# Learning Priors for Error-related Decoding in EEG data for Brain-Computer Interfacing

**Lernen von Priors zur fehlerbezogenen Dekodierung in EEG-Daten für Gehirn-Computer-Schnittstellen**

Bachelor-Thesis von Matthias Georg Schultheis
Tag der Einreichung:

Prüfer: Prof. Ph. D. Jan Peters
Betreuer: M. Sc. Natalie Faber

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Learning Priors for Error-related Decoding in EEG data for Brain-Computer Interfacing
Lernen von Priors zur fehlerbezogenen Dekodierung in EEG-Daten für Gehirn-Computer-Schnittstellen

Vorgelegte Bachelor-Thesis von Matthias Georg Schultheis

Prüfer: Prof. Ph. D. Jan Peters
Betreuer: M. Sc. Natalie Faber

Tag der Einreichung:

# Erklärung zur Bachelor-Thesis

Hiermit versichere ich, Matthias Georg Schultheis, die vorliegende Bachelor-Thesis ohne Hilfe Dritter und nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die aus Quellen entnommen wurden, sind als solche kenntlich gemacht worden. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen. In der abgegebenen Thesis stimmen die schriftliche und elektronische Fassung überein.

Darmstadt, den 28. Oktober 2016

(Matthias Georg Schultheis)

# Abstract

Brain-Computer Interfaces (BCIs) offer the possibility of controlling external devices by modulating brain activity. A user can send commands to control the BCI by performing mental tasks while wearing an electroencephalography (EEG) cap. EEG-based systems can already be used as a way for basic communication; however, due to non-stationarity of the brain signals, a system is only valid for a short amount of time after training. To provide remedy for this problem, an adaptive system can be built which adapts to changed brain signals and therefore stays suitable for a long period of time. In this thesis, an adaptive approach is presented introducing a prior over commands. In the decoding phase, this prior balances commands in the case where a traditional BCI becomes biased to specific commands. For applying such a system, brain signals need to be labeled online, i. e. the correct commands for control have to be known. Often, this information is only known by the user and unknown for the system. In the case of unknown labels, the system can infer feedback information by analyzing the user's brain signals shortly after presenting visual feedback about decoded commands. Neurobiological research found error detection mechanisms in the human brain which can be used for decoding feedback. Feedback indicating incorrect decoding of the signals leads to particular time-bound deflections in EEG signals. These deflections can function as features to infer correct commands for establishing an adaptive BCI.

This thesis investigates feedback decoding in the case of individuals playing a video game. This game is currently used for testing BCI systems for the use in daily tasks. In the game the user is supposed to send commands using an EEG system and receives feedback about the decoded commands. An experiment where subjects were supposed to play this game shows that the information whether any feedback was perceived can be decoded with an accuracy of 76.6% on average. In the presented experimental setup, the decoding whether correct or wrong commands were sent reaches an accuracy of 66.3%. For the discrimination analysis, subsampled signals of one electrode in the time domain are used as features and linear discriminant analysis and support vector machines are used as classification methods. Additionally, variations of the analysis such as using spatial and frequency-based features and additional preprocessing are considered.

# Zusammenfassung

Gehirn-Computer-Schnittstellen (GCS) bieten die Möglichkeit, externe Geräte durch Veränderung von Gehirnaktivität zu steuern. Ein Benutzer kann Befehle zur Steuerung der GCS senden, indem er mentale Aufgaben ausführt, während er eine Elektroenzephalografie(EEG)-Kappe trägt. EEG-basierte Systeme können bereits zur einfachen Kommunikation eingesetzt werden; allerdings ist ein System nach dem Trainieren nur für einen kurzen Zeitraum anwendbar aufgrund der Nichtstationariät des Gehirns. Um Abhilfe zu schaffen, können adaptive Systeme entwickelt werden, die sich an veränderte Gehirnsignale anpassen und daher über einen langen Zeitraum hinweg angemessen bleiben können. In dieser Thesis wird ein adaptiver Ansatz vorgestellt, welcher einen Prior über die Befehle einführt. In der Dekodierungsphase gleicht dieser Prior die Befehle aus für den Fall, in welchem eine traditionelle GCS einen Bias zu gewissen Befehlen entwickelt. Um ein solches System einzusetzen, müssen die Gehirnsignale zur Laufzeit gekennzeichnet sein, d. h. die korrekten Steurerungsbefehle müssen bekannt sein. Meist ist diese Information nur dem Benutzer und nicht dem System bekannt. Im Fall von unbekannten Kennzeichnungen kann das System Informationen über Rückmeldungen ableiten, indem Gehirnsignale des Benutzers kurz nach Präsentation von visuellen Rückmeldungen über entschlüsselte Befehle betrachtet werden. In der neurobiologischen Forschung wurden Fehlerdetektionsmechanismen im menschlichen Gehirn gefunden, die benutzt werden können, um Rückmeldungen zu entschlüsseln. Feedback, welches den Benutzer auf inkorrekte Entschlüsselung der Signale hinweist, führt zu bestimmten zeitgebundenen Ausschlägen in den EEG-Signalen. Diese Ausschläge können als Merkmal verwendet werden, um auf korrekte Befehle zu schließen und damit eine adaptive GCS zu schaffen.

Diese Thesis untersucht die Entschlüsselung von Rückmeldungen in einem Anwendungsfall, in dem Testpersonen ein Videospiel spielen. Dieses Spiel wird aktuell zum Testen von GCS zum Einsatz in täglichen Aufgaben eingesetzt. In dem Spiel sendet der Benutzer Befehle mittels einer GCS und erhält Rückmeldung über die entschlüsselten Befehle. Ein Experiment, in dem Testpersonen dieses Spiel spielten, zeigt, dass die Information, ob irgendeine Rückmeldung wahrgenommen wurde, mit einer Treffsicherheit von 75,6% durchschnittlich erkannt werden kann. Im vorgestellten Experimentaufbau erreicht die Entschlüsselung, ob korrekte oder inkorrekte Befehle gesendet wurden, eine Treffsicherheit von 66,3%. Für die Unterscheidungsanalyse wurden stichprobenartige Signale von zwei Elektroden im Zeitbereich als Unterscheidungsmerkmale verwendet und lineare Diskriminanzfunktionen sowie eine Stützvektormaschine (Support Vector Machine) zur Klassifikation verwendet. Zusätzlich wurden unterschiedliche Varianten der Analyse betrachtet wie beispielsweise räumliche und frequenzbasierte Unterscheidungsmerkmale und zusätzliche Vorverarbeitung der Signale.

# Acknowledgments

Foremost, I would like to offer my special thanks to my supervisor Natalie Faber for all her continuous advice and encouragement throughout my thesis.
I would also like to express gratitude to Professor Jan Peters for providing me the opportunity to work on such an exciting topic.
Not least, I would like to thank the rest of the cybathlon team for creating such a great working atmosphere in the lab and assisting me in my experiments. My special appreciation is herby dedicated to Thomas Hesse for many inspiring discussions which pushed me into the right direction.

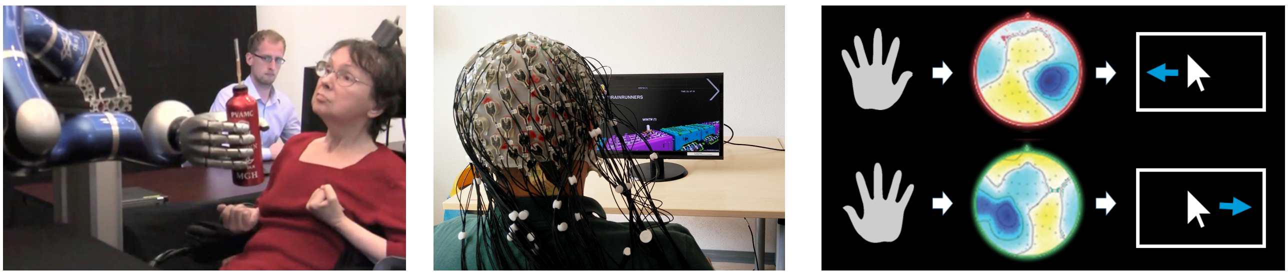
# Contents

# 1 Introduction

## 1.1 Motivation

Coupling of brains with machines has been an imagination in humans' minds for a long time. The opportunities to augment humans' physical and mental capabilities by means of external devices seem nearly limitless. Until recently, realizations have been restricted to Hollywood movies and science fiction books [1]. Current advantages in the fields of neuroscience and engineering could turn the vision into reality soon. By now, various methods have been developed in order to build an interface for connecting a brain with an external device [2]. Such an interface is called Brain-Computer Interface (BCI). The relatively young research field of BCIs combines findings of neuroscience, signal processing, machine learning, and information technology. Some application examples in daily life can already be spotted in the clinical sector where BCIs provide brain stimulation for parkinson disease and prosthesis control for amputees. Also applications for able-bodied individuals seem conceivable in the near future like for use in security, lie detection, alertness monitoring, and gaming. Figure 1.1a and 1.1b illustrates applications for disabled and able-bodied individuals.

In this thesis BCIs which are used to control external devices, are considered. The task is hereby to translate brain signals into commands that can control external devices or into messages for communication. To measure brain activity a typical option is to use an electroencephalography (EEG) device that records voltage on top of the scalp. To define the form of mental communication, a paradigm has to be designed. A well-studied paradigm is to use motor imaginations to encode commands where the user imagines movements with different body parts [3]. In a typical study human subjects have learned to move a cursor in two dimensions on a computer screen using EEG. Control was managed by left and right hand movement imagination leading to brain rhythm changes [4]. The control in the study is depicted in Figure 1.1c.



**(a)** **(b)** **(c)**

**Figure 1.1: Applications for BCIs. (a)** A stroke patient controls a robot arm by means of a BCI to pick up a cup[1]. **(b)** An individual is playing a video game by using a BCI. **(c)** Workflow to control a cursor on a screen in an experiment. Left hand imagination leads to brain rhythm changes that are interpreted to move the cursor to the left. Right hand imagination leads to a different pattern in brain signals which causes the cursor to be moved to the right.

A common problem with EEG is that recorded signals are contaminated by noise and comprise a mixture of numerous sources in the brain. Interpretation of signals therefore becomes often ambiguous and the user's intentions are not clearly readable. Due to non-stationarity of the brain [5, 6], patterns change already after a few hours which makes it even more difficult to discriminate reliably between users' intentions. For reliable communication it would be plausible to add a control mechanism that requires sending a command several times and only to forward it if their interpretations in decoding match. More elaborated approaches present the decoded command to the user in

[1] `http://www.theatlantic.com/health/archive/2012/05/the-brain-computer-interface-that-let-a-quadriplegic-woman-move-a-cup/257275/` (09/21/2016)

order to receive feedback. By taking the subject's brain signals into account, it can be derived whether the subject agrees with the decoded command. If a wrong command is presented, the activity of an error detection mechanism in the subject's brain signals can be observed. The command will finally be only executed if the mechanism is found to be inactive, i. e. the user agrees with the command. A promising approach seems to be using the error detection mechanism in order to update the decoding component of the system [7]. Due to non-stationarity of the brain, decoding is likely to become biased to one command. If error signals can be reliably detected, they can be used to introduce a prior over commands that is taken into account by the decoder and compensates for the bias.

In this thesis the improvement of a BCI system for a new application is addressed. A video game is taken into account which demands the player to send appropriate commands to the right time by means of an EEG system. If wrong commands given by the players can reliably be detected, a crucial step in order to establish a prior over commands is achieved and a bias in decoding can be compensated. If for example a wrong command has been sent, the probability of this command to be sent again can be decreased. The main novelty in this thesis is that the game represents almost a real world scenario. The BCI user has to process a lot of visual information and the time when feedback is presented is not predictable. Also, different paradigms are applied during the use of the BCI.

## 1.2 Overview of Chapters

This section provides an overview of the chapters contained in this thesis.

**Chapter 2** introduces the foundations of working with BCIs. The four common processing stages (i. e. signal acquisition, signal preprocessing, decoding, and interaction) are explained in detail.

**Chapter 3** presents work dealing with feedback decoding in BCIs which is related to the work in this thesis. It also describes how the work of this thesis complements existing research efforts.

**Chapter 4** presents the application addressed in this thesis, the video game, and the role of feedback detection in an adaptive BCI system. The conduction of a study for investigating feedback detection in the video game is elaborated and obtained results are analyzed.

**Chapter 5** concludes this thesis by summarizing the results and showing applications for future adaptive BCI systems. Also directions for continuing this work in the future are presented.

# 2 Brain Computer Interfaces

A Brain-Computer Interface (BCI) is an interface allowing communication between a brain and an external device such as a computer. The purpose is to translate brain activity into commands which can control devices or into messages for communication.

The process of a BCI is usually divided into several processing stages. In Figure 2.1 a traditional BCI pipeline is presented to provide an overview of the functioning of a BCI. The pipeline consists of the following steps:

- **Signal acquisition:** Signals of the brain are recorded. This stage is further explained in section 2.1.
- **Signal processing:** The raw signals are preprocessed and methods to reduce artifacts are applied. More information about the preprocessing step can be found in section 2.2.
- **Decoding:** This stage consists of two steps: At first, features are extracted containing relevant information for interpretation of the signal. Secondly, machine learning classification methods are applied in order to generate control signals based on patterns in the features. Both steps are presented in section 2.3.
- **Interaction:** The control signals cause changes in the environment of the user. Usually the user can perceive them and receives feedback about his performance. For example a robot arm may be actuated or a command in a video game may be executed.

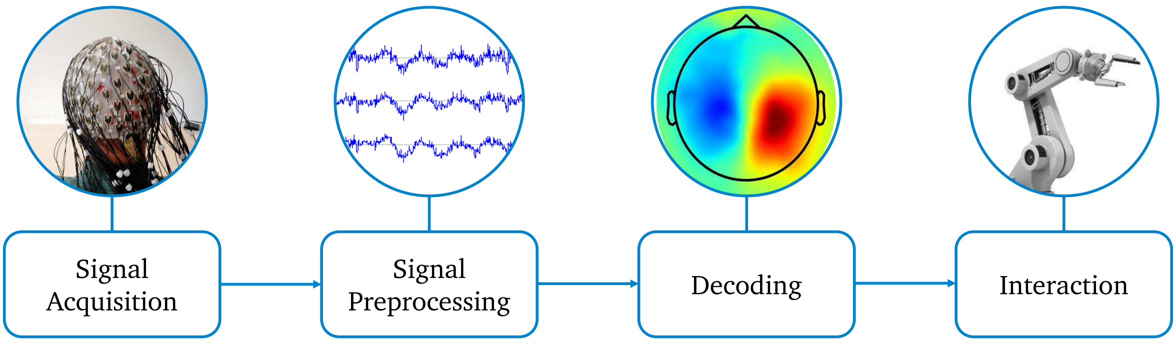



**Figure 2.1: A common BCI pipeline.** The system is usually divided into the four processing stages *signal acquisition*, *signal processing*, *decoding*, and *interaction*.

In sections 2.1, 2.2, 2.3, the presented pipeline steps (i. e. *signal acquisition*, *signal processing*, *decoding*, and *interaction*) are further explained. Ways of communication by means of a BCI are introduced in section 2.4. At the end of this chapter in section 2.5, a slightly modified adaptive pipeline is presented, which is addressed in this thesis.

## 2.1 Signal Acquisition

For interpreting recorded brain signals, it is important to understand how signals are acquired. At first, a general overview about available methods is given. Afterwards, signal recording by means of electroencephalography (EEG) is explained in detail as this method is used in this thesis for recording brain data. BCIs can be grouped into three main classes by considering their signal acquisition type: invasive, partially-invasive and non-invasive BCIs.

In case of invasive BCIs, electrodes are directly implanted into the grey matter of the brain by neurosurgery. This method yields the best quality of signals as there is a direct connection to the brain cells. However, this method comes along with high risks due to infections and is therefore ethically questionable.

Partially-invasive BCIs are implanted inside the skull but rest outside the brain. Thus, they require also an implantation by neurosurgery. Partially-invasive methods do not provide such a high signal quality as invasive methods as they are only placed on top of the brain and receive mixed information of whole neural populations. A widespread technology is *Electrocorticography* where electrical activity of the brain is measured by electrodes from beneath the skull under the dura matter [8]. This technology is depicted in Figure 2.2a.
Non-invasive methods do not require any form of surgery as these BCIs measure brain activity from above the scalp. These methods minimize risks of infections and are easy to use. However, as the skull dampens brain signals, recorded signals are heavily contaminated by artifacts and the spatial resolution is rather low [1]. In order to employ this method, signal processing techniques such as spatial filters have to be applied in order to enhance signal quality. Common non-invasive techniques are *Magnetoencephalography* recording magnetic fields around the scalp, and *Functional magnetic resonance imaging* measuring changes associated with blood oxygenation and flow in the brain [9]. Both approaches are depicted in Figure 2.2b and 2.2c. *Electroencephalography (EEG)* is another non-invasive method recording voltage fluctuations which is further explained in the following section.

**(a)** **(b)** **(c)**

**Figure 2.2: Methods to acquire brain signals. (a)** Electrocorticography is an semi-invasive method. A pad mounted with electrodes is implanted on top of the dura matter of the brain.[1] **(b)** A person sitting in a Magnetoencephalography machine. In this non-invasive method, magnetic fields around the scalp are recorded.[2] **(c)** A machine for functional magnetic resonance imaging recordings. This non-invasive technique works by detecting changes in blood flow.[3]

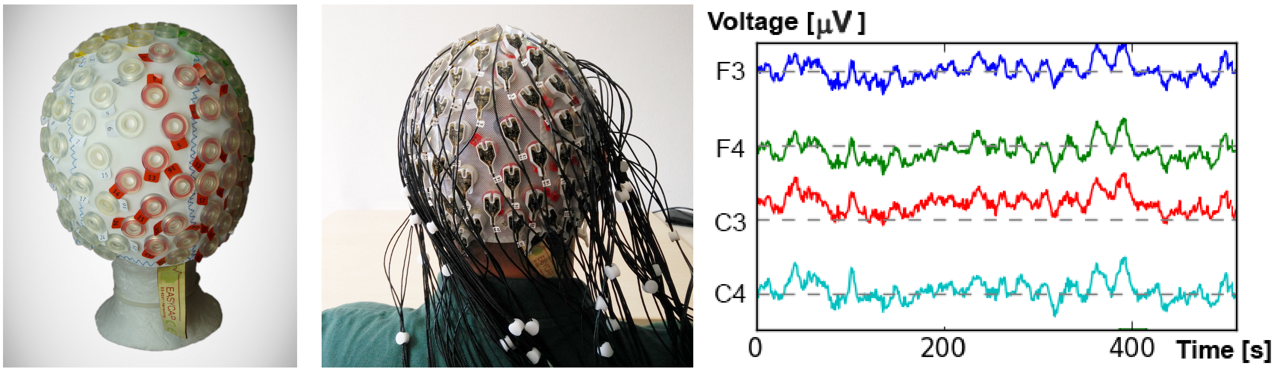
### 2.1.1 Electroencephalography

Electroencephalography (EEG) is a common BCI method for non-invasive signal acquisition. In order to obtain these signals, multiple electrodes are placed on top of the scalp and record electrical activity. Usually, the electrodes measure voltage against a reference electrode and one electrode operates as ground. Electrodes are usually first plugged into a cap (see Figure 2.3a) that is subsequently put onto the scalp of the BCI user (see Figure 2.3b). Herewith it is ensured that electrodes are spatially distributed in a structured manner across the scalp. As the electrodes are placed on the scalp and not in immediate proximity to the brain, they can only capture a mixed activity of neuron population. EEG systems therefore achieve a relatively small spatial resolution and recorded signals are contaminated by artifacts. Concurrent ocular and muscular activity affects the signals so that advanced preprocessing techniques (see section 2.2) are needed to reveal brain activity patterns. Exemplary EEG signals are depicted in Figure 2.3c

Compared to other non-invasive methods, EEG systems are simple to use, portable and have low hardware costs. They also provide a good trade-off between temporal and spatial resolution and are therefore suitable for many cases. Besides clinical applications such as detecting brain disorders, EEG methods provide a way to communicate with a external devices. As an example it may be used by entertainment industry customers in the near future, e. g. in order to control video games [10]. More importantly, EEG systems can be applied to paralyzed people or those who suffer from severe muscle disabilities. As most communication methods need voluntary muscle control,

[1] `http://www.neurofisiologia.net/wp-content/uploads/2009/07/corticografia.jpg` (09/22/2016)
[2] `http://www.supraconductivite.fr/en/index.php?p=applications-medical-meg` (09/22/2016)
[3] `http://www.trueimpact.ca/tag/neuroimaging/` (09/22/2016)

a BCI can provide the possibility to communicate and perform tasks like controlling an electric wheelchair for those people. In this thesis EEG is used as method for recording brain signals.




(a) (b) (c)

**Figure 2.3: EEG signal acquisition. (a)** An EEG cap where electrodes can be plugged into. **(b)** An individual during EEG recordings. **(c)** Typical EEG signals at four electrodes.

## 2.2 Signal Preprocessing

Once recorded signals have been acquired, they are preprocessed in order to enhance relevant brain information. Within this thesis, signals are centered and temporal as well as spatial filters are applied. Temporal filters are used to restrict the signal to specific frequency bands. Spatial filtering techniques such as common average reference (CAR) and the Laplace filter are applied for enhancing local brain information at the electrodes. In the following the applied preprocessing methods are presented.

### 2.2.1 Centering

As brain signals are shown be non-stationary [5, 6] and recording conditions are always slightly different, the range of recorded signals varies in every session. For the application of spatial filters it is often assumed that the signals have a mean equal to zero. In order for signals to be comparable and to apply spatial filters, it is a common step to subtract from each electrode signal its average over time. This method is referred to as *Centering*.

### 2.2.2 Temporal Filters

Temporal filters are applied to signals to remove components lying on specific frequency bands. Frequently they are used to attenuate frequency-specific noise. The applied filters are mostly of one of the following types:

- **High-pass filter:** attenuates low frequencies and lets high frequencies pass
- **Low-pass filter:** attenuates high frequencies and lets low frequencies pass
- **Band-pass filter:** lets only frequencies in a specific band pass
- **Band-stop filter:** attenuates frequencies that lie in a specific band
- **Notch filter:** rejects one specific frequency

A common realization of these filters is the *Butterworth* filter [11, 12]. One application of temporal filters for EEG data is to remove drifts in the signals. Low frequency components of the signals result in drifts that makes the signals difficult to compare. By applying a high-pass filter, these drifts are attenuated. Another example is alternating current influencing recorded signals at one specific frequency. In case of Europe, alternating current usually oscillates at a frequency of 50 Hz. A 50 Hz notch filter can be used to remove these influences of the signal.

### 2.2.3 Common average reference

EEG signals are usually recorded against a reference electrode and are therefore reference-dependent. Common average reference (CAR) is a common spatial filtering technique making the signals reference-free [13]. For applying a CAR, the mean of all channels is subtracted from each channel for each time step. Due to re-referencing the average brain activity is suppressed such that the local information from sources below the electrodes are enhanced.

### 2.2.4 Laplace Filter

The Laplace Filter is a spatial filtering technique used to enhance local information at electrodes. From each electrode signal, a linear combination of the signals of surrounding electrodes is subtracted. This method suppresses the activity of the sources beneath other electrodes causing the local information to be intensified.
A simple way to realize a Laplace filter is to subtract the mean of the signals of direct neighboring electrodes from a electrode signal. This procedure is illustrated in Figure 2.4.

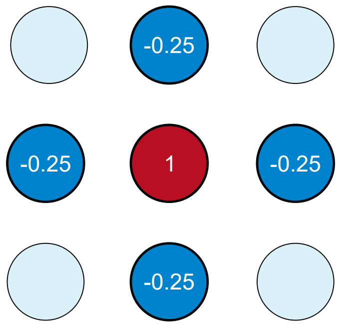



**Figure 2.4: Laplace filter.** The figure presents a clipping of the relevant EEG electrodes for application of the filter on the electrode depicted in red. Local information at this electrode is enhanced by subtracting the mean of the signal of the neighboring electrodes depicted in blue.

## 2.3 Decoding

In the Decoding phase the preprocessed signals are interpreted to derive a command for control. This phase is divided into two steps. In the first one, which is called Feature extraction, relevant information of the signal for discrimination is extracted. In the second step which is called Classification, the command for control encoded in the signal is derived based on previously extracted information.

Formally, the goal in the feature extraction part is to generate a $p$-dimensional feature vector $x \in \mathbb{R}^p$ containing the important properties of the signal to discriminate between distinct mental commands. The information to extract from the signal and the form of representation of the feature vector hereby depends on the application. If clear patterns are visible in the time domain, it is suitable to subsample the signal in the time domain to reach a lower dimensional representation. A lower dimensionality of features prevents the classification of becoming too much adapted to the data when there is a small number of samples. These circumstances exist usually in the case of EEG recordings as the number of samples is mostly rather low. In other applications where frequency components have to be taken into account, the signals are transformed into the frequency domain. By methods like *Fast Fourier Transform* [14], a power spectrum of frequency components can be obtained. The power of specific frequency bands can then be used as feature components. Two applications incorporating features in the time and frequency domain, respectively, are presented in section 2.4.

After having generated features of the signal, the classification step consists of finding patterns in the data that can be regarded for determining the encoded mental action. Classification methods of the field of Machine Learning can be applied to find those patterns and to decode signals by considering generated features. For decoding of brain signals, linear classification methods were shown to be suitable [15]. In the following, two linear approaches for classification, i. e. Linear Discriminant Analysis (LDA) and Support Vector Machines (SVMs), are explained in detail.

### 2.3.1 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is a simple classification method for classifying BCI data. LDA is a linear binary classifier which projects a $p$-dimensional input vector $\mathbf{x}$ onto a hyperplane. This hyperplane acts as decision boundary and divides the input space into two halfspaces where each represents one of the classes $(+1, -1)$. The decision boundary can be stated using its normal vector $\mathbf{w}$ and some threshold $w_0$:

$$\mathbf{w}^T\mathbf{x} + w_0 = 0$$

A new input vector $\mathbf{x} \in \mathbb{R}^P$ can then be classified by computing its predicted class label $y = \text{sign}(\mathbf{w}^T\mathbf{x} + w_0)$. The formula assigns $\mathbf{x}$ to class $+1$ if $\mathbf{w}^T\mathbf{x} + w_0$ is negative and to class $-1$ otherwise.

In order to compute $\mathbf{w}$ and $w_0$, it is assumed that the class conditionals $P(\mathbf{x}|c)$ are Gaussian distributed with mean $\boldsymbol{\mu}_c$ and covariance $\boldsymbol{\Sigma}_c$ for $c \in \{+1, -1\}$:

$$P(\mathbf{x}|c) = \frac{1}{\sqrt{(2\pi)^p \det(\boldsymbol{\Sigma}_c)}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_c)^T \boldsymbol{\Sigma}_c (\mathbf{x} - \boldsymbol{\mu}_c)\right)$$

The optimal decision strategy is to assign $\mathbf{x}$ to class $+1$ if the log likelihood ratio $\log(P(\mathbf{x}|c = +1)/P(\mathbf{x}|c = -1))$ is above a threshold and otherwise to assign $\mathbf{x}$ to class $-1$. As both class conditional distributions are Gaussian, this comparison can be reformulated to

$$(\mathbf{x} - \boldsymbol{\mu}_{+1})^T \boldsymbol{\Sigma}_{+1}^{-1}(\mathbf{x} - \boldsymbol{\mu}_{+1}) - (\mathbf{x} - \boldsymbol{\mu}_{-1})^T \boldsymbol{\Sigma}_{-1}^{-1}(\mathbf{x} - \boldsymbol{\mu}_{-1}) > k$$

where $k \in \mathbb{R}$ is the threshold. By making the assumption that the class covariances are equal, i. e., $\boldsymbol{\Sigma}_{+1} = \boldsymbol{\Sigma}_{-1} = \boldsymbol{\Sigma}$ and have full rank, the following classification criterion is obtained:

$$\mathbf{w}^T\mathbf{x} > k$$

where $\mathbf{w} = \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_{+1} - \boldsymbol{\mu}_{-1})$. This choice of $\mathbf{w}$ leads to a decision boundary maximizing the distance between means $m_1$ and $m_2$ of the projected data $\mathbf{w}^T\mathbf{x}$ for each class and at the same time minimizes the within-class variance of the projected data. The projection within LDA for a two-dimensional example is depicted in Figure 2.5.
A common choice for $k$ is to set it to the average of the projection of both class means. This way $k$ can be obtained by

$$k = \frac{\mathbf{w}^T(\boldsymbol{\mu}_{+1} + \boldsymbol{\mu}_{-1})}{2}$$

LDA is a popular classification method in BCI research as it is simple to implement and computation is fast enough in order to apply it online. Further information about LDA can be found in [16].

### 2.3.2 Support Vector Machine

As explained in section 2.3.1, LDA uses a separating hyperplane of the form $\mathbf{w}^T\mathbf{x} + w_0$ in order to discriminate between two classes. The resulting hyperplane is only one among a potentially infinite number of hyperplanes that can be used to separate the input data if the data is linear separable (see Figure 2.6a). According to Vapnik [18] best generalization is achieved if we select the hyperplane yielding the largest margin between both classes. The margin describes the maximal width of the slab which has no interior data points and lies parallel to the hyperplane.

A Support Vector Machine (SVM) is a classifier that finds the separating hyperplane which maximizes the margin between the samples of two classes. The data points which lie closest to the decision boundary and form the boundary of the slab are named *support vectors*. An two-dimensional example of the application of a SVM is

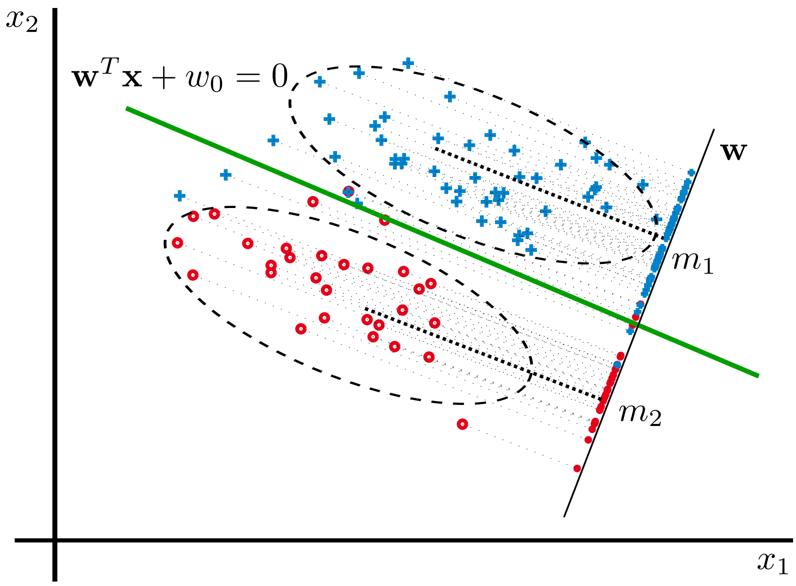



**Figure 2.5: Linear Discriminant Analysis (LDA).** In LDA each of two classes is modeled as being generated by a Gaussian. The plot depicts these Gaussians as dotted ovals around a set of two-dimensional data points. Class 1 is represented by red circles while class 2 is represented by blue crosses. LDA finds a vector **w** maximizing the distance between $m_1$ and $m_2$ of the projected data while minimizing the within-class variance. This vector is orthogonal to the separating hyperplane (depicted in green). The projected data points onto **w** are represented by smaller circles and crosses. [1, 17]

depicted in Figure 2.6b. In case of SVMs the equation for the separating hyperplane is like in the LDA case of the form

$$f(x) = \mathbf{w}^T\mathbf{x} + w_0.$$

A new data point $\mathbf{x}$ can be classified by computing $y = \operatorname{sgn}(f(\mathbf{x}))$. In order to obtain $\mathbf{w}$ and $w_0$, a set of samples $\{(\mathbf{x}_i, y_i) \mid i = 1, \ldots, K\}$ is regarded consisting of data points $\mathbf{x}_i \in \mathbb{R}^P$ and their corresponding category $y_i \in \{-1, 1\}$. For now it is assumed that the data is linear separable. The distance of a point to the separating hyperplane is given by

$$d(\mathbf{x}) = \frac{\mathbf{w}^T\mathbf{x}_i + w_0}{||\mathbf{w}||}.$$

The optimal hyperplane can be stated in an infinite number of different ways by scaling of **w** and $w_0$. It is convenient to set

$$|\mathbf{w}^T\mathbf{x}_s + w_0| = 1$$

for every support vector $\mathbf{x}_s$. For the support vectors the distance to the separating hyperplane $\frac{1}{||\mathbf{w}||}$. As the margin $m$ is twice the distance to the closest data points, m is obtained by

$$m = \frac{2}{||\mathbf{w}||}.$$

Maximizing the margin $m$ is equivalent to minimizing $||\mathbf{w}||$. The objective can thus be stated as minimizing $||\mathbf{w}^2||$ subject to the constraint that all data points are correctly classified. This task can be framed as a quadratic constrained optimization problem which can be efficiently solved:

$$\begin{array}{ll} \min\limits_{\mathbf{w},w_0} & \frac{1}{2}||\mathbf{w}||^2 \\ \text{s.t.} & y_i(\mathbf{w}^T\mathbf{x}_i + w_0) \;\; \geq \;\; 1, \quad i = 1, \ldots, K \end{array}$$

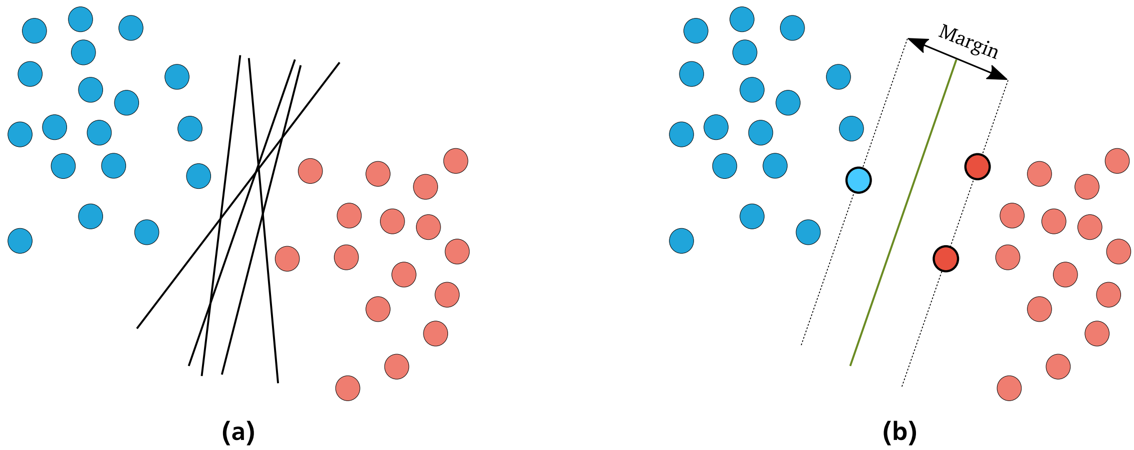



**Figure 2.6: Support Vector Machine (SVM). (a)** A case where two classes (blue and red points) are linear separable. There exist an infinite number of linear decision boundaries which can be used to discriminate between both classes. **(b)** An SVM chooses the optimal decision boundary which maximizes the margin between both classes. The data points lying closes to the boundary are called support vectors (depicted in more saturated colors). Only these points are needed for determining the boundary.

**Soft Margin Support Vector Machine**

The SVM as previously derived, is only applicable to linear separable data due to the definition of the margin. However, in most cases EEG data cannot be perfectly separated. For this case the data is supposed to be separated by making a minimum amount of error. Therefore *slack variable* $\xi_i$ are introduced for each data point which accounts for the error and makes the data linear separable. Additionally to the objective previously defined, it has to be ensured that the influence of the slack variables is kept as small as possible. Therefore, an additional term can be added to the objective in order to minimize the sum of slack variables multiplied by a penalty parameter $C$. This method is referred to as *Soft margin SVM* [19]. Formally we obtain the following optimization problem:

$$\begin{aligned} \min_{\mathbf{w},\xi,w_0} \quad & \tfrac{1}{2}||\mathbf{w}||^2 + C\sum_{i=1}^{K}\xi_i \\ \text{s.t.} \quad y_i(\mathbf{w}^T\mathbf{x}_i + w_0) &\geq 1, \quad i = 1,\dots,K, \\ \xi_i &\geq 0, \quad i = 1,\dots,K \end{aligned}$$

So far only SVMs having a linear decision boundary were considered. By using the kernel trick [20], an effective non-linear mapping from our data to a high-dimensional space can be achieved where the data becomes linear separable. However, for EEG data non-linear classification methods only lead to small performance gains [15] and are therefore not used in this thesis. More information about non-linear SVMs can be found in [21].

## 2.4 Communication by means of a BCI

In this thesis two common types of patterns in signals are regarded which can be used to establish communication by using a BCI. event-related potentials (ERPs) are responses visible as deflections in the time domain of EEG, event-related desyncronizations (ERDs) and event-related syncronizations (ERSs) are visible as frequency specific changes. In the following sections both types are presented. Afterwards, the use of paradigms in the context of BCIs is explained to provide an overview about communication in a typical BCI.

### 2.4.1 Event-related Potential

Several kinds of events induce time-locked electrophysiological responses in the brain that can be measured in the time domain of EEG signals. These responses are referred to as event-related potentials (ERPs). By detection of ERPs, the presence or absence of a stimulus can be concluded which can be used for communication. After presenting the characteristics of ERPs, an example frequently used for communication, the P300 ERP, is given. error-related negativity (ERN), another type of a ERPs which is mainly addressed in this thesis, is covered in the related work section in chapter 3.

ERPs are mostly triggered by sensory stimuli but can also be responses to cognitive or motor events. They manifest in form of deflections in the time domain of EEG and arise after a fixed time delay to the stimulus. As electrical fields propagated from the brain to the electrode have to be large enough for being able to be measures by EEG, many neurons have to be jointly responsible for the initiation of an ERP. ERPs therefore represents a common activity of neuron ensembles [22, 23, 24, 25].

In single trial signals the ERP characteristics are mostly not clearly visible as they are shadowed by background brain activity and other noise. As the noise effects the signals approximately additively [26], averaging techniques can be applied to visualize ERP patterns. By conducting multiple trials and averaging over those, background brain activity as well as other noise cancels out and the relevant waveforms become visible. We can derive the reasonability of the averaging technique if we model that our measured signal $x_k(t)$ of trial $k$ at time $t$ is additively composed of a fixed source signal $s(t)$ responsible for the ERP and Gaussian noise $n_k(t) \sim \mathcal{N}(0, \sigma_0)$:

$$x_k(t) = s(t) + n_k(t)$$

We can show that by collecting an increasing number of trials containing the ERP, the mean of the trial signals converges to the pure ERP source signal.

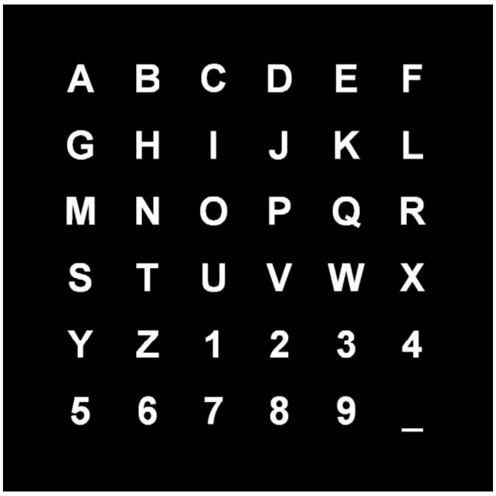
$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} x_k(t) &= \lim_{N\to\infty} \frac{1}{N} \left( \sum_{i=1}^{N} s(t) + \sum_{i=1}^{N} n_k(t) \right) && (2.1) \\
&= s(t) + \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} n_k(t) && (2.2) \\
&= s(t) + \lim_{N\to\infty} \frac{1}{N} \cdot N \cdot 0 && (2.3) \\
&= s(t) && (2.4)
\end{aligned}$$

In step (2.3) we have used the central limit theorem [27]. $N$ denotes the number of trials taken into account.

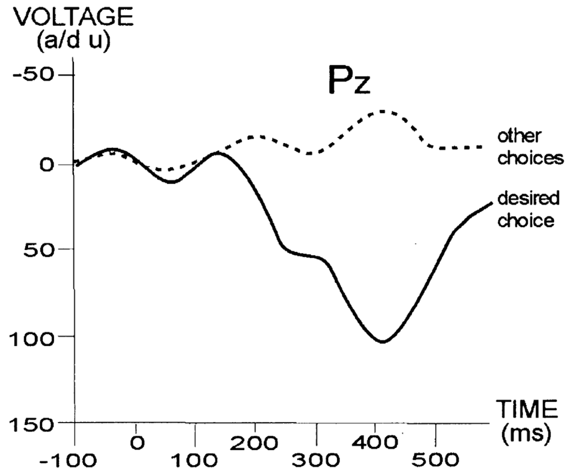
In this thesis, ERPs play a crucial role for checking the presence of specific stimuli.

**P300 Evoked Potential**

A well-studied ERP is the *P300 evoked potential* [28] which is depicted in Fig. 2.7b. It manifests as positive deflection of voltage in EEG approximately 300 ms subsequent to a stimulus. The deflection is mostly distributed in the scalp region over the parietal lobe. It surfaces when an infrequent target stimulus is applied to a subject. An example for the appearance of this ERP is the presentation of auditory stimuli in which one tone different from the other ones is given. The P300 also arises when a subject perceives an infrequent stimulus that is relevant for the task he is performing. A well-known application where P300 is used is the *P300 speller* [29]. By observing a matrix of figures and letters (see Figure 2.7a) where rows and columns are successively highlighted in a random order a subject can spell by focusing on the desired letter. Each time the highlighted row or column contains the chosen letter, a mechanism in the brain which is based on focus triggers an ERP in the brain signals. In cases where other rows or columns not containing the desired letter are highlighted, the described deflection cannot be observed. After testing all rows and columns the position of the desired letter on the matrix can be identified.




(a)




(b)

**Figure 2.7: P300 event-related potential. (a)** The P300 speller can be used for spelling by using a BCI. Rows and columns are successively highlighted and a P300 event-related potential can be observed in the EEG when the letter focused on is lighted[1]. **(b)** The P300 event-related potential manifests as positive deflection in the EEG. It appears in case of the P300 speller when the row or column of the desired letter is highlighted (desired choice). Other choices will lead to no deflection (other choices). Note that the ERP is plotted with negative voltages upward, a common practice in ERP research. [9]

### 2.4.2 Event-related Desynchronization/Synchronization

Sensory and cognitive processing as well as motor behavior is not only visible in form of ERPs in EEG but can also lead to frequency specific changes of EEG rhythms. These changes generally consist of a short-lasting decrease or increase of power in a specific frequency band. The change of power is caused by a decrease or increase of the underlying neural populations, respectively. The former is referred to as event-related desyncronization [30, 31], the latter as event-related syncronization [32] (see Figure 2.9). In this thesis the decoding whether positive or negative feedback was perceived by a subject is investigated also by considering frequency-specific changes (see chapter 4).

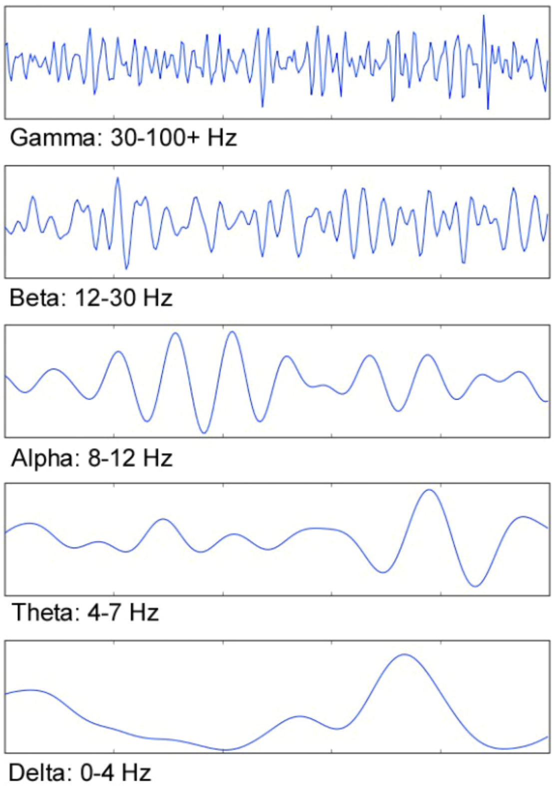



**Figure 2.8:** An EEG signal can be decomposed into several frequency components.[2]

Compared to ERPs, ERDs and ERSs are also time-locked but are not directly visible in the time domain. Thus, they cannot be extracted by a simple linear method such as averaging but require more elaborated techniques. For frequency analysis, methods like *Fast Fourier Transform* are applied to obtain a power spectrum of frequency components. The power in specific frequency bands can then be used as feature in the decoding step. Generally, the power within a specific frequency band is displayed relative to the power obtained in a baseline which is recorded shortly before the

[1] http://www.bci2000.org/wiki/index.php/File:P3SpellerMatrix.PNG (09/21/2016)
[2] http://neurosky.com/wp-content/uploads/2015/05/Screen-Shot-2015-05-14-at-10.41.17-AM.png (09/21/2016)

event.

In Fig. 2.8 common frequency bands used in EEGs are presented. Generally, the amplitude of brain oscillations is negatively correlated with their frequency. Thus with increasing frequency the amplitude of fluctuations decreases. A large number of interconnecting neurons which lead to many coherently activated neurons lets the amplitude increase and the frequency decrease [33, 34].

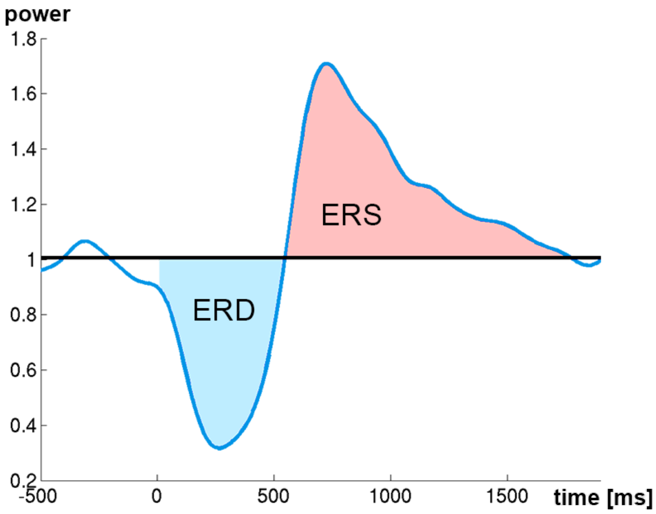



**Figure 2.9: Event-related Desyncronization/Syncronization (ERD/ERS).** ERDs/ERSs manifest as decreased or increased power of EEG signals, respectively. The plot depicts the signal band power (blue) in case of an event occurring at 0 ms leading to an ERD (0 — 500 ms) followed by an ERS (500 — 2000 ms). The signal power is compared against a constant baseline level (solid black). The baseline is usually estimated as averaged power across a reference interval prior to the event.[1]

### Common ERD/ERS Patterns

Activations of frequency bands in EEG signals are often taken into account for establishing communication by using a BCI. The following section explains common ERD and ERS patterns for communication to provide insights to the reader about the importance of this method in BCI research.

Performed and imagined body movements typically activate premotor and primary sensorimotor areas and are a common way for interacting with a BCI. These activations result in amplitude changes in low frequency bands such as the mu (8–12 Hz) or beta band (13–30 Hz). During the movement, an ERD can be observed in form of reduced power in these frequency bands.
A study of Wolpaw and McFarland [4] presents an application for ERD patterns in BCI communication. Subjects were supposed to move a cursor on a screen to a desired position by using a motor-imagery paradigm. For control the amplitude of the mu-band over the sensorimotor cortex was coupled to a movement of the cursor. After a few trials, the subjects learned to modulate their mu-rhythm activity to generate a change in the power and were able to control the cursor. Modern systems currently reach an accuracy of 80 - 95 % for discrimination between two commands by using motor-imagery paradigms [1].

Besides motor-imagery paradigms, several other mental tasks are found to modulate specific brain rhythms. Cognitive and memory performance often results in observable changes in the alpha and theta band. As an example, performing judgment tasks was found to cause ERDs in the alpha band [35]. Long- and short term memory tasks also cause frequency specific changes in the brain; while the former leads to an ERD in the upper alpha band, the latter triggers an ERS in the theta band [36, 37]. Moreover, slower alpha frequencies reflect attentional demands such as alertness and expectancy [38]. In spatial navigation tasks, theta oscillations can be observed [39, 40].

[1] http://www.bbci.de/supplementary/conditionalERD/conditionalERD.htm (10/02/2016)

### 2.4.3 Paradigm

For communication by means of a BCI, numerous experimental paradigms have been introduced [41, 42]. This section provides an overview of the employment of paradigms for the example pipeline which was presented in Figure 2.1. Moreover, in this section the reader gains an understanding of the procedure for selecting relevant parts of recorded brain signals for decoding.

A crucial step for employing a paradigm is to discover mental actions leading to responses in the brain which can be accurately detected in EEG. One class of paradigms frequently used is the motor-imagery paradigm demanding the user to imagine body movements for control [1]. To obtain knowledge about the manifestation of patterns in the brain signals linked to the paradigm, it is helpful to take neurobiological knowledge into account. For sensorimotor actions and imaginations there exists a rough topographical mapping from sensory and motor body parts to brain regions. This mapping was initially introduced by Dr. Wilder Penfield and is referred to as *cortical homunculus* [43]. The sensory and motor cortical homunculus are depicted in Figure 2.10. They provide an overview about activated brain regions in case of sensing touch or executing movements. Considering motor-imagery paradigms, a decoder can deduce performed movement imaginations by regarding activation patterns in these specific brain regions.

A common way of encoding commands, which a BCI system can discriminate between, is imagination of left and right hand movements. Right hand movement is processed mainly in the left brain hemisphere and triggers ERDs in left primary sensorimotor areas whereas left hand movement causes such patterns in right primary sensorimotor areas [3]. By detecting ERDs in these regions, the performed imagination and therewith the desired command can be deduced.

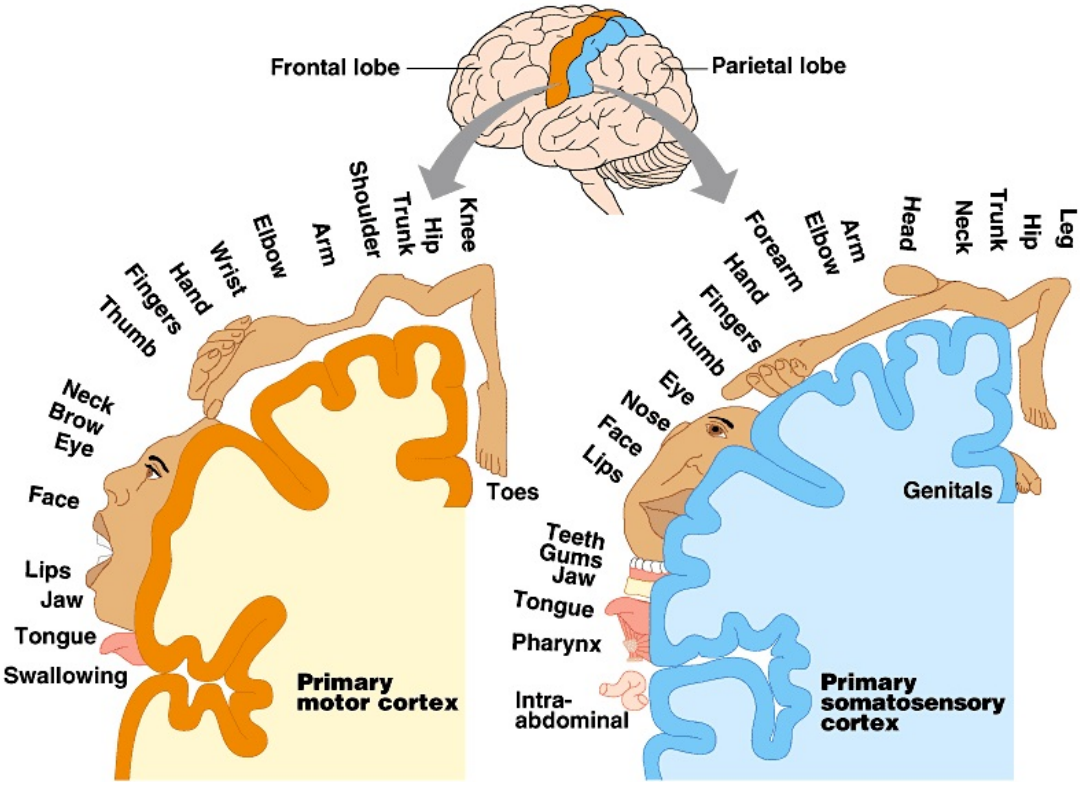



**Figure 2.10: Sensory and motor cortical homunculus.** The homunculus represents a rough topographical mapping from sensory and motor body parts to brain regions. The sensory cortical homunculus shows regions of the primary motor cortex responsible for processing the sense of touch in body parts (left). The motor cortical homunculus shows regions in the primary somatosensory cortex becoming active in case of planning and executing body movements (right).[1]

[1] http://cogsci.bme.hu/~ikovacs/latas2005/prepI_4_2.html (08/10/2016)

Including prior knowledge in the feature selection process is important for decoding. Taking the whole set of channels and full signal information leads to a high feature space for the classifier which makes classification difficult [44]. In this case, the classification method likely makes use of artifacts. Muscular or ocular activity patterns may be used for better discrimination between mental tasks but are not beneficial for the goal to analyze brain data. For feature selection, promising channels often are preselected which have been found to capture brain activity patterns suitable for the paradigm.

Besides motor-imagery paradigms, various other mental tasks have been found to be applied for controlling a BCI system. In the following, promising mental tasks are described [41, 42]:

- **Mental rotation:** visualize a three dimensional figure to rotate in three-dimensional space
- **Association:** list words in a chain of associations, e. g. generate words beginning with a specific letter
- **Auditory imagery:** imagine listening to a familiar tune and focus on the melody
- **Spatial navigation:** imagine navigation through a familiar place like such as walking from room to room in a house and focus on the orientation
- **Calculation:** perform calculations, e. g. successively subtracting numbers or enumerating number sequences

## 2.5 Adaptive BCIs

Traditional BCI systems are trained by using offline data which was recorded beforehand. Afterwards learned functions are kept fixed during use. However, brain signals change over time, even within a single session, which makes the BCI system no longer optimal. Changes in brain signals can be caused by internal factors such as strategy changes by the BCI user or fatigue, or by external factors such as change of electrode impedance.

One solution for this problem is to periodically update the classifier offline by using newly collected data. However, this approach lacks data efficiency and the interval for classifier retraining is unclear. A promising approach is to use feedback information to adapt the system continuously on an ongoing basis to keep it optimal for longer amounts of time. Such a BCI system is referred to as *adaptive BCIs* and is depicted in Figure 2.11. In an adaptive BCI system, the classifier is usually initialized with respect to previous sessions. During use in the current session the classifier adapts to current brain signals. The adaptation has to be done with great care as rapid and unpredictable changes are likely to confuse the user. Furthermore, too large incorrect adaptations will decrease the performance.

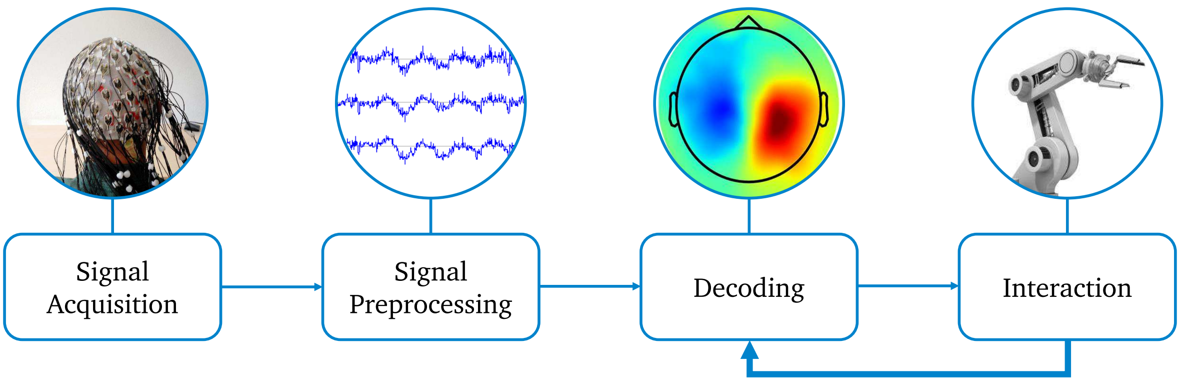



**Figure 2.11: Example adaptive EEG pipeline.** In an adaptive BCI, information of interaction is used in the decoding step to adapt the system to changed EEG signals.

In this thesis, foundations for the use of an adaptive BCI system are laid by decoding information about feedback presented to the user.

# 3 Related Work

Event-related potentials (ERPs) occurring after the perception of feedback indicating erroneous performance were first found by Hohnsbein and Falkenstein in 1989. They provided evidence for the existence of a neural system in the human brain that is responsible for error detection and correction [45, 46]. After negative feedback was given to a subject, a negative deflection in the time domain in the fronto-central region could be observed. Gering et al. also observed this type of ERP approximately 100 ms after negative feedback was given and named it error-related negativity (ERN) [47, 48].

Miltner et al. conducted a study to investigate situations where subjects were not able not derive feedback by performing the task but only received external feedback [49]. Hereby subjects had to estimate and indicate the end of a short time interval by pressing a button at the right time. Shortly after indication, they got a visual, auditory, or somatosensory stimulus showing whether their estimation had been correct. In the case that an incorrect estimation was indicated, an ERP in form of a negative deflection between 230 and 330 ms following the feedback with a duration of about 260 ms was observed.

In a study by Schalk et al. subjects had to control a cursor vertically to a given word on a screen showing two words [50]. The cursor control was established by means of an Brain-Computer Interface (BCI) using a motor-imagery paradigm. When the cursor reached a word, it was displayed in large size. In cases where the wrong word was selected, an ERN was observed. The ERN appeared centered above the fronto-central region of the brain and appeared approximately 180 ms after the word selection. By this study it was shown that the observed ERN actually depended on the subject's interpretation of the feedback as the selected word was not explicitly displayed as being wrong.

Parra et al. were able to detect ERN in single trials of choice reaction tasks. By considering a signal window of 100 ms directly after feedback was given they achieved a decoding accuracy up to 79% for discriminating between erroneous and correct trials [51]. By means of an ocular artifact removal technique and an increased window size of 200 ms they reached accuracies up to 91% for similar decoding [52]. As features subsampled signals of all electrodes in the time domain were chosen and for classification an Linear Discriminant Analysis (LDA) classifier (see section 2.3.1) was used. In the first 100 ms following negative feedback a fronto-central negativity was observed. Additionally, in case of correct trials, a more prolonged bilateral positivity appeared within the first 200 ms. The detected ERNs were subsequently used for correcting erroneous responses in order to minimize the overall human-machine error rate.

The question whether ERNs appear also if an error is not caused by the subject but clearly by the interface during recognition was investigated by Ferrez et al. [53, 54]. In an experiment subjects had to control an arrow on a screen by pressing buttons indicating the desired direction of movement. For the subjects it was straightforward to press the correct button, however the control was randomized up to a certain degree so that the interface was clearly responsible for errors which occurred with a probability of 20%. Even in this setup, ERNs followed negative trials and were able to be detected with an average recognition rate of correct and erroneous single trials of 82.4% and 79.9% respectively. They also discovered that ERNs are almost constant over time per subject so that trained classifiers reached almost the same performance in sessions taking place after a three months period. Moreover, they investigated different error rates of the interface. Error rates of 20% and 50% led to the same ERN characteristics but in the latter case the amplitudes were smaller. In this study the ERN was characterized by a positive peak appearing 200 ms after feedback, followed by a negative and positive peak at 250 and 320 ms after feedback, respectively. Finally, a second broader peak showed up about 450 ms after feedback. The detected ERNs are depicted in Figure 3.1. The signals were filtered by a $1-10$ Hz bandpass filter and the common average reference (CAR) (see section 2.2.3) for spatial filtering was applied. The features consisted of subsampled signals in the time domain from 150 ms to 650 ms after feedback of the channels FCz and Cz lying in the fronto-central region. No artifact rejection or reduction algorithm was applied.

In an experiment by Chavarriaga and Millán, subjects observed an agent without giving commands at all [55]. The task hereby was to assess the quality of the agent. Even in this case ERPs could be observed in the case that the agent behaved erroneously. The ERP consisted of a small positive peak approximately 200 ms after feedback and was followed by a negative deflection at around 260 ms and a second positive peak at around 330 ms. For preprocessing also a CAR and a $1-10$ Hz bandpass filter was applied. Single trials were able to be discriminated with mean classification accuracies of 75.81% and 63.21% for correct and error trials, respectively if the error

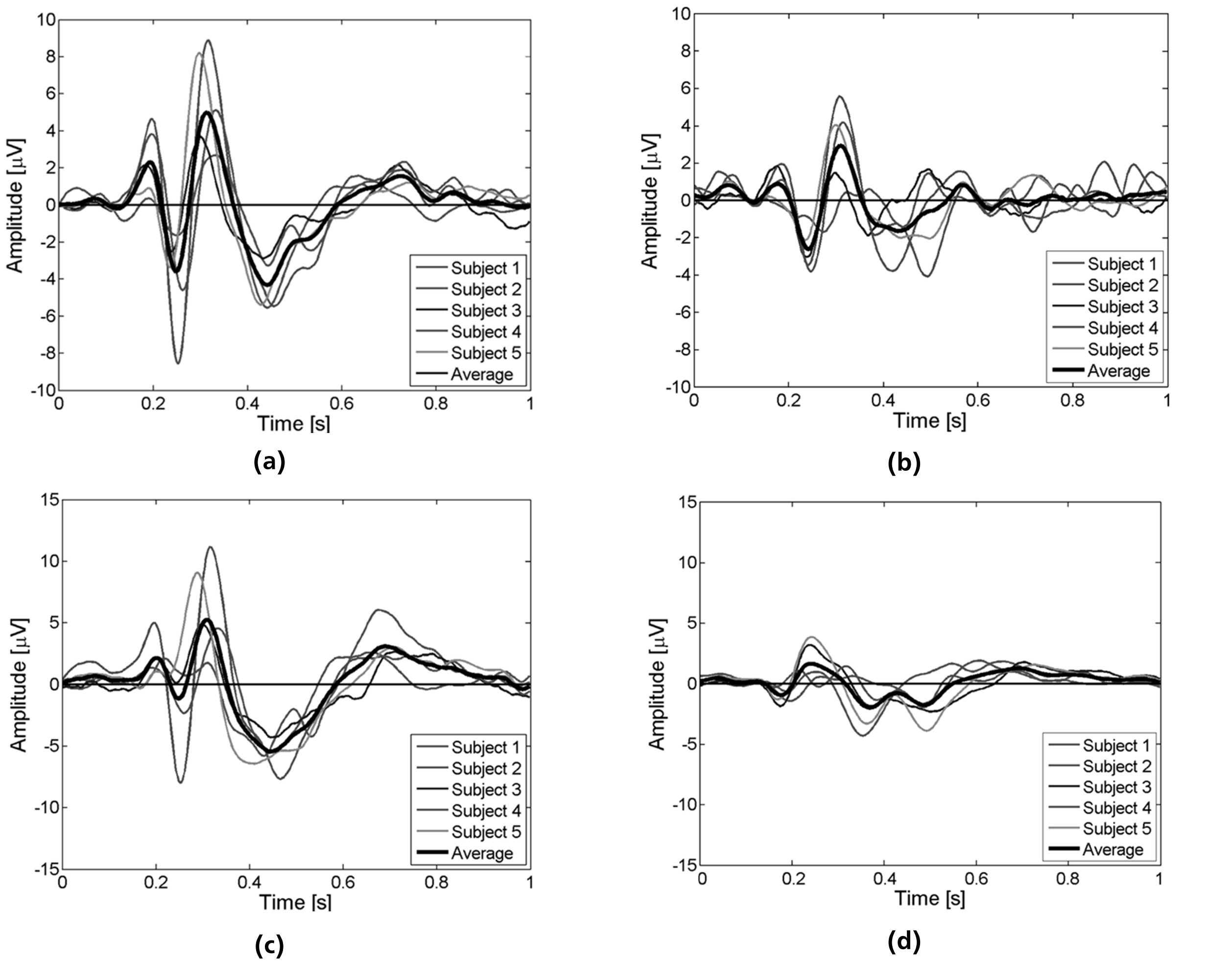



**Figure 3.1:** ERNs occurring in a control task with a control interface which is partly randomized (by Ferrez et al.) [54]. Averages of single trials at channel FCz after feedback was perceived for five subjects and the average of them. **(a)** Difference between signals of correct and erroneous trials in case of errors occurring with a probability of 20%. **(b)** Difference between signals of correct and erroneous trials in case of errors occurring with a probability of 50%. **(c)** Signals of erroneous trials in case of an error probability of 20%. **(d)** Signals of correct trials in case of an error probability of 20%.

probability of the agent was 20%. In the case when the error probability was set up to 50%, the respective accuracies were 64.42% and 59.36%. The decoding of feedback type was finally successfully used in order to learn the optimal policy for the agent within 50 trials.
Online training of a decoder by means of ERNs detection was realized by Buttfield et al. in 2006 [7]. They showed that after decoding the ERNs, stochastic gradient decent approaches (i. e. basic stochastic gradient decent and stochastic meta descent) can be successfully used to train an LDA classifier which even outperformed offline learning methods.

The previously introduced error decoding approaches were all kept rather simple where no other continuous visual stimuli were presented. In all cases the time when feedback was given was predictable for the subject. So was feedback the direct response of an action of the subject (Ferrez et al.) or the subject was able to see how far he is away from getting feedback (Miltner et al., Schalk et al., Parra et al., Chavarriaga and Millán). Also, if systems were controlled by means of a BCIs, a fixed motor-imagery paradigm was used (Miltner et al.).
In this thesis, feedback of a video game presenting continuous visual information is tried to be decoded. In an experiment, the subject has to keep focus on several aspects of the game and the time when feedback will be given is not predictable for him. The paradigm, which he is supposed to apply, can freely be chosen at the beginning of each session. This thesis herewith investigates ERNs for the case of different paradigms which may not be motor-imagery based.

# 4 Error-related Decoding in a Video Game

In order to use an adaptive Brain-Computer Interface (BCI), brain signals need to be labeled, i. e. the correct commands for control have to be known. Frequently in daily tasks, this information is not directly available for the BCI. In this case labels can be inferred during the use of a BCI system by regarding the user's brain signals and decoding whether the command was perceived as wrong. Related work herefore provides evidence for detecting specific error-related negativitys (ERNs) when errors are perceived. To investigate error decoding in daily tasks during the use of a BCI, in this thesis an experiment is conducted in which individuals played a video game.
In the beginning of this chapter in section 4.1, it is shown how feedback detection can be used for an adaptive BCI. Afterwards the details about the game (section 4.2) and the experimental setup (section 4.3) will be presented. The outcomes will be analyzed in section 4.4 and the results will be presented in section 4.5. Finally, in section , variations of the analysis and their results are considered.

## 4.1 Error-related Decoding for an Adaptive BCI

The information whether a subject agrees with a decoded command of the BCI can be used to obtain a prior over commands. The prior can subsequently be used in the decoding process together with the current brain signals for generating new commands. While the classifier of a traditional BCI system may become biased over time to one command [5, 6], the prior leads to decreased probabilities for commands which were mistakenly sent. Other possible commands, on the other hand, obtain a higher probability and will therefore be slightly preferred.

The following modified adaptive pipeline (depicted in Figure 4.1) takes the prior into account to decode commands:

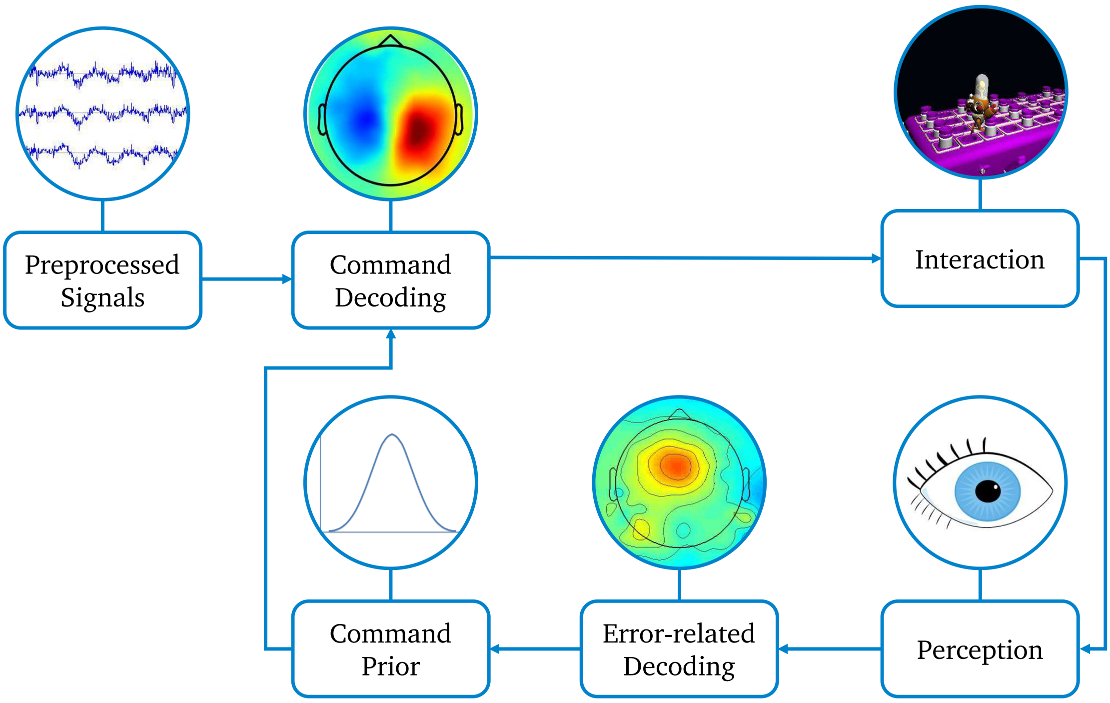



**Figure 4.1: Adaptive BCI using error-related decoding.** This BCI decodes errors perceived by the subject of previously decoded commands. This information is used to obtain a prior which is incorporated in the command decoding phase for balancing the commands.

- **Command Decoding:** The decoder creates a command based on the command prior and current preprocessed brain signals of the user. In the beginning the prior is usually assumed to be uniform.

- **Interaction:** The command is executed and leads to changes in the environment.
- **Perception:** The changes in the environment are observed by the user.
- **Error-related Decoding:** The information whether the user agrees with the executed command is decoded based on patterns in the user's brain signals.
- **Command Prior:** The information whether the executed command was correct is used to learn a prior over commands. Commands which were correctly sent, get a higher prior probability whereas for commands wrongly sent the prior probability is decreased.

For error-related decoding, the user's brain signals after observing the interaction of the BCI are regarded. In case of executed commands which are erroneous, the occurrence of an ERN as presented in related work is expected. Such an event-related potential (ERP) is tried to be decoded within this thesis. The decoding process for the considered game is presented in section 4.4.

## 4.2 Brain-Runners Game

In the *Brain-Runners* game, a player has to control an avatar by sending the appropriate commands at the right time by using specific toughs. The game was created for the Cybathlon to provide a way to test BCI methods for daily tasks. The Cybathlon is a competition for individuals with physical disabilities participate using assistive technologies such as BCIs. It takes place the first time in 2016 in Switzerland. The Brain-Runners game "simulates the control of assistive devices that are currently being developed for future use". More information about the Cybathlon can be found at the corresponding website[1].

During the gameplay, the avatar is running by itself through a track and the goal is to reach the end of the track in shortest time. The game is depicted in Figure 4.2. The subject can speed up the avatar by sending the appropriate command while the avatar is walking on dedicated areas indicated by colored pads (see Figure 4.3). Incorrect commands that are sent will lead to deceleration (see Figure 4.4). There exist three types of pads:

- **Rotate** (cyan): The action *Rotate* has to be sent in order to spin the avatar for acceleration.
- **Jump** (magenta): The action *Jump* has to be sent to make the avatar jumping for acceleration.
- **No-Command** (grey): Every command being sent leads to deceleration.

When a command is sent, a lightning colored according to the command type appears above the avatar and the avatar changes its running behavior (rotating, jumping). Thus, the subject receives direct feedback about the command that was sent. When the avatar reaches a new pad its running behavior is reset to simple running at basic speed.
For the whole study the same track was used. It consisted of eight *Rotate*, six *Jump*, and three *No-Command* pads. If no commands are sent, the duration of one round is 176 seconds.

## 4.3 Experimental Setup

In order to investigate the ERPs of the feedback in the Brain-Runners game, a study was conducted. Ten individuals (between 22 and 32 years old; two female; nine were healthy, one paraplegic) participated and were seated in front of a screen where they were supposed to play the *Brain-Runners* game. Although they were made to believe that they were controlling the avatar by sending mental commands, all interaction commands were predefined. Prior to the study, the individuals had to perform mental actions (left and right hand movement, feet movement, calculation, spatial navigation, word association, and listening to music) where they got used to the application of mental commands. In this phase, they got no feedback how well they performed in the tasks. All but two individuals had no experiences with electroencephalography (EEG) recording and were therefore new to EEG. Each individual participated twice in the study with the same setup. After the session, the subject was supposed to

[1] http://www.cybathlon.ethz.ch/en/ (09/10/2016)

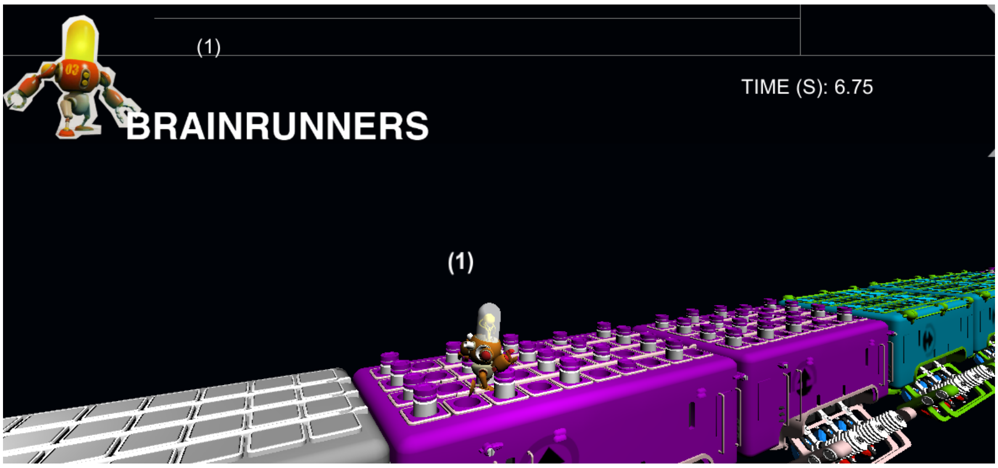



**Figure 4.2: The Brain-Runners game that is played in the experiment and at the Cybathlon.** The avatar runs on a *Jump pad* (magenta) where it can be accelerated if the correct command is imagined. Previously, the avatar has passed the grey *No Command* field.

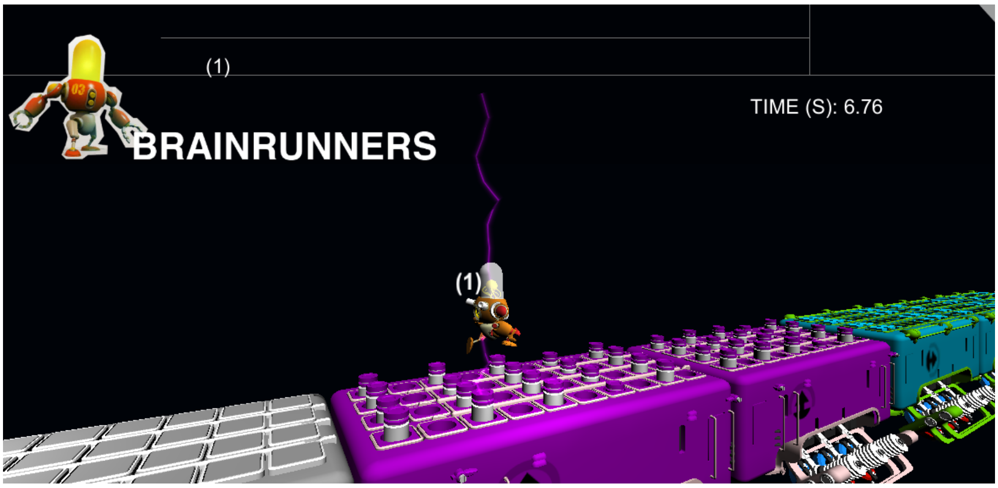



**Figure 4.3: Correct command in the Brain-Runners game.** The avatar runs on a *Jump pad* and the correct command (*Jump*, shown by a magenta flash) was sent. The avatar begins to jump and accelerates.

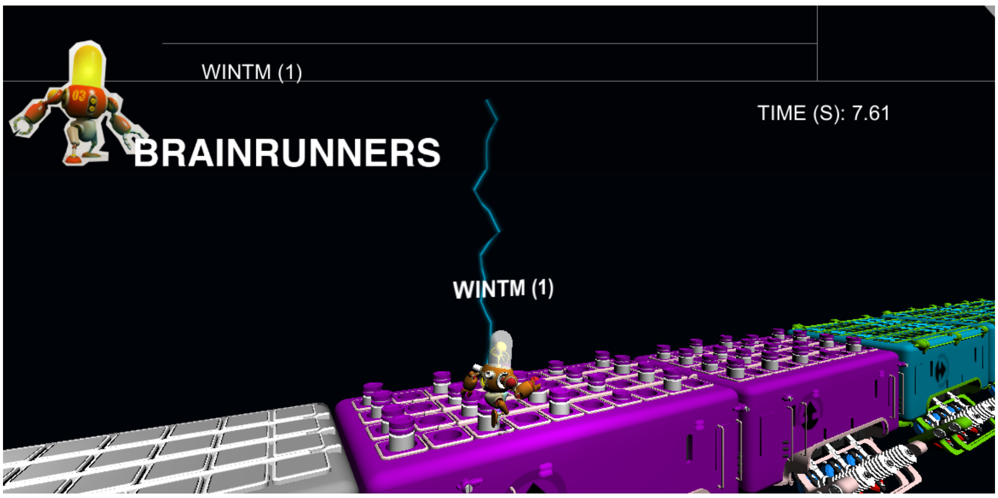



**Figure 4.4: Wrong command in the Brain-Runners game.** The avatar runs on a *Jump pad* and the wrong command (*Rotate*, shown by a blue flash) was sent. The avatar begins to rotate and decelerates.

answer a questionnaire for self-assessment.

Signals were recorded at a sampling rate of 500 Hz using an BrainProducs actiCHamp EEG system with 128 active electrodes. The distribution of electrodes is depicted in Figure **??**. For recording and signal labeling the software *BCI2000* [56] was used.

### 4.3.1 Gameplay

In each session each subject played five times the same track with different predefined commands for each time. In order to control the rate of correct commands, all commands were predefined beforehand to achieve desired success rates. However, the subject was lead to believe that he controlled the avatar himself by using mental commands. This believe was given to achieve a case most closely to the real world of users communicating by means of a BCI. Even if related work by Ferrez et al. [54] showed that ERNs occur even if subjects only observe faulty behavior of the user interface, the application of a mental command might influence the ERNs. To keep the subject motivated, the success rate was increasing during the session (50%, 66%, 66%, 75%, 75% of the signals that were sent were correct and led to acceleration). Commands were sent in an interval of six or twelve seconds.

The subject was supposed to control the avatar using an implicit paradigm. Before playing the game, he selected two of eight mental commands presented prior to the game for controlling the avatar. Each was used to encode one of the actions *Rotate* and *Jump*. The subject had to apply the demanded mental command shortly before the avatar reached a pad until either the correct command was executed or almost a new pad was reached. In case that a wrong command was executed in the game, the subject had to continue performing the mental command. For sending no command, the subject was supposed to relax.

## 4.4 Experiment Analysis

Each trial consisted of the signals from 200 ms prior to the command to 4200 ms after the command. The recorded signals were preprocessed by subtracting the baseline and using the common average reference (CAR) method. Signals exceeding 100 $\mu$V were rejected to remove signals heavily contaminated by artifacts and a Butterworth bandpass filter [12] from 1 to 20 Hz was applied as ERPs are generally found in lower frequencies [48].

To investigate ERPs occurring after a command was executed, the preprocessed signals were averaged per command type. For each time step a topographic map [56] of the signals was created and analyzed to detect deflections indicating an ERP. Herewith suitable electrodes and time periods were regarded which can be used for the detection of the ERP. Also the difference between the averaged signals after correct and wrong commands was examined using topographic maps to assess their discriminablility.

First, the discriminability between the signals of the cases *command sent / baseline* was taken into account to investigate the ability to detect the found ERP. In order to obtain reference signals for the *baseline* case random parts of trials between 3000 ms and 6000 ms after executing a command were extracted. For generating features for classification, time domain signals of the channel FCz in the period from 200 ms to 2200 ms after command execution were regarded. The extraction of trials is depicted in Figure 4.5. The channel and period selection was based on the results of the topographic map analysis in section 4.5. The selected signal parts were divided into 25 ms bins and the mean of each bin was determined to subsample the signal to 40 Hz. The subsampled signals in the time domain were used as features. Using the subsampled signals instead of the original signals as features offers the advantage of a lower feature dimensionality as well as compensation of small measurement inaccuracies. For discrimination, Linear Discriminant Analysis (LDA) and linear Support Vector Machine (SVM) classifiers were trained and evaluated for each subject individually. Additionally, cross validation has been used, a technique to assess how well the results of a classifier will generalize to a new independent data set. In 10-fold cross validation, the data set is divided into 10 folds where in every combinations 9 folds are used for training and the remaining one for evaluation. Afterwards the mean and variance of the results is taken into account. For balancing, the larger class was 60 times randomly subsampled so that classification was realized with an equal number of samples.
To investigate the discriminability between correct and wrong commands in the signals, the same feature generation as in the case for discrimination between *command sent* and *baseline* was used. As before, the classes were subsampled for balancing and 10-fold cross validation LDA and SVM classifiers were trained and evaluated.

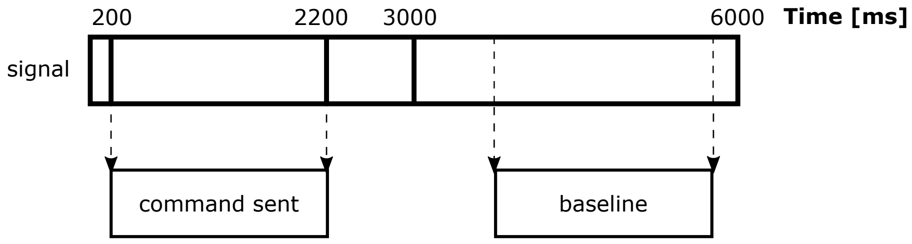



**Figure 4.5: Trial extraction of the recorded signals.** The time signal of the period from 200 ms to 2200 ms after avatar control is extracted as trial of the class "command sent". A random part of the signal of the period from 3000 ms to 6000 ms after avatar control is extracted as baseline trial.

## 4.5 Experimental Results

The averaged preprocessed signals over all subjects and sessions in a topographic map are shown in Figure 4.6 and 4.7. According to related work, ERNs are expected to occur in the fronto-central region of the scalp for the erroneous control case. In the topographic maps one can actually find deflections in this region indicating an ERP after both correct and wrong avatar control. They are best visible at the electrode FCz and diminish towards the edges of the cap. Different deflections can also be spotted at electrodes close to edges of the cap which are increasing towards the edges. These, however, likely represent artifacts as the edges cover rather muscular and ocular movements than brain activity. The deflections at frontal electrodes as Fp1 and Fp2 for example are likely caused by ocular activity as they are increasing towards the front.
The deflections at electrodes FCz are further analyzed by means of topographic maps. The time signals in detail as well as topographic maps for the timepoints of deflections are depicted in Figure 4.9 and 4.10. 500 ms after after avatar control, in both cases a negative deflection occurs, followed by a positive deflection after 800 ms. After 1200 ms a second, smaller negative deflection takes place and after 1870 ms there appears a second, smaller positive deflections.
At each of the four time points, in the topographic plots the deflection specifically appears in the fronto-central region centered at electrode FCz. These specific deflections diminish towards the edges of the cap which makes it unlikely to be caused by ocular and muscular artifacts. This observation strengthens the assertion that the deflections are caused by brain activity and represent a specific ERP. However, the deflections appear much later compared to the ERNs found in related work where the most prolonged ERPs take place until 650 ms after the stimulus [54].

For investigation of the discriminability between correct and wrong commands, the difference between both average signals in a topographic map is investigated. The topographic map of the difference signal is depicted in Figure 4.8. In the fronto-central region, small differences in deflection are visible in the period of 300 ms to 1000 ms after command execution which provide possible patterns for discrimination. The difference signal at electrode FCz and topographic maps at the time steps of the ERP deflections previously found are depicted in Figure 4.11. At time step 500 ms, the differences in the deflections can be observed to be centered in the fronto-central region whereas in the topographic plots at 800 ms, 1200 ms, and 1870 ms no clear differences can be found. Towards the edges of the cap, especially in the frontal region, large signal differences between the correct and erroneous control case can be observed. As the whole signal shows large differences, they are most likely not connected to the correctness of the commands but come from large variations in these signals caused by artifacts.

As the period which shows differences between correct and wrong signals is contained in the period of the previously found ERP, features of the whole interval from 200 to 2200 ms are used as features for decoding. In section 4.6, however, also the smaller interval is considered for decoding.
The classification results per subject for both the cases *command sent* vs. *baseline* and *correct command* vs. *wrong command* are depicted in Table 4.1. It can be seen that trials of the types "command sent" and "baseline" can be discriminated with an average accuracy of 72.7% with a standard deviation of 7.2% in case of LDA classification and 71.1% with a standard deviation of 7.5% in case of an SVM. Trials of "correct command" and "wrong command" can be discriminated with an average accuracy of 58.7% ± 13,1% and 57.7% ± 13.2% by LDA and SVM classification, respectively.

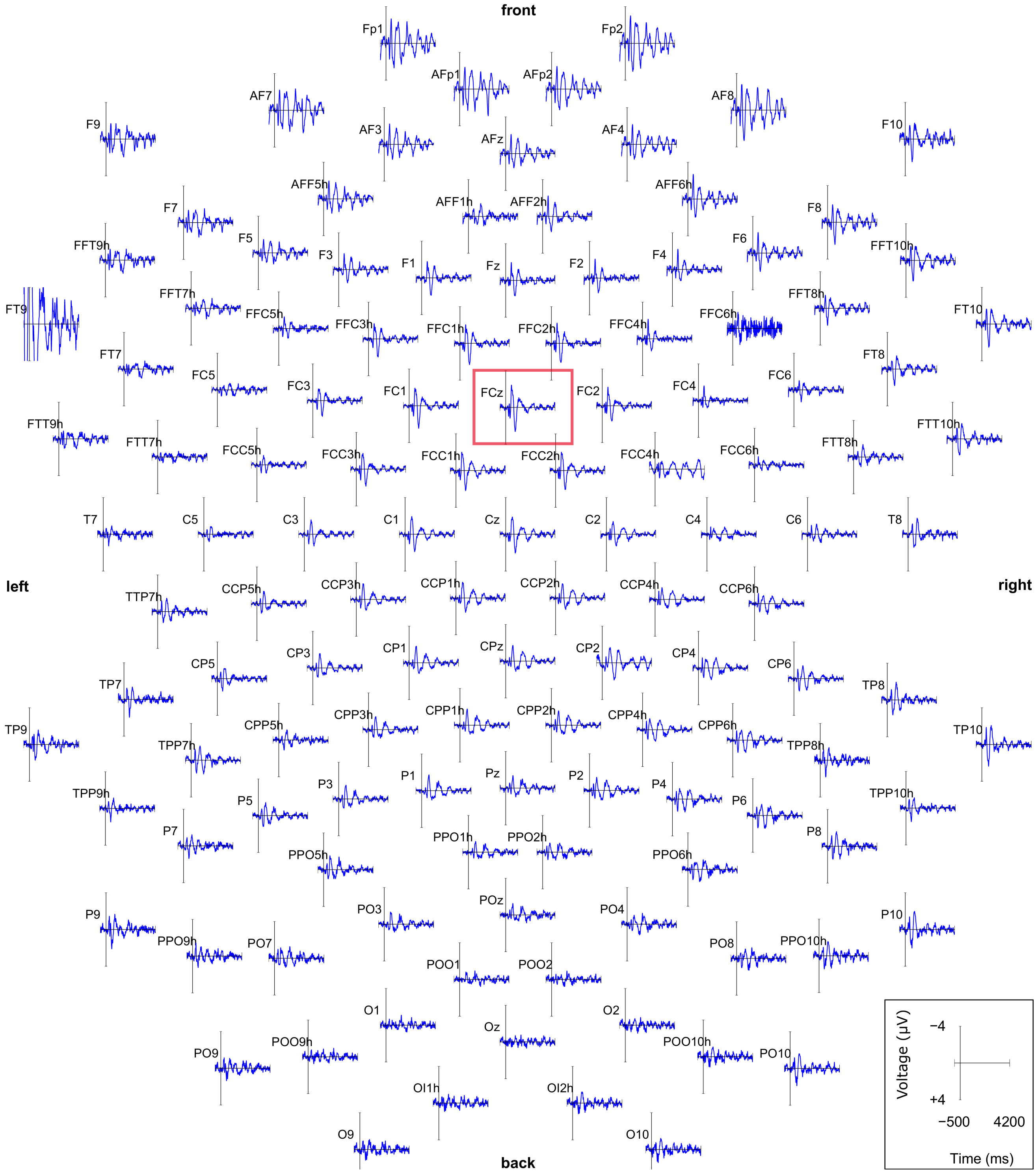



**Figure 4.6: Topographic map of the averaged preprocessed signals over all subjects after erroneous avatar control.** Characteristic ERP deflections appear in the fronto-central region (e. g. at electrode FCz, highlighted by a red box). Signal patterns at electrodes located at the edge of the cap likely represent artifacts (E. g. deflections at electrodes Fp1 and Fp2 at the front are caused by ocular movements.)

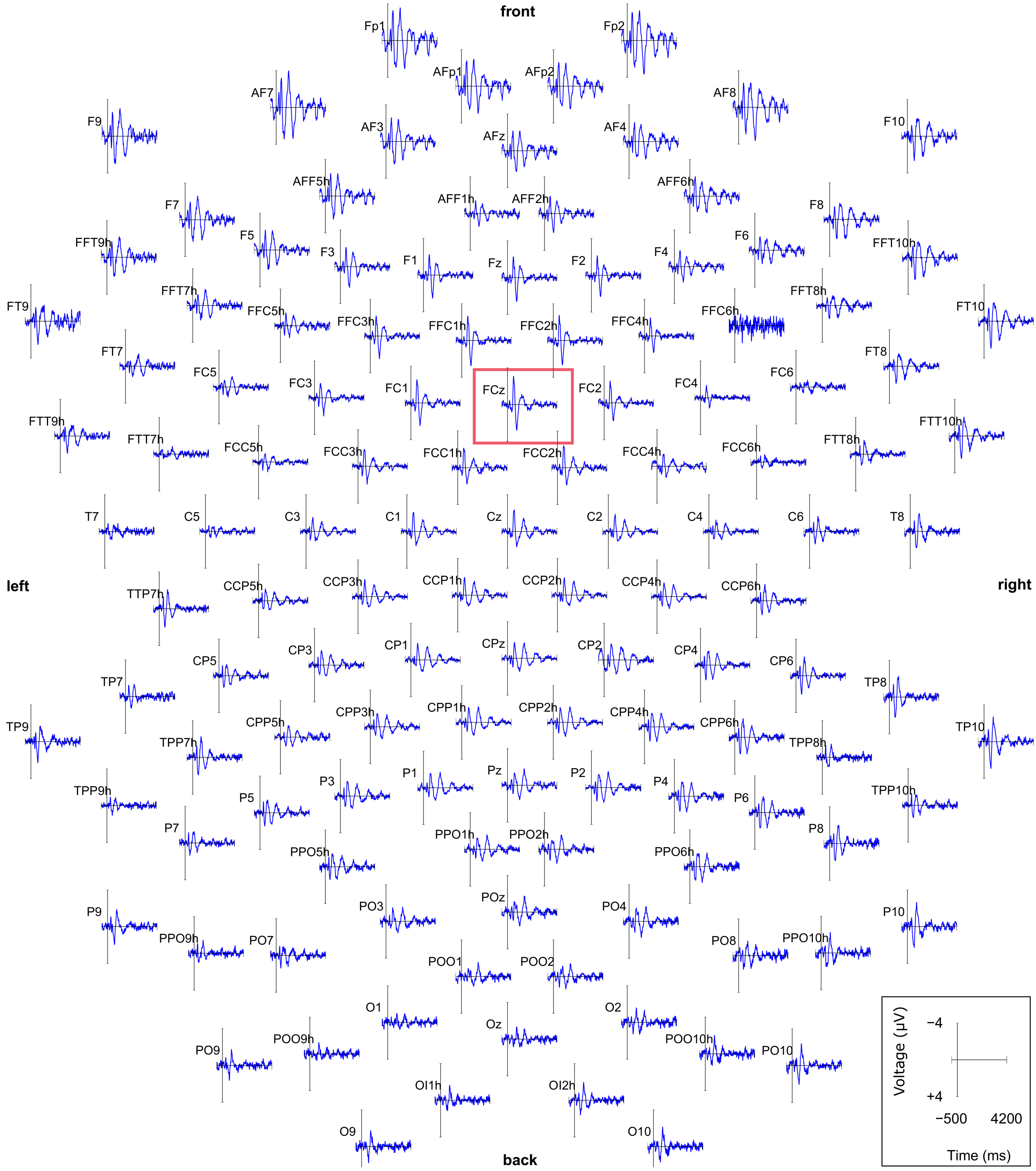



**Figure 4.7: Topographic map of the averaged preprocessed signals over all subjects after correct avatar control.** ERP deflections in the fronto-central region (e. g. at electrode FCz, highlighted by a red box) resemble the ones in the negative feedback case in Figure 4.6.

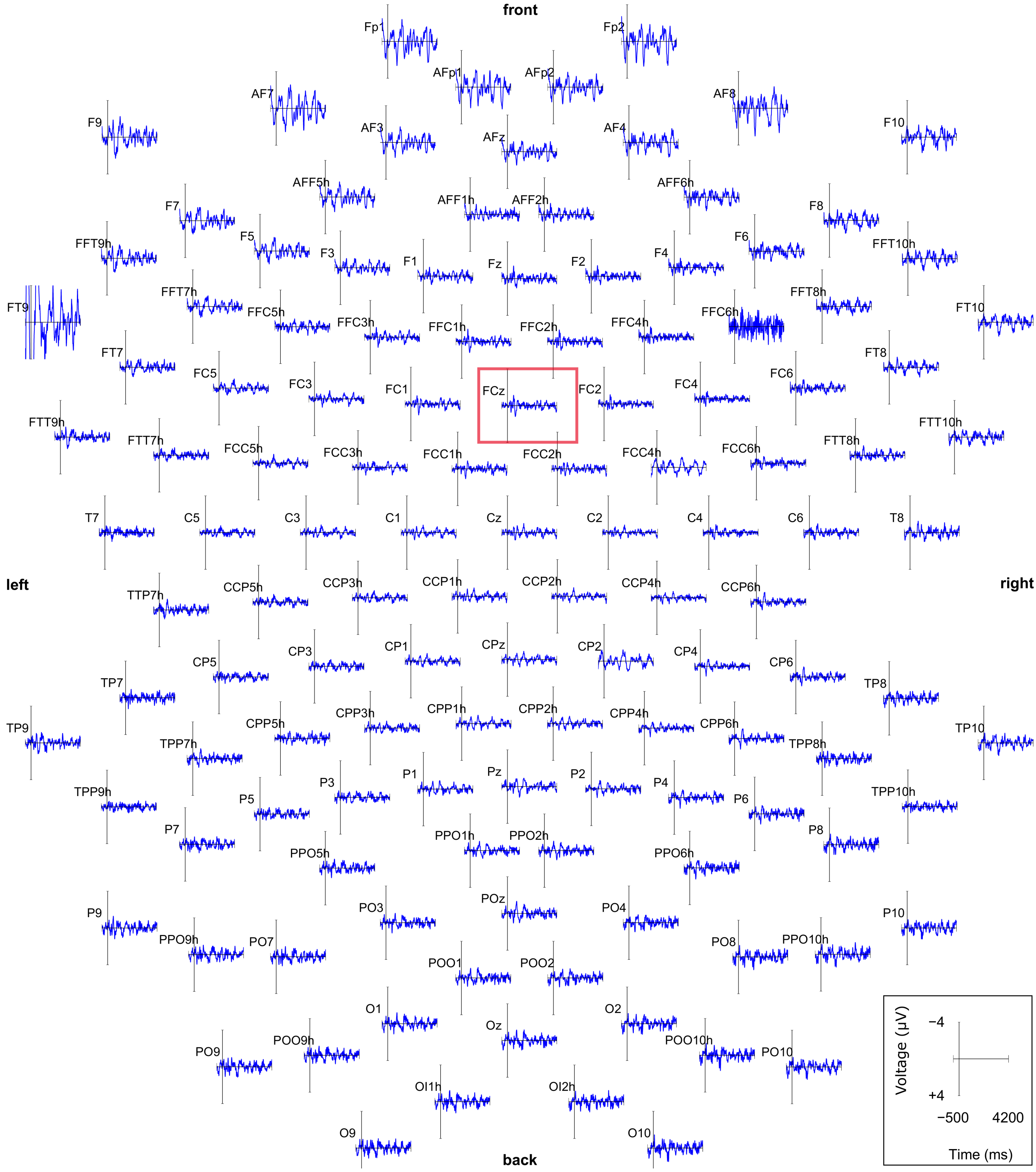



**Figure 4.8: Topographic map of the difference between the averaged preprocessed signals over all subjects of correct and erroneous avatar control.** In the fronto-central region (e. g. at electrode FCz, highlighted by a red box) there occur small deflection differences shortly after a command was sent.

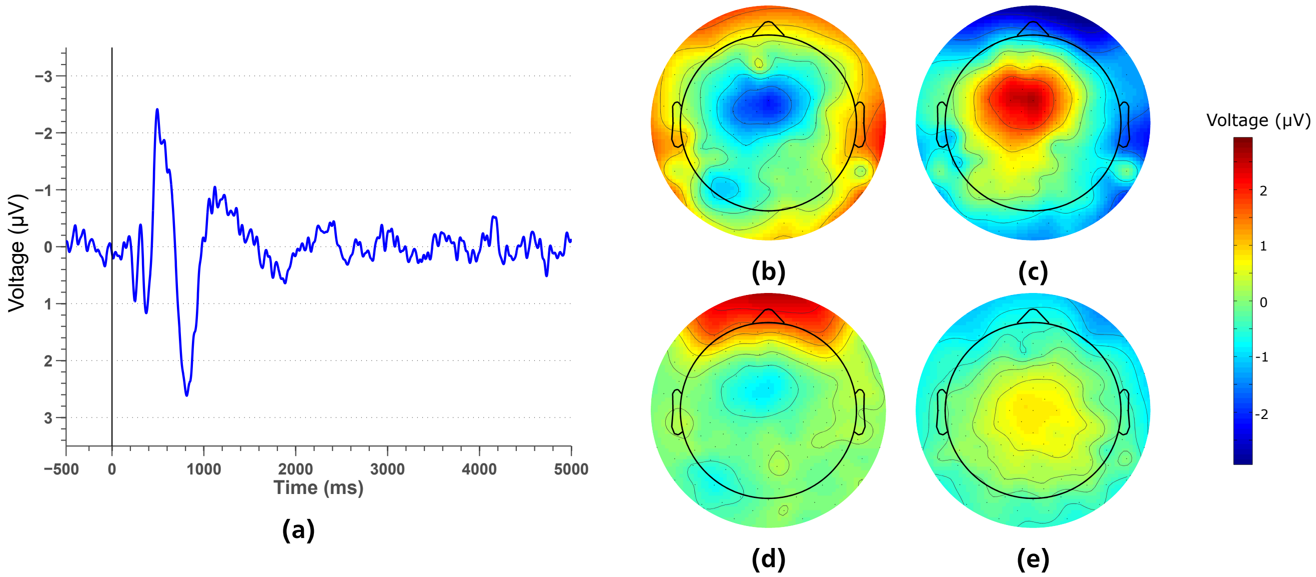



**Figure 4.9: Average signals over all subjects and sessions after erroneous avatar control.** **(a)** Time signals at electrode FCz. Deflections in the interval from 200 ms to 2000 ms after avatar control indicate an ERP. **(b - e)** Topographic plot of the time signals at specific time steps after avatar control: **(b)** 500 ms, **(c)** 800 ms, **(d)** 1200 ms, **(e)** 1870 ms. The deflections appear centered in the fronto-central region.

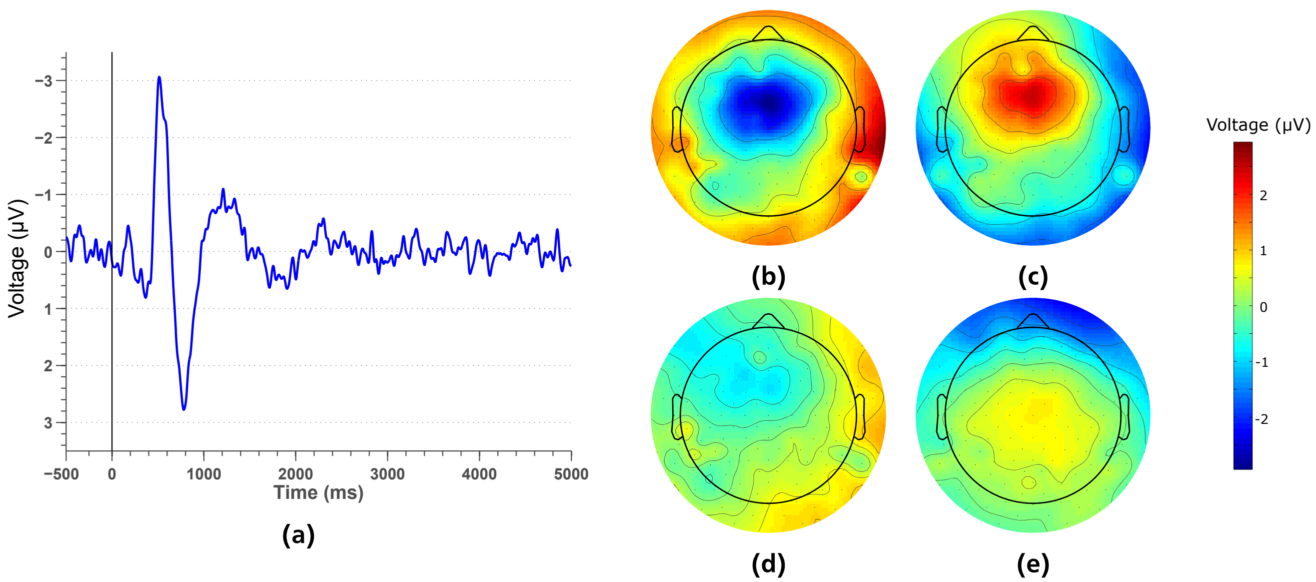



**Figure 4.10: Average signals over all subjects and sessions after correct avatar control.** **(a)** Time signals at electrode FCz. Like in the erroneous control case (4.9), deflections indicate an ERP. **(b - e)** Topographic plot of the time signals at specific time steps after avatar control: **(b)** 500 ms, **(c)** 800 ms, **(d)** 1200 ms, **(e)** 1870 ms. The deflections appear centered in the fronto-central region.

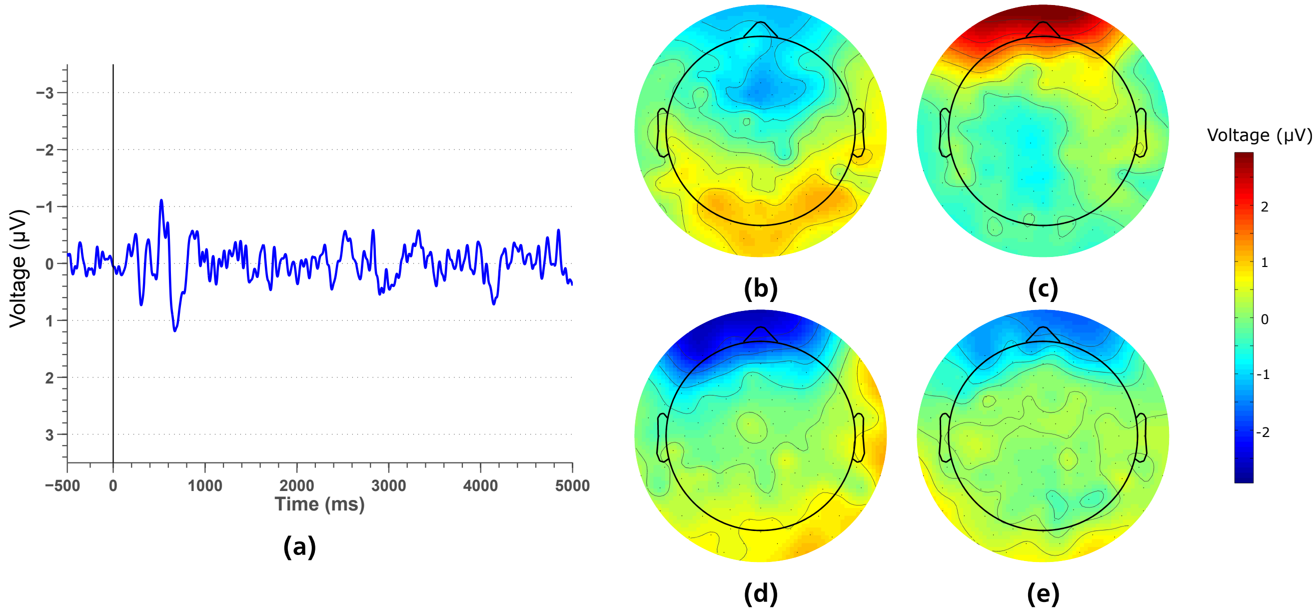



**Figure 4.11: Average signals over all subjects and sessions of the difference between correct and erroneous avatar control (a)** Time signals at electrode FCz. Deflections between 300 ms and 1000 ms after avatar control show small differences between correct and erroneous avatar control. **(b - e)** Topographic plot of the time signals at specific time steps after avatar control: **(b)** 500 ms, **(c)** 800 ms, **(d)** 1200 ms, **(e)** 1870 ms. Only in the plot at 500 ms, clear deflection differences in the fronto-central region can be found which show a difference of the ERPs between correct and erroneous avatar control.

| | command vs. baseline | | correct vs. wrong command | |
|---|---|---|---|---|
| | LDA | SVM | LDA | SVM |
| S1 | 92.1 ± 5.0 | 88.9 ± 5.9 | 58.6 ± 14.0 | 57.6 ± 14.5 |
| S2 | 62.8 ± 8.7 | 60.2 ± 8.9 | 66.1 ± 14.3 | 66.2 ± 13.4 |
| S3 | 71.6 ± 5.4 | 70.5 ± 5.6 | 51.5 ± 9.1 | 52.4 ± 9.3 |
| S4 | 62.4 ± 9.2 | 61.4 ± 9.8 | 49.3 ± 14.4 | 52.2 ± 14.2 |
| S5 | 74.3 ± 6.5 | 72.9 ± 6.8 | 59.8 ± 11.3 | 57.4 ± 11.9 |
| S6 | 77.3 ± 7.0 | 75.2 ± 7.2 | 55.5 ± 13.9 | 54.4 ± 12.7 |
| S7 | 71.7 ± 7.4 | 69.3 ± 7.4 | 61.8 ± 12.6 | 59.4 ± 13.9 |
| S8 | 76.1 ± 7.4 | 73.8 ± 7.8 | 61.6 ± 14.3 | 58.6 ± 14.4 |
| S9 | 61.3 ± 9.0 | 61.4 ± 8.9 | 67.7 ± 14.6 | 66.4 ± 14.8 |
| S10 | 77.8 ± 6.7 | 76.8 ± 6.9 | 55.1 ± 11.9 | 52.1 ± 12.4 |
| Mean | 72.7 ± 7.2 | 71.1 ± 7.5 | 58.7 ± 13.1 | 57.7 ± 13.2 |

**Table 4.1: Classification results.** For each subject (S1 - S10) LDA and linear SVM classifiers were trained and evaluated in a 10-fold cross validation setup. "command vs. baseline" denotes the discrimination between signals directly after avatar control and those where no command was sent. The "correct vs. wrong command" case differentiates between signals after correct and erroneous avatar control. The numbers present classification accuracies in percent with their corresponding standard deviation.

## 4.6 Experiment Analysis Variations

Different analysis of the experiment can lead to improved results or provide additional insights. In the following section, variations of the analysis presented in section 4.4 are investigated and their results are presented.

### 4.6.1 Spatial Features

Decoding may be improved if not only the signal of one electrode is considered but signals of an pair of electrodes are regarded for considering spatial changes of the ERP. Additionally to electrode FCz, the subsampled signals of the electrode Cz are added to the features. Cz is an electrode in the central region where the deflection at 1870 ms after avatar control is centered on. The deflections between 200 ms to 1500 ms are smaller but also clearly visible (see Figure 4.6, 4.7, and 4.8).
The analysis of section 4.4 is repeated with the adapted feature sets. The decoding accuracies are depicted in table 4.4.

| | command vs. baseline | | correct vs. wrong command | |
|---|---|---|---|---|
| | LDA | SVM | LDA | SVM |
| Mean accuracy | 76.6 ± 4.9 | 76.2 ± 4.9 | 66.3 ± 8.8 | 66.1 ± 8.9 |

**Table 4.2: Classification results in case of spatial features.** The decoding analysis was repeated by adding signals of the electrode Cz to the feature sets. The average decoding accuracies are increased compared to the analysis in table 4.1.

Adding the signals of electrode Cz to the feature sets, increases the average accuracy in "command vs. baseline" discrimination by 3.9% (LDA) and 5.1% (SVM). Discrimination between correct and erroneous trials was improved in average by 7.6% (LDA) and 8.5% (SVM). This type of features also leads to a smaller standart deviation (-2.5% for "command vs. baseline" and -4.3% for correct vs. wrong command").
By considering two electrodes, the spatial development of the ERP can be taken into account, which leads to a more accurate decoding.

### 4.6.2 Laplace Filter

The application of the laplace filter (see section 2.2.4) increases the spatial resolution and thus might improve the decoding of the ERPs. The previous analysis of decoding with spatial features (section 4.6.1) is extended by the application of the laplace filter. The filter is applied on both electrodes FCz (by subtracting signals of electrodes Fz, FC2, Cz, and FC1) and Cz (by subtracting signals of electrodes FCz, C2, CPz, and C1) after applying the bandpass filter in the signal preprocessing step. The results of the analysis are depicted in table 4.3.

| | command vs. baseline | | correct vs. wrong command | |
|---|---|---|---|---|
| | LDA | SVM | LDA | SVM |
| Mean accuracy | 64.7 ± 5.5 | 65.1 ± 5.5 | 55.5 ± 9.5 | 55.5 ± 9.2 |

**Table 4.3: Classification results in case of applying a laplace filter and using spatial features.** To increase spatial resolution at the electrodes, a laplace filter was added as preprocessing step. The application of the filter leads to decreased averaged decoding accuracies.

The application of the laplace filter decreases average decoding accuracies by 11.9% (LDA) and 11.1% (SVM) for discrimination between "command sent" and "baseline". For the "correct vs. wrong command" case, the accuracy is decreased in average by 10.8% (LDA) and 10.6% (SVM). Most likely, as the ERP appeared in the whole fronto-central region, the electrodes subtracted within the filter captured almost the same ERP deflections as the electrodes the filter was applied on. The application of the filter hence led to the deflections being cancelled at the signals taken into account for decoding.

### 4.6.3 Reduced Time Interval

In the averaged signals of the difference between correct and erroneous avatar control, differences in deflections are only visible in the interval of 300 ms to 1000 ms. To investigate the importance of the deflections taking place after 1 second after avatar control for decoding, a reduced time interval is considered. Hereby additionally the feature dimensionality is decreased which facilitates finding patterns for decoding. The analysis using spatial features (section 4.6.1) is adapted to take into account only signals in the period of 300 ms to 1000 ms. The average results of decoding with this method are depicted in table 4.4.

| | command vs. baseline | | correct vs. wrong command | |
|---|---|---|---|---|
| | LDA | SVM | LDA | SVM |
| Mean accuracy | 71.4 ± 5.2 | 71.3 ± 5.2 | 61.4 ± 9.0 | 62.0 ± 9.1 |

**Table 4.4: Classification results for spatial features with a reduced time interval.** The shorter time interval of 300 ms to 1000 ms was regarded in the decoding. The decoding accuracies are decreased by approximately 5% compared to the analysis of the signal interval of 200 ms to 2200 ms.

On the basis of the results, it can be seen that the reduced time interval of the features leads to decreased performance in both the "command sent vs. baseline" (-5.2% for LDA and -4.9% for SVM) and "correct vs. wrong command" cases (-4.9% for LDA and -4.1% for SVM). This drop in performance shows the importance of the smaller ERP components appearing only one second after a sent command for discrimination. Even if in the average signal no large deflections in this period are visible, the rather small deflections seem to contain important information for decoding.

### 4.6.4 Frequency Features

Perceived avatar control may cause event-related desyncronizations (ERDs) or event-related syncronizations (ERSs) (see section 2.4.2) which have not been taken into account so far. To investigate the occurence of frequency specific changes, the power spectrum of the signals is analyzed. Herefore, *Fast Fourier Transform* [14] is applied on each signal in the time period of 200 - 2200 ms and the resulting power spectra are averaged per command type. The average spectrum for the baseline and correct and erroneous control case is depicted in Figure 4.12. As small differences between wrong and correct commands are mainly visible at 8 Hz and between 15 and 22 Hz, the power spectrum between 5 and 22 Hz was divided into 3 Hz bins and their power taken as features for classification. The results for decoding whether a command was sent and the discrimination between correct and negative trials is depicted in table 4.5.

| | command vs. baseline | | correct vs. wrong command | |
|---|---|---|---|---|
| | LDA | SVM | LDA | SVM |
| Mean accuracy | 52.6 ± 7.9 | 52.4 ± 7.7 | 55.4 ± 12.3 | 55.6 ± 12.3 |

**Table 4.5: Classification results in case of features consisting of power in the frequency spectrum.** For both the decoding whether a control was sent and whether the control was correct, decoding accuracies sparingly exceed chance level (50%). Features based on frequency components in this form therefore seem not to be suitable for decoding information about the command.

As the resulting average accuracies are only sparingly above chance level (50%), the frequency-based features chosen here seem not to be suitable for decoding information about commands. For the discrimination between

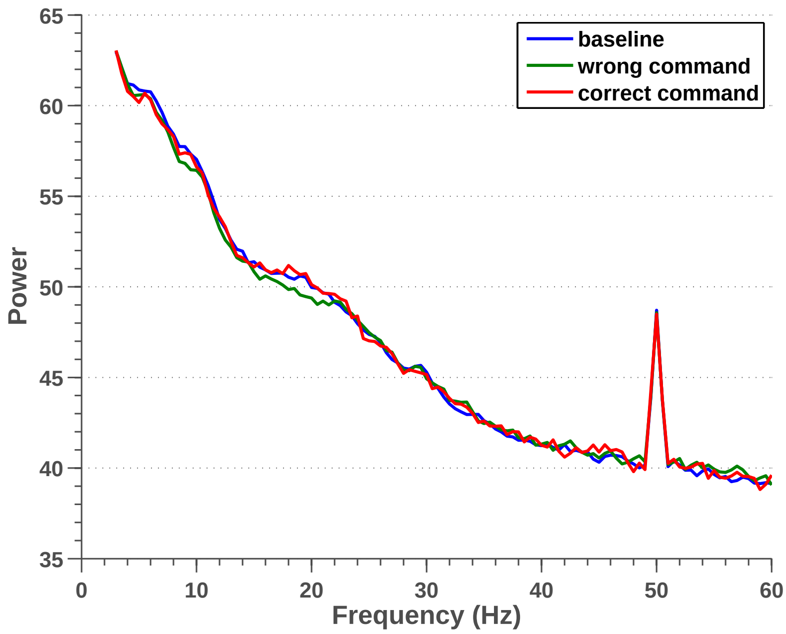



**Figure 4.12: Mean frequency spectra for different trials.** The spectra of the cases in which no command was sent ('baseline') and a correct command was sent, are very similar. The case where a wrong command was sent and an error was perceived by the user is slightly different from the other cases at 8 Hz and between 15 and 22 Hz.

"command sent" and "baseline" the resulting accuracy of 2.5% above chance level is not surprising as average frequency of "wrong command" and "baseline" are very similar. In case of discrimination between correct and erroneous commands (5.5% above chance level), there are small differences visible in the power spectra in Figure 4.5. An explanation for the lack of discriminability would be that differences in the power spectrum are too small in single trials to be reliably used for discrimination.

# 5 Conclusion and Future Work

Brain-Computer Interfaces (BCIs) offer the possibility of controlling external devices by specific thoughts. Whereas decoding in a traditional BCI system is not optimal for a longer time, adaptive approaches are able to adapt online to changed brain signals. In this thesis an adaptive BCI system was approached which compensates for the bias towards specific commands introduced by changed brain signals. For realization, signal labels need to be known in an online setup which can be gathered from the knowledge, whether an command executed by the BCI was correct. This information could be decoded from the user's brain signals after the presentation of visual feedback. Within this thesis a new scenario was addressed which can be transferred to daily tasks of future BCI systems. An experiment was conducted in which individuals had to control a video game and received feedback after performing tasks. During the gameplay, the subjects were supposed to choose two mental tasks for game control (movement imagination of left and right hand, feet movement, mental rotation, association, auditory imagery, spatial navigation, and calculation). The concrete goal was to analyze how the correctness of commands sent in the game are able to be decoded from brain signals for using this information for an adaptive BCI system.

Signals were recorded by means of an electroencephalography (EEG) system, preprocessed, and analyzed. The occurrence of an event-related potential (ERP) in the fronto-central region was detected which takes place 200 ms to 2200 ms after game control in both the correct and erroneous control case. The found ERP, however, is timed differently from the error-related negativity (ERN) found in related work and shows no large differences between correct and erroneous trials. Linear Discriminant Analysis (LDA) and linear Support Vector Machine (SVM) classifiers were trained in a 10-fold cross validation setup to investigate the detection of the ERP in single trials. By using subsampled signals in the time domain of one electrode as features, an average accuracy of 72.7% and 71.1% was reached for detection whether a command was sent by using LDA and SVM classification, respectively. Correct and wrong commands could be discriminated with an accuracy of 58.7% (LDA) and 57.7% (SVM). In both decoding cases, the LDA classifier performed slightly better than the SVM.

Furthermore, variations of the analysis were considered. Incorporating spatial features by using the signals of an additional electrode increased decoding accuracies by 3.9% (LDA) and 5.1% (SVM) for detection whether a command was sent. Discrimination between correct and erroneous was improved in average by 7.6% (LDA) and 8.5% (SVM).

Applying a laplace filter as additional preprocessing step, however, led to a lower performance. By using two electrodes in this setup, the decoding accuracies were decreased by 11.5% ("command sent" vs. "baseline") and 10.7% ("correct" vs. "erroneous commands"). Taking into account a reduced time period led also to decreased performance. With this analysis, the importance of the whole longer time period was shown, even if the difference between correct and erroneous trials exhibits no clear difference pattern in the additional time period in average. Also, the use of frequency components as features did not seem to be a promising approach. By using the power of frequency bands as features, decoding accuracies only slightly above chance level were reached. As a result, the approach of using spatial filters of two electrodes seems to be best suitable for decoding whether a command was correct.

In the future, this information may be used for realizing an adaptive BCI. Herefore a prior based on the decoding results still has to be designed and included into the decoding process of a BCI pipeline. The prior balances the decoding of commands and thus prevents a bias towards specific commands. To increase decoding performance further, approaches to transfer decoding across subjects [57] might be included. Furthermore, variations of the experiment design may be investigated. As in the experiment conducted within this thesis the subject was not able to predict the time when a command was sent in the game, the detected ERP might be partly caused by the experience of surprise by the subjects. In a modified experiment, one could announce the execution of a command for examining this influence.

# Bibliography


[1] R. P. Rao, *Brain-computer interfacing: an introduction*. Cambridge University Press, 2013.

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

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

[4] J. R. Wolpaw and D. J. McFarland, "Multichannel eeg-based brain-computer communication," *Electroencephalography and clinical Neurophysiology*, vol. 90, no. 6, pp. 444–449, 1994.

[5] N. Kawabata, "A nonstationary analysis of the electroencephalogram," *IEEE Transactions on Biomedical Engineering*, vol. 6, no. BME-20, pp. 444–452, 1973.

[6] G. Ferber, "Treatment of some nonstationarities in the eeg," *Neuropsychobiology*, vol. 17, no. 1-2, pp. 100–104, 1987.

[7] A. Buttfield, P. W. Ferrez, and J. R. Millán, "Towards a robust bci: error potentials and online learning," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 164–168, 2006.

[8] M. Serruya and J. Donoghue, "Design principles of a neuromotor prosthetic device," *Neuroprosthetics: Theory and Practice*, pp. 1158–1196, 2003.

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

[10] M. van Vliet, A. Robben, N. Chumerin, N. V. Manyakov, A. Combaz, and M. M. Van Hulle, "Designing a brain-computer interface controlled video-game using consumer grade eeg hardware," pp. 1–6, 2012.

[11] S. Butterworth, "On the theory of filter amplifiers," *Wireless Engineer*, vol. 7, no. 6, pp. 536–541, 1930.

[12] T. W. Parks and C. S. Burrus, *Digital filter design*. Wiley-Interscience, 1987.

[13] K. A. Ludwig, R. M. Miriani, N. B. Langhals, M. D. Joseph, D. J. Anderson, and D. R. Kipke, "Using a common average reference to improve cortical neuron recordings from microelectrode arrays," *Journal of neurophysiology*, vol. 101, no. 3, pp. 1679–1689, 2009.

[14] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex fourier series," *Mathematics of computation*, vol. 19, no. 90, pp. 297–301, 1965.

[15] D. Garrett, D. A. Peterson, C. W. Anderson, and M. H. Thaut, "Comparison of linear, nonlinear, and feature selection methods for eeg signal classification," *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 11, no. 2, pp. 141–144, 2003.

[16] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2012.

[17] D. Barber, *Bayesian reasoning and machine learning*. Cambridge University Press, 2012.

[18] V. N. Vapnik, "The nature of statistical learning theory," 1995.

[19] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[20] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the fifth annual workshop on Computational learning theory*, pp. 144–152, ACM, 1992.

[21] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Data mining and knowledge discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[22] E. D. Adrian and K. Yamagiwa, "The origin of the berger rhythm.," *Brain: A Journal of Neurology*, 1935.

[23] C.-l. Li, H. McLennan, and H. Jasper, "Brain waves and unit discharge in cerebral cortex.," *Science*, 1952.

[24] S. J. Luck, "Ten simple rules for designing erp experiments," *Event-related potentials: A methods handbook*, vol. 262083337, 2005.

[25] P. L. Nunez and R. Srinivasan, *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, USA, 2006.

[26] G. F. Woodman, "A brief introduction to the use of event-related potentials in studies of perception and attention," *Attention, Perception, & Psychophysics*, vol. 72, no. 8, pp. 2031–2046, 2010.

[27] S. Bernstein, "Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes," *Mathematische Annalen*, vol. 97, no. 1, pp. 1–59, 1927.

[28] E. Donchin, "Surprise!… surprise?," *Psychophysiology*, vol. 18, no. 5, pp. 493–513, 1981.

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

[30] G. Pfurtscheller, "Graphical display and statistical evaluation of event-related desynchronization (erd)," *Electroencephalography and clinical neurophysiology*, vol. 43, no. 5, pp. 757–760, 1977.

[31] G. Pfurtscheller and A. Aranibar, "Event-related cortical desynchronization detected by power measurements of scalp eeg," *Electroencephalography and clinical neurophysiology*, vol. 42, no. 6, pp. 817–826, 1977.

[32] G. Pfurtscheller, "Event-related synchronization (ers): an electrophysiological correlate of cortical areas at rest," *Electroencephalography and clinical neurophysiology*, vol. 83, no. 1, pp. 62–69, 1992.

[33] F. L. Da Silva, A. Van Rotterdam, P. Barts, E. Van Heusden, and W. Burr, "Models of neuronal populations: the basic mechanisms of rhythmicity," *Progress in brain research*, vol. 45, pp. 281–308, 1976.

[34] G. Pfurtscheller and F. L. Da Silva, "Event-related eeg/meg synchronization and desynchronization: basic principles," *Clinical neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.

[35] G. Pfurtscheller and W. Klimesch, "Functional topography during a visuoverbal judgment task studied with event-related desynchronization mapping.," *Journal of Clinical Neurophysiology*, vol. 9, no. 1, pp. 120–131, 1992.

[36] W. Klimesch, H. Schimke, and J. Schwaiger, "Episodic and semantic memory: an analysis in the eeg theta and alpha band," *Electroencephalography and clinical Neurophysiology*, vol. 91, no. 6, pp. 428–441, 1994.

[37] W. Klimesch, "Memory processes, brain oscillations and eeg synchronization," *International Journal of Psychophysiology*, vol. 24, no. 1, pp. 61–100, 1996.

[38] W. Klimesch, M. Doppelmayr, H. Russegger, T. Pachinger, and J. Schwaiger, "Induced alpha band power changes in the human eeg and attention," *Neuroscience letters*, vol. 244, no. 2, pp. 73–76, 1998.

[39] M. J. Kahana, R. Sekuler, J. B. Caplan, M. Kirschen, and J. R. Madsen, "Human theta oscillations exhibit task dependence during virtual maze navigation," *Nature*, vol. 399, no. 6738, pp. 781–784, 1999.

[40] J. B. Caplan, J. R. Madsen, A. Schulze-Bonhage, R. Aschenbrenner-Scheibe, E. L. Newman, and M. J. Kahana, "Human $\theta$ oscillations related to sensorimotor integration and spatial learning," *The Journal of Neuroscience*, vol. 23, no. 11, pp. 4726–4736, 2003.

[41] E. V. Friedrich, R. Scherer, and C. Neuper, “The effect of distinct mental strategies on classification performance for brain–computer interfaces,” *International Journal of Psychophysiology*, vol. 84, no. 1, pp. 86–94, 2012.

[42] E. V. Friedrich, C. Neuper, and R. Scherer, “Whatever works: a systematic user-centered training protocol to optimize brain-computer interfacing individually,” *PloS one*, vol. 8, no. 9, p. e76214, 2013.

[43] W. Penfield and E. Boldrey, “Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation.,” *Brain: A journal of neurology*, 1937.

[44] C. M. Bishop, “Pattern recognition,” *Machine Learning*, vol. 128, 2006.

[45] J. Hohnsbein, M. Falkenstein, J. Hoormann, *et al.*, “Error processing in visual and auditory choice reaction tasks,” *Journal of Psychophysiology*, vol. 3, p. 32, 1989.

[46] M. Falkenstein, J. Hohnsbein, J. Hoormann, L. Blanke, *et al.*, “Effects of errors in choice reaction tasks on the erp under focused and divided attention,” *Psychophysiological brain research*, vol. 1, pp. 192–195, 1990.

[47] W. Gehring, M. Coles, D. Meyer, and E. Donchin, “The error-related negativity: an event-related brain potential accompanying errors,” *Psychophysiology*, vol. 27, p. 34, 1990.

[48] W. J. Gehring, B. Goss, M. G. Coles, D. E. Meyer, and E. Donchin, “A neural system for error detection and compensation,” *Psychological science*, vol. 4, no. 6, pp. 385–390, 1993.

[49] W. H. Miltner, C. H. Braun, and M. G. Coles, “Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a “generic” neural system for error detection,” *Journal of cognitive neuroscience*, vol. 9, no. 6, pp. 788–798, 1997.

[50] G. Schalk, J. R. Wolpaw, D. J. McFarland, and G. Pfurtscheller, “Eeg-based communication: presence of an error potential,” *Clinical Neurophysiology*, vol. 111, no. 12, pp. 2138–2144, 2000.

[51] L. Parra, C. Alvino, A. Tang, B. Pearlmutter, N. Yeung, A. Osman, and P. Sajda, “Linear spatial integration for single-trial detection in encephalography,” *Neuroimage*, vol. 17, no. 1, pp. 223–230, 2002.

[52] L. C. Parra, C. D. Spence, A. D. Gerson, and P. Sajda, “Response error correction-a demonstration of improved human-machine performance using real-time eeg monitoring,” *IEEE transactions on neural systems and rehabilitation engineering*, vol. 11, no. 2, pp. 173–177, 2003.

[53] P. W. Ferrez and J. d. R. Millán, “You are wrong!—automatic detection of interaction errors from brain waves,” in *Proceedings of the 19th international joint conference on Artificial intelligence*, no. EPFL-CONF-83269, 2005.

[54] P. W. Ferrez and J. d. R. Millán, “Error-related eeg potentials generated during simulated brain–computer interaction,” *IEEE transactions on biomedical engineering*, vol. 55, no. 3, pp. 923–929, 2008.

[55] R. Chavarriaga and J. d. R. Millán, “Learning from eeg error-related potentials in noninvasive brain-computer interfaces,” *IEEE transactions on neural systems and rehabilitation engineering*, vol. 18, no. 4, pp. 381–388, 2010.

[56] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, “Bci2000: a general-purpose brain-computer interface (bci) system,” *IEEE Transactions on biomedical engineering*, vol. 51, no. 6, pp. 1034–1043, 2004.

[57] V. Jayaram, M. Alamgir, Y. Altun, B. Scholkopf, and M. Grosse-Wentrup, “Transfer learning in brain-computer interfaces,” *IEEE Computational Intelligence Magazine*, vol. 11, no. 1, pp. 20–31, 2016.

# Abbreviations, Symbols and Operators

## List of Abbreviations

| Notation | Description |
| --- | --- |
| BCI | Brain-Computer Interface |
| CAR | common average reference |
| EEG | electroencephalography |
| ERD | event-related desyncronization |
| ERN | error-related negativity |
| ERP | event-related potential |
| ERS | event-related syncronization |
| LDA | Linear Discriminant Analysis |
| SVM | Support Vector Machine |

## List of Operators

| Notation | Description | Operator |
| --- | --- | --- |
| det | the determinant | $\det(\bullet)$ |
| log | the natural logarithm | $\log(\bullet)$ |
| $\|\| \ \|\|$ | the euclidean norm | $\|\bullet\|$ |

# Figures and Tables

## List of Figures



## List of Tables