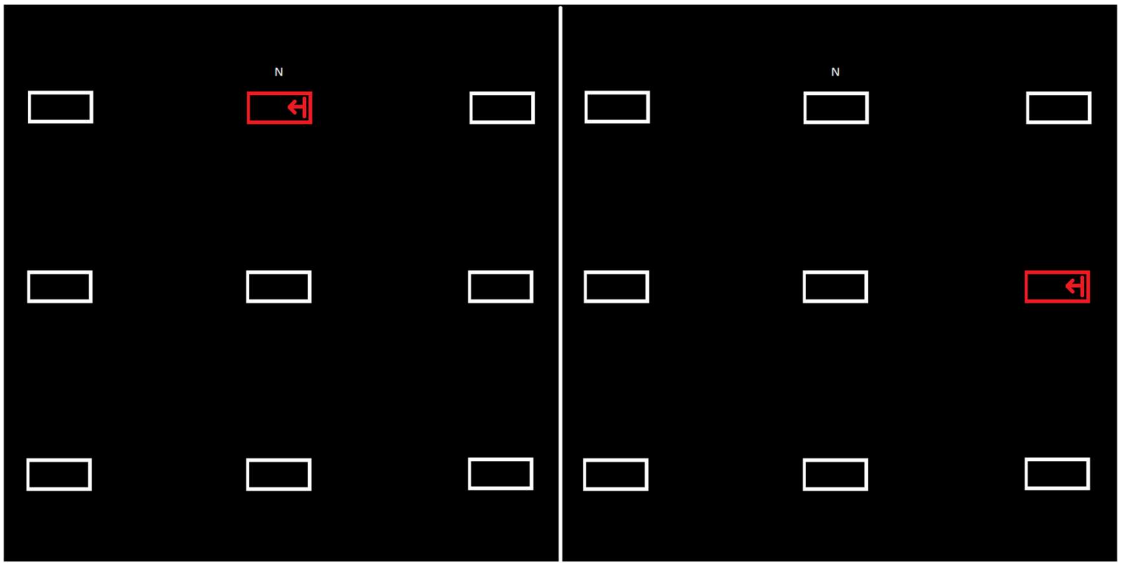

Figure 14: experiment with LT option combined with the P300 option. N is the target and stimulus passes through the target (left) and N is target and stimulus does not pass through the target (right).

2.6. On-line spelling console

The on-line spelling consoles used for comparison utilize a free-spelling setting [17], [18]. To make comparison fair and relevant the on-line spelling console must have a similar interface. For the on-line spelling console the schematic described in [1] is implemented. The first proposal interface shows 36 targets (1.5 x 1.5 cm) and were set 3 cm apart from their closest neighbor with a unique symbol depicted above each target. The second proposal showed 36 target (1x1 cm) and we set 4 cm apart from its nearest neighbor. To increase the number of possible selections per minute the stimuli were shown per column and row. A visual representation is presented in Figure 15 for training and Figure 16 for the spelling interface

The training data is stored on the system and training of the classifier is performed in a similar on-line capture was achieved by storing the recorded data of the entire training experiment and splitting the data into epochs similarly to the method described above. The on-line target selection was achieved by storing the recorded data per trial into a buffer, splitting the data into epochs, averaging and classifying the averaged epochs.

The classifier was trained by letting each target be the cued target and stimulating each row and column 5 times in a pseudorandom order per target resulting in 30 on target stimuli for each row and column per target.

To calculate the accuracy of the trained classifier the subject was asked to spell the following set of words which collectively contain each symbol on the interface:

'1984','607','253','XRAY','WIVES','QUEUE','CLASH','BRAIN','JAZZ',
'MOPED','FAKE','GOAT'.

The subject was asked to focus on the box below the desired symbol. Each row and column were stimulated three times in a pseudorandom order and resulting data from each stimulation was averaged. The classifier was adapted to select the cross point

between the highest row and the highest column as the intended target. The predictions made by the classifier were stored and later compared to the actual prediction to calculate accuracy.

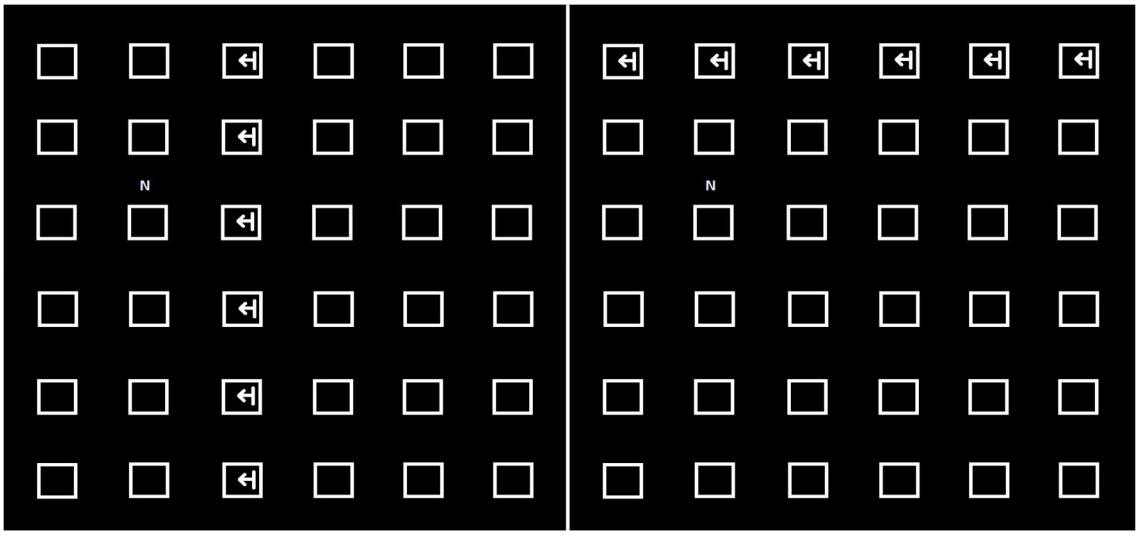

Figure 15: spelling console training session for LT option. N is the focus target. Stimuli are depicted column wise (left) or row wise (right). Scale of image is not accurate.

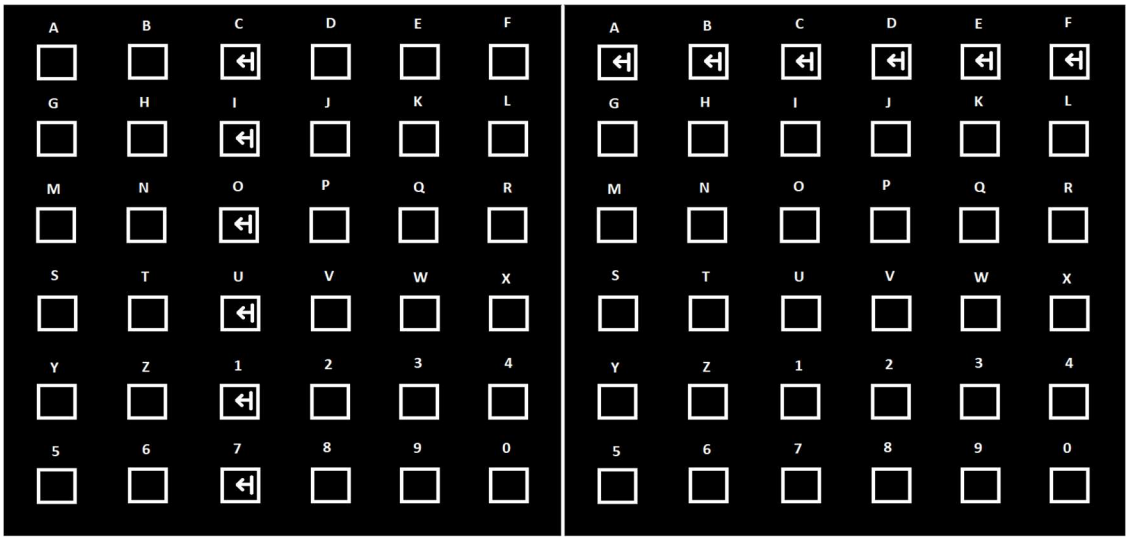

Figure 16:spelling console test session for LT option. Focus target is the desired letter. Stimuli are depicted column wise (left) or row wise (right). Scale of image is not accurate.

2.7. Conclusion

In this chapter, the materials and methods utilized in the experiments and for the implementation of the on-line spelling console were described. All interfaces were created in Matlab and utilized the same recording equipment, electrode layout and data analysis methods.

Three experiments were designed to measure the feasibility of classification, effect of the movement direction and effect of the addition of other ERPs to the mVEP. The experiments were all similar and varied in a single detail to ensure relevance of the

results. 12 subjects performed each one of these experiments and allow for individual and population comparison.

The on-line spelling console BCI based on mVEP was implemented in Matlab and 2 subjects tested the functionality of two separate proposals. The first proposal had the target closer to their neighbours than the second. Several test words were chosen to measure accuracy of the spelling console BCI to allow for the calculation of the ITR for comparison to the previously implemented N200 mVEP spelling console BCIs.

3. Results

3.1. Introduction

In this chapter, the results of the previous experiments and on-line spelling console BCI will be presented in several figures and graphs.

3.2. mVEP classification for left and right movement

The decoding accuracies of stBF, SVM, PP and SWLDA classifiers for the left- and rightward line movements are shown in Figure 17. The averaged (across all subjects) activation pattern of the spatiotemporal beamformer for the left and right movements are shown in Figure 18. The average time required to train and perform classification for each decoder is listed in Table 2. The significance of the difference between the classification results for the classifiers for the left- and rightward movements are given in Tables 3 and 4. The significance of the difference between the classification results for the left- and rightward movements for the same classifier types is given in Table 5.

From Figure 17 it is observed that classification performances are distinctively different from that of a random classifier and that, for both the left- and rightward movements, the performance of all classifiers further improve with the number of stimulus repetitions used. The stBF requires 3 repetitions to reach the 70% accuracy threshold often deemed necessity to establish reliable communication [96].

From table 3 and 4, it can be concluded that the stBF and SVM classifiers do not show a significant difference in accuracy but there is a significant difference between the stBF and SWLDA classifiers for the case of 3 repetitions for the leftward movement and for 2 repetitions for the rightward movement. The PP classifier performed significantly worse than the other classifiers. When comparing the accuracy of the spatiotemporal beamformer to the accuracy of the SWLDA classifier, we notice a significant drop in accuracy for a small number of stimulus repetitions. An explanation for this can be the presence of peaks not caused by the stimuli but by previously mentioned artefacts (e.g. movement itself rather than the onset of it). The accuracies of the stBF obtained by the left- and rightward movement were not significantly different ($p > 0.0083$) for any of the numbers of stimulus repetitions.

When studying the activation patterns in Figure 18, an interesting result is observed. The mVEP appears present throughout the scalp for the rightward movement case. The N200 peaks appear to shift 50 ms between the left and right side of the scalp areas. In contrast, the activation patterns for the left translation show only dominant mVEPs on the left side of the scalp. The N200 peak appears less dominant in the responses measured on the right side of the scalp. According to previous works, the leftward movement should be predominantly eliciting activations on the left side of the scalp while the right movement should be most dominant on the right side of the scalp [9], [19], [67].

Judging from the results in Table 2, the average time required to train each classifier shows a vast increase in computation time for the SVM and SWLDA compared to the beamformer. The PP could be trained faster. The ratio depicted in the table shows the

division of the average computing time as divided by the average computer time of the spatiotemporal beamformer.

The results in Table 5 show no statistically significant difference between the classification of signal resulting from the leftward or rightward movement. The direction of the movement of the on-line spelling console will thus be chosen according to the handedness of the person, giving left-handed subjects the RT and right-handed subjects the LT variation.

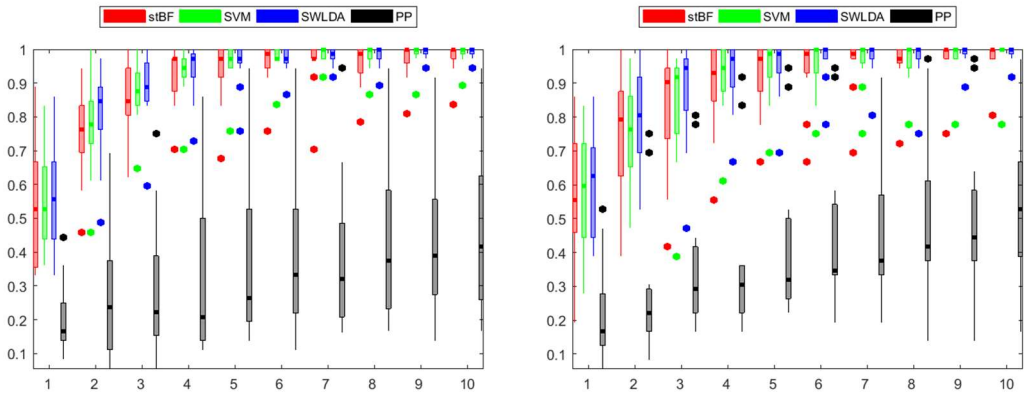

Figure 17: boxplots of the accuracies for stBF (red), SVM (green), SWLDA (blue) and PP (black) for RT (left) and LT (right). Outliers are marked by dots. x-axis depicts the number of averaged blocks used for testing. y-axis depicts the accuracy ranging between [0,1].

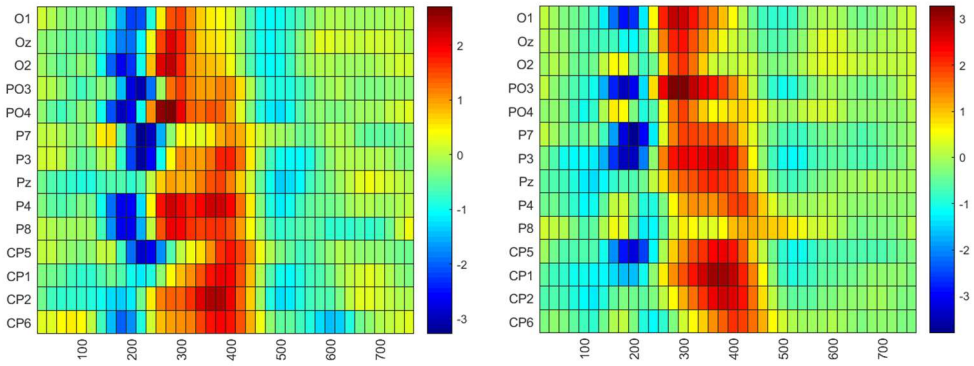

Figure 18: averaged activation pattern over all subjects for the RT (left) and the LT (right). x-axis depicts the milliseconds after motion-onset. y-axis depicts the recording location.

Table 2:summed training and testing times for each classifier in seconds together with the standard deviation. Ratio is the time divided by the average stBF time.

time in sec	stBF	SVM	SWLDA	PP
average	163.71	6982.29	1607.38	15.03
std dev	65.66	4186.80	885.94	2.26
ratio	1	42.650426	9.8184943	0.091821

Table 3: Wilcoxon sign rank test between the accuracies of the classifiers for the RT option.

	Number of averaged epochs									
right translation	1	2	3	4	5	6	7	8	9	10
stBF - SVM	0.877	0.197	0.426	0.672	0.266	0.266	0.156	0.063	0.125	0.250
stBF - SWLDA	0.473	0.002	0.078	0.250	0.078	0.531	0.375	0.063	0.125	0.250
stBF - PP	< 0.001	<0.001	< 0.001	< 0.001	<0.001	< 0.001	< 0.001	<0.001	< 0.001	< 0.001
SWLDA - PP	< 0.001	<0.001	< 0.001	< 0.001	<0.001	< 0.001	< 0.001	<0.001	< 0.001	< 0.001
SVM - PP	< 0.001	<0.001	< 0.001	< 0.001	<0.001	< 0.001	< 0.001	<0.001	< 0.001	< 0.001
SWLDA - SVM	0.246094	0.015625	0.335938	0.8125	1	0.5	0.5	1	1	0.75

Table 4: Wilcoxon sign rank test between the accuracies of the classifiers for the LT option.

	Number of averaged epochs									
left translation	1	2	3	4	5	6	7	8	9	10
stBF - SVM	0.219	0.547	0.410	0.313	0.063	0.250	0.500	0.531	0.625	0.375
stBF - SWLDA	0.146	0.020	0.004	0.031	0.063	0.125	0.125	0.125	0.250	0.250
stBF - PP	< 0.001	<0.001	< 0.001	< 0.001	<0.001	< 0.001	< 0.001	<0.001	< 0.001	< 0.001
SWLDA - PP	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
SVM - PP	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
SWLDA - SVM	0.307	0.011	0.007	0.117	0.625	0.250	0.250	1.000	0.500	1.000

Table 5: Wilcoxon sign rank test between the accuracies of the LT and RT options for each classifier.

	Number of averaged epochs									
left vs right	1	2	3	4	5	6	7	8	9	10
stBF	0.502441	0.834473	0.606445	0.828125	0.953125	0.4375	1	1	0.960938	0.796875
SVM	0.348633	0.834473	0.476563	0.5	0.375	0.265625	0.515625	0.34375	0.5	1
SWLDA	0.223145	0.87793	0.927734	0.847656	0.4375	0.9375	1	0.671875	0.75	0.75
PP	0.598633	0.554688	0.097656	0.203125	0.180664	0.121094	0.081055	0.055664	0.196289	0.202148

3.3. mVEP and P300

The results from the previous section showed no significant difference between the prediction accuracy of the left and right translation for any classification method. However, the activation patterns of the spatiotemporal beamformer did show a difference between left and right. Hence, the two cases will be handled separately.

The decoding accuracies of stBF, SVM, PP and SWLDA classifiers for the right- and P300 right line movements are shown in Figure 19 whilst the accuracies for the left- and P300 left line movements are depicted in Figure 20. The averaged (across all subjects) activation pattern of the spatiotemporal beamformer for the right- and P300 right line movements and for the left- and P300 left line movements are shown in Figure 21 and Figure 22 respectively. The significance of the difference between the classification results for the classifiers for the right- and P300 right movements are given in Table 6 whilst the significance of the difference between the classification results for the classifiers for the left- and P300 left movements are given in Table 7.

From Figure 19 and 20 it is observed that classification performances are distinctively different from that of a random classifier and that, for both translation directions and their P300 counterparts, the performance of all classifiers further improve with the number of stimulus repetitions used. It is apparent that for the rightwards movement P300 variant the spatiotemporal beamformer requires 2 averages to reach the aforementioned 70% accuracy mark. However, the leftwards movement P300 variant doesn't reach the threshold until 6 averages are included.
From table 6 and 7, it can be concluded that the leftwards movement nor the rightwards movement and their P300 variant do not show a significant difference in accuracy.

When studying the activation patterns in Figure 21, the same phenomena as before can be observed. The mVEP appears present throughout the scalp for the rightward movement case. The N200 peaks appear to shift 50 ms between the left and right side of the scalp areas. In contrast, the activation patterns for the left translation show only dominant mVEPs on the left side of the scalp. A more dominant positive peak is present in the activation pattern for the rightwards movement P300 compared to the normal rightwards movement. However, for the leftwards translation both the positive and negative peaks are more dominant compared to the leftwards movement P300 counterpart. Due to the instability and the increase in fatigue of subjects resulting from the addition of the contrast variation the P300 variant was not utilized in the development of the spelling console.

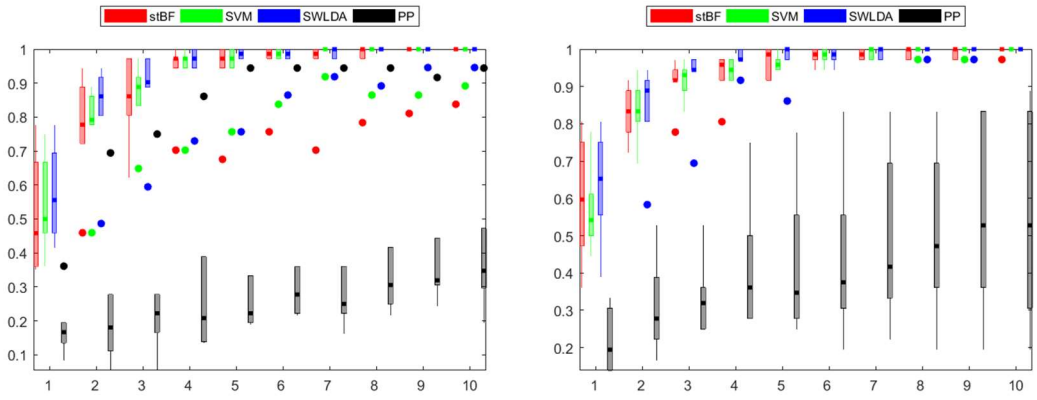

Figure 19: boxplots of the accuracies for stBF (red), SVM (green), SWLDA (blue) and PP (black) for RT (left) and RT and P300 (right) Outliers are marked by dots. x-axis depicts the number of averaged blocks used for testing. y-axis depicts the accuracy ranging between [0,1].

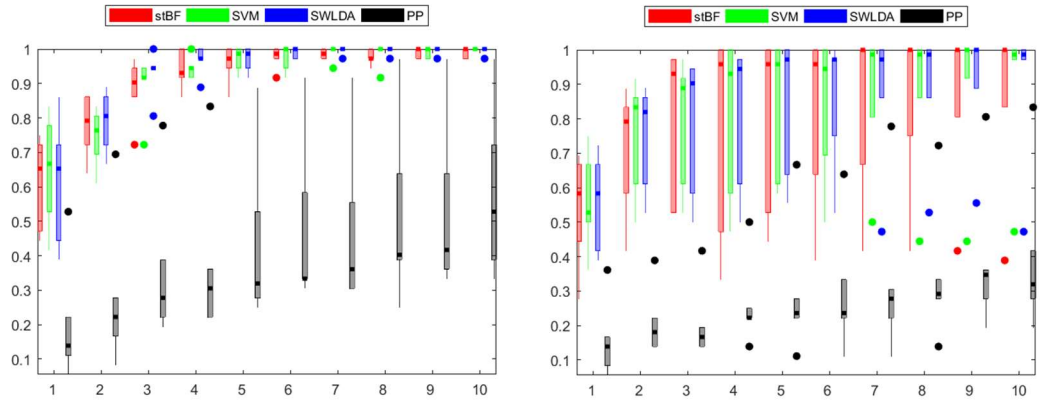

Figure 20: boxplots of the accuracies for stBF (red), SVM (green), SWLDA (blue) and PP (black) for LT (left) and LT and P300 (right). Outliers are marked by dots. x-axis depicts the number of averaged blocks used for testing. y-axis depicts the accuracy ranging between [0,1].

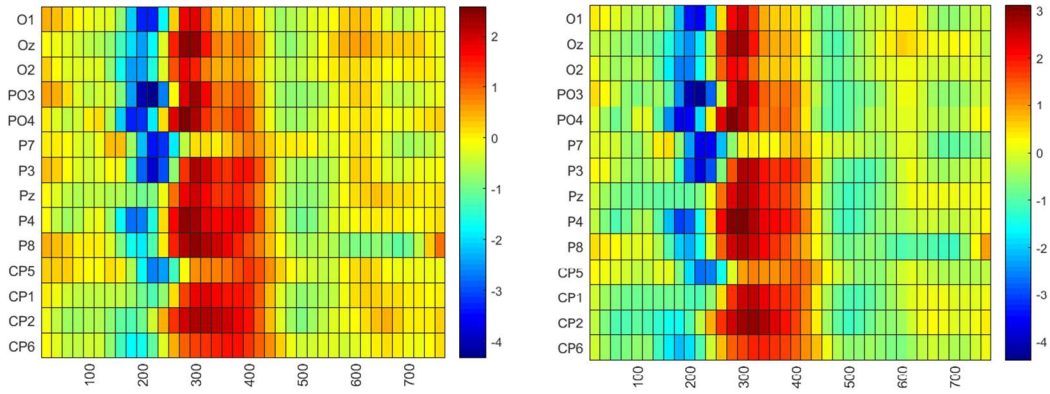

Figure 21: averaged activation pattern over all subjects for the RT (left) and the RT and P300 (right). x-axis depicts the milliseconds after motion-onset. y-axis depicts the recording location.

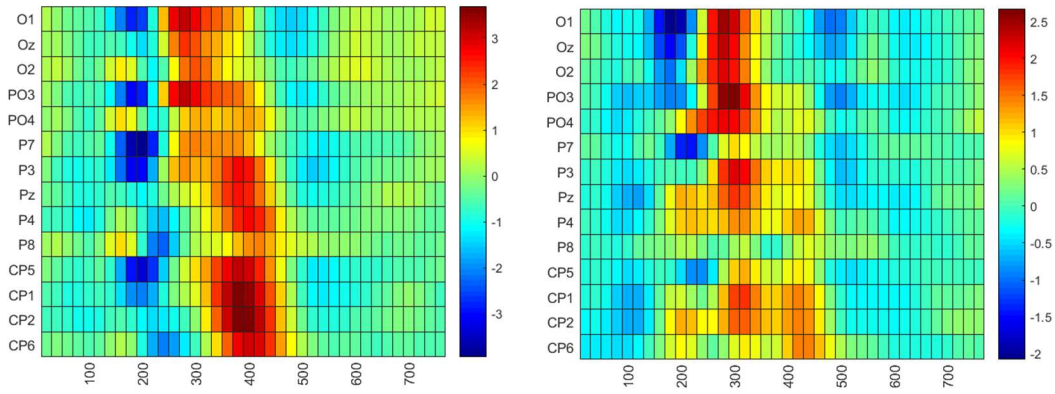

Figure 22: averaged activation pattern over all subjects for the LT (left) and the LT and P300 (right). x-axis depicts the milliseconds after motion-onset. y-axis depicts the recording location

Table 6: Wilcoxon sign rank test between the accuracies of the different classifiers for the RT and RT and P300 options.

	Number of averaged epochs									
Right vs P300R	1	2	3	4	5	6	7	8	9	10
stBF	0.71875	1	0.375	0.25	0.625	0.25	0.25	0.375	0.5	0.5
SVM	0.5625	0.15625	0.25	1	1	1	1	1	1	1
SWLDA	0.46875	0.5	0.21875	0.5	0.5	0.75	0.75	1	1	1
PP	0.25	0.3125	0.5625	0.21875	0.3125	0.625	0.25	0.34375	0.21875	0.40625

Table 7: Wilcoxon sign rank test between the accuracies of the different classifiers for the RT and RT and P300 options.

	Number of averaged epochs									
Left vs P300L	1	2	3	4	5	6	7	8	9	10
stBF	0.34375	0.875	0.8125	0.8125	0.5	0.375	0.625	0.75	0.5	0.5
SVM	0.3125	0.8125	0.625	0.5	0.5625	0.5	0.5	0.5	0.625	0.25
SWLDA	0.71875	1	0.375	0.25	0.625	0.25	0.25	0.375	0.5	0.5
PP	1	0.28125	0.1875	0.5	0.625	0.3125	0.3125	0.34375	0.625	0.3125

3.4. On-line spelling console

The previous results for the accuracy of the left and right translation allow for the calculation of theoretical offline ITR. Given the stimulus time was 200ms per stimuli and was run on 12 targets, this gives us 2400 ms per target. Calculate in 600ms for target swapping gives us an even 3000ms per target. In other words, at maximum efficiency 20 targets per minute can be selected (=M). The number of possible targets was 36 in the spelling console. Given an accuracy of 80% we can now calculate the ITR:

$$ITR = 20 * (\log_2 36 + 0.8 * \log_2 0.8 + 0.2 * \log_2 \frac{0.2}{35}) = 64.44\ bits/min.$$

Table 8 and 9 and Figure 23 and 24 represent the results of the confusion of subject 1 and subject 2 respectively. The accuracy of the on-line spelling console proved much lower for the first proposal. Showing an accuracy of 3.92% for subject 1 and 13.7% for subject 2. Analysis of the results show a large chance of the faulty target to be within 2 squares of the selected target. The target column of subject 1 was predicted correctly 19.61% of the cases whilst the rows were predicted correctly 3.92% of the cases. Subject 2 had a correct prediction for the row 29.41% of the cases and correct the column correctly 39.22% of the cases. The results of these initial experiments caused a second proposal to be constructed where the targets were smaller and moved further apart to counter the confusion between row and column.

Table 10 and Figure 25 depict the results of subject 2 on the second proposed interface. The accuracy has effectively doubled to 27.45%. The row was predicted correctly 33% of the cases whilst the column was predicted correctly 47.06% of the cases. The second interfaces results showed a decrease in faulty predictions by 27.28% showing not only an effect on the confusion by the classifier but also resolving any confusion resulting from the training session of the classifier.
The on-line information transfer rate was calculated from the data of subject 2 for the second proposed method. The number of possible targets was set a 36. 51 targets were selected during 10.26-minute session. Thus, the number of selections per minute becomes 4.97. The accuracy was calculated to be 27.45% resulting in the ITR being:

$$ITR = 4.97 * (\log_2 36 + 0.2745 * \log_2 0.2745 + 0.7255 * \log_2 \frac{0.7255}{35}) = 17.08\ bits/min.$$

It is important to note with this result that firstly, one test using one subject is not a valid evaluation basis for any test. Secondly, BCI commonly require the subject to be trained before effective communication can be established.

Table 8: confusion of subject 1 on the first proposed interface.

	correct row	wrong row
correct column	2	8
wrong column	0	41

Table 9: confusion of subject 2 on the first proposed interface.

	correct row	wrong row
correct column	7	13
wrong column	8	33

Table 10: confusion of subject 2 on the second proposed interface.

	correct row	wrong row
correct column	14	10
wrong column	3	24

%

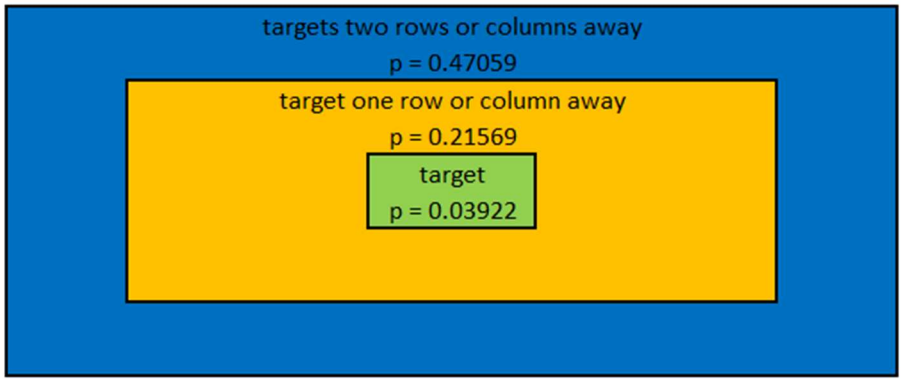

Figure 23: chance of the predicted target being in the vicinity of the actual target for subject one for the first proposed interface. The remaining predicted targets (52.94%) were not in proximity of the intended target.

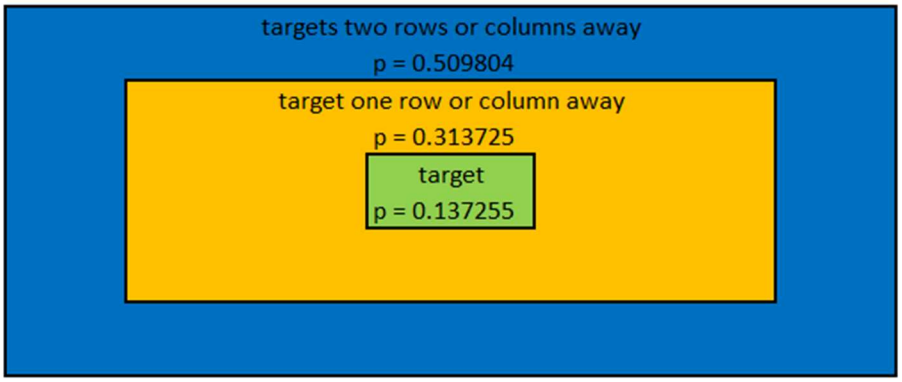

Figure 24: chance of the predicted target being in the vicinity of the actual target for subject 2 for the first proposed interface. The remaining predicted targets (49.02%) were not in proximity of the intended target.

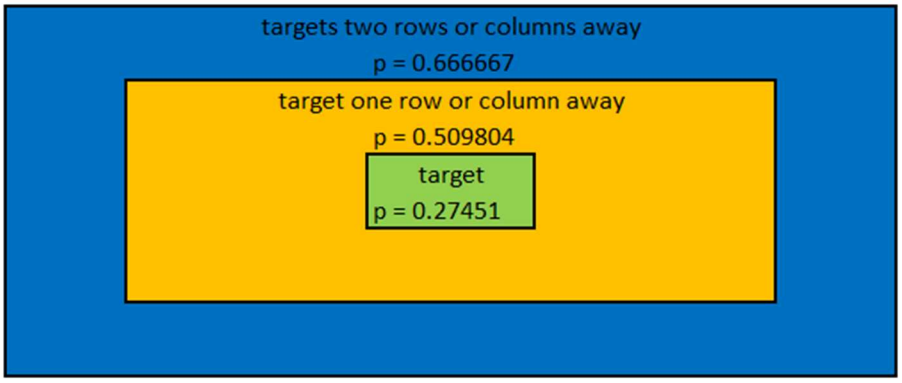

Figure 25: chance of the predicted target being in the vicinity of the actual target for subject 2 for the second proposed interface. The remaining predicted targets (33.33%) were not in proximity of the intended target.

3.5. Conclusion

In this chapter, three questions were answered. The spatiotemporal beamformer is as accurate as other used classification methods in other works for the classification of the mVEP whilst showing a significant classifier training speed increase resulting in the possibility for on-line application.

No significant difference was observed in the difference of the translation direction for classification accuracy for any of the classifiers. The on-line spelling console was set to correspond with the handedness of the subject.

No significant difference was observed by the addition of a P300 component to the classification accuracy for any of the classifiers whilst an increase in fatigue rate was noted. The P300 variation was not used in the on-line spelling console.

The on-line spellings console showed an ITR of 17.08 bits/min. The mVEP appears to show a strong sensitivity to peripheral stimulation. The distance between targets is thus

a crucial factor in the accuracy of the mVEP BCI. Note that more test must be performed on different subject to create a valid evaluation of the BCI.

4. Conclusion

4.1. Discussion

In this work, a spatiotemporal beamformer was used to classify mVEP responses in a multi-target BCI context. From the processed data we can conclude that the stBF shows no significant difference in accuracy compared to the SVM and SWLDA classifiers, previously used for mVEP classification. However, the big advantage is that the spatiotemporal beamformer can be training significantly faster. This renders the stBF a viable candidate for online BCI use. The peak picking method was reported to yield much higher accuracies which may be used as an indicator to how trained the subjects of the experiments were to mVEP-based BCI [9], [17]. Recently, deep neural networks and adaptive on-line classifiers have been used for the decoding of mVEP which may be compared to the stBF in an on-line setting [8], [10], [70].

The mVEP template of the stBF for the leftward movement showed prominent responses over the left hemisphere; for the right hemisphere there was a dominant positive peak 300-350 ms after stimulus onset. For the rightward movement, the mVEP was present over both hemispheres. The N200 response on the left side of the scalp showed a decrease in peak amplitude and a 50ms delay. The delay has, to my knowledge, not been documented before since most studies on mVEP focus on a set time point to depict the scalp signals [9]–[11]. The difference in accuracies between left and right translation proved insignificant. The choice was therefore concluded to be up to preference or handedness of the subject. A third option for translation exists in the contraction of the stimuli. The contraction was reported to elicit the strongest mVEP [9]. In the future this variant could be explored to study the significance between LT, RT and contraction.

The combination of the mVEP and P300 paradigms did not show any significant improvement for accuracy of any classifier. The activation patterns for the right translation show and increase in positive amplitude and minor change in negative value when adding the P300; the activation pattern for the left translation showed a decrease in both positive and negative amplitudes. The accuracy for the addition of the P300 to the normal translation was insignificant. These results, combined with the increase in fatigue caused by the flashing stimuli of the P300 oddball, makes us conclude that a combination of the mVEP and P300 paradigms is best avoided.

The results of the spelling console showed mVEP to be a sensitive paradigm to peripheral stimulation. A minimum distance between possible targets is advised when creating a mVEP-based BCI. The proposed BCI system would not require flashing which poses little discomfort and less fatigue on BCI users. Comparing the results of the on-line BCI has proven difficult since most on-line mVEP BCI utilized a completely different interface layout for the presentation of the stimuli [2], [10], [17], [70]. Comparing the result of our offline ITR makes it one of the best variants for this interface [17], [18]. The comparison of the on-line mVEP based spelling console BCI to other variants using their interface would prove an interesting future work. The clever adaptation of the interface is a way to improve the ITR of the BCI.

4.2. Conclusion

In this study, we have shown that the spatiotemporal beamformer is at par with the best mVEP classifiers for a larger quantity of averaged epochs. The difference in accuracy showed significant for a lower number for both the left and the right translation with the SWLDA classifier. The stBF performed on par with the SVM and significantly better than PP. However, the spatiotemporal beamformer requires a dramatically smaller training time making it a viable option for online BCI.

The movement direction of the stimuli does not lead to a significant difference in accuracy when accounting for the difference in mVEP scalp distribution. The addition of the P300 paradigm to the experiment did not show the expected results and is therefore ill-advised.

The creation of a mVEP-based spelling console can be achieved provided the interface keeps a minimum distance between targets due to the sensitivity of the paradigm. The implemented console can reach a higher information transfer rate when comparing it to similar mVEP-based spelling consoles for offline use. No online console using a similar interface was found for comparison. The speller needs more subjects and some alterations in interface before becoming a communication method for the disabled.

Future works can be the testing of the spelling console on more subjects and eventual implementation on a tablet platform to be used by patients in hospitals. Several other interface layouts can be explored to study the effect on ITR. Other applications using mVEP BCI can also be explored.

5. References

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, “Brain-computer interfaces for communication and control.,” *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–91, 2002.

[2] D. Marshall, S. Wilson, and D. Coyle, “Motion-Onset Visual Evoked Potentials for Gaming: A pilot study,” *24th IET Irish Signals Syst. Conf. (ISSC 2013)*, pp. 56–56, 2013.

[3] I. Martišius and R. Damaševičius, “A prototype SSVEP based real time BCI gaming system,” *Comput. Intell. Neurosci.*, vol. 2016, 2016.

[4] J. Jin, H. Zhang, I. Daly, X. Wang, and A. Cichocki, “An improved P300 pattern in BCI to catch user’s attention,” *J. Neural Eng.*, vol. 14, no. 3, 2017.

[5] P. Stawicki, F. Gembler, and I. V. B, “A User-Friendly Dictionary-Supported SSVEP-based BCI,” vol. 9961, pp. 168–180, 2017.

[6] X. Gao, D. Xu, M. Cheng, and S. Gao, “A BCI-based environmental controller for the motion-disabled,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 137–140, 2003.

[7] J. J. Wilson and R. Palaniappan, “Augmenting a SSVEP BCI through single cycle analysis and phase weighting,” *2009 4th Int. IEEE/EMBS Conf. Neural Eng. NER ’09*, pp. 371–374, 2009.

[8] T. Ma, F. Li, P. Li, D. Yao, Y. Zhang, and P. Xu, “An Adaptive Calibration Framework for mVEP-Based Brain-Computer Interface,” vol. 2018, 2018.

[9] F. Guo, B. Hong, X. Gao, and S. Gao, “A brain-computer interface using motion-onset visual evoked potential,” *J. Neural Eng.*, vol. 5, no. 4, pp. 477–485, 2008.

[10] T. Liu, L. Goldberg, S. Gao, and B. Hong, “An online brain-computer interface using non-flashing visual evoked potentials,” *J. Neural Eng.*, vol. 7, no. 3, 2010.

[11] S. Schaeff, M. S. Treder, B. Venthur, and B. Blankertz, “Exploring motion VEPs for gaze-independent communication,” *J. Neural Eng.*, vol. 9, no. 4, 2012.

[12] B. Wittevrongel and M. M. Van Hulle, “Faster P300 Classifier Training Using Spatiotemporal Beamforming,” *Int. J. Neural Syst.*, vol. 26, no. 03, p. 1650014, 2016.

[13] B. Wittevrongel, E. Van Wolputte, and M. M. Van Hulle, “Code-modulated visual evoked potentials using fast stimulus presentation and spatiotemporal beamformer decoding,” *Sci. Rep.*, vol. 7, no. 1, pp. 1–10, 2017.

[14] M. Van Vliet *et al.*, “Single-trial ERP component analysis using a spatio-temporal LCMV beam former.,” *IEEE Trans. Biomed. Eng.*, vol. 63, no. 1, pp. 55–66, 2016.

[15] B. Wittevrongel and M. M. Van Hulle, “Spatiotemporal beamforming: A transparent and unified decoding approach to synchronous visual brain-computer interfacing,” *Front. Neurosci.*, vol. 11, no. NOV, pp. 1–12, 2017.

[16] B. Wittevrongel and M. M. Van Hulle, “Frequency- and Phase Encoded SSVEP Using Spatiotemporal Beamforming,” pp. 1–18, 2016.

[17] B. Hong, F. Guo, T. Liu, X. Gao, and S. Gao, “Clinical Neurophysiology N200-speller using motion-onset visual response,” *Clin. Neurophysiol.*, vol. 120, no. 9, pp. 1658–1666, 2009.

[18] D. Zhang, H. Song, H. Xu, W. Wu, S. Gao, and B. Hong, “An N200 speller integrating the spatial profile for the detection of the non-control state,” *J. Neural Eng.*, vol. 9, no. 2, 2012.

[19] A. Bruce, J. Alexander, L. Julian, R. Martin, R. Keith, and W. Peter, “Molecular Biology of the Cell: The nervous system,” 2002.

[20] “https://simple.wikipedia.org/wiki/Neuron#/media/File:Neuron.svg: image of a neuron.” [Online]. Available: https://simple.wikipedia.org/wiki/Neuron#/media/File:Neuron.svg.

[21] “http://www.interactive-biology.com/49/proteins-in-the-membrane-of-the-neuron-and-their-functions/: image of cell membrane.” [Online]. Available: http://www.interactive-biology.com/49/proteins-in-the-membrane-of-the-neuron-and-their-functions/.

[22] “http://physiologyplus.com/wp-content/uploads/2017/06/Propagation-of-action-potential.png: action potential propagation.”

[23] “https://en.wikipedia.org/wiki/Neurotransmission: detailed image of synapse.”

[24] D. Cic, “Brain Waves,” no. 66, pp. 1–4, 2013.

[25] W. J. freeman, “The Basics of Brain Waves,” 1930.

[26] D. R. S. E. D. and B. A. Hill, *Coping with Concussion and Mild Traumatic Brain Injury*. 2013.

[27] J. L. Fannin, “Understanding Your Brainwaves,” *Underst. your brainwaves*, pp. 1–20.

[28] S. S. and V. K. Sinha, “Event-related potential: an overview,” *Ind. Psychiatry J.*, 2009.

[29] J. D. Kropotov, *Functional Neuromarkers for Psychiatry Applications for Diagnosis*. 2016.

[30] M. G. H. Coles and M. D. Rugg, “Event-related brain potentials : an introduction.”

[31] M. Kuba, Z. Kubová, J. Kremláček, and J. Langrová, “Motion-onset VEPs: Characteristics, methods, and diagnostic use,” *Vision Res.*, vol. 47, no. 2, pp. 189–202, 2007.

[32] A. K. Smith *et al.*, “Processes in Schizophrenia Conclusions :,” no. October, pp. 1264–1275, 2010.

[33] J. Y. Bennington and J. P. U, “Comparison of P300 from passive and active tasks for auditory and visual stimuli,” pp. 171–177, 1999.

[34] S. J. Golubic, A. Susac, R. Huonker, J. Haueisen, and S. Supek, “Early attentional modulation of the neural network evoked with the auditory paired-click paradigm : An MEG study,” *Procedia - Soc. Behav. Sci.*, vol. 126, pp. 195–196, 2014.

[35] C. W. Robinson, “Proceedings of the Annual Meeting of the Cognitive Science Evidence for auditory dominance in a passive oddball task,” 2010.

[36] R. Hashimoto, F. Homae, and K. Nakajima, “Functional Differentiation in the Human Auditory and Language Areas Revealed by a Dichotic Listening Task,” vol. 158, pp. 147–158, 2000.

[37] E. Halgren, K. Marinkovic, and P. Chauvel, “Generators of the late cognitive potentials in auditory and visual oddball tasks,” vol. 106, pp. 156–164, 1998.

[38] H. Erhan, J. C. Borod, C. E. Tenke, and G. E. Bruder, “Identification of Emotion in a Dichotic Listening Task : Event-Related Brain Potential and Behavioral Findings and,” vol. 307, no. 37, pp. 286–307, 1998.

[39] G. G. Briggs and R. D. Nebes, “NOTE THE EFFECTS OF HANDEDNESS , FAMILY HISTORY AND SEX ON THE PERFORMANCE OF A DICHOTIC LISTENING TASK,” vol. 14, no. 2, pp. 129–133, 1976.

[40] “http://brainandlimbicsystem.weebly.com/the-brain.html: image of the right half of the brain.”

[41] E. R. J. Samuel Sutten, Margery Braren, Joseph Zubin, “Evoked-Potential Correlates of Stimulus Uncertainty,” *Science (80-.).*, vol. 150, no. 3700, pp. 1187–1188, 1965.

[42] C. Neurophysiology, “The P300 Wave of the Human Event-Related Potential,” no. November 1992, 2017.

[43] L. Storeoencophalographte, U. De Recherches, U. Eptlepste, C. P. Broca, and P. France, “The intracranial topography of the P3 event-related potential elicited during auditory oddball 1 Recordings,” 1990.

[44] D. E. J. LINDEN, “The P300 : Where in the Brain Is It Produced and What Does It Tell Us ?,” vol. 11, no. 6, pp. 563–576, 2005.

[45] J. Y. Bennington and J. Polich, “Comparison of P300 from passive and active tasks for auditory and visual stimuli,” vol. 34, pp. 171–177, 1999.

[46] I. J. H. Studies, Y. Sivarajah, E. Holden, R. Togneri, G. Price, and T. Tan, “Quantifying target spotting performances with complex geoscienti fi c imagery using ERP P300 responses $,” *J. Hum. Comput. Stud.*, vol. 72, no. 3, pp. 275–283, 2014.

[47] G. Juckel *et al.*, “NeuroImage Age effects on the P300 potential and the corresponding fMRI BOLD-signal,” *Neuroimage*, vol. 60, no. 4, pp. 2027–2034, 2012.

[48] G. Sara, E. Gordon, C. Kraiuhin, S. Coyle, A. Howson, and R. Meares, “jourual of The P300 ERP component : an index of cognitive dysfunction in depression ?,” 1994.

[49] B. N. Gangadhar, J. Ancy, and N. Janakiramaiah, “P300 amplitude in non-bipolar , melancholic depression,” vol. 28, pp. 57–60, 1993.

[50] J. M. Ford, D. H. Mathalon, S. Kalba, and L. Marsh, “N1 and P300 Abnormalities in Patients with Schizophrenia , Epilepsy , and Epilepsy with Schizophrenialike

Features," no. Lm, 2001.

[51] S. S. Kristina Tokic, Marina Titlic, Amira Beganovic-Petrovic, Enra Suljic, Rinaldo Romac, "P300 Wave changes in patients with Parkinson's Disease," 2016.

[52] M. Wade, Y.-C. Li, and G. M. Wahl, "Toward Enhanced P300 Speller Performance," *Nat. Rev. Cancer*, vol. 13, no. 2, pp. 83–96, 2013.

[53] V. Guy, C. Desnuelle, and M. Clerc, "ScienceDirect Brain computer interface with the P300 speller : Usability for disabled people with amyotrophic lateral sclerosis," vol. 61, pp. 5–11, 2018.

[54] C. Guan, M. Thulasidas, and J. Wu, "Performance p300 speller for brain-computer interface," pp. 13–16, 2004.

[55] D. J. Krusienski *et al.*, "A Comparison of Classification Techniques for the P300 Speller," 2010.

[56] N. V Manyakov, N. Chumerin, A. Combaz, and M. M. Van Hulle, "Comparison of Classification Methods for P300 Brain-Computer Interface on Disabled Subjects," vol. 2011, 2011.

[57] P. Lee, "Brain computer interface using flash onset and offset visual evoked potentials," vol. 119, pp. 605–616, 2008.

[58] Y.-T. W. Po-Lei Lee, Jen-Chuen Hsieh, Chi-Hsun Wu, Kuo-Kai Shyu, Shyan-Shiou Chen, Tzu-Chen Yeh, "The Brain Computer Interface Using Flash Visual Evoked Potential and Independent Component Analysis," vol. 34, no. 10, pp. 1641–1654, 2006.

[59] B. Z. Allison, D. J. Mcfarland, G. Schalk, S. Dong, M. Moore, and J. R. Wolpaw, "Towards an independent brain – computer interface using steady state visual evoked potentials," vol. 119, pp. 399–408, 2008.

[60] J. Xie, G. Xu, J. Wang, F. Zhang, and Y. Zhang, "Steady-State Motion Visual Evoked Potentials Produced by Oscillating Newton ' s Rings : Implications for Brain- Computer Interfaces," vol. 7, no. 6, 2012.

[61] N. V Manyakov, N. Chumerin, A. Robben, A. Combaz, M. Van Vliet, and M. M. Van Hulle, "Sampled Sinusoidal Stimulation Profile and Multichannel Fuzzy Logic Classification for Monitor-based Phase-coded SSVEP Brain-Computer Interfacing."

[62] M. Cheng, X. Gao, S. Gao, S. Member, and D. Xu, "Design and Implementation of a Brain-Computer Interface With High Transfer Rates," vol. 49, no. 10, pp. 1181–1186, 2002.

[63] G. Bin, X. Gao, and Y. Wang, "VEP-Based Brain-Computer Interfaces: Time, Frequency, and Code Modulations 1.," no. November, pp. 22–26, 2009.

[64] K. Momose, "Evaluation of an eye gaze point detection method using VEP elicited by multi-pseudorandom stimulation for brain computer interface," *Annu. Int. Conf. IEEE Eng. Med. Biol. - Proc.*, vol. 2, pp. 5063–5066, 2007.

[65] S. I. Dimitriadis and A. D. MARIMPIS, "Enhancing Performance and Bit Rates in

a Brain–Computer Interface System with Phase-to-Amplitude Cross-Frequency Coupling: Evidence from Traditional c-VEP, fast c-VEP and SSVEP Designs," *Front. Neuroinform.*, vol. 12, no. April, p. 19, 2018.

[66] M. Kuba and Z. Kubova, "Visual evoked potentials specific for motion onset," pp. 83–89, 1992.

[67] Fei Guo, Bo Hong, Xiaorong Gao, and Shangkai Gao, "A brain computer interface based on motion-onset VEPs," *2008 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 4478–4481, 2008.

[68] M. Kuba, *Motion-onset Visual Evoked Potentials and their Diagnostic Applications*. .

[69] T. Martin, K. Huxlin, and V. Kavcic, "Motion-onset visual evoked potentials predict performance during a global direction discrimination task," *Neuropsychologia*, vol. 48, no. 12, pp. 3563–3572, 2010.

[70] T. Ma *et al.*, "The extraction of motion-onset VEP BCI features based on deep learning and compressed sensing," *J. Neurosci. Methods*, vol. 275, pp. 80–92, 2017.

[71] C. Raitta, U. Karhunen, A. Maria, and M. Naukkarinen, "Graefes Archiv Ophthalmologie Potentials During General Anaesthesia," vol. 144, pp. 139–144, 1979.

[72] A. S. Richard R. Uhl, Kenneth C. squires, David L. Bruce, "Effect of halothan Anesthesia on the Human Cortical Visual Evoked Response," *J. Anesthesiol.*, vol. 53, no. 3, 1980.

[73] D. H. York, M. W. Pulliam, J. G. Rosenfeld, and C. Watts, "Relationship between visual evoked potentials and intracranial pressure," *J. Neurosurg.*, vol. 55, no. 6, pp. 909–916, 1981.

[74] W. Tsang and R. F. Hampson, "Acute visual evoked potential changes in hydrocephalus," *Natl. Inst. Stand. Technol.*, no. m, pp. 331–333, 1986.

[75] E. L. Reilly, C. Kondo, J. A. Brunberg, and D. B. Doty, "Visual evoked potentials during hypothermia and prolonged circulatory arrest," *Electroencephalogr. Clin. Neurophysiol.*, vol. 45, no. 1, pp. 100–106, 1978.

[76] T. W. and J. E.O., "Evoked cortical potentials in patients with 'isoelectric' EEGs," *Electroencephalogr. Clin. Neurophysiol.*, vol. 35, pp. 301–309, 1973.

[77] P. L. Lee *et al.*, "An SSVEP-actuated brain computer interface using phase-tagged flickering sequences: A cursor system," *Ann. Biomed. Eng.*, vol. 38, no. 7, pp. 2383–2397, 2010.

[78] C. Jia, X. Gao, B. Hong, and S. Gao, "Frequency and phase mixed coding in SSVEP-based brain - Computer interface," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 200–206, 2011.

[79] N. V. MANYAKOV, N. CHUMERIN, and M. M. VAN HULLE, "Multichannel Decoding for Phase-Coded Ssvep Brain–Computer Interface," *Int. J. Neural Syst.*, vol. 22, no. 05, p. 1250022, 2012.

[80] P. G. H. Cla, “Comparison of Visual Evoked Potentials l o Stationary and to Moving Patterns,” vol. 164, pp. 156–164, 1973.

[81] W. Yan *et al.*, “Steady-State Motion Visual Evoked Potential (SSMVEP) based on equal luminance colored enhancement,” *PLoS One*, vol. 12, no. 1, pp. 1–18, 2017.

[82] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T.-P. Jung, and S. Gao, “High-speed spelling with a noninvasive brain–computer interface,” *Proc. Natl. Acad. Sci.*, vol. 112, no. 44, pp. E6058–E6067, 2015.

[83] D. Marshall, R. Beveridge, S. Wilson, and D. Coyle, “Interacting with multiple game genres using Motion Onset Visual Evoked Potentials,” *Proc. CGAMES 2015 USA - 20th Int. Conf. Comput. Games AI, Animat. Mobile, Interact. Multimedia, Educ. Serious Games*, pp. 18–27, 2015.

[84] S. S. Keerthi and D. DeCoste, “A modified finite newton method for fast solution of large scale linear SVMs,” *J. Mach. Learn. Res.*, vol. 6, pp. 341–361, 2005.

[85] N. R. Draper and H. Smith, “Applied Regression Analysis.” pp. 335–339, 1998.

[86] O. L. Mangasarian, “A finite Newton method for classification,” *Optim. Methods Softw.*, vol. 17, no. 5, pp. 913–929, 2002.

[87] A. J. Izenman, “Linear Discriminant Analysis,” in *Modern Multivariate Statistical Techniques*, 2008, pp. 237–263.

[88] N. Draper and H. Smith, “Applied Regression Analysis: stepwise regression procedure,” in *Applied regression analysis*, 1968, pp. 171–173.

[89] B. D. Van Veen and K. M. Buckley, “Beamforming: a versatile approach to spatial filtering,” *IEEE ASSP Mag.*, vol. 5, no. 2, pp. 4–24, 1988.

[90] B. D. Van Veen, W. Van Drongelen, M. Yuchtman, and A. Suzuki, “Localization of brain electrical activity via linearly constrained minimum variance spatial filtering,” *IEEE Trans. Biomed. Eng.*, vol. 44, no. 9, pp. 867–880, 1997.

[91] G. Van Dijck, A. Jezzini, S. Herwik, and S. Kisban, “Neural Information Processing. Models and Applications,” vol. 6444, no. November, 2010.

[92] “http://www.fieldtriptoolbox.org/template/layout: scalp recording layout.”

[93] M. M. Van Hulle, “neural recording: invasive vs non-invasive.”

[94] D. B. and D. Pelli, “Psychtoolbox.” .

[95] “Curry 7 compumedics neuroscan.” [Online]. Available: https://compumedicsneuroscan.com/curry-7-signal-processing-basic-advanced-source-analysis/.

[96] N. B. Andrea Kübler, Nicola Neumann, Barbara Wilhelm, Thilo Hinterberger, “Predictability of Brain-Computer Communication,” *J. Psychophysiol.*, vol. 18, pp. 121–129, 2004.

Master's thesis filing card

Student: Arno Libert

Title: Spatiotemporal beamforming for decoding motion-onset Visual Evoked Potentials: a BCI study.

Dutch title: Spatiotemporale beamforming voor het decoderen van motion-onset Visual Evoked Potentials: een BCI studie.

UDC: 681.3*I20

Abstract: Patients suffering from amyotrophic lateral sclerosis (ALS), other muscular and neurodegenerative diseases or spinal cord injuries have difficulty communicating their intention to people. The recent developments in the field of technology and artificial intelligence give the patients a method to express themselves using brain-computer interfacing (BCI). BCI is aimed at establishing a direct communication channel between the brain and an external device (often a computer) bypassing the need for muscular control, including speech and gestures. One of the less common electroencephalography (EEG)-based paradigms relies on the motion-onset visual evoked potentials (mVEP), a lateralized potential over the parieto-occipital scalp area evoked in sync with the onset of a briefly moving stimulus to which the subject pays attention. In this work, the effect of the stimulus direction, the combination of the moving stimulus with the P300 paradigm and the feasibility for the use of mVEP in BCI is investigated in terms of both scalp activations and decoding performance. For the first time, a novel spatiotemporal beamforming algorithm is utilized, and its performance is compared to that of existing mVEP decoders. While a clear effect of stimulus direction on scalp activation is observed, this does not seem to be the case for BCI performance for any of the considered decoders. The combination of the mVEP and P300 paradigm does not show any influence on the scalp activation. This, combined with the increase in fatigue for the subject by focusing on flashing stimuli, makes the combination of mVEP and P300 unadvisable. Spatiotemporal beamforming has the advantage of being significantly less computationally expensive, and in this way more suitable for online BCI applications. Finally, a mVEP spelling console BCI was developed and tested. Test results show a large sensitivity to peripheral stimulation for mVEP. Thus, making the designed layout of the spelling interface a key factor in the creation of mVEP based BCI.

Thesis submitted for the degree of Master of Science in Artificial Intelligence, option Engineering and Computer Science

Thesis supervisor: Prof. dr. ir. M. Van Hulle

Assessors: Prof. L. Geurts K. B. Wittevrongel

Mentor: B. Wittevrongel