

# **Brain Computer Interfaces based on imaginary hand movement using EEG beamforming**

**Celine De Vreese**

Promotoren: prof. dr. Stefaan Vandenberghe, dr. ir. Hans Hallez  
Begeleiders: Gregor Strobbe, Pieter van Mierlo

Masterproef ingediend tot het behalen van de academische graad van  
Master of Science in Biomedical Engineering

Vakgroep Elektronica en Informatiesystemen  
Voorzitter: prof. dr. ir. Jan Van Campenhout  
Faculteit Ingenieurswetenschappen en Architectuur  
Academiejaar 2011-2012





# **Brain Computer Interfaces based on imaginary hand movement using EEG beamforming**

**Celine De Vreese**

Promotoren: prof. dr. Stefaan Vandenberghe, dr. ir. Hans Hallez  
Begeleiders: Gregor Strobbe, Pieter van Mierlo

Masterproef ingediend tot het behalen van de academische graad van  
Master of Science in Biomedical Engineering

Vakgroep Elektronica en Informatiesystemen  
Voorzitter: prof. dr. ir. Jan Van Campenhout  
Faculteit Ingenieurswetenschappen en Architectuur  
Academiejaar 2011-2012



De auteur geeft de toelating deze masterproef voor consultatie beschikbaar te stellen en delen van de masterproef te kopiëren voor persoonlijk gebruik.

Elk ander gebruik valt onder de beperkingen van het auteursrecht, in het bijzonder met betrekking tot de verplichting de bron uitdrukkelijk te vermelden bij het aanhalen van resultaten uit deze masterproef.

The author gives permission to make this master dissertation available for consultation and to copy parts of this master dissertation for personal use.

In the case of any other use, the limitations of the copyright have to be respected, in particular with regard to the obligation to state expressly the source when quoting results from this master dissertation.

Ghent, June 2012

The author

Celine De Vreese

# Prologue

This thesis was written as the final completion of my studies of Master of Science in Biomedical Engineering. Though it was not always easy, I have definitely enjoyed the past 5 years. And to be able to conclude with a thesis in - to my opinion - one of the most interesting subjects in the field of signal processing and biomedical engineering, most certainly was a nice reward.

However I couldn't have done it without the help of my supervisor ir. Gregor Strobbe whom I want to thank for his profound guidance and the occasional encouragement. Furthermore I want to thank my promoters prof. ir. Stefaan Vandenberghe en dr. ir. Hans Hallez for introducing and engaging me into the fascinating subject of BCIs.

Lastly I also want to thank my family and close friends for their seemingly infinite support.

Celine De Vreese

June, 2011

# Brain Computer Interfaces based on imaginary hand movement using EEG beamforming

By

Celine DE VREESE

Promotor: prof. dr. Stefaan VANDENBERGHE, dr. ir. Hans HALLEZ

Supervisor: ir. Gregor STROBBE, ir. Pieter VAN MIERLO

Master's Thesis submitted to obtain the degree of  
MASTER OF SCIENCE IN BIOMEDICAL ENGINEERING  
Faculty of Engineering and Architecture  
Ghent University  
Academic year 2011-2012

Department ELECTRONICS AND INFORMATION SYSTEMS

President: prof. dr. ir Jan CAMPENHOUT

## Summary

In this thesis an EEG-based BCI system is developed using beamforming spatial filtering techniques for feature extraction. We start by optimizing the system by comparing the accuracy of the system for different values of each parameter. The system is applied on both a synchronous and asynchronous dataset. The results are then compared. In a next step we use this optimized system to research the influence of the head model on the classification accuracy. We compare a single shell and 3-shell spherical head model, a template BEM model and a personalized BEM head model. Furthermore we combine this personalized BEM head model with fMRI data obtained from the same subject. Results show that for increasingly complex head models, the classification accuracy tends to increase as well.

## Keywords

Brain-Computer Interface - Electroencephalography - Motor Imagery - Beamformer - Head models

# Brain-Computer Interfaces based on imaginary hand movement using EEG beamforming

Celine De Vreese

Supervisor(s): Dr. ir. S. Vandenberghe, ir. G. Strobbe and ir. P. van Mierlo

**Abstract**— In this article an elektroencephalograph (EEG)-based BCI which uses Linear Constrained Minimum Variance (LCMV)-beamforming as a spatial filtering technique, was designed. In a first step we optimized the system's parameters and at the same time looked into its robustness. Next we evaluated the BCI for real-time applications. Lastly we investigated how the use of different head models influences the performance of the system. We constructed a BCI analysis method based on beamforming techniques with a lot of potential for real-time applications. Performance can be increased by using more complex and personalized head models.

**Keywords**— Brain-Computer Interfaces - Beamformer - Motor imagery - Head models - Electroencephalography

## I. INTRODUCTION

Brain-Computer Interfaces (BCIs) are systems developed to measure brain activity and use this information to address an external device. They foresee in a direct communication channel between an individual and a machine. Some people are completely paralysed after a severe stroke or by a disease such as Amyotrophic Lateral Sclerosis (ALS). These conditions are purely physical, leaving the mental capacities perfectly intact. In these cases BCIs might offer a solution for partly restoring communication possibilities.

We developed a BCI which uses brain activity measured at the scalp. We want to be able to distinguish two different states, namely right and left hand motor imagery (MI). The corresponding activity is generated in the motorisensor cortex and can typically be found in a frequency band between 8 and 14 Hz ( $\mu$  rhythm). During motor imagery Event Related Desynchronization (ERD), followed by Even-Related Synchronization (ERS) of signals takes places in the contra-lateral hemisphere. It is this information we want to extract from the signals using spatial filtering. We opted to use a beamformer as spatial filter.

Problems frequently occurring with BCIs are a lack of robustness and low accuracy for real-time applications. Therefore we first optimized the system in order to make the system more robust. Next we investigated the performance of the system on an asynchronous dataset. Furthermore BCIs are often prone to inter-subject differences. We investigated to which extent different head models can address this problem.

## II. MATERIALS

### A. Data

For the different parts of the research we used different datasets. For the calibration of the designed BCI we used the datasets provided by the BCI Competition IV.[1] To see how the different parameters influence the system, we used the calibration dataset of subject 1E and 1G. The final system was also tested on the calibration datasets of subjects 1B, 1C and 1D. Every calibration dataset consisted of 100 trials right MI and 100

trials left MI. Each MI task lasted 4 seconds and was always proceeded by 4 seconds of rest. Besides the calibration dataset also an evaluation dataset was provided per subject. This was an asynchronous dataset, which consisted of trials of variable length. Since it was an asynchronous dataset, no markers to indicate the start and the end of the trials were added.

To explore the influence of the head models, we used a second dataset which was recorded at Ghent University. This dataset contained 50 trials per state and each trial lasted 4 seconds. [2]

### B. Head models

We used 5 different head models, namely a single shell spherical head model, a 3-shell spherical head model, the ICMB 152 MNI template BEM model [3] and a personalized BEM model constructed based on MRI scans (Magnetom Trio, A Tim System 3T, Siemens, Germany). For the same subject, also fMRI data was recorded during a fingertapping experiment.

## III. METHODS

### A. Source localization based on Beamformers

EEG source localization is the estimation of the active sources in the brain, based on the measured potentials on the scalp. The localization problem consists of a forward and inverse problem. The forward problem incorporates the characteristics of the head, the electrode positions and the dipole source distribution, i.e. electrical dipoles distributed on a 3D grid in the head. The inverse problem tries to find a filter which explains the measured potentials optimally, assuming the forward model. The inverse problem is however an ill-posed problem. The resultant filter depends on the imposed constraints. In case of the beamformer filter, a separate filter per grid point is constructed. In the ideal case the filter only passes the signal at that grid point and attenuates all the signals coming from other locations, leaving us with the dipole moment at that grid point. [4]

### B. The translation process

The preprocessing step consisted of a Common Average Filter, baseline correction and a bandpass filter. Bandpass frequencies were determined per subject, based on the time frequency analysis of the averaged trials. Trials with artefacts were visually detected and completely rejected. Next we chose a certain amount of grid points per hemisphere and applied spatial filtering using beamformers. We separated the signals obtained after filtering in subsignals from 1 second. The feature vector was then built by taking the logarithmic power of these subsignals. For the final classification we used a standard Fisher LDA classifier. This classifier was trained by 80% of the training trials,

and resulted in a vector B which contained the coefficient of the optimal linear combination of features and an offset value c. To classify a new signal the feature vector of this signal was calculated, and linearly combined using B and c. Based on its final value, the signal was then classified as right or left hand MI. [5]

#### IV. RESULTS

##### A. Robustness of the system and real-time performance

We already raised three frequent problems with BCIs. A first problem is the lack of robustness for some BCIs. To estimate the robustness of our system and to maximize the performance, we investigated the influence of different parameters of the system and optimized them. An overview of the parameters and their optimal values is given in Table I. The parameter influencing the performance most is the amount of grid points (features) used for classification.

Parameter	Final value
Baseline Correction	Applied
Artefact removal	Applied
# Trials for filter training	80/100
Regularization Parameter	0.05
Time window	0.5 - 3 sec
# Grid points per hemisphere	15
Trial step size	1 sec
Training Percentage	80 %

TABLE I  
PARAMETER LIST: FINAL VALUES

The results obtained with these parameters for the different datasets of BCI Competition IV can be found in Table II. The parameters were optimized based on the datasets of subjects 1E and 1G. This is reflected in a higher accuracy for these subjects in comparison with the accuracy obtained for the other subjects. But also differences between subjects, e.g. the ability to handle BCIs, will be at the origin of the difference. Regarding subject 1C we want to mention that this is an artificial dataset (as is 1E), and thus more likely to show less profound the ERD and ERS patterns, on which our system is based.

Subject	1B	1C	1D	1E	1G
Performance	58.2%	51.9%	64.7%	79.3%	79.5%

TABLE II  
CLASSIFICATION ACCURACY FOR DIFFERENT SUBJECTS

A second problem with the design of BCIs is that the system often works well when it can classify a complete trial, but when applied in real-time the performance leaves much to be desired. However, since the ultimate goal of BCI research is to be able to handle an external device, real-time performance should be one of our main concerns. Therefore we tested our system on the evaluation datasets of subject 1E and 1G. The results are respectively an accuracy of 83.65% and 80.81%. We thus even notice a slight increase in the accuracy. So the system definitely seems

to be applicable in real-time. This is a system with a window of 500 samples or a delay of 0.5 seconds (sample frequency = 1000 Hz).

##### B. Influence of the head model and fMRI data

Beamformer filters are obtained in the inverse problem of source localization. Therefore the chosen head model is an important factor during the calculation of the filter. Results of the classification of the dataset recorded at the university of Ghent for different head models is given in Figure 1. We see that the accuracy increases for head models with a higher complexity. The added value of including fMRI data appears to be rather limited. This might be due to small deviations in the correct alignment of the head model, the grid points and the electrode cap. It is also possible that the region active during fingertapping doesn't exactly correspond with the region activated during MI. Another interesting conclusion we can draw is that including information about the orientation of the dipoles (i.e. we assume them to be perpendicular to the cortex) improves the accuracy considerably.

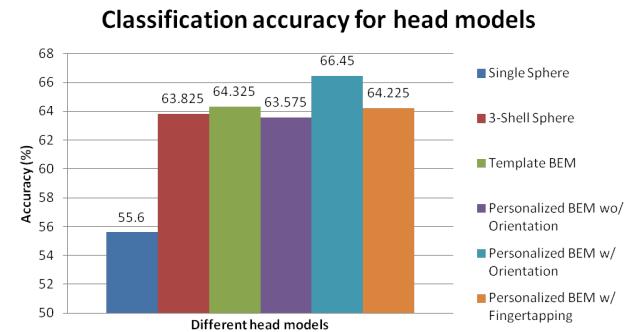


Fig. 1. Classification accuracy for different head models

#### V. CONCLUSION

We can conclude that beamforming techniques are a valid alternative for standard spatial filtering techniques. In general these BCI systems deliver a good performance, especially when its parameters are optimized for the used dataset. The most surprising result was the excellent performance on an asynchronous dataset. This implies that beamforming techniques might work well in real-time applications. Finally we conclude that the use of a more complex head model, leads to better results. Preferably a personalized head model with predefined dipole orientations is used. The added value of the fMRI data appeared to be rather limited.

#### REFERENCES

- [1] BCI competition IV, dataset 1, <http://www.bbci.de/competition/iv>
- [2] G. Stroobie, and R. Van de Vaerd, *Geschiktheidsstudie van connectiviteitsmaten in het electroencephalogram voor Brain-Computer Interfaces*, Ghent University, 2010.
- [3] ICBM Atlases, [www.loni.ucla.edu/ICBM](http://www.loni.ucla.edu/ICBM).
- [4] B.D. Van Veen, W. van Drongelen, M. Yuchtman and A. Suzuki, *Localization of brain electrical activity via linearly constrained minimum variance spatial filtering*, IEEE Trans Biomed Eng, 1997.
- [5] G. Dornhege, J.R. Millán, T. Hinterberger, D.J. McFarland and K.R. Müller, *Toward Brain-Computer Interfacing*

# Brein-Computer Interfaces voor imaginaire handbewegingen op basis van EEG beamformers

Celine De Vreese

Supervisor(s): Dr. ir. S. Vandenbergh, ir. G. Strobbe and ir. P. van Mierlo

**Abstract**— In dit artikel werd een elektro-encephalografie(EEG)-gebaseerde Brein-Computer Interface (BCI) ontwikkeld die gebruik maakt van spatiale filtering aan de hand van Linear Constrained Minimum Variance (LCMV)-beamforming. Eerst optimaliseerden we de parameters van het systeem en onderzochten zo zijn robuustheid. Vervolgens evaluateerden we de BCI voor real-time applicaties. Ten slotte onderzochten we hoe het gebruik van verschillende hoofdmodellen de performantie beïnvloedde. We construeerden een BCI analyse methode gebaseerd op beamformer technieken, met veel potentieel voor real-time applicaties. De performantie kan verhoogd worden door complexere en gepersonaliseerde hoofdmodellen te gebruiken.

**Keywords**— Brein-Computer Interfaces - Beamformer - Imaginaire motoriek - hoofdmodellen - elektro-encephalografie

## I. INTRODUCTIE

BCIs zijn systemen ontworpen om hersenactiviteit op te meten. Ze gebruiken deze informatie om een extern apparaat aan te sturen. Ze voorzien dus in een rechtstreeks communicatiekanaal tussen een individu en een machine. Sommige personen zijn compleet verlamd na een beroerte of door een aandoening zoals vb. amyotrofe laterale sclerose (ALS). Deze aandoeningen zijn puur fysisch, en de mentale capaciteiten zijn nog perfect intact. Hier kunnen BCIs mogelijk een oplossing bieden.

Wij gebruikten een BCI die de hersenactiviteit opneemt ter hoogte van de hoofdhuid. We wilden 2 toestanden van elkaar kunnen onderscheiden, namelijk imaginaire rechterhand en linkerhand beweging. De corresponderende activiteit wordt gegenereerd ter hoogte van de sensorimotorische cortex en bevindt zich typisch in een frequentieband tussen 8 en 14 Hz ( $\mu$ -ritme). Gedurende de imaginaire bewegingen, kan men in deze frequentieband een gebeurtenis gerelateerde synchronisatie en desynchronisatie van de signalen in de contralaterale hemisfeer onderscheiden. Het is deze informatie die we wensten uit de signalen te extraheren met behulp spatiale filtering. We hebben geopteerd voor een beamformer spatiale filter.

Vaak voorkomende problemen bij BCIs zijn een gebrek aan robuustheid en slechte performantie in real-time applicaties. Eerst optimaliseerden we de BCI om hem zo robuster te maken. Vervolgens onderzochten we hoe het systeem presteerde voor real-time data. Verder zijn BCIs vaak ook gevoelig voor inter-subject verschillen. We onderzochten in welke mate verschillende hoofdmodellen hier een antwoord op boden.

## II. MATERIALEN

### A. Data

Voor de verschillende delen van het onderzoek gebruikten we verschillende datasets. Voor het calibreren van het ontwikkelde BCI systeem maakten we gebruik van de datasets voorzien door

de BCI Competition IV [1]. Om de invloed van verschillende parameters te onderzoeken, maakten we in eerste instantie gebruik van de calibratie dataset voor subject 1E en 1G. Het finale systeem werd vervolgens ook uitgetest op de data van subject 1B, 1C en 1D. Elke dataset bestond uit 100 trials per imaginaire handbeweging, die elk 4 seconden duurden. Hierbij werd elke trial voorafgegaan door 4 seconden rust. Per subject was ook een evaluatiedataset voorzien. Dit is een asynchrone dataset met trials van variabele lengte. De data was niet voorzien van markers die het begin en eind van de trial aangaven.

Om de invloed van de hoofdmodellen te onderzoeken, gebruikten we een tweede dataset die reeds vroeger aan de universiteit van Gent werd opgenomen. Deze bevatte ook 4 seconden durende trials. Hier waren 50 trials per toestand voorzien [2].

### B. Hoofdmodellen

De hoofdmodellen die we gebruikten, zijn een 1-laag sferisch hoofdmodel, een 3-lagen sferisch hoofdmodel, het ICMB 152 MNI template BEM model [3] en een gepersonaliseerd BEM model gebaseerd op MRI scans (Magnetom Trio, A Tim System 3T, Siemens, Germany). Bij deze persoon werd ook fMRI data opgenomen tijdens een fingertapping experiment.

## III. METHODEN

### A. Bronlokalisatie gebaseerd op beamformers

Bronlokalisatie is een schatting van de actieve bronnen in de hersenen, gebaseerd op de gemeten EEG-potentiaLEN. Dit probleem wordt opgedeeld in een voorwaarts en een invers probleem. Het voorwaartse probleem incorporeert de kenmerken van het hoofd, de electrode posities en de brondistributie, i.e. de elektrische dipolen gedistribueerd over een 3D grid in het hoofd. Het invers probleem probeert een filter te vinden die, uitgaande van het voorwaartse model, de gemeten EEG-signalen optimaal kan verklaren. De filter is afhankelijk van voorwaarden die opgelegd worden. Bij een beamformer construeren we per gridpunt een filter die het signaal op deze locatie volledig doorlaat, en de signalen van alle andere locaties volledig onderdrukt. In het ideale geval resulteert dit in het dipoolmoment van de bron op het beschouwde gridpunt.[4]

### B. Het translatieproces

In de voorbewerking pasten we achtereenvolgens een Common Average filter, baseline correctie en een banddoorlaatfilter toe. De banddoorlaatfrequenties werden per subject bepaald op basis van een tijds-frequentie analyse van het gemiddelde over alle trials. Ten slotte werden de trials die artefacten bevatten uit de dataset verwijderd. Dan kozen we een aantal gridpunten per

hemisfeer en pasten daar de corresponderende beamformer filters op toe. We deelden de signalen verkregen na filtering dan op in aparte signalen van 1 seconde. De feature vector werd dan opgebouwd door het logaritmisch vermogen van deze signalen te nemen. Voor de uiteindelijke classificatie gebruikten we een Fisher LDA classifier. Deze leverde ons de lineaire combinatie van features voor optimale classificatie op. Nieuwe signalen werden geclasseerd door deze lineaire combinatie toe te passen op de bekomen feature vector. Het resultaat bepaalde dan welke klasse aan het signaal werd toegekend.[5]

#### IV. RESULTATEN

##### A. Robuustheid van het systeem en real-time performantie

We hebben reeds drie vaak voorkomende problemen bij BCIs aangekaart. Een eerste probleem is het gebrek aan robuustheid bij sommige BCIs. Om een idee te krijgen van de robuustheid van ons systeem en om de performantie te maximaliseren, hebben we de invloed van verschillende parameters van het systeem onderzocht en deze trachten te optimaliseren. Daartoe hebben we dataset 1E en 1G gebruikt. Een overzicht van de parameters en hun finale geoptimaliseerde waarde is te vinden in Tabel I.

Parameter	Finale waarde
Baseline Correctie	Toegepast
Artefacts verwijderen	Toegepast
# Trials voor training filter	80/100
Regularisatie Parameter	0.05
Tijdsvenster	0.5 - 3 sec
# Gridpunten per hemisfeer	15
Trial stapgrootte	1 sec
Training Percentage	80 %

TABLE I

PARAMETERLIJST: FINALE WAARDEN

De resultaten bekomen met deze parameters voor de verschillende datasets zijn te vinden in Tabel II. Hier zien we dat de subjecten voor wie het systeem geoptimaliseerd werd, een heel goede performantie hebben. Voor de andere subjecten ligt de performantie iets lager. Dit kan verschillende oorzaken hebben. Ten eerste werd het systeem niet geoptimaliseerd voor deze subjecten, o.a. de keuze van de optimale gridpunten kan tussen personen heel verschillend zijn. Tussen subjecten kan er ook een groot verschil in aanleg zijn voor het aansturen van BCIs.

Subject	1B	1C	1D	1E	1G
Performantie	58.2%	51.9%	64.7%	79.3%	79.5%

TABLE II

CLASSIFICATIE NAUWKEURIGHEID VOOR VERSCHILLENDEN DATASETS

Een tweede probleem bij het ontwikkelen van BCIs is dat het systeem vaak wel werkt bij het classificeren van volledige trials, maar indien het in real-time wordt toegepast, daalt de accuraatheid dramatisch. Gezien het uiteindelijke doel van een BCI

is om een extern device aan te sturen, is real-time toepasbaarheid nochtans een must. Daarom hebben we ons systeem ook op de evaluatie datasets van subjecten 1E en 1G toegepast. Dit resulteerde respectievelijk in een nauwkeurigheid van 83.65% en 80.81%. We nemen een kleine stijging in de performantie waar. Het systeem zal dus hoogst waarschijnlijk goed toepasbaar zijn in real-time applicaties. Bij de evaluatie werd met 500 samples, of een delay van 0.5 seconde gewerkt.

##### B. Invloed van het hoofdmodel

Daar de filter berekend wordt op basis van het voorwaarts model, speelt het hoofdmodel een grote rol bij de berekening van de filter. De resultaten voor de verschillende hoofdmodellen zijn gegeven in Figuur 1. We zien dat de performantie stijgt als het hoofdmodel realistischer wordt. De toegevoegde waarde van fMRI data blijkt eerder beperkt. Dit komt mogelijk door kleine afwijkingen tussen de reële elektrodenposities en de veronderstelde elektrodenposities. Er kan ook een klein verschil zitten op de regio corresponderend met imaginaire handbewegingen en de regio geactiveerd tijdens actieve vingerbewegingen.

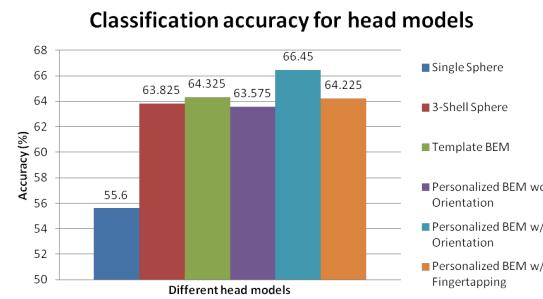


Fig. 1. Classificatie nauwkeurigheid voor verschillende hoofdmodellen

#### V. CONCLUSIE

We concluderen dat beamforming technieken een goed alternatief kunnen vormen voor meer standaard spatiale filter technieken. Ze leveren over het algemeen goede resultaten, die geoptimaliseerd kunnen worden door de parameters per subject te optimaliseren. Een heel belangrijke factor is ook dat ze in real-time kunnen toegepast worden zonder al te veel aan nauwkeurigheid in te boeten. Ten slotte merken we nog op dat een meer complex hoofdmodel gebruiken, leidt tot betere resultaten. Bij voorkeur wordt een gepersonaliseerd hoofdmodel gebruikt, waarbij men de veronderstelde dipoolorientaties meegeeft. De toegevoegde waarde van fMRI data blijkt eerder beperkt.

#### REFERENCES

- [1] BCI competition IV, dataset 1, <http://www.bbci.de/competition/iv>
- [2] G. Stroobie, and R. Van de Vaerd, *Geschiktheidsstudie van connectiviteitsmaten in het electroencephalogram voor Brain-Computer Interfaces*, Ghent University, 2010.
- [3] ICBM Atlases, [www.loni.ucla.edu/ICBM](http://www.loni.ucla.edu/ICBM).
- [4] B.D. Van Veen, W. van Drongelen, M. Yuchtman and A. Suzuki, *Localization of brain electrical activity via linearly constrained minimum variance spatial filtering*, IEEE Trans Biomed Eng, 1997.
- [5] G. Dornhege, J.R. Millán, T. Hinterberger, D.J. McFarland and K.R. Müller, *Toward Brain-Computer Interfacing*

# Contents

<b>Prologue</b>	<b>iv</b>
<b>Prologue</b>	<b>iv</b>
<b>Summary</b>	<b>vi</b>
<b>Extended abstract</b>	<b>viii</b>
<b>Extended abstract - Dutch</b>	<b>x</b>
<b>Acronyms</b>	<b>xiv</b>
<b>1 Introduction to BCIs</b>	<b>1</b>
1.1 A general introduction . . . . .	1
1.1.1 The principle of BCIs . . . . .	1
1.1.2 Different kinds of BCIs and their (dis-)advantages . . . . .	3
1.2 Signal Acquisition: a BCI based on scalp EEG . . . . .	4
1.2.1 What are we measuring with scalp EEG? . . . . .	4
1.2.2 EEG analysis . . . . .	6
1.2.3 Artefacts . . . . .	7
1.3 Signal Processing . . . . .	8
1.3.1 Preprocessing . . . . .	8
1.3.2 Feature Extraction . . . . .	9
1.3.3 Signal Classification . . . . .	10
1.4 Source localization . . . . .	10
1.4.1 Introduction . . . . .	10
1.4.2 Forward model . . . . .	11
1.4.3 Inverse model . . . . .	12
1.5 Applications and current state-of-the-art . . . . .	12
1.5.1 Communication . . . . .	12
1.5.2 Environmental control and virtual worlds . . . . .	13
1.5.3 Neural prosthetics . . . . .	13

<b>2 Overview of this Master's Thesis</b>	<b>15</b>
2.1 The goal of BCI Research . . . . .	15
2.2 Main problems encountered in BCI research . . . . .	15
2.3 Objectives . . . . .	16
2.3.1 Understanding beamforming and its possibilities . . . . .	16
2.3.2 Design of an accurate BCI system . . . . .	17
2.3.3 Fine-tuning of parameters . . . . .	18
2.3.4 Real-time application . . . . .	19
2.3.5 Influence of the head model . . . . .	19
<b>3 Beamforming</b>	<b>20</b>
3.1 Introduction . . . . .	20
3.2 Forward Model . . . . .	21
3.3 The inverse model: Filter Design . . . . .	22
3.3.1 Introduction . . . . .	22
3.3.2 Definitions and notations . . . . .	23
3.3.3 Mathematical description . . . . .	24
3.4 Some remarks . . . . .	25
<b>4 Input: Material and Experimental Set-up</b>	<b>27</b>
4.1 Datasets . . . . .	27
4.1.1 BCI competition IV, dataset 1 . . . . .	27
4.1.2 Data recorded at Ghent University . . . . .	29
4.1.3 Data Protocol . . . . .	30
4.2 Head Models . . . . .	30
4.2.1 Standardized head models: 1-shell 3-shell spherical head model . . . . .	31
4.2.2 Standardized head models: Realistically shaped head model . . . . .	32
4.2.3 Individual head model and neurophysiological data . . . . .	34
<b>5 Data processing, Classification and Evaluation</b>	<b>37</b>
5.1 Preprocessing . . . . .	37
5.1.1 Objectives of the preprocessing step . . . . .	37
5.1.2 Common average reference . . . . .	37
5.1.3 Trial extraction . . . . .	38
5.1.4 Artefact rejection . . . . .	38
5.1.5 Baseline correction . . . . .	39
5.1.6 Time-Frequency analysis and band pass filtering . . . . .	39
5.2 Feature Extraction . . . . .	40
5.2.1 Objectives of feature extraction . . . . .	40
5.2.2 How to determine the beamformer filter . . . . .	40

5.2.3	Determine grid points corresponding with ROI . . . . .	42
5.2.4	Feature calculation . . . . .	43
5.3	Classification . . . . .	44
5.3.1	Objectives of classification . . . . .	44
5.3.2	Characteristics of Fisher LDA classifier . . . . .	44
5.4	Evaluation . . . . .	46
5.4.1	Objectives and evaluation criteria . . . . .	46
5.4.2	Calibration data set . . . . .	46
5.4.3	Evaluation data set . . . . .	48
<b>6</b>	<b>Results: Building a robust BCI</b>	<b>49</b>
6.1	Input: BCI competition IV Data specifications . . . . .	49
6.2	Fine-tuning some parameters . . . . .	49
6.2.1	Standard parameters and the corresponding classification error . . . . .	49
6.2.2	Preprocessing . . . . .	51
6.2.3	Beamformer parameters . . . . .	52
6.2.4	Feature extraction . . . . .	54
6.2.5	Classification . . . . .	55
6.3	Evaluation of the calibration datasets . . . . .	56
6.3.1	Final parameter selection . . . . .	56
6.3.2	Final results . . . . .	57
6.4	Evaluation of the evaluation datasets . . . . .	58
6.4.1	Specifics of the evaulation . . . . .	58
6.4.2	Influence of threshold . . . . .	59
6.4.3	Influence of the window size . . . . .	61
6.4.4	Idle state . . . . .	61
<b>7</b>	<b>Results: Influence of the head model</b>	<b>64</b>
7.1	Introduction . . . . .	64
7.2	Description . . . . .	65
7.3	Results . . . . .	65
7.4	Remarks: Amount of electrodes and position of the electrodes . . . . .	68
<b>8</b>	<b>Conclusion</b>	<b>69</b>
<b>9</b>	<b>Discussion and Future Work</b>	<b>72</b>
9.1	Introduction . . . . .	72
9.2	Building an accurate BCI System . . . . .	72
9.3	Real-time classification . . . . .	73
9.4	Head models . . . . .	74

<i>Contents</i>	xiii
<b>Bibliography</b>	<b>75</b>
<b>List of Figures</b>	<b>78</b>
<b>List of Tables</b>	<b>81</b>

# Acronyms

**ALS** Amyotrophic Lateral Sclerosis.

**BCI** Brain-Computer Interface.

**BMI** Brain-Mind Interface.

**CLIS** Complete Locked-In Syndrome.

**CSF** Cerebrospinal Fluid.

**EEG** Electroencephalographic signal.

**ERD** Event Related Desynchronization.

**ERP** Event Related Potential.

**fMRI** functional Magnetic Resonance Imaging.

**HMM** Hidden Markov Model.

**ICA** Independent Component Analysis.

**ITR** Information Transfer Rate.

**LCMV** Linearly Constrained Minimum Variance.

**LDA** Linear Discriminant Analysis.

**LIS** Locked-In Syndrome.

**MI** Motor Imagery.

**NAI** Neural Activity Index.

**PCA** Principal Component Analysis.

**ROI** Region Of Interest.

**SCI** Spinal Cord Injury.

**SCP** Slow Cortical Potential.

**SMR** Sensory Motor Rhythm.

**SNR** Signal-to-Noise Ratio.

**SVM** Support Vector Machine.

# Chapter 1

## Introduction to BCIs

### 1.1 A general introduction

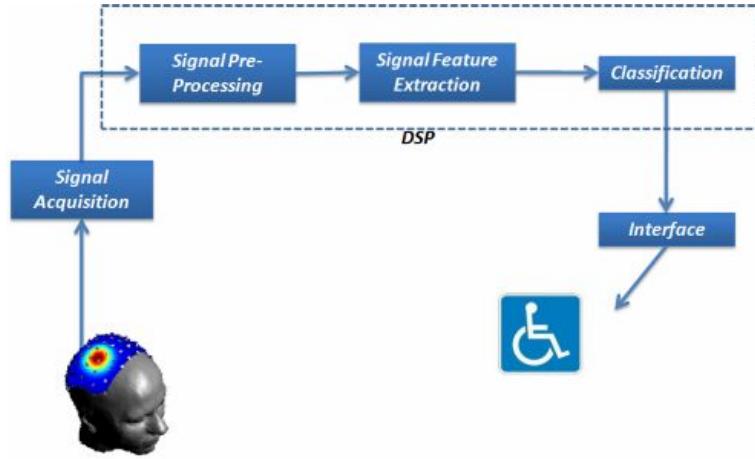
#### 1.1.1 The principle of BCIs

Brain-Computer Interfaces (BCIs) - also called Brain-Mind Interfaces (BMIs) - are systems developed to record brain activity and to use this information to address an external device. Jeff Hawkins, a pioneer in computer engineering, whose attention in the late nineties was drawn by neuroengineering, stated during one of his presentations: "We are our brains. My brain is talking to your brain; our bodies are hanging along for the ride." If one wants to capture the main idea behind brain computer interfaces, this quote probably says it all.

Since every thought or action is accompanied by a change of state in our brains, it is possible to use these alterations as a source of knowledge to discover what is going on in a person's mind. Brain activity can be captured as electric or magnetic signals directly at the level of the neurons or via electrodes placed on the scalp (invasive vs non-invasive BCIs). Given that different thoughts or acts generate different signals in different parts of the brain, it is possible to process this input to a functional signal that can activate a mechanical device, e.g. a bionic limb, or can control a special computer.

There are different diseases which can keep people from communicating with the outside world. Some patients are completely paralysed after having a severe stroke. Also people suffering from Amyotrophic Lateral Sclerosis (ALS) are locked in their own body during the last phases of their disease. These disorders are often purely physical, leaving the mental capacities and the senses perfectly intact. In those cases a BCI might offer a solution to partly replace the muscular channels and can be used as an alternative way to communicate. Some specific examples of output devices are thought controlled motorized wheel chairs, various neuroprosthetic devices and mental mouse applications in technology products. Additionally BCI technology is about to jump from the medical sector into the consumer gaming world.

The challenge of BCI design is to find a robust method which can translate with high accuracy the captured brain signals to a command for an external device. In order to achieve an optimal accuracy a lot of processing has to be performed on the recorded signals.



**Figure 1.1:** Basic block diagram of a BCI system incorporating signal acquisition, processing and deployment [10]

Figure 1.1 shows the different components of a typical BCI system:

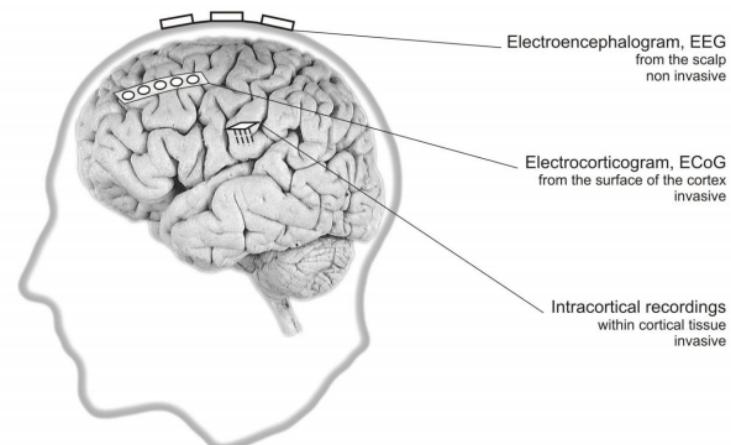
- **Signal acquisition:** The first step is to capture the brain activity. The brain activity can be measured in an invasive or non-invasive manner (see Section 1.1.2). Brain activity can be recorded as Electroencephalographic signal (EEG), functional Magnetic Resonance Imaging (fMRI), Positron Emission Tomography (PET) or via other methods.[1] In this dissertation we will use scalp EEG measured with an electrode cap. This is one of the most common acquisition methods. After the acquisition of the signals, the signals are amplified and digitized.
- **Signal preprocessing:** Raw EEG data are extremely noisy signals. The goal of preprocessing is to increase the Signal-to-Noise Ratio (SNR). Preprocessing can include (re-)referencing, artefact rejection and band-pass filtering.
- **Feature Extraction:** After the preprocessing step, we want to extract the features of the signals. These should contain the relevant information of the signals. A typical procedure during feature extraction is spatial filtering. This step is used to reduce the dimensionality of the problem.
- **Classification:** Based on the features a decision regarding the intention of the user has to be made in the final classification step. The classifier will translate the feature vector into a simple command.

- **Output:** Based on the classification outcome we can now give an instruction to an external device.

In the following sections we will discuss the different steps into more detail. Research concerning BCIs has become so wide-ranging that we will have to focus this discourse to the systems most relevant for this thesis.

### 1.1.2 Different kinds of BCIs and their (dis-)advantages

A distinction between BCIs can be made based on differences in the implementation of every step of the system. One can make a distinction based on the acquisition method, how the subjects are trained, how the signal is processed or based on the output. [1] The characteristics, advantages and disadvantages depend on the type of BCI used, and its implementation. An overview of the most common BCI systems available is given below. We make a principal distinction based on the place of recording, see Figure 1.2, [3].



**Figure 1.2:** Three different ways to detect the brain's electrical activity: EEG, ECoG, and intracranial recordings

- **Invasive BCIs:** The electrodes are placed directly in the grey matter. These BCIs are thought to record the most pure signals, since they are directly connected to single neurons. The direct connection ensures that there will be no attenuation nor spreading of the signal. Indeed, in practice some good results have been obtained concerning vision repair. But in case an invasive BCI is applied, there's a high risk of creating scar tissue around the electrodes which might lead to malfunction. As a result of the invasive procedure and the need for a personalized system, the overall cost will be much higher than the cost of a non-invasive BCI.
- **Partially Invasive BCIs:** The electrodes are still placed under the skull. Instead of placing them inside the grey matter, they are now placed at the surface of the grey

matter. There they can record the Electrocorticography (ECoG), the signals at the cerebral cortex. The main advantage here is that on the one hand there is a lower risk of scar-tissue formation, and on the other hand the attenuation and spreading of the signal is still limited.

- **Non-Invasive BCIs:** The interfaces used nowadays are in most cases non-invasive methods.[1] These use an electrode cap placed over the head to record the brain potentials. This reduces the risk of medical problems significantly. The high temporal resolution is preserved, making real time applications possible. On the contrary, the spatial resolution of non-invasive BCIs is quite low. This is due to the fact that the signals now first have to pass the low conductive skull before being measured. The system however is wearable and not too expensive with no medical risks. One of the main disadvantages is the extensive training often necessary before the user can use the interface optimally. Even after training, accuracy might still leave much to be desired.

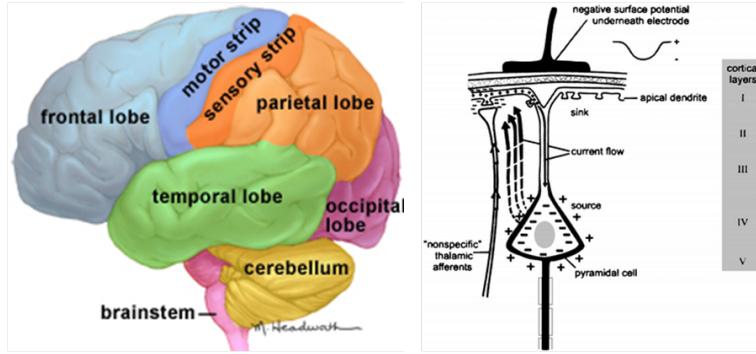
So on the whole one has to remark that the advantages and disadvantages strongly depend on the kind of system applied (e.g. invasive vs non-invasive, real time,...). In this thesis we will only address non-invasive BCIs based on scalp EEG.

## 1.2 Signal Acquisition: a BCI based on scalp EEG

### 1.2.1 What are we measuring with scalp EEG?

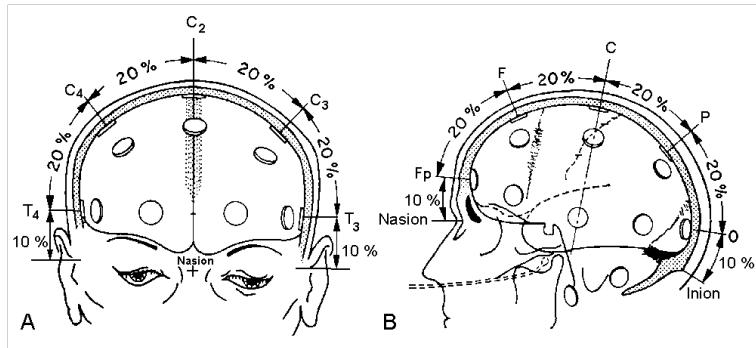
The brain is the centre of the nervous system and the most complex and important organ of our body. It consists of different structures. The cerebral cortex is the most important one for information processing and execution of sensor and motor tasks. It is divided into a left and a right hemisphere, which are connected by the corpus callosum. Each part of the hemisphere is responsible for a different function. This neurophysiological knowledge can be very useful when processing the EEG recordings, since this gives a first idea of where to look (i.e. at which electrode) for interesting signals.

The brain is composed of millions of neurons. The functions of the brain depend on the ability of neurons to transmit electrochemical signals to other cells, and their ability to respond appropriately to electrochemical signals received from other cells. When this happens a voltage change arises across the neuron cell. The voltage change is accompanied by an ion redistribution inside and outside the cell, which can be measured as an electrical current. The electrical activity can thus be modelled as a current dipole.[30] The current flow causes an electric field and also a potential field inside the human head, which then extends to the scalp. However the magnitude of these fields are very small and in order to have a measurable potential, the neurons have to fulfil two conditions. First of all they have to be simultaneously active. Secondly, the potential fields have to be aligned, such that instead of cancelling



**Figure 1.3:** Left: Functional areas of the brain: the cortical area particularly important for motor imagery based BCIs are the motor are and the somatosensory cortex, Right: Pyramidal cell

each other out, the potential fields amplify each other. At the surface of the cerebral cortex, one can find such kind of cell, namely the pyramidal cells. Neighbouring pyramidal cells are organized so that the axes of their dendrite tree are parallel with each other and normal to the cortical surface. Since these cells are also at the edge of the grey matter, they are close to the scalp and thus easy to measure. They are suggested to be the generators of the EEG, see Figure 1.3.[1, 30]



**Figure 1.4:** Electrode placement as defined by the International 10-20 system. The names of the electrodes are based on the area of the cerebral cortex corresponding with the electrode position, i.e. Frontal, Central, Parietal and Occipital. Electrodes on the left hemispheres are given odd numbers, whereas electrodes on the right hemisphere are given even numbers.

The EEG consists of a recording in time of the potential differences between a certain amount of electrodes (e.g. 31 electrodes) placed over the scalp and a reference electrode. If one wants to build a generalizable BCI system, it is important that the acquisition of the signals happens in a uniform manner. Therefore a specific system of electrode placement called International

10-20 system is used, see Figure 1.4.[12] The electrodes are positioned such that the actual distances between adjacent electrodes are either 10% or 20% of the total front-back or right-left distance of the skull. The system allows the use of additional electrodes for achieving a higher spatial resolution.

### 1.2.2 EEG analysis

#### Frequency bands of the EEG

Since different frequency bands of the EEG-signals correspond with different states of the human brain, a frequency analysis of the signals is considered very important. It gives us a first indication of what state the user is in. Furthermore if abnormalities are recorded, this can be an indication of a certain pathology, such as epilepsy. [26] Generally one distinguishes five or six different frequency bands, also called brain rhythms.

The  $\delta$ -rhythm covers all frequencies up to 4 Hz. These low frequencies tend to have relatively high amplitudes up to 100  $\mu V$ . Usually it is seen in adults in a deep sleep. The  $\vartheta$ -rhythm ranges from 4 to 7 Hz. This band is mainly present in young children. It might also be measured with adults during the first stages of sleep. If so, the signals show no organization. The  $\alpha$ -rhythm ranges from 8-12 Hz. Presence of this rhythm implies closing of the eyes and general relaxation. When the eyes are open, the signal attenuates. The  $\beta$ -rhythm ranges from 12-30 Hz. The  $\beta$ -band becomes deactivated while motor tasks are performed. Low amplitude in the  $\beta$ -band is often associated with busy or anxious thinking. The highest frequency band ranges from circa 30-100 Hz and is called the  $\gamma$ -rhythm.  $\gamma$ -rhythms are thought to represent binding of different populations of neurons together into a network for the purpose of carrying out a certain cognitive or motor function.[30, 27]

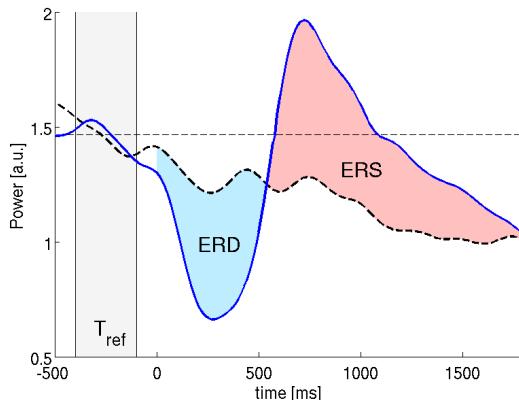
Beside these five subsequent frequency bands, researches generally also distinguish the  $\mu$ -rhythm. This rhythm corresponds to the frequency range between 8 and 13 Hz, and thus overlaps the  $\alpha$ - and partially the  $\beta$ -rhythm. An amplitude change of the signals in this frequency band might reflect synchronous active motor neurons. When analysing EEG-recordings of Motor Imagery (MI) tasks, this will thus be the brain rhythm of interest. [30, 11]

#### Sensory motor rhythms

The Sensory Motor Rhythm (SMR) is an oscillatory idle rhythm of synchronized electromagnetic brain activity recorded specifically over the sensorimotor cortex. For most individuals, the frequency of the SMR is in the range of 8 to 12 Hz. It is often mixed with a  $\beta$ - and  $\gamma$ -component, respectively at around 20 and 40 Hz. The electrodes corresponding with the sensorimotor cortex are mainly the C3 (left hemisphere) and C4 (right hemisphere) electrodes. As the name reveals, this part of the cortex is activated or deactivated by tasks concerning

movement and sensory stimuli. This is described as Event Related Desynchronization (ERD) for a decrease of oscillatory activity and Event Related Synchronization (ERS) for an increase of oscillatory activity. However these effects are very small with respect to the ongoing spontaneous EEG activity. So in order for these effects to be measurable, it is important that the area where they take place is large enough or very close to the cortex. The areas corresponding with left and right hand MI fulfils these conditions. Therefore left and right hand MI is interesting to use as mental tasks for the input of a BCI.

The effect of imagining movement compared to actually performing a movement is very similar for the brain activity, i.e. they provoke a similar ERD and ERS pattern in the same brain area. This is a very useful aspect for the development of BCIs for Locked-In Syndrome (LIS)-patients. However one has to note that if one mentions MI-tasks, this is not just visualizing movements. MI implies imagining the movement kinesthetically, meaning to feel and experience the movement. Both motor imagery or the actual execution of the movement will lead to a decrease in the rhythm. This means that during an idle state the signal is actually bigger in amplitude than during the movement itself, see Figure 1.5.[16] The ERD corresponding with initiation of movement is most prominent over the contralateral sensorimotor areas and extends bilaterally with movement initiation. [1]



**Figure 1.5:** ERD and ERS

### 1.2.3 Artefacts

The electrodecap used during EEG-recordings is designed to record the electrical activity of the underlying scalp. However it may also record activity that is not of cerebral origin or artefacts. The sources of these artefacts can lay outside (extraphysiological artefacts) or inside (physiological artefacts) the body.

A common example of an extraphysiological artefact is the influence of the electricity grid which fluctuates at 50 Hz. This interference can cause an obvious peak at 50 Hz and is thus easy to filter out, if not already done so by a notch filter.

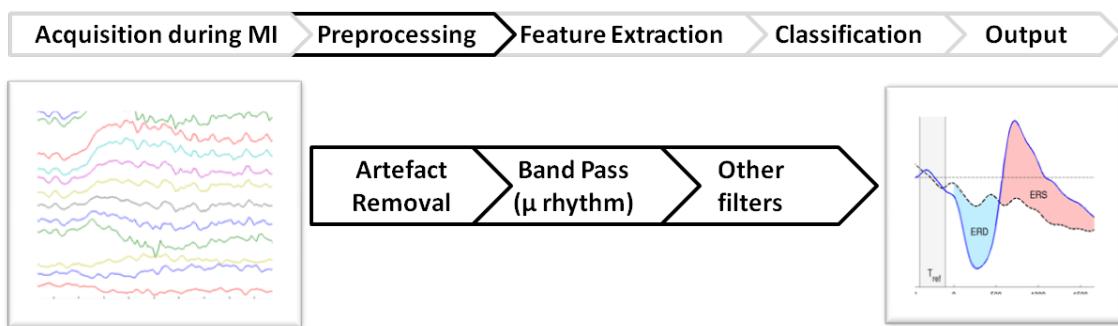
The physiological artefacts might pose a bigger problem, since identifying these artefacts has to be based on duration, morphology and neuron activity. The most frequent artefacts are artefacts due to muscle movement or eye movements. The latter artefacts are usually detected by the electrodes near the eye. They are the result of the polarization of the eyeball. The cornea corresponds with a positive pole oriented anteriorly, the retina acts as a negative pole oriented posteriorly. When the globe rotates around its axis, it generates a large amplitude alternate current field, causing serious disturbances in the EEG.

It is clear that these artefacts first have to be eliminated before any further processing can be done. In most cases a simple bandpass filter will be sufficient to cancel out the biggest influences. If not, one can just reject the trials contaminated by artefacts. Or one can use more sophisticated artefact removal techniques like Independent Component Analysis (ICA). [1]

## 1.3 Signal Processing

After measuring the EEG signals, the signal first has to pass through an amplifier and will then be digitized by an A/D conversion. The digitized signals are then preprocessed in order to prepare them for spatial filtering and feature extraction. The obtained feature vector is then translated by the classification algorithm into a command for an external device.

### 1.3.1 Preprocessing



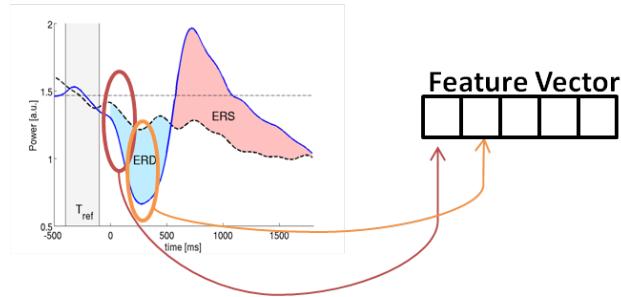
**Figure 1.6:** General preprocessing steps

The objective of the preprocessing step is to increase the SNR and to filter out the useful parts of the signals, i.e. the parts that contain the ERD and ERS corresponding with left

and right hand MI. First of all, in order to increase the SNR, the signals contaminated with artefacts are rejected. Generally as a second step a high pass or band pass filtering is applied. Since the EEG-signals are quite slow, this will already eliminate a lot of the noisy signals. Since in the case of MI the information is mainly restricted to the  $\mu$  brain rhythm, a bandpass filter corresponding with this frequency band will be applied.

If wanted, further preprocessing can be performed using more specialized techniques. In some techniques the signals obtained at different electrodes will be combined. This is called spatial filtering. These techniques are based on the idea that the important information sources usually are all localized in a specific region of the brain, whereas the sources of the noisy signal are distributed all over the brain. The most known spatial filter is the Laplacian. This method subtracts the mean of the signals of the neighbouring electrodes from the signal at the electrode of interest. This method is known to be far more efficient than subtracting the mean of all electrodes, also called common averaging. In the specialized literature, one can find much more elaborate spatial filtering methods. Some of them have to be adapted to the subject, others are generalizable for all subjects. Ultimately, they need to be compared in on-line experiments that measure speed and accuracy.

### 1.3.2 Feature Extraction

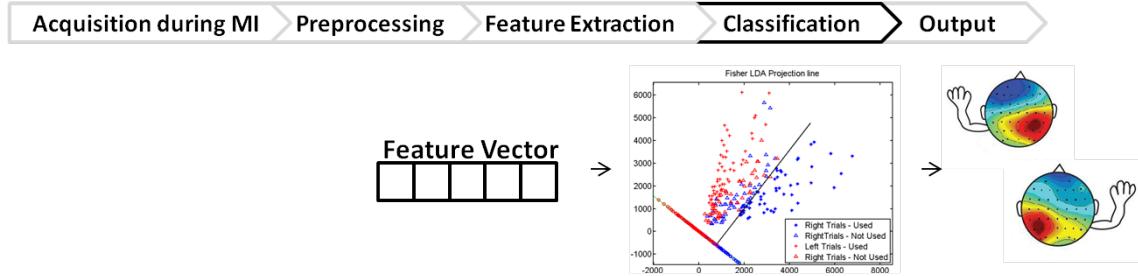


**Figure 1.7:** Feature Extraction steps

After the preprocessing step the amount of data is still too much to be useful. Therefore it is important to subtract the relevant information and discard the redundant data. This is the goal of the feature extraction step. One of the most common methods to do this is the use of Principal Component Analysis (PCA) or ICA. These systems try to convert the set of probably correlated features in a set of uncorrelated features. Decorrelating the features decreases the chance of being redundant. Other methods search for the features that are best separable. These techniques are typically supervised methods. This means that one needs a

training dataset of which one already knows to which classes (e.g. left or right) the features correspond. In all cases the goal is to reduce the dimensionality of the dataset without loosing too much information. This is done by selecting the most discriminating features. We will thus end up with a distinct set of features for each mental task.

### 1.3.3 Signal Classification



**Figure 1.8:** Signal Classification steps

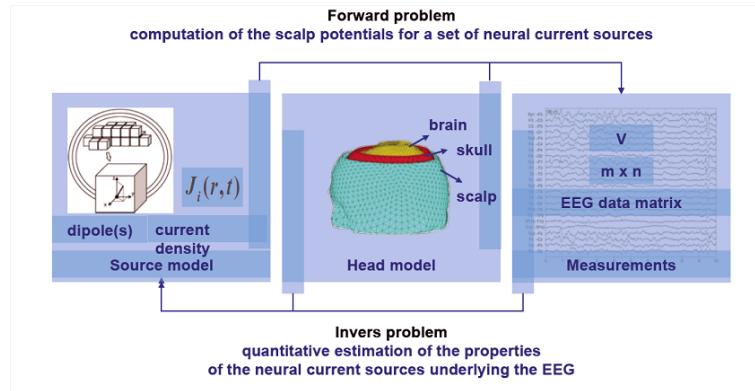
The features extracted in the previous stage are the input for a classifier. The goal of the classification step is to determine the mental state of an individual. Based on that classification a command can be given to an external device. So the classification algorithm takes the abstract feature vector that reflects specific aspects of the current state of the user EEG and transforms that vector into an application-dependent device command. In certain cases the classification can simply be done by comparing the signal resultant from the preprocessing step to a threshold. Other possibilities are the use of linear classifiers such as Linear Discriminant Analysis (LDA) or Fisher LDA classifiers. Another very popular method is to use neural network methods. These are more complex non-linear techniques. The most common examples are Support Vector Machines (SVMs) and Hidden Markov Models (HMMs). Furthermore one can choose between an adaptive or a non-adaptive classifier. In the first case machine learning techniques will be applied. The choice of classifier will be dependent on the needs of the application. But in all cases the goal is to maximize performance and practicability. ??

## 1.4 Source localization

### 1.4.1 Introduction

Before we go any further we want to introduce the reader to the concept of source localization. The goal of source localization is to locate the active sources in the brain based on the measured EEG potentials. The process of source localization typically consists of two parts, namely the forward model and the reverse model. The reverse model is the model we're

interested in, since this designs the filter we can apply on the measured potential to estimate the source localization. This filter will be applied as the first step in feature extraction. However, the reverse model can't be determined without the forward model. The forward model estimates the EEG recording assuming a certain dipole distribution and activity. The solution to the reverse model is then found by calculating the spatial filter which minimizes the difference between the forward estimations of the signal and the actual measured potentials, see Figure 1.9.



**Figure 1.9:** Relation between forward and inverse problem of EEG source localization

#### 1.4.2 Forward model

To be able to make an estimation of how the potential at the measuring electrodes will look like, one has to take into account different factors:

- First of all one has to make an assumption on the dipole model. this model shows how many dipoles one assumes and the distribution of these dipoles.
- A second factor is the head model. This takes into account the shape and the conductivity of the different layers. this is important since this determines how the signal generated by a dipole will be spread and attenuated from inside the skull towards the scalp.
- And finally one also needs the electrode distribution to know where the potentials are measured.

These three factors are all taken into account in the lead field matrix, which is used during the forward calculations as followed:  $\mathbf{V}_{\text{est}} = \mathbf{L}(\mathbf{r}) \cdot \mathbf{d}$ ,  $\mathbf{V}_{\text{est}}$  being the estimated potentials,  $\mathbf{L}(\mathbf{r})$  being the lead field matrix at position  $\mathbf{r}$  and  $\mathbf{d}$  being a dipole determined by its position and moment.

### 1.4.3 Inverse model

So the objective of the inverse model is to find a filter which can determine the active sources in the brain based on the measured potentials. Since the amount of possible dipole distributions is much higher than the amount of electrodes used, for none of the admissible output voltages a unique solution can be found. Therefore one needs to make some constraints with respect to the filter, or the source distribution, or the forward model. Applying different constraints lead to different reverse models and thus gives us different spatial filters.

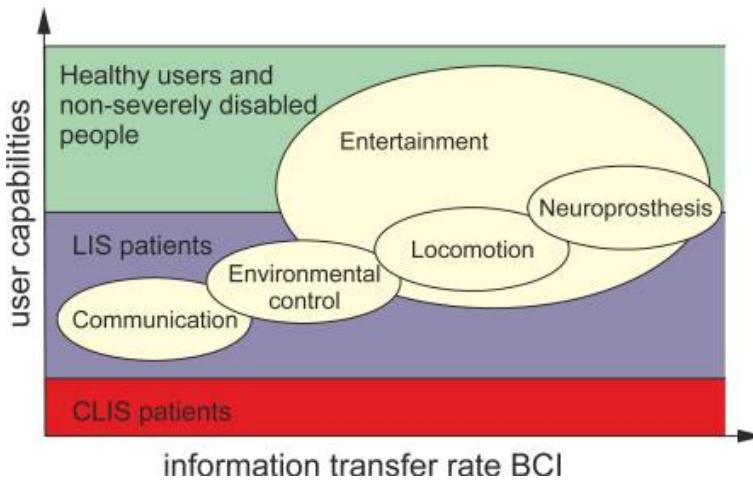
Throughout this dissertation we will use beamformer techniques. Per possible dipole position (grid point) a filter is designed. Ideally the filter passes the signal at the specified location and attenuates the signals from all other locations. The advantage of this filter is that it can easily incorporate neurophysiological information concerning the regions where we expect the interesting signals to be. Furthermore it can easily be applied in real-time. And another important feature is that one doesn't need to define the expected amount of dipoles on beforehand.

## 1.5 Applications and current state-of-the-art

In general the objective of a BCI is to detect small differences in brain signals and use these to steer an external device. In principle this external device can be anything, as so can be the input causing the change in brain signal. However the input is generally limited to some typical tasks intended for subject training. These tasks include (limited) cursor control, motor imagery, tracking a moving object or selecting a target. The results of these tasks can then be translated into more useful applications in the field of communication, environmental control or neural prosthetics. As shown in Figure 1.10 the kind of application will on the one hand depend on the severity of the locked-in state. A distinction is made between Complete Locked-In Syndrome (CLIS) and LIS patients, and healthy subjects. On the other hand it will depend on the Information Transfer Rate (ITR) of the BCI-system. This is a measurement for how often in time an accurate decision can be made.[20]

### 1.5.1 Communication

Since spoken and written language are the main forms of communication nowadays, an impairment in this field be a huge loss for a person's social life. Therefore communication is the most interesting objective of BCIs. In most cases applications addressing communication problems include a virtual keyboard on a screen. Typically a letter can be selected by going through a three step procedure. The task one has to perform to guide the system might differ among the various types of systems. For example, there are keyboards driven by voluntary Slow Cortical Potentials (SCPs) or by visual Event Related Potentials (ERPs). The input



**Figure 1.10:** Relationship between BCI application areas, BCI information transfer rates and user capabilities

can also be delivered by performing three different motor imagery tasks, e.g. left, right and both feet MI.

Another communication pathway that is being more and more addressed by BCIs, due to its importance in everyday life, are internet browsers. The selection methods for links are similar to the ones used for the virtual keyboard. Of course the possible sites that can be browsed are rather limited, since the selection process would otherwise be too extensive.

### 1.5.2 Environmental control and virtual worlds

Environmental control applications focus on the control of domestic devices such as TV, lights or temperature regulation. Giving disabled people the opportunity to perform these relatively simple tasks on their own can give them the opportunity to be at least partly independent. The methods applied to control these tasks are again very similar to the methods used for communication applications. For example, a cursor and a keyboard will be available. Depending on the severity of the immobility of the person, also an eye or head tracking system can be incorporated. Again excessive training will precede the use of these applications. Virtual reality can offer a safe but realistic environment for training and tuning neurally controlled interfaces to real-world devices, see Figure 1.11 [4].

### 1.5.3 Neural prosthetics

The use of a BCI as a neural pathway between the brain and the limbs is one of the most to the imagination appealing application. However there is still a long way to go towards a useful and reliable signal for controlling a hand neuroprosthesis. Nevertheless there are a few success stories of patients who learned to handle an electrical driven hand orthosis. Since motor corti-



**Figure 1.11:** BCI and virtual reality

cal activation in Spinal Cord Injury (SCI) patients seems to change with BCI training, success is only obtained after months of intensive exercising. For example, after a training period of some months a patient was able to close the hand orthosis by both-feet-movement-imagery and to open the orthosis by right-hand-movement-imagery nearly error-free with classification accuracy close to 100%. This performance was reached mainly by the ability to induce voluntarily specific beta oscillations close to the foot area, see Figure 1.12. [22]

Good results are also obtained when the BCI is combined with robot-assisted physical therapy. After a stroke patients are often not completely paralysed but they do suffer from moderate motor impairment. In those cases brain signal based reinforcement of the patient's intent to move his/her own arm using a robot arm may have a positive effect in recovering his mobility, as it is likely to result in increased cortical plasticity. [13]



**Figure 1.12:** BCI as interface for hand orthosis

# **Chapter 2**

## **Overview of this Master's Thesis**

### **2.1 The goal of BCI Research**

As stated in the previous chapter a BCI is a system which tries to translate the brain activity of a user to an input or command for an external device. It captures the signals generated by the brain during a task, and based on these signals it tries to decide what the intention of the user exactly is. The possible applications are numerous and are mainly interesting for people suffering from LIS. The main difference between BCIs and other assisting technologies is the fact that there is a direct communication channel between the natural brain output (under the form of small electric or magnetic potentials) and the device.

In each step of the translation process there is a broad range of possibilities to achieve a certain output. The research already performed with respect to BCIs is plentiful, however with quite varying results. The ideal BCI still hasn't been found, and probably doesn't exist. Requirements for a commercial BCI are among others: safe, long lasting, not too expensive, reliable with an acceptable accuracy rate (in general one accepts a classification error of maximal 30%), workable in real time and a fast training protocol. Some of these requirements are already addressed pretty well, but for others there is still a long way to go.

### **2.2 Main problems encountered in BCI research**

After reading the previous chapter one might have the impression that the process from input to output is quite straightforward. Since we know where to find the signal, and in which frequency band to find it, it shouldn't be too difficult to extract it from the complete EEG. However when dealing with BCIs we are dealing with people. And of course every person is different, both in anatomy as in physiology. Even more, we're dealing with the human brain, the most complicated and unresolved part of the human body. This increases the complexity of the problem dramatically and leads us directly to the principle problems encountered during

BCI research. The main challenge of BCI research is to improve the robustness of the current systems. We're going to discuss the most frequent problems encountered in BCI research in the following sections. In Section 2.3 we describe in detail how we want to address these problems in this thesis.

- **Inter-subject differences:** One could build a general BCI system, but the odds that this will perform as good as necessary for each person are quite low. So the first problem we have to deal with is inter subject variability. First of all, everybody has its own anatomy of the skull, and the differences between persons can be quite pronounced. Secondly the physiology might differ between persons. And lastly, as with everything in everyday life, there are people that just know better how to handle the device than others. However in general performance increases with increasing experience. So one of the main challenges in BCI research is to develop a system that addresses or can handle these inter-subject differences.
- **Intra-subject differences:** Even for the same person, there might be changes in between sessions. This can be a change of active region, a change of magnitude of the ERD-ERS signal (maybe because of a lower or higher attention rate), or a difference in the ERD - ERS pattern. It isn't evident to take these intra-subject into account during signal processing. So these intra-subject differences pose the second challenge in the development of a BCI system.
- **Real-time processing:** Another problem is that in order to have useful applications, everything has to be processed in real time. This asks first of all for strong computation power. This also requires real time detection of artefacts. And one has to be able to classify the signal based on a very limited amount of data.

## 2.3 Objectives

### 2.3.1 Understanding beamforming and its possibilities

We know that our main objective is to build a BCI system which can fulfil the tasks and requirements described in Section 2.1, while in the meantime it addresses the problems described in Section 2.2. We decided to test if we can achieve this goal by using a BCI system based on beamforming spatial filtering. Of course this is not just a random choice. So now the first question is: Why do we believe spatial filtering by beamforming might be an answer to the system's requirements?

We already addressed shortly the construction of beamformers in Section 1.4.3. We explained that in order to calculate the filter, one first has to design a forward model. This forward model consists of a dipole model, a head model and an electrode distribution. The electrode

distribution will be determined by the available acquisition system. But the other parameters can be adjusted. So to incorporate inter-subject differences, one can for example adjust the head model.

Another characteristic of beamformers is that one actually has a separate filter for every possible dipole position. In the ideal case the filter only passes the signal at that position and attenuates all the signals coming from other locations, leaving us with the dipole moment at that position. The advantage of this is that one doesn't have to make assumptions on the dipole distribution beforehand, or in other words the dipole distribution is another parameter we can play with to address inter-subject differences.

Furthermore once the filter is calculated based on a training dataset, it won't change in time. This makes it easy to use the filter in real-time applications.

It is now our task to test if these characteristics indeed can be translated into a good working BCI system. We will do this in four steps

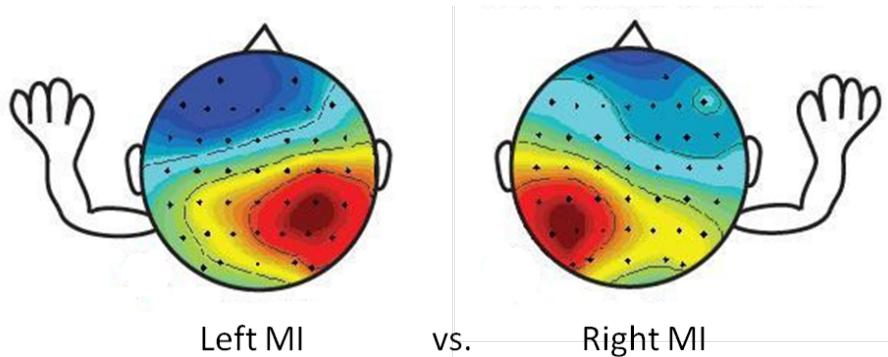
- First we're going to design a quite straightforward BCI system based on beamforming. This BCI system has to become the backbone of our further research.
- Secondly, we're going to fine-tune the parameters of this system. By fine-tuning these parameters, we hope to address the problem of intra-subject differences and make the system more robust.
- After the design of the system, we will test if beamforming technology indeed is a valid method for real-time applications.
- And finally we will investigate what the incorporation of personalized anatomical and neurophysiological data does with the .

### 2.3.2 Design of an accurate BCI system

So the first step is the design a working BCI system. The idea is to have a BCI system that can distinguish between two states with an adequate accuracy. The states we'll work with during this thesis are the mental performance of left hand MI and right hand MI. As illustrated in the previous chapter, in order to have a working BCI system one needs to provide data, a preprocessing and feature extraction step and a final classification. We want this BCI system to be robust, so that it can be the backbone of the further of the research.

#### Which data are we going to use?

The data used is the data provided by the BCI-IV competition. This competition has put multiple datasets on-line [6]. 7 different datasets, of which some are artificial, recorded during



**Figure 2.1:** Left and right hand motor imagery

the correct mental states are made available. Moreover these datasets contain both calibration as evaluation data. They will be described into more detail in Chapter 4.

### How are we going to process the data?

The processing and feature extraction step are partly based on code provided by the FieldTrip toolbox. FieldTrip is a Matlab software toolbox for EEG analysis which can be downloaded on-line [21].

The most important component during feature extraction is the spatial filtering. In this thesis we will apply Linearly Constrained Minimum Variance (LCMV) beamforming. The specific implementation of the beamforming procedure is discussed in Chapter 3. A more detailed overview of both the preprocessing as the feature extraction can be found in Chapter 5. Finally the classification and evaluation step has been implemented in Matlab ourselves and is based on quite straightforward linear classification algorithms. A more complete implementation is reported in chapter 5. All the results will be reported in chapter 6.

#### 2.3.3 Fine-tuning of parameters

In a second phase we want to fine tune the parameters used by the process. In each step from data to classification a lot of parameters are included. Each one will have a certain influence on the final result and can thus be optimized. In order to be able to draw appropriate conclusions, one has to make sure to compare results obtained by changing only one parameter at a time. Fine tuning the parameters will end up in a more robust system and it will help us to address the problems cited in the previous section about inter subject and inter trial differences.

#### Which parameters are we going to fine-tune?

In the preprocessing step we will consider whether or not to apply baseline correction and complete artefact rejection. In the feature extraction step, fine-tuning parameters will mainly

concern the fine-tuning of the parameters of the filter. These are the regularization parameter, the amount of trial data, the time window,... Also part of the feature extraction step is the determination of which and how many dipole positions to use. And finally for the classification one has to determine the training percentage for classification and the size of the signals one wants to classify.

#### **2.3.4 Real-time application**

To test the ability of real-time classification, we will again use the data from the BCI Competition IV, but this time the evaluation dataset. This is an asynchronous dataset and therefore a good representation of real-time data.

#### **2.3.5 Influence of the head model**

In the last part of this thesis we want to explore two more topics. First we want to see if using an individualised head model outperforms the use of a simple spherical model or a more advanced though not person-specific head model. Secondly we want to see what happens if we not only adapt the head model to the person, but also use personal neurophysiological information about the functional areas related with the tasks. This additional information is recorded during a fMRI experiment. On the one hand an MRI scan of the head is made. This makes it possible to build a personalized head model. On the other hand one can distinguish the regions in the brain activated during finger tapping. The coordinates of these areas can then be translated to a Region Of Interest (ROI) for the processing. The results can again be found in Chapter 6.

# Chapter 3

## Beamforming

### 3.1 Introduction

Systems designed to receive signals which can propagate spatially often end up not only recording the desired signal but also the presence of interfered signals. If both the desired signal and the interfered ones exist in the same temporal frequency band, it is impossible to distinguish the different signals solely by applying temporal filtering. However the desired signal often originates from another location than the interfered signal. This spatial separation can then be exploited by a spatial filter in order to retrack the aspired signal any way.

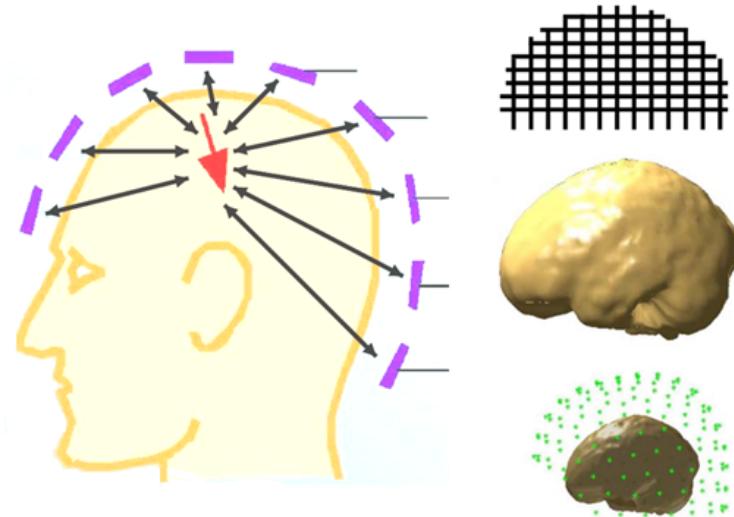
EEG signals are a typical example in which spatial filtering is necessary. The spatial filter used during feature extraction therefore tends to be a critical choice for acquiring an accurate BCI system. The possible spatial filters are plentiful. In this thesis we decided to focus on beamforming techniques.

This spatial filtering technique has been applied in radar technology for decades. The term is derived from the fact that originally the filter was designed to form pencil beams in order to receive a signal radiating from a specific location and attenuate signals from other locations. Nowadays this technique is being applied in very diverse fields such as seismology, wireless communications and biomedical signal processing. In the latter the power at the output of the spatial filter is the estimate of the neural power originating within the spatial passband of the filter.

Applying beamforming techniques makes it possible to use this neurophysiological information about both the location and the direction of the signal during filtering of the data. So we can state that the beamformer doesn't try to interpret the complete data set, but it only wants to bring out the signal coming from a specific position in the brain, also called the ROI. One of the advantages of beamforming is that no prior assumptions have to be made with

respect to the number or distribution of the active sources. In theory, one can identify both signal or distributed sources, provided a high SNR.

### 3.2 Forward Model



**Figure 3.1:** Elements determining lead field matrix for forward problem

The forward model estimates the EEG recording assuming a certain dipole distribution and activity. As we have seen in 3.7 an accurate forward solution is a required prerequisite for beamforming, since it is a determining factor for the final filter. As shown in figure 3.1, the forward model is influenced by 3 main factors.

- The **grid** determines the positions that are being scanned. Choosing a smaller grid size might influence the accuracy of the solution. There is a trade-off between the ability to resolve sources that are close together and the possibility of introducing spurious sources. Therefore it is important to use similar grids when comparing different head models.
- The **head model** determines the shape and the conduction of the volume containing the sources. The head consists of different layers, among which the Cerebrospinal Fluid (CSF), the skull and the skin. Each one has a different conductivity, e.g. the skull has a conductivity that is about 80 times higher than the skin or CSF. Whether or not taking this fact into account, will have an unmistakable influence on the final accuracy of the system. The shape of the head will likewise have a distinct effect on the measurements. On the one hand the signals will propagate differently, on the other hand there might be a shift in the region of interest.

- And lastly the **electrode distribution** determines where the resulting potentials are measured. Since the electrodes are always placed according to the international 10-20 system, it will mainly be the amount of electrodes used that will have an impact on the obvious recordings and on the final resolution of the system, determined by the passband of the filter.

All these items will thus define the forward model. Mathematically these considerations are expressed in what we call the lead field matrix. Each grid point has one lead field matrix. This matrix is the equivalent of the filter in the inverse model.

### 3.3 The inverse model: Filter Design

#### 3.3.1 Introduction

In the forward problem, one assumes a known mathematical and physical model of the brain activity. Based on these models one tries to determine what the corresponding EEG-recording would be. The forward model consists of a source model and a head model. In this thesis we'll assume a dipole source model, in which the dipoles are uniformly distributed over the head model in x-, y- and z-direction. If the dipole orientation is known on beforehand, this can also be added to the model. Every dipole corresponds to a possible electrical source en is fully determined by its position and its moment. The head model, for its part, takes into account the volume conduction of the sources throughout the head. Attenuation and smearing typically takes places at the level of the interfaces between layers of the head with a different conductivity. These thus will play an important role in the final signal distribution. So the more accurate these models are, the better the estimated results will correspond to the reality.

However for real-life applications the measured EEG data is the given, and it is the dipole model that has to be determined. This is what one calls the inverse problem. The inverse problem is, though, an ill-posed problem. Since the amount of possible dipole locations is much higher than the amount of electrodes used, for none of the admissible output voltages a unique solution can be found. Moreover solutions are typically unstable: a small change in the noisy data, can have a considerable high influence on the final result. This is why it is so necessary to impose some well-considered constraints on the forward model. Essentially one tries to find the source distribution that results in the biggest overlap between the measured data and the theoretical signal.

#### 3.3.2 Definitions and notations

In the following sections we are going to describe the derivation of the beamformer mathematically. But in order to prevent misunderstandings we're going to start by defining the

notations and elements used. Since the beamformer is implemented in FieldTrip, we will use the derivation as proposed by Van Veen et al. [29].

## Notation

For notation, lower- and upper-case boldface symbols represent respectively vector and matrix quantities. Superscript  $T$  denotes matrix transpose and superscript  $-1$  matrix inverse. The trace of the matrix  $\mathbf{A}$  is written as  $\text{tr}\{\mathbf{A}\}$ .

## Definitions

$N$  = amount of electrode sites

$L$  = amount of active dipoles

$\mathbf{x} = N \times 1$  vector, composed of the potentials measured at the  $N$  electrode sites at a given instant in time associated with a single dipole source

$\mathbf{q} = 3 \times 1$  vector, source location

$$\mathbf{x} = \mathbf{H}(\mathbf{q})\mathbf{m}(\mathbf{q})$$

with

$\mathbf{m}(\mathbf{q}) = 3 \times 1$  vector, moment of source at location  $\mathbf{q}$  at the instant in time  $x$  is measured

$\mathbf{H}(\mathbf{q}) = N \times 3$  matrix, transfer or leadfield matrix of which the columns represent the solutions to the forward problem. The first column of  $\mathbf{H}(\mathbf{q})$  is the potential at the electrodes due to a dipole source at location  $\mathbf{q}$  having unity moment in the  $x$  direction and zero moment in  $y$  and  $z$  directions. Similarly the second and third columns represent the potential due to sources with unity moment in  $y$  and  $z$  directions respectively.

Only the elements of  $\mathbf{H}(\mathbf{q})$  depend on the particular sensing modality. In a physical sense,  $\mathbf{H}(\mathbf{q})$  represents the material and geometrical properties of the medium in which the sources are submerged.

The medium is linear so the potential at the scalp is the superposition of the potentials from many active neurons. Suppose  $\mathbf{x}$  is composed of the potentials due to  $L$  active dipole sources at locations  $\mathbf{q}_i, i = 1, 2, \dots, L$  and a measurement noise  $\mathbf{n}$ . Then the relationship between dipole models and the surface recordings can be calculated as in 3.1.

$$\mathbf{x} = \sum_{i=1}^L \mathbf{H}(\mathbf{q}_i)\mathbf{m}(\mathbf{q}_i) + \mathbf{n} \quad (3.1)$$

Note that  $\mathbf{x}$  does not contain any temporal information since it is obtained by sampling all electrodes at a single time instant. It represents the spatial distribution of potential at the

measurement sites at the sampling time.

The electrical activity of an individual neuron is assumed to be a random process influenced by external inputs to the neuron. Hence, we model the dipole moment as a random quantity and describe its behaviour in terms of mean and covariance.

$$\bar{\mathbf{m}}(\mathbf{q}_i) = \mathbf{E}\{\mathbf{m}(\mathbf{q}_i)\} \quad (3.2)$$

$$\mathbf{C}(\mathbf{q}_i) = \mathbf{E}\{[\mathbf{m}(\mathbf{q}_i) - \bar{\mathbf{m}}(\mathbf{q}_i)][\mathbf{m}(\mathbf{q}_i) - \bar{\mathbf{m}}(\mathbf{q}_i)]^T\} \quad (3.3)$$

$$\mathbf{C}(\mathbf{q}_i) = \mathbf{E}\{[\mathbf{m}(\mathbf{q}_i) - \bar{\mathbf{m}}(\mathbf{q}_i)][\mathbf{m}(\mathbf{q}_i) - \bar{\mathbf{m}}(\mathbf{q}_i)]^T\} = \mathbf{0} \quad (3.4a)$$

Then

$$\begin{aligned} \bar{\mathbf{m}}(\mathbf{x}) &= \mathbf{E}\{\mathbf{x}\} \\ &= \sum_{i=1}^L \mathbf{H}(\mathbf{q}_i) \bar{\mathbf{m}}(\mathbf{q}_i) \end{aligned} \quad (3.4b)$$

$$\begin{aligned} \mathbf{C}(\mathbf{x}) &= \mathbf{E}\{[\mathbf{x} - \bar{\mathbf{m}}(\mathbf{x})][\mathbf{x} - \bar{\mathbf{m}}(\mathbf{x})]^T\} \\ &= \sum_{i=1}^L \mathbf{H}(\mathbf{q}_i) \mathbf{C}(\mathbf{q}_i) \mathbf{H}^T(\mathbf{q}_i) + \mathbf{Q} \end{aligned} \quad (3.4c)$$

### 3.3.3 Mathematical description

The problem of designing a filter to localise EEG sources based on recorded data, doesn't have a unique solution. Therefore one will have to apply some constraints during the discourse of the mathematical solution. In the following we will determine these constraints. During this mathematical discourse we will apply the notations and definitions from the previous section consistently in both the following section, as in the attached appendix 1.

Active neurons are typically characterized by having a high variance of the power of the dipole moment. So in order to distinguish the active sources from the surrounding neural tissue, we need a filter that translates the EEG-recording in an expression for the dipole moment  $\mathbf{m}(\mathbf{q}_0)$  at each location within the tissue. This is expressed in 3.5.

$$\mathbf{y} = \mathbf{W}^T(\mathbf{q}_0) \mathbf{x} \quad (3.5)$$

And thus  $\mathbf{W}(\mathbf{q}_0)$  is the  $N \times 3$  filter for location  $\mathbf{q}$ , in which each column corresponds with the filter for one of the main directions of the dipole moment. In the ideal case  $\mathbf{y} = \mathbf{m}(\mathbf{q}_0)$  applies. In order for this to be true, based on 3.1, 3.6 has to be fulfilled.

$$\mathbf{W}^T(\mathbf{q}_0)\mathbf{H}(\mathbf{q}_0) = \begin{cases} \mathbf{I} & q = q_0 \\ 0 & q \neq q_0, q \in \Omega \end{cases} \quad (3.6)$$

$$\mathbf{W}^T(\mathbf{q}_0)\mathbf{H}(\mathbf{q}_0) = \mathbf{I} \quad (3.7)$$

In the case of an LCMV filter, the filter coefficients  $\mathbf{W}(\mathbf{q}_0)$  are determined such that 3.7 is satisfied for the location at interest, while at the same time the variance of the filter output is minimized 3.8.

$$\min_{\mathbf{W}(\mathbf{q}_0)} \text{tr}[\mathbf{W}^T(\mathbf{q}_0)\mathbf{C}(\mathbf{x})\mathbf{W}(\mathbf{q}_0)] \quad \text{subject to} \quad \mathbf{W}^T(\mathbf{q}_0)\mathbf{H}(\mathbf{q}_0) = \mathbf{I} \quad (3.8)$$

The constraints insure that the signals of interest are passed by the filter. Minimization of variance optimally allocates the stop band response of the filter to minimize the contribution to the filter output due to signals in the stop band.

An explicit expression for  $\mathbf{W}(\mathbf{q}_0)$  in function of  $\mathbf{x}$  and  $\mathbf{H}(\mathbf{q}_0)$  can be found by applying Lagrange multipliers. For an extensive mathematical implementation I refer to appendix 1 or ???. The final solution for  $\mathbf{W}(\mathbf{q}_0)$  is then given by 3.9.

$$\mathbf{W}(\mathbf{q}_0) = [\mathbf{H}^T(\mathbf{q}_0)\mathbf{C}^{-1}(\mathbf{x})\mathbf{H}(\mathbf{q}_0)]^{-1}\mathbf{H}^T(\mathbf{q}_0)\mathbf{C}^{-1}(\mathbf{x}) \quad (3.9)$$

### 3.4 Some remarks

#### Correlated sources

During calculation of the beamformer filter one assumes the sources outside the ROI to be uncorrelated with respect to the sources inside the ROI. However this assumption is not necessarily fulfilled. Definitely not for sources not inside, but still near the ROI. If there is indeed a fully or even partly correlated source included, the correlated sources will cancel each other out. And thus the source inside the ROI won't be reconstructed. However this is a problem that will occur more often in evoked potentials. Adapted algorithms to take into account possible correlations already exist, but are not implemented here. [2, 5]

#### Influence of noise

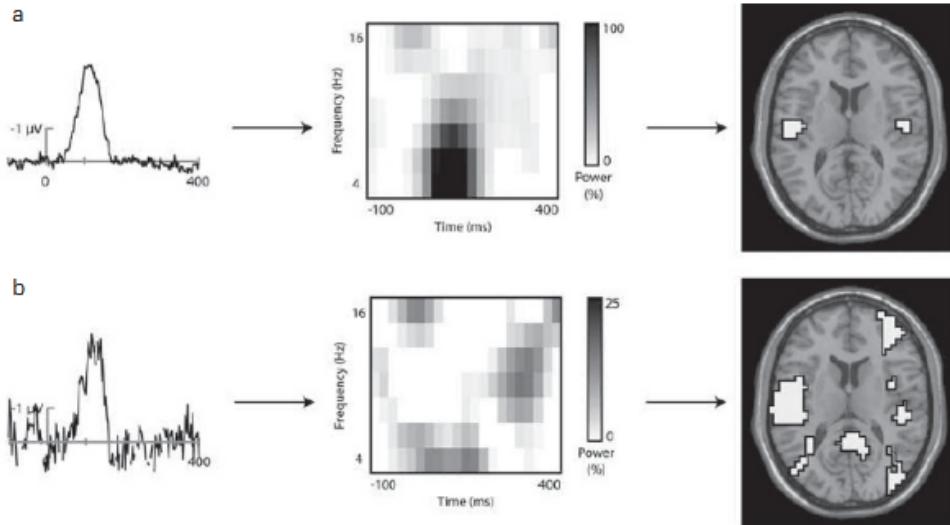
Just like for all other signal processing methods, the performance of beamforming techniques also strongly depends on the SNR of the signals. One of the most important steps in filter computation is the estimation of the covariance matrix of the signal  $\mathbf{x}$ . However  $\mathbf{x}$  is not only the wanted signal, but is a combination of the wanted signal and noise, i.e.  $\mathbf{x} = \mathbf{s} + \mathbf{n}$ . The lower the SNR, the more difficult it will become to fully separate different sources. The

noise blurs out the sources after reconstruction. A nice example of this is seen in Figure 3.2 [2].

An often used method to reduce the influence of the noise is to calculate the Neural Activity Index (NAI). This is done by normalizing the estimated spatial spectrum of the data by the estimated noise spatial spectrum to obtain the normalized estimate, see Equation 3.10.

$$\widehat{\text{Var}}_N(\mathbf{q}_0) = \frac{\text{tr}[\mathbf{H}^T(\mathbf{q}_0)\mathbf{C}^{-1}(\mathbf{x})\mathbf{H}(\mathbf{q}_0)]^{-1}}{\text{tr}[\mathbf{H}^T(\mathbf{q}_0)\mathbf{Q}^{-1}\mathbf{H}(\mathbf{q}_0)]^{-1}} \quad (3.10)$$

If noise covariance matrix  $\mathbf{Q}$  is unknown, one can assume the noise to uncorrelated and thus say  $\mathbf{Q} = \mathbf{I}$ . If one uses the NAI, one can localise sources by looking for the maxima in the NAI.



**Figure 3.2:** Effect of noise on beamformer source reconstruction. (a) The average ERP, timefrequency plot, and beamformer output for a pair of bilateral sources in auditory cortex for data with a high signal-to-noise ratio. (b) The same signal with increased background noise. Information about the activity of interest is difficult to extract from the background noise, resulting in lower power in the timefrequency domain and spurious sources in the beamformer output

# Chapter 4

## Input: Material and Experimental Set-up

### 4.1 Datasets

#### 4.1.1 BCI competition IV, dataset 1

##### **General Information**

The data used during this thesis are a part of the datasets made available for the BCI Competition IV [6]. The goal of the BCI Competition IV is to validate signal processing and classification methods for BCIs. Therefore they provided different kinds of datasets. We will use the first dataset. This one contains data acquired during right and left hand motor imagery. A distinction is made between a calibration and a evaluating dataset. The calibration dataset is mainly used for training a certain algorithm or the complete system, whereas the evaluation dataset is solely intended for evaluating the performance of a certain algorithm. The difference between both will be explained later on.

We have access to datasets of seven different subjects. Of those seven subjects, 5 subjects performed left and right hand motor imagery. The other two subjects performed left hand and both foot motor imagery. We will only use the dataset of left and right hand motor imagery. Furthermore, four datasets are recorded from real persons, and three datasets are generated artificially. The idea is to have a mean for generating artificial EEG signals with specified properties which are so realistic that they can be used to evaluate and compare analysis techniques. For artificial data to be useful, it is important that techniques perform equally good as for the real data. The 'real' data sets were recorded from healthy subjects. In the whole session motor imagery was performed without feedback.

### Technical aspects

The recordings were made using BrainAmp MR plus amplifiers and a Ag/AgCl electrode cap. Signals from 59 electrode positions were measured. The electrodes were most densely distributed over sensorimotor areas. Signals were bandpass filtered between 0.05 and 200 Hz and then digitized at 1000 Hz with 16 bit (0.1 uV) accuracy. Also a downsampled version of the data has been provided. The downsampling happens with a factor 10, so the new sampling frequency is 100 Hz. However, since the available computational power is high enough, the latter won't be used in this thesis.

### BCI competition IV, dataset 1, calibration data : protocol

We are now going to describe the protocol used during data acquisition. The protocol used to generate the datasets was for all considered subjects the same. The subjects were placed in front of a computer screen. Arrows pointing left or right were presented as visual cues on the computer screen. Cues were displayed for a period of 4s during which the subject was instructed to perform the cued motor imagery task. These periods were interleaved with 2s of blank screen and 2s with a fixation cross shown in the centre of the screen. The fixation cross was superimposed on the cues, i.e. it was shown for 6s, see Figure 4.1.

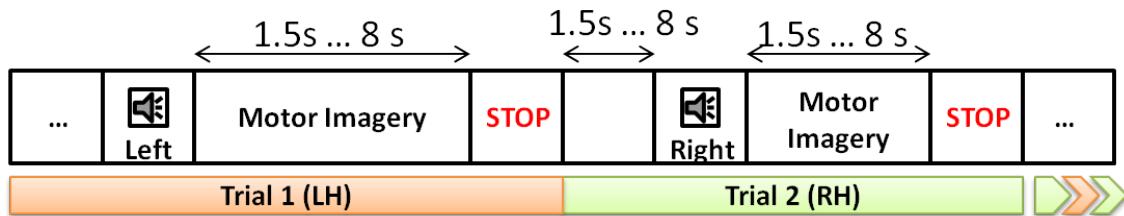


**Figure 4.1:** Acquisition protocol for BCI competition IV calibration datasets

### BCI competition IV, dataset 1, evaluation data a-g : protocol

In a second phase, evaluation data was recorded. Most demonstrations of algorithms on BCI data are evaluating classification of complete EEG trials, i.e. windowed EEG signals with a fixed length, where each trial corresponds to a specific mental state. But in more realistic BCI applications one is faced with the problem that the classifier has to be applied continuously to the incoming EEG without having cues of when the subject is switching her/his intention. This is what we call an asynchronous BCI. The evaluation dataset now poses the challenge of applying a classifier to continuous EEG, so no cue information is added. Another issue that is being addressed in this data set is that the evaluation data contains periods in which the user has no control intention. During those intervals the classifier is supposed to return 0 (no affiliation to one of the target classes).

The protocol used for evaluating the submissions to the competitions is slightly different. Here, the motor imagery tasks were cued by soft acoustic stimuli (words left or right) for periods of varying length between 1.5 and 8 seconds. The end of the motor imagery period was indicated by the word 'stop'. Intermittent periods had also a varying duration of 1.5 to 8s, see Figure 4.2. Note that in the evaluation data, there are not necessarily equally many trials from each condition.



**Figure 4.2:** Acquisition protocol for BCI competition IV evaluation datasets

#### 4.1.2 Data recorded at Ghent University

The second dataset we are going to use is a dataset that was recorded within the context of another thesis. The data was measured with one of the authors. Reusing this dataset has multiple advantages. First of all it is easy because it is already available. Secondly the mental states performed are nearly the same as for the BCI Competition data. And lastly we also have data of the same subject regarding anatomy and neurophysiology at our disposal. [28]

## Technical aspects

The acquisition equipment contains an electrode cap, an amplifier with power pack and a USB-adapter. This arrangement was designed by Brainproducts GmBH.

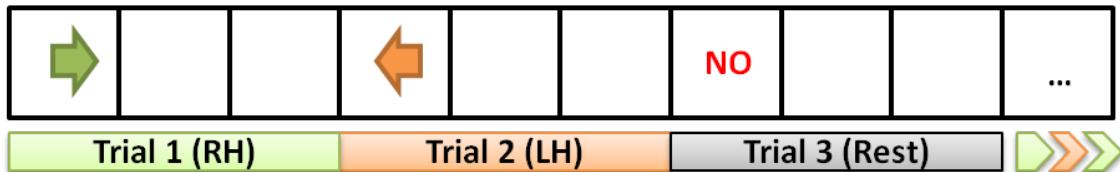
- **Electrode cap:** The brain signals are captured by the electrode cap. The electrode cap is elastic and can be pulled over the head. It can accommodate up to 64 electrodes. However in this set-up only 32+1 were used. The last electrode was used to register the Electrocardiography (ECG) and is thus not positioned on the cap itself. The other electrodes are positioned according the international 10-20 system. The connection between the scalp and the electrodes is filled with an abrasive electrolyte gel, reducing the resistance of each electrode to less than  $5\text{ k}\Omega$ .
  - **Amplifier:** The signals are then sent to a Brainamp MR plus amplifier. This is an MR-compatible amplifier with an analog-to-digital conversion and a frequency range up to DC-1000 Hz. The amplifier is powered by a power pack to reduce the noise at 50 Hz induced by the electricity grid.

- USB-adapter: Next the digital signal is sent through an optical fibre from the amplifier to a USB-adapter. The latter converts the signal to an electrical signal that can be interpreted by the computer.

Data acquisition is implemented in the BCI2000 platform. BCI2000 is a general-purpose system for BCI research and delivers software tools for diverse areas of real-time biosignal processing.

#### 4.1.3 Data Protocol

The protocol is similar to the one used for BCI competition. However there is one difference. In this dataset three different states are distinguished instead of two states: an explicit rest state is added to the common left hand MI and right hand MI. In practice each trial consists of 3 slots of 2 seconds. During the first slot the instruction for the upcoming slots is shown. This slot can also be used for relaxation and mental preparation. After the first two seconds, the instruction disappears from the screen. Now the user has to perform the instruction during 4 seconds, preferably without introducing artefacts such as eye blinks. The possible instructions are an arrow to the left for left hand MI, an arrow to the right for right hand MI or the word 'NO' for the rest state, see Figure 4.3. Every state shows up 50 times, in a random sequence. [28]



**Figure 4.3:** Acquisition protocol for own recorded dataset

## 4.2 Head Models

As stated in the previous chapter the choice of head model can be of big influence to the final performance of a BCI system. Therefore we want to investigate what the consequence is of using a less or using a more detailed head model. Each one has its advantages and disadvantages. Since it is not possible to measure the volume conductivities of each person's brain, skull or scalp, it is interesting to work with estimates of both the conductivities and the shape of the different anatomical layers in the head. Moreover the simpler the model, the more computationally efficient the forward problem becomes. And one also has to take into account that not every research institute has the money or the equipment to take the scans necessary for a personalized approach. However we also know the more accurate forward solutions become, the less the probable inaccuracy in the source localization. So a personalized

approach can certainly add a certain value.

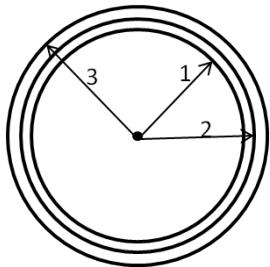
In the following we will present the details of the head models we'll use throughout this dissertation. The head models of choice are a 3-shell spherical head model, a realistic but generic head model and a personalized head model. When defining the conductivity for a certain layer, we assume isotropic layers.

#### 4.2.1 Standardized head models: 1-shell 3-shell spherical head model

The spherical head model is the most easy head model available. In this case we use a 1-shell 3-shell spherical head model. The dissertation of the 1-shell spherical head model is completely analogue to the dissertation of a 3-shell spherical head model. One just has to remove the 2 inner shells.

Each shell represents one of the main compartments of the head, namely the brain itself, the skull and the scalp. By using a 3-shell head model we keep the calculations quite simple, without ignoring the transitions and smearing at the boundaries between different tissues. Nevertheless, anatomically a sphere is of course far off of the real shape of a human head. Anyhow a 3-shell spherical model can be a reliable reference model due to its easy resolvability. In [19] it is shown that the analytical solution and Boundary Element Method (BEM) give very similar results.

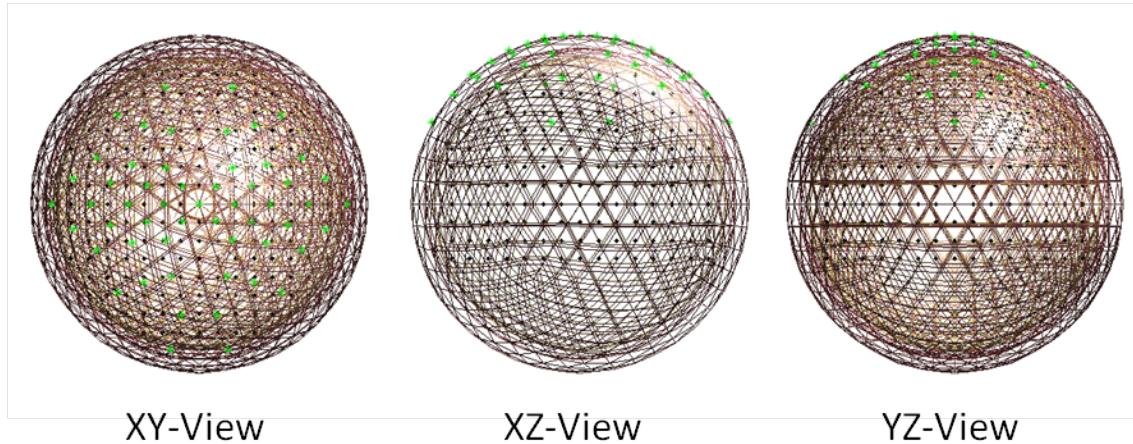
In our model the brain, skull and scalp have respectively a radius of 8, 8.5 and 9.2 cm. The corresponding conductivities are  $\sigma = 1$ ,  $\sigma = 0.0125$ ,  $\sigma = 1$ . The amount of grid points used equals 1703. These are equally divided over the space. A schematic overview of the model and the addressed values can be found in respectively Figure 4.4 and Table 4.1. How the model finally looks like is shown from different points of view in Figure 4.5.



**Figure 4.4:** Schematic view of the 3-shell spherical head model

	Name	Radius [cm]	Conduc- tivity
1	Brain	8	1
2	Skull	8.5	0.0125
3	Scalp	9.2	1

**Table 4.1:** Specifications 3-shell spherical model

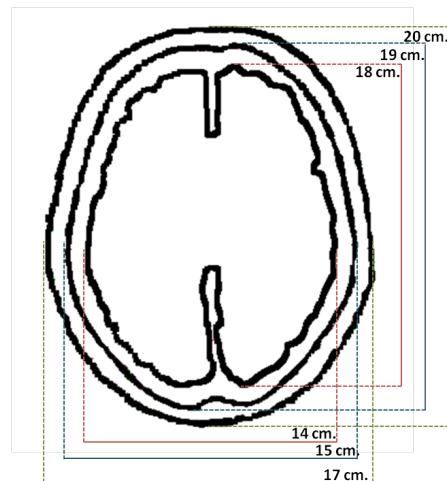


**Figure 4.5:** BEM model of 3-shell spherical head model

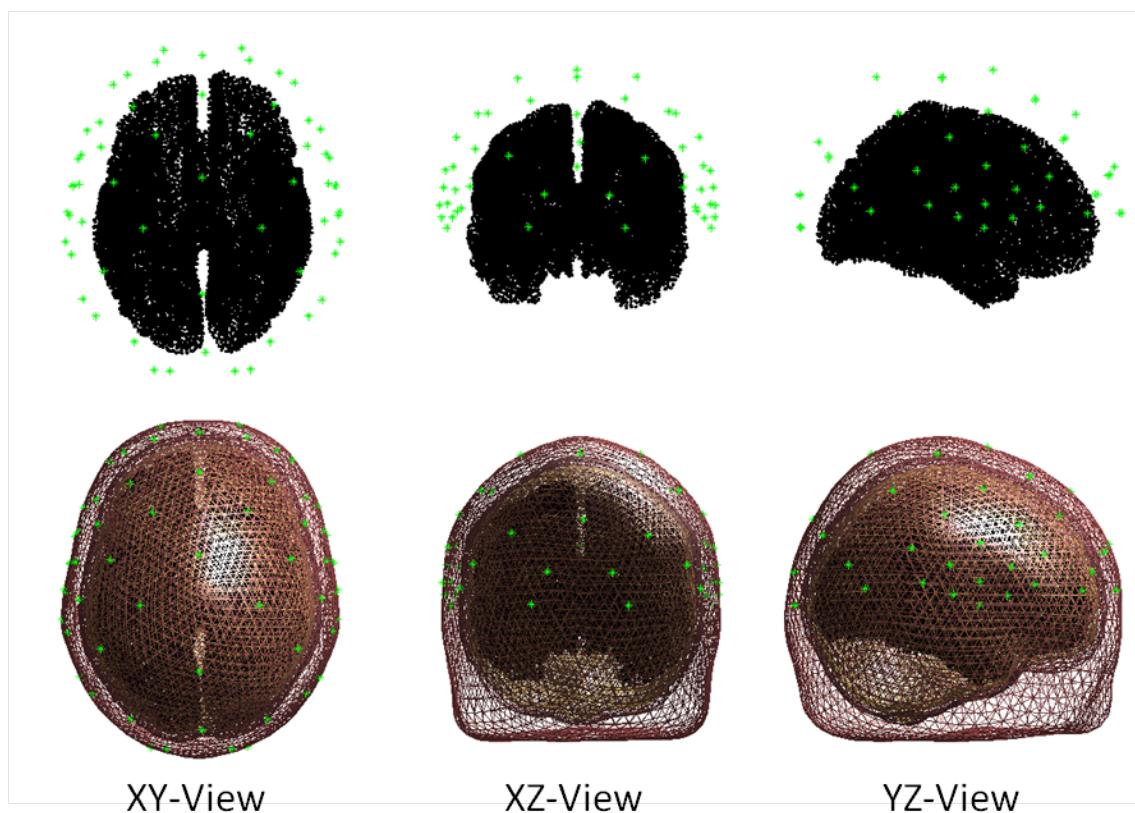
#### 4.2.2 Standardized head models: Realistically shaped head model

Instead of approximating the head by a sphere, one can use a more realistic head model. Realistic heads are available for standard heads. The realistic head model is often constructed from averaged MRI- or CT-scans. With this model we still avoid the burden of imaging a specific person's head, but it will provide a more accurate estimate of the anatomy. The cost one has to pay is the increase in computational load. There no longer exists an analytical solution. In order to solve the problem numerically, the surfaces of the model have to be divided into triangles. In other words, here one is obliged to use BEM techniques.

In this specific model we again distinguish three different layers, corresponding with the brain, skull and scalp. They have respectively a radius of approximately 9.5, 10 and 11 cm. The corresponding conductivities are  $\sigma = 0.33$ ,  $\sigma = 0.0041$ ,  $\sigma = 0.33$ . The amount of grid points used equals 20484. These are equally divided over the space. There is an obvious increase in the amount of grid points, the distribution shall therefore be more dense. A schematic overview of the model and the addressed values can be found in respectively Figure 4.6 and Table ???. How the model finally looks like in Matlab is shown from different points of view in Figure 4.7.



**Figure 4.6:** Schematic view of the 3-shell realistic head model



**Figure 4.7:** BEM model of standardized head model, up: faces, down: edges, green cross = electrode position

### 4.2.3 Individual head model and neurophysiological data

#### Technical aspects

Both the anatomical imaging as the fMRI experiments were performed on a 3 Tesla whole body scanner (Magnetom Trio, A Tim System 3T, Siemens, Germany) using a conventional circular polarized head coil. The protocol used was an EPI sequence, with TR = 3000 ms, TE = 27 ms and FA=90 degrees.

#### Head model

Based on an MRI-scan one reconstruct the anatomical structure of the different layers in the head. On an MRI scan one can distinguish the scalp, the skull, the CSF, the white matter and the grey matter. For the modelling and calculations one can use different methods. We could opt to use again the BEM method. However we prefer to use Finite Difference Method (FDM) modelling this time. This is a numerical method for approximating the solutions of differential equations using finite difference equations to approximate the derivatives. In previous research it was already shown that this is a solid method to use in forward modelling. We prefer the use of FDM here, because it is known to give a better image of the cortical structures than BEM.

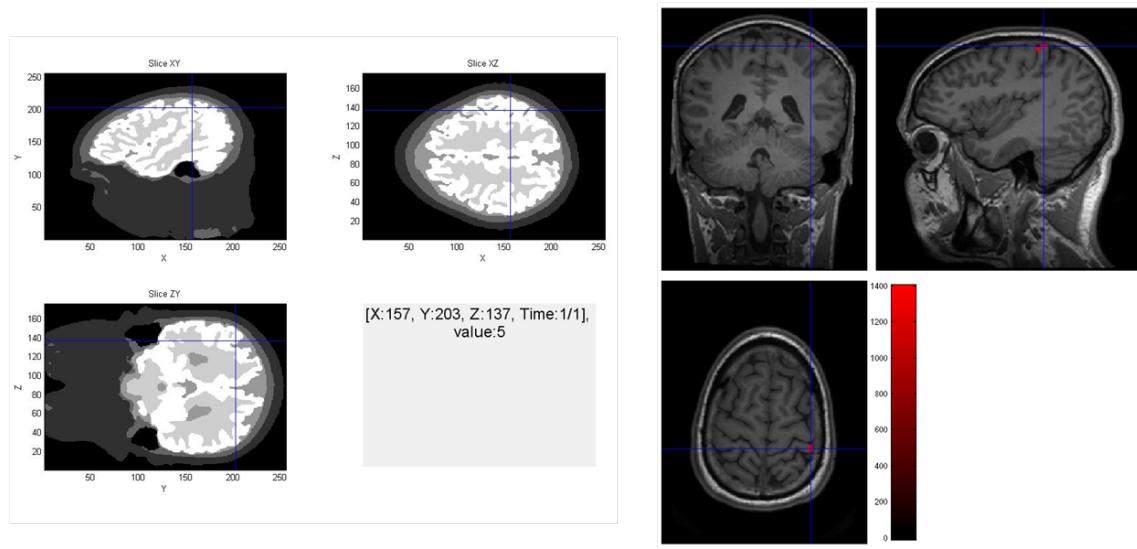
For the anatomy we no longer have to care about approximating the real head, since we already use the anatomical correct data. But we do still have to address a certain conductivity to each one of the structures. These are estimated values and are assumed to be constant within one layer. These will thus differ from the real values, and will still induce some error in the final results. *The values used are . The amount of grid points used is 8212. The MRI image can be found in Figure .*

#### Neurophysiological data

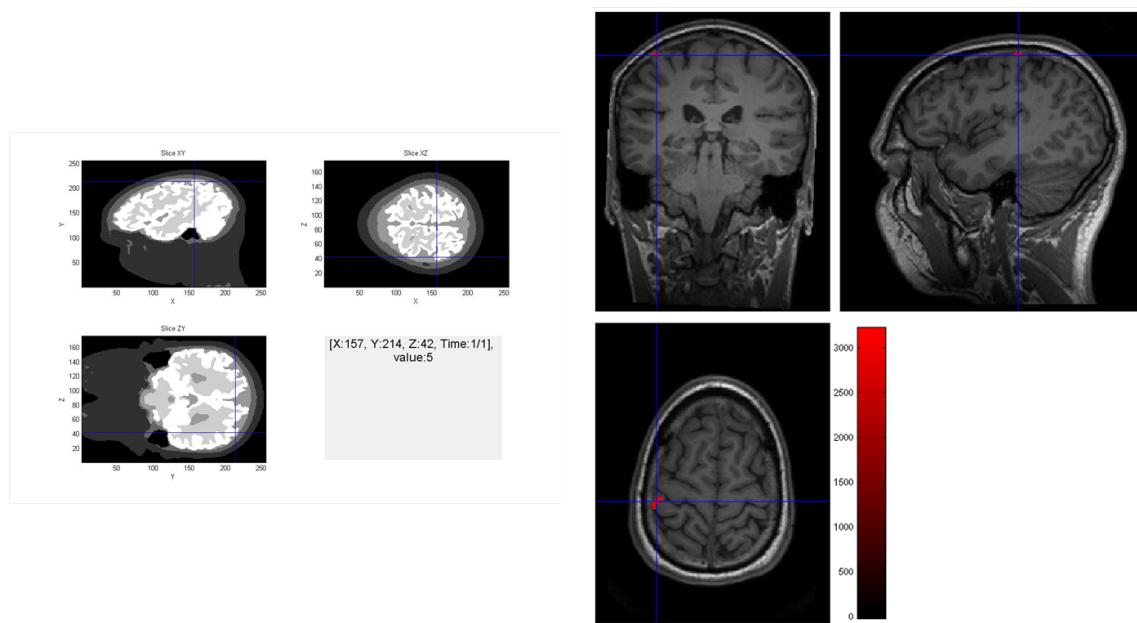
Neurophysiological data can be recorded in an MRI scanner by using functional Magnetic Resonance Imaging (fMRI). fMRI is based on the fact that during performance of a certain task, the blood flow in the corresponding area will increase. A higher blood flow leads to a change in the ratio oxygenated to deoxygenated hemoglobin. These two forms of hemoglobin have different magnetic characteristics. And thus will a change in the ratio between both induce a change in the signal captured by the MRI scanner. After processing the region that was activated during the task can then be detected. This non-invasive method has therefore become very popular to detect cortical activity and map functional areas.

In this case the user was asked to perform a motor tapping experiment. He was ordered to bring his right index finger to his left thumb and tap both fingers at a self-paced rate. And thereby activating the contra-lateral sensorimotor cortex. Next he was asked to do the same,

but now with the left index finger and the left thumb. From the resulting images we can now deduce the functional ROIs, see Figure 4.8 for the results of left MI and Figure 4.9 for the results for right MI.



**Figure 4.8:** Determination of left ROI based on fMRI data



**Figure 4.9:** Determination of right ROI based on fMRI data

# Chapter 5

# Data processing, Classification and Evaluation

## 5.1 Preprocessing

### 5.1.1 Objectives of the preprocessing step

In the previous chapter we described the input data to build and test our BCI system. One of these datasets is the BCI Competition IV dataset. The training set consists of 200 trials (containing both left and right hand MI) which last each 8 seconds. The data is recorded at 59 different electrodes and at a sampling frequency of 1000 Hz. This makes that for each trial we have  $59 \times 8 \times 1000$  samples or 472000 raw features. In other words, we have way too much features to work with. And most of them won't contain useful information. This is also called the 'Curse of Dimensionality'. So a first objective of the preprocessing step is to reduce the dimensionality of the data by discarding non discriminative information. Moreover the trials are delivered as a continuous flow of data. So a second objective will be to transform the continuous set of samples in a usable data configuration. Finally, we also want to keep the possibility of biased results as low as possible. This includes referencing the data and rejecting artefacts as much as possible.

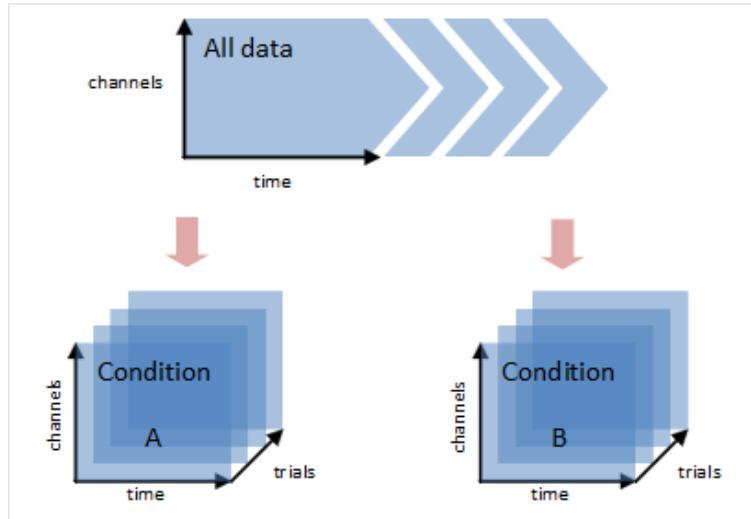
We will discuss the different preprocessing steps in the order as performed in the Matlab implementation.

### 5.1.2 Common average reference

Common Average Reference (CAR) is one of the most straightforward spatial filtering techniques. At each time point it subtracts from each sample the average over the samples at that time point of all the electrodes. By doing this the SNR can be increased, but it will also cause some spatial smearing of e.g. artefacts. The data is now re-referenced to a non-arbitrary zero value.

### 5.1.3 Trial extraction

Besides the continuous data itself, also some markers and labels are added to the data. These markers show the start of the trials. The labels reveal which kind of movement the trial is about. Based on this extra information we can divide the continuous data flow into separate trials. Dependent on the matching label one can then add the trial to the corresponding set of left or right trials. This is graphically shown in Figure 5.1, [21].



**Figure 5.1:** From continuous data to trials separated per condition

### 5.1.4 Artefact rejection

Some artefacts are more difficult to eliminate than others. Some artefacts are already cancelled out by averaging the data, e.g. when they take place at random time points in a trial. High frequency artefacts will be removed by applying a low pass filter on the data. It will be mainly eye movement and eye blinking artefacts that we will have to be worried about. This is because they appear in the same frequency band as the ERPs we're interested in, but they have a much higher amplitude than the ERPs. Moreover their occurrence might be linked with the onset of a trial.

One way to deal with these artefacts could be by visual inspection and delete the contaminated epochs. However this might decrease the amount of data available substantially. This in turn will decrease the SNR and the quality of the average ERP. Another possibility is to use ICA in combination with visual inspection. ICA is a computational method for separating a multivariate signal into additive subcomponents. We can then look at the different components and only reject the component corresponding most likely with the artefacts. We tried to apply this method, but a lot of trials were still contaminated, so we did reject the

complete trials

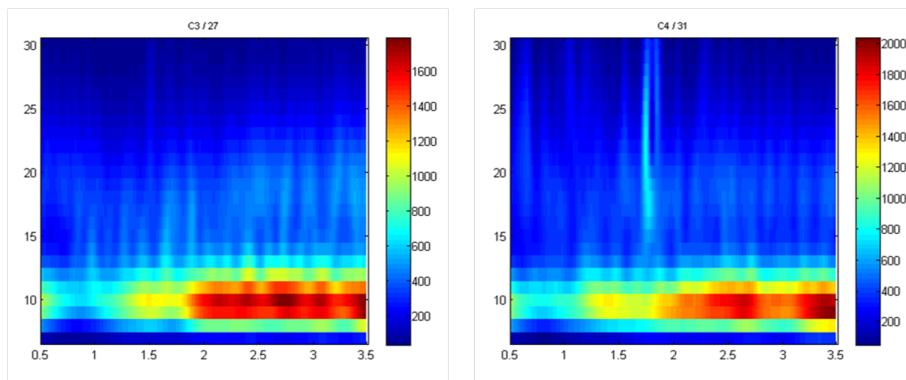
### 5.1.5 Baseline correction

During the beginning and the end of a session of data acquisition quite some time can pass by. During this time the noise level might change due to a change in environment noise, change in brain activity, etc. Over time this might lead to slow shifts such that the zero level might differ considerably across channels. To take this into account we will per trial and per channel calculate the mean of the pretrial data and withdraw this value per channel from the complete following trial. This is what we call baseline correction. The method is based on the idea that pretrial data doesn't contain stimulus correlated activity and is therefore a good measurement of the noise level at this time point.

### 5.1.6 Time-Frequency analysis and band pass filtering

When applying a time-frequency analysis one wants to analyse the time evolution of the frequency content of a signal. But first one has to find a balance between the resolution on the time domain and the resolution in the frequency domain. This time-frequency analysis is very interesting, since it gives us a clear image of the active frequency band during motor imagery. The signals in this frequency band are the ones we want to use for feature extraction.

In Figure 5.2 the time-frequency analysis of dataset 1g can be found. At the left we can see the result of the time-frequency analysis of the signal at electrode C3 averaged over all left MI trials. Analogous, we find at the right the time-frequency analysis of the signal at electrode C4 averaged over all right MI trials. We see that in both cases the signal starts off quite strong in a frequency band of 8 to 12 Hz. After half a second the signal starts to decrease, to amplify again after 1.5 second. This behaviour is indeed similar to that of an ERP signal.



**Figure 5.2:** Time-Frequency analysis between 7 and 30 Hz, left: for right MI trials at C3, right: for left MI trials at C4 for dataset 1g

Now that we know the frequency band of interest, we will use this information to band pass filter the data. The frequency band of interest will be determined per subject. In the case of subject g the data will be filtered between 8 and 12 Hz. In FieldTrip this is performed with a twopass fourth order Butterworth filter. Since it is a twopass filter, there will be no phase change. But the signal is filtered twice, so there will be double as much attenuation.

## 5.2 Feature Extraction

### 5.2.1 Objectives of feature extraction

The goal of feature extraction is to select the correct elements from a digital signal, the correct elements being the parts of the data in which the wanted information is coded. So in this step one is looking for a subset of variables which the system can use to make a deliberate decision. Feature extraction will further reduce the dimensionality of the data. Unlike dimensionality reduction techniques such as ICA, feature selection doesn't transform the variables, but keeps the essence of the variables. This makes reasoning and calculating with the features much easier.

Feature extraction is a two phase procedure. In a first phase one needs to determine how one wants to define and calculate the features. After the first phase the feature set will still be quite large and definitely redundant. Therefore it is necessary to look in the second phase for an optimal subset of features.

I also want to emphasize here the fact that during the feature extraction step, we only use the calibration data set of BCI Competition IV.

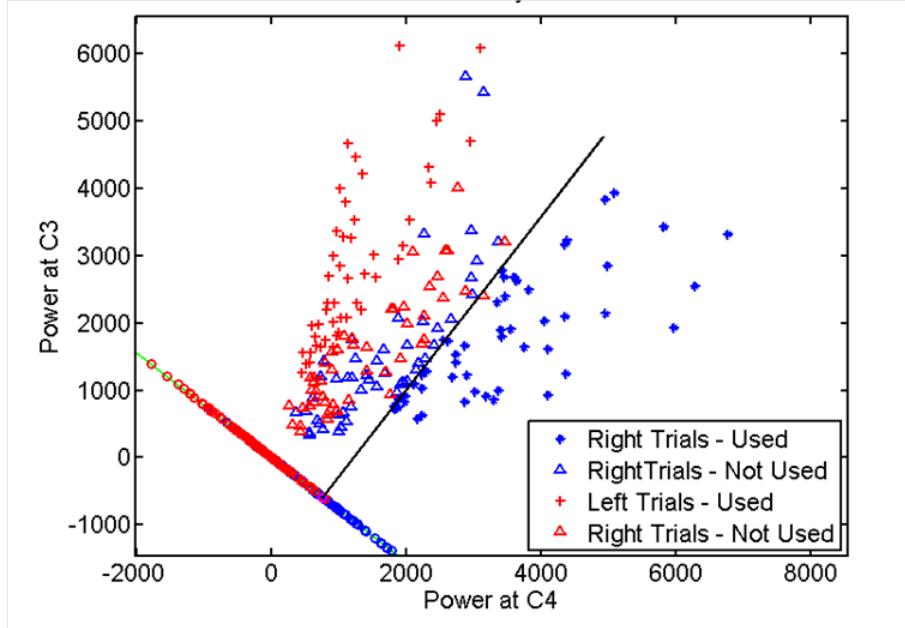
### 5.2.2 How to determine the beamformer filter

One of the most important parts of the feature extraction step is the decision of which spatial filter to use. In our case we will use a filter based on LCMV beamforming. This filtering approach has been discussed in detail in Chapter 3 and won't be repeated here. This part will mainly focus on the data used for calculations of the beamformer and on the specifics used during the calculation. Beamforming can be trained with unlabelled data. Distinction between different states is then solely based on the chosen ROIs. But since we do have the labelled information, we decided to use this as an extra asset. We will create two filters, one based on trials of left MI and one based on trials of right MI.

#### Training data for filter

The first step in the calculation of filter is to determine which data one wants to use to train this filter. A logic assumption would be to use all the data available, especially since we know

that an increase in amount of trials results in an increase in SNR. However while testing filters based on all trials, we noticed the results turned out to be quite disappointing. Therefore we have decided to only use a part of the trials. The amount of trials per set used for training is up to the user. But the selection procedure itself is not random and has been implemented. The implementation is based on the idea that the best trials for training are those that can be most easily separated based on the power at the C3 and C4 electrodes. These trials can easily be found by applying Fisher LDA. This is a linear discriminant analysis technique that will be explained into more detail later on. The result of the selection procedure can be found in Figure 5.3.



**Figure 5.3:** Trial selection for training of beamformer filter

### Filter calculations

Once we have decided which data to use for the training of the filter, we can start building the filter. This is completely implemented in FieldTrip, following the mathematics as described in Chapter 3. But we still have to fill out some parameters before the calculations can start. These parameters mainly concern the calculation of the correlation matrix. This is namely the value that can change if the head model and grid and the data are kept the same.

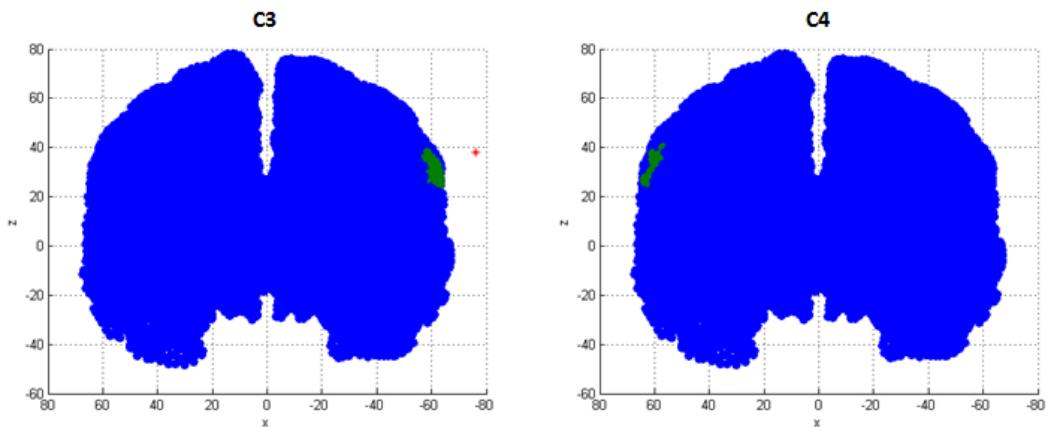
One of these parameters is the (length of) the time window that is used for calculating the correlation matrix. In general one assumes the longer the time window, the more robust the resultant filter will be. Since our trial is rather short, we can't vary the time window much. Nevertheless we tried to capture the influence of this parameter. The results of this

is depicted in chapter 6. As a standard value we chose a time window from 0.5 seconds after the start of the trial, until 2.5 seconds into the trial. We don't start from the very beginning of the trial, because a subject always needs some time to process the instruction and to start the execution of the instruction. The end of the time window is based on the information we can retrieve from the time-frequency analysis, see Figure 5.2.

The second parameter we have to assign a value to, is the regularisation parameter. In Equation 3.9 we see that at a certain point the correlation matrix has to be inverted. In order for the matrix to be invertible, the matrix can't be singular. However the correlation matrix will often have nearly-dependent columns, leading to instability during inversion. One can solve this problem by adding a small regularization parameter at the diagonal elements. If the parameter is chosen too small, the problem of instability won't be resolved completely. If the parameter is chosen too big, the beamformer will become less focused, leading to a blurred activity reconstruction. The influence of this parameter will also shortly be investigated. Again the result can be found in chapter 6.

### 5.2.3 Determine grid points corresponding with ROI

The determination of the grid points is of great importance for the feature extraction step. Both the amount and the position of the grid points will define the first subset of features. The grid points are determined based on neurophysiological data. We know that in general MI is presented at the C3 and C4 electrodes. We also know that the pyramidal cells are the generators of the EEG. These lay at the cortical surface of the brain. Based on this we decided to use the grid points which are closest in distance to the C3 and C4 electrodes. This is visualized in Figure 5.4.



**Figure 5.4:** Trial selection for training of beamformer filter

### 5.2.4 Feature calculation

After the calculations in the previous step we obtain a filter  $\mathbf{W}(\mathbf{q})$  for each grid point  $\mathbf{q}$ . If we now apply this filter  $\mathbf{W}$  on a trial  $\mathbf{x}(t)$ , we get a result  $\mathbf{y}(t)$ .

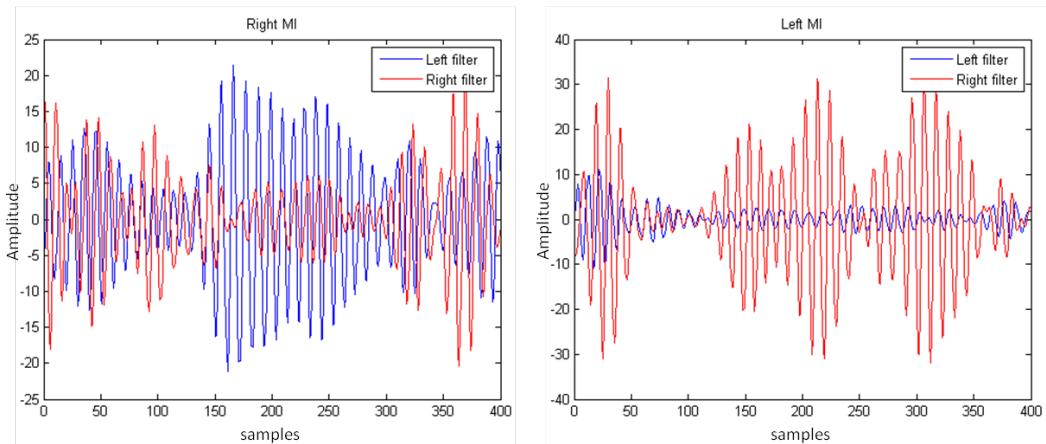
$$\mathbf{y} = \mathbf{W}(\mathbf{q}_0)\mathbf{x} \quad (5.1)$$

This matrix  $\mathbf{y}(t)$  contains the estimate of the amplitude of the dipole at grid point  $\mathbf{q}$  for each direction. In order to obtain the amplitude of the complete signal sent out by the dipole, one needs to take the norm of the dipole at each time point. We will call this signal  $\mathbf{s}(t)$ . This is now the estimate of the signal generated during MI at the corresponding location in the brain. After averaging over all trials of right or left MI, we see indeed the ERP appear, see Figure 5.5.

$$\mathbf{s}(t) = \|\mathbf{y}(t)\| \quad (5.2)$$

Since we want our system to be workable in real-time, we cannot use the complete trial. We will divide each signal into separate parts with a certain step size. If for example the step size is equal to 1 second, the signal will be divided into 4 non-overlapping parts. We will call them  $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3$  and  $\mathbf{s}_4$ . A feature is now defined as the logarithm of the power of a subsignal.

$$F = \log(\mathbf{s}_1 \mathbf{s}_1^T) \quad (5.3)$$



**Figure 5.5:** Right: Combination of averaged signals at 11 nearest grid points at C3 (blue) and C4(red) for right hand MI. Left: Combination of averaged signals at 11 nearest grid points at C3 (blue) and C4(red) for left hand MI. We indeed see a decrease of the signal power at the contra-lateral electrode.

The signal at each grid point can be used for feature calculations. So we will have to decide how many grid points per electrode for building the feature vector. In case we have  $N$  trials,

4 subsignals per trial, and we choose to use  $nGP$  grid points per electrode (i.e. C3 and C4), the final size of the feature vector will be  $4N \times 2nGP$ .

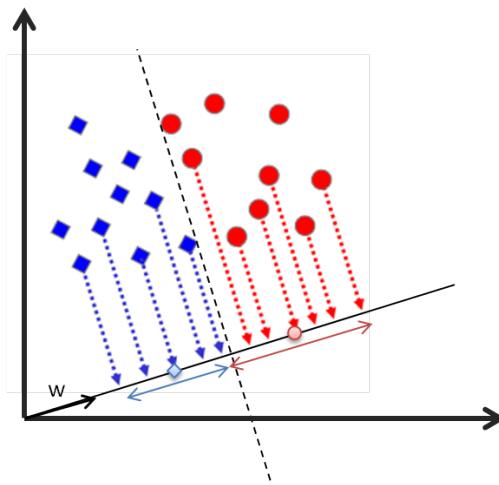
## 5.3 Classification

### 5.3.1 Objectives of classification

The feature extraction step gives a feature vector of a predefined length. In the classification step we will now translate this feature vector to a decision that can be interpreted by the computer or another machine. The classification algorithm actually tries to detect some kind of pattern corresponding with one of the classes in the feature vector. Which classification algorithm to use is dependent on the characteristics of the features, the BCI and the preferences of the user. A clear overview of the most popular classification algorithms, their characteristics and applications and their performances is given by Lotte et al [17].

### 5.3.2 Characteristics of Fisher LDA classifier

We've opted to use a linear discriminant classifier, more specifically the Fisher LDA. A linear classifier calculates the parameters of a hyperplane in the feature space, such that it separates the features of the different classes optimally. Classification is then done by calculating at what side of the hyperplane the feature vector is on. The best separating hyperplane is the one for which the distance between the mean of the features is maximized, while at the same time the variance within the distribution of the feature for each class is minimized. This is depicted in Figure 5.6 for a 2 dimensional feature vector. This technique makes some assumptions regarding the data. The covariance matrices of the two classes is assumed to be equal, and the features have to follow a normal distribution.



**Figure 5.6:** Example of Fisher LDA for a two dimensional feature vector

Using Fisher LDA has both some advantages and disadvantages, both often the consequence of its simplicity. Some of the advantages are:

- The computational power necessary is quite limited. This is definitely a benefit if in time we would like to use the system for real-time classification procedures.
- The classifier is easy to understand, easy to use and easy to implement.
- In general the Fisher LDA classifier performs quite good. Specifically interesting for our study is that success was already booked with the classifier for MI based BCIs and asynchronous BCIs
- It is also a stable classifier. This means that variations in the training data not necessarily have a big influence of the classifier and its final performance
- Also the fact that it is linear is interesting. The more complex a classifier, the more training data one needs to sufficiently train it. Since our training set is rather limited, it will need minimum training.

Some disadvantages are:

- Sometimes the classifier is just too simple, especially if the EEG data or its corresponding feature vectors show non linear behaviour
- The classifier is not regularized, which makes it more prone for outliers

Beside the obvious advantages of using Fisher LDA as classifying system, also some other factors played a role in our choice:

- Some classifiers do have the possibility to take time information of the data into account. These are called dynamic classifiers. Using time information might definitely be beneficial for final classification results. However our evaluation data set is presented as asynchronous data. Therefore we don't have any time information. So our system won't gain anything by using this more complex classifier
- There are also generative classifiers that would fulfil most of our needs. These classifiers calculate the probability that a feature vector relates to a certain class, instead of making a real decision. However these probability based classifiers are often more prone to noise.

For the mathematical implementation we refer to more specialized literature. [1]

## 5.4 Evaluation

### 5.4.1 Objectives and evaluation criteria

The last step of the design process is the evaluation of the proposed BCI system. This implies we need to find an appropriate measure for the performance of a BCI. In the end it will be this measure that will be used to compare the influence of different parameters on the achievement of the system.

A typical measure for the performance of a BCI is the accuracy. The accuracy tells us how much of the data was classified correctly. The result can be presented as a percentile, or can be presented in the form of a confusion matrix. The confusion matrix contains beside the amount of true positives and true negatives (i.e. the correct classified samples), also the false positives and the false negatives. Imagine a classifier has to decide whether or not a sample belongs to class 1. The true positives are then the samples of class 1 that are also classified as class 1. The true negatives are the samples not belonging to class 1 and that are accordingly not classified as class 1. The false positives are the samples that don't belong to class 1 but are nevertheless classified as class 1. And eventually the false negatives are the samples that do belong to class 1 but are classified differently. An example of a confusion matrix for a 2 class problem is presented in Table 5.1. If the results in the confusion matrix are given as percentiles, the sum of all elements should add up to 100%. The sum of diagonal elements is equal to the total accuracy, as defined at the beginning. In case of a low accuracy, the advantage of using a confusion matrix is that one might get an idea of what goes wrong.

Due to the difference in nature between the calibration and the evaluation data set of the BCI competition, we will have to apply a different method of evaluation. The calibration data set is the result of a synchronous BCI whereby there is clear a distinction between periods with relevant and irrelevant brain activity. The main objective of this dataset is to fine-tune the parameters of the system. The evaluation dataset however is the result of an asynchronous BCI. This means that we have a continuous flow of data, without any knowledge about the beginning or end of the trials. Moreover the in between trial period is defined as the idle state and should also be recognized. Evaluation of this dataset will give a more realistic image of the real performance and the generalization ability of the system.

### 5.4.2 Calibration data set

A first step in the evaluation of the calibration data set is to divide the calibration data in a training set and a testing set. The test set is built by randomly choosing a certain amount of trials out of the calibration data set. The amount is determined by the training percentage. The complementary set is then the training set. The training set is used to train the classifier. Spatial filtering and feature extraction will be performed. The classifier will then

	Class 1 (Predicted)	Class 2 (Predicted)
Class 1 (Actual)	35%	15%
Class 2 (Actual)	13%	37%

**Table 5.1:** Example of confusion matrix for a two class problem

adapt its parameters to classify the resultant features optimally. It is of great importance that the testing set is being ignored completely in this phase. In the second step we will now also apply spatial filtering and feature extraction on the test set. The features will then be classified by the classifier obtained after the previous step. This procedure is repeated ten times. Each time another test set is randomly chosen.

For the evaluation we will take both the error on the classification of the training set and the error on the classification of the testing set into account. The first one is called the training error, the latter is called the testing error. The standard deviation will also be calculated for both errors. To determine whether or not the system is reliable, the testing error is the most important factor. This one gives the most reliable idea of how well the system performs. By also taking the training error and the standard deviations into account, we can see if the classifier is prone to overfitting. Since we use a linear classifier the chance of overfitting is rather small. But if the classifier would be inclined to overfit, this is quite a big problem. Overfitting is characterized by a testing error that is (much) higher than the training error. The overfitting will namely ensure that a big amount of the training samples are correctly classified, but the classifier will be too specific adapted to this training set and consequently the testing set will be badly classified causing a larger testing error. Also the standard deviation of the testing set will be large, because sometimes a testing dataset will fit the classifier anyway and sometimes the classifier will completely fail.

Since the final goal is to classify data coming from an asynchronous BCI, we want our system to be able to classify different parts of a trial, without knowing to which time frame in the trial the part corresponds. Therefore we will separate each trial in 4 parts. Each part then corresponds with 1 second of the trial, or 1000 samples. So instead of 200 trials per data set, we will now have 800 'trials'. In case of an 80% training percentage, 80% of these parts will be used to train the classifier, the other 20% will be used for testing. I also want to remark that the training percentage might also influence the performance of the classifier.

We discussed all this for the calibration data sets of BCI Competition IV. A complete ana-

logue procedure will be applied for the own recorded data, since this data set has a similar structure.

### 5.4.3 Evaluation data set

It is the evaluation of the evaluation data set that will give us the best idea of the performance and robustness of the classification possibilities of the BCI system. To do this the data set will first be preprocessed in the same manner as the other data sets. However baseline correction won't be possible, since we don't know on beforehand which parts have or don't have relevant brain activity. Next spatial filtering, feature extraction and classification will be performed. In this case the classifier will be trained by the complete calibration data set. Other parameters will be chosen as fine-tuned by the calibration data set.

Since the evaluation data set doesn't have labels to mark the beginning or end of different phases in the data, classification will be performed on the part of data selected out by a sliding window of 1 second or 1000 samples. In every time step the window will move 1 samples. In the end every sample will be classified 1000 times. Each time a sample is accredited a value of 1 if classified as a right hand movement or -1 if classified as a left hand movement. The final class assigned to a certain sample is based on the average of the values assigned during classification. If the average is between -1 and -0.25, the sample gets a final classification of left hand movement. If the average is between 0.25 and 1, the sample gets a final classification of right hand movement. All samples with a value in between are considered undefined or to be in the idle state. We use -0.25 as the standard threshold, but this threshold can be adapted to the wishes of the user. Also the size of the sliding window is a variable. The size of the window will determine the ITR. Because the system has to weight until the window has completely 'passed by' to make a final decision on the class, the window will introduce a delay. The longer the window, the bigger the delay.

Final classification accuracy is only based on the samples that were indeed recorded during MI. The final classification is then compared with the actual class. Results will be represented in the form of a confusion matrix, as described earlier.

# Chapter 6

## Results: Building a robust BCI

### 6.1 Input: BCI competition IV Data specifications

The first objective of this thesis is to build a robust and well functioning BCI system based on spatial filtering using beamformer techniques. The details of this design are found in the previous chapter. In this master's thesis we have to test and fine-tune the system. Therefore we use the BCI Competition IV dataset. This dataset actually contains 7 datasets from 7 different subjects. Of those 7 subjects 2 subjects (1a and 1d) didn't perform right and left hand MI, but left hand and both feet MI. These datasets won't be used, i.e. all trials are rejected. The details of the datasets are found in Chapter 4.

Before we start the actual fine-tuning we want to give an overview of the different datasets and their characteristics. The characteristics most relevant for our system are:

- **Frequency band.** This is the frequency band most active during motor imagery. This band determines the band-pass filter applied during preprocessing. The band is selected after visual inspection of the time-frequency distribution.
- **Rejected trials.** These are the trials severely contaminated with artefacts. These are also selected by visual inspection and rejected during preprocessing.

The details are given in Table 6.1.

### 6.2 Fine-tuning some parameters

#### 6.2.1 Standard parameters and the corresponding classification error

During the description of the design process we noted that we can choose a lot of parameters. Each parameter will have a certain influence on the final system. To ensure an optimal working system, it is thus important to fine-tune these parameters. Therefore we listed all

Name	Frequency band	MI Classes	Artificial	Rejected left trials	Rejected right trials
1b	10-15 Hz	Left hand, right hand	No	86, 76, 95, 96	30, 77, 97, 98
1c	8-12 Hz	Left hand, right hand	Yes	2, 51, 65, 70	4, 34, 60, 64, 63, 71, 87, 95
1e	11-15 Hz	Left hand, right hand	Yes	None	92
1f	8-12 Hz	Left hand, right hand	No	80	90
1g	8-12 Hz	Left hand, right hand	No	14, 21, 22, 23, 28, 35, 43, 89, 99, 100	38, 39, 40, 44, 67, 86, 89, 95, 99, 100

**Table 6.1:** The specifics of the available and used data sets

variables, chose a specific well-thought-out standard value for each variable (see Chapter 5) and tried to find the optimal value for each parameter. This was done by varying only one parameter at a time and leaving the other parameters constant. An overview of the variables, and their standard value is given in Table 6.2.

Parameter	Standard Value	Possible Values
Baseline Correction	0	0, 1
Artifact Removal	1	0, 1
# Training Trials	50	30, 40, 50, 60, 70, 80, 90
Regularization Parameter	0.05	0.005, 0.01, 0.05, 0.1, 0.2
Time window	0.5 - 2.5	see Figure 6.4
# Grid points	20	10, 15, 20, 25, 30, 50, 75
Trial step size	1000	800, 1000, 2000, 4000
Training Percentage	0.8	0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9

**Table 6.2:** Parameters list: standard values and possible values

In order to be sure that the results are not a coincidence, we decided to use 2 datasets for the fine-tuning step, namely dataset 1E and 1G. In Table 6.3 and Table 6.4 we find the confusion matrix of the standard system of respectively dataset 1E and 1G. For 1E we obtain a classification accuracy of 79% and for 1G we obtain a classification accuracy of 78.5 %.

	Right MI (Predicted)	Left MI (Predicted)
Right MI (Actual)	0.357	0.1195
Left MI (Actual)	0.09075	0.43275

**Table 6.3:** Confusion matrix - Standard parameters - 1E

	Right MI (Predicted)	Left MI (Predicted)
Right MI (Actual)	0.397778	0.095
Left MI (Actual)	0.119722	0.3875

**Table 6.4:** Confusion matrix - Standard parameters - 1g

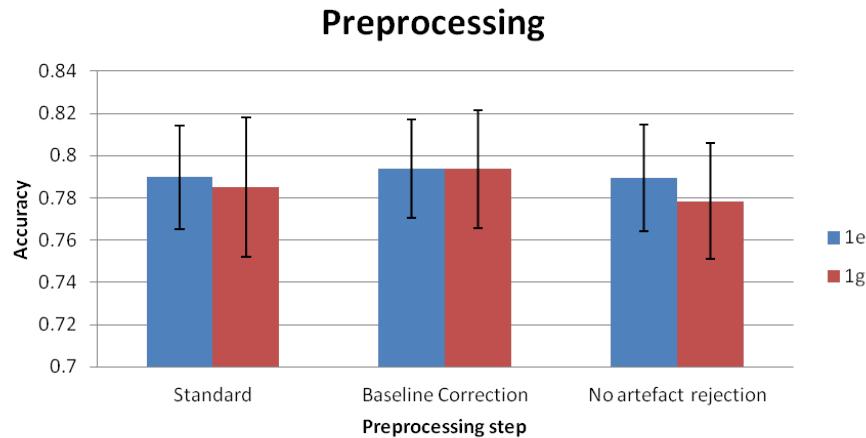
## 6.2.2 Preprocessing

### Baseline Correction

As mentioned before, the objective of baseline correction is to take into account changes in noise level over different trials. One thus expects that performing this baseline correction, will increase the accuracy. From Figure 6.1 we can indeed deduce that there is a small increase of the accuracy due to baseline correction for both dataset 1E and 1G. Although the increase only ranges from 0.5% to 1%, it is advisable to apply this correction. However, we have to note that in order to be able to perform baseline correction, one needs to have marked trial parts corresponding with an idle state. This won't be the case for the evaluation dataset which consists of asynchronous data.

### Artefact Removal

In the standard method we apply artefact rejection, but before we go on we wanted to see what the influence is of this on the artefact rejection. Since we reject the complete trials, it is possible that accuracy decreases due to a loss of information. For these datasets we see that the influence on 1E is hardly noticeable. It is more likely that the decrease is due to variations in the training and testing sets than due to artefact rejection. This can simply be explained by noticing in Table 6.1 that we only identified one contaminated trial. Rejecting this artefact won't have much influence on the final beamformer. For dataset 1G the situation is completely different. Here up to 10% of the trials are contaminated with artefacts. And Figure 6.1 suggests that by taking these trials also into account the accuracy will decrease. So despite the decrease in data, it is preferable to apply artefact rejection.



**Figure 6.1:** The influence of baseline correction and artefact removal on the accuracy

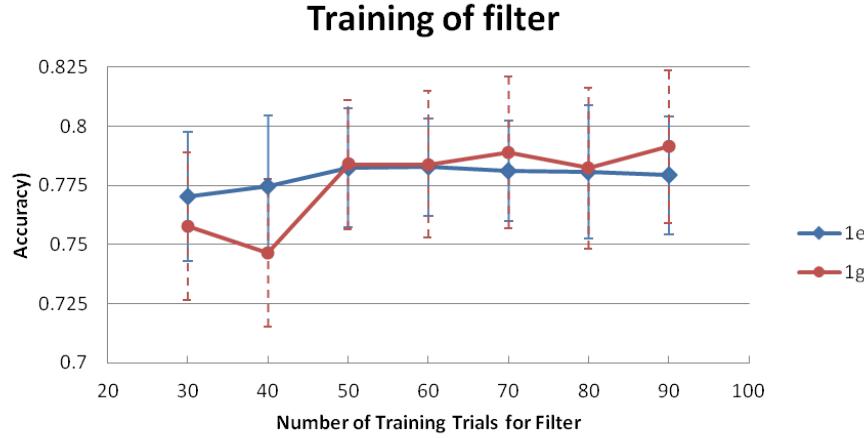
### 6.2.3 Beamformer parameters

#### Training trials

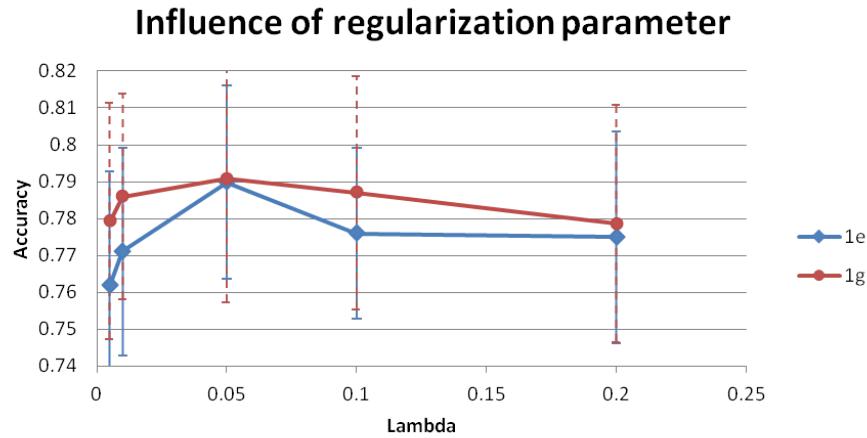
In general one assumes that the more data one has for training the filter, the better or the more robust the filter becomes. But during the first tests of the system we noticed this wasn't necessarily always the case. Therefore we decided to investigate the influence of the amount of trials used for training on the final accuracy. The results are depicted in Figure 6.2. The selection of trials was done as described in section 5.2.2. We can remark that in the case the amount of trials is really small (<40 trials), the accuracy will indeed decrease. However from 50 trials per filter on, the accuracy seems to stabilize. Since the extra computational power needed for taking into account more trials is rather limited, we prefer to use in the final system 80 trials. This will on the one hand ensure us that there is enough data for a good estimation of the correlation matrix, on the other hand by not using all the trials we still exclude the worst differentiating trials.

#### Regularisation parameter

The regularisation parameter  $\lambda$  is a term added to the diagonal of the correlation matrix in order to stabilize calculations. If the parameter is chosen too little, the stabilisation problem is not resolved completely. If the parameter is chosen too big, the system loses specificity. By varying the parameter between rather small and quite big values, we want to determine which parameters is the best trade-off between specificity and stability. In Figure 6.3 we see that for both dataset 1E and 1G the curve reaches a peak for  $\lambda = 0.05$ . This will thus become the value of choice in the final system.



**Figure 6.2:** The influence of the amount of trials for training of the filter on the accuracy

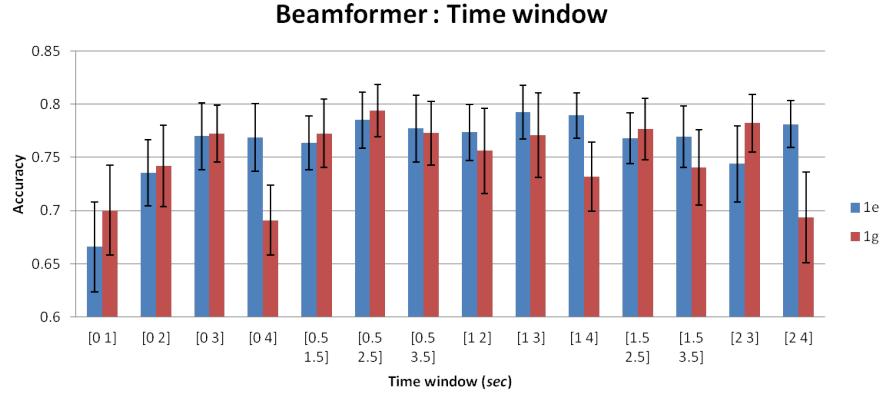


**Figure 6.3:** From continuous data to trials separated per condition

### Time window of interest

The time window is the last parameter that determines the correlation matrix, and thus the filter itself. We can choose the start and the end of the time window, and vary both the position and the length of the window. In essence the main importance is to capture the ERD. In Figure 6.4 the classification accuracy for different time windows is presented. We notice that the behaviour is different for dataset 1E than for 1G. For dataset 1E the length of the window seems to be the determinant factor, whereas for dataset 1G it seems particularly important that the last second is not included in the window. The optimal values are obtained for respectively a time window from 1 to 3 seconds and from 0.5 to 2.5 seconds. Based on

these values we will use a time window from 0.5 to 3 seconds in the final system. On the one hand, by using a relatively long time window, we make sure the complete ERD is included. On the other hand the beginning and the end of the trials are excluded, this guarantees us that not too much inactive moments are present.

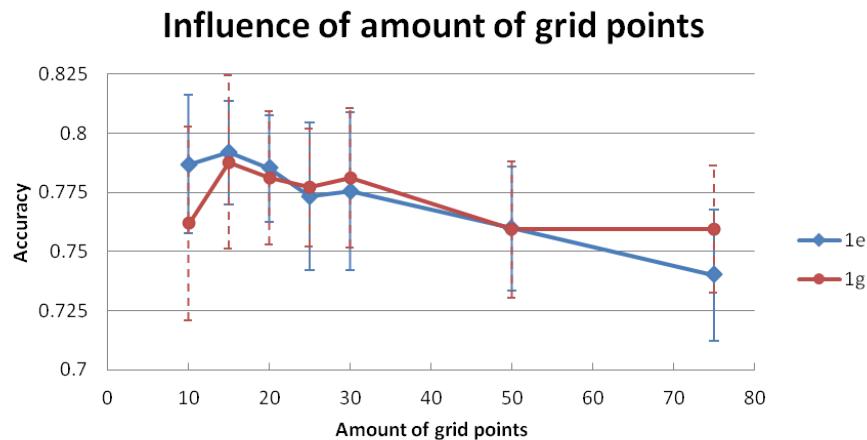


**Figure 6.4:** Influence of the time window on the classification accuracy

#### 6.2.4 Feature extraction

##### Number of grid points

The extracted features are determined by the feature selection algorithm. In our case this is really simple, the feature is the power of the signal at grid points near the C3 and C4 electrode. So the only parameter we have to determine is the amount of grid points we want to use. We varied the amount of grid points from 10 to 75 per hemisphere, see Figure 6.5.



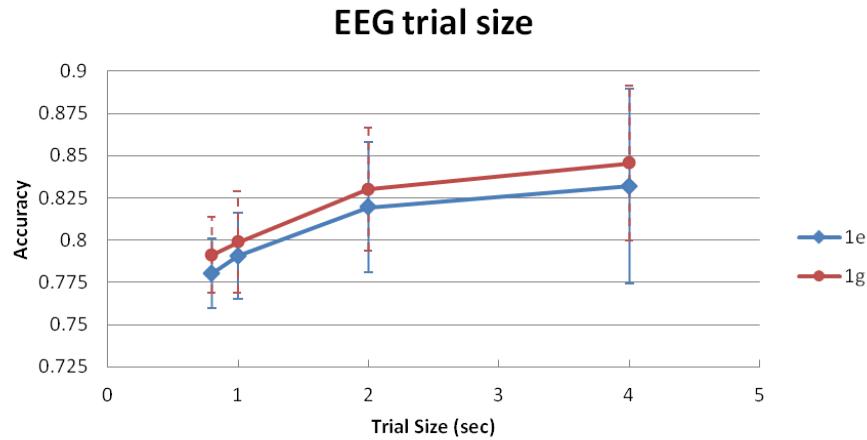
**Figure 6.5:** Influence of the number of grid points / features used on the classification accuracy

We see for both datasets a peak at 15 grid points. This means that for each electrode optimally 15 features are used. In total this brings us to 30 features. This is slightly higher than the value ( $\pm 20$ ) we predicted in Chapter 5, but since this value occurs for both the datasets this will be the standard value used for the final system.

### 6.2.5 Classification

#### (Sub-)Trial size

The size of the trial, or the amount of samples used for classification, is an important factor regarding the classification in real time. If we can get a good accuracy for small parts of the trial, this is very promising for real-time evaluation of data. In Figure 6.6 we see the accuracy increasing for an increasing trial size. This is what we expected. We see that the accuracy for a trial size of 1 sec (or 1000 samples) has an accuracy of near 80% for both datasets. This thus seems a workable value.



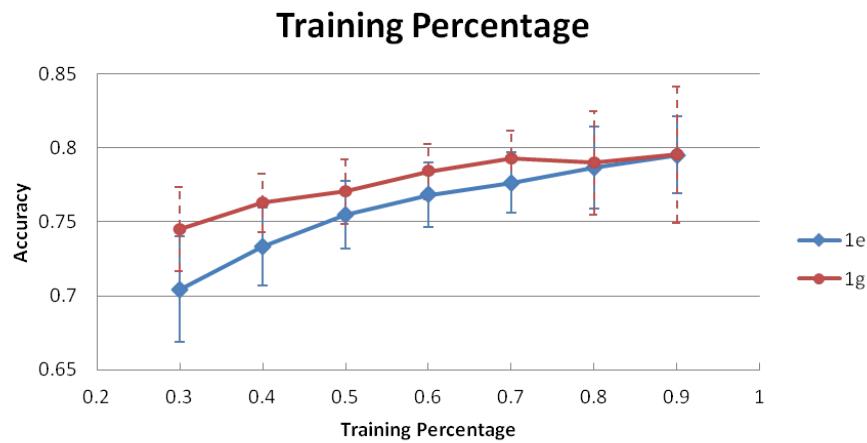
**Figure 6.6:** Influence of the trial size on the classification accuracy

#### Training percentage

And the final parameter we want to discuss is the training percentage. This is the ratio of amount of trials used for training of the classifier to the total amount of trials available. The more data used for training the classifier, the more robust the classifier will become. However if the training percentage is too high, the classifier will be prone to overfitting. Moreover we want the test set to be more or less representable for the complete dataset. Therefore the test set can't be too small.

We changed the training percentage over a range from 30% to 90%, see Figure 6.7. We notice that the results fulfil our expectations quite good: the higher the training percentage,

the better the classification results. Furthermore we see that even for very high training percentages, the classification accuracy keeps on increasing. This makes us suspect that the classifier won't be prone to overfitting at all. This is not that surprising, since we opted to use a simple linear classifier. Despite the very good results for a training percentage of 90%, we do prefer to use a training percentage of 80% in our final system. And this because of the reason discussed earlier, namely we want the training set to be representable for the complete dataset.



**Figure 6.7:** The influence of the training percentage on the classification accuracy

## 6.3 Evaluation of the calibration datasets

### 6.3.1 Final parameter selection

Parameter	Final Value
Baseline Correction	1
Artefact Removal	1
# Training Trials	80
Regularization Parameter	0.05
Time window	0.5 - 3
# Grid points	15
Trial step size	1000
Training Percentage	0.8

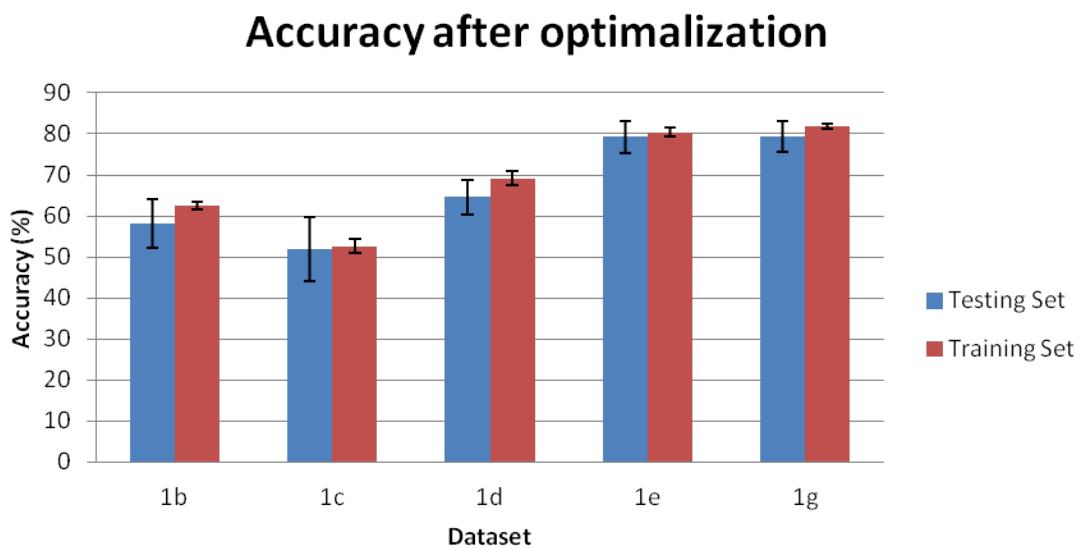
**Table 6.5:** Parameters list: final values

After determining what seem to be the optimal values for the different parameters of the system, we are now going to test the complete system. Therefore we use the 5 different datasets provided by the BCI Competition IV. Before giving the results, we give an overview of the final chosen values in Table 6.5.

### 6.3.2 Final results

In a previous step we explored the results obtained by varying some parameters in the pre-processing, spatial filtering and classification step. Based on these results we decided on what the parameters of our final system should be. An overview is given Table 6.5. In Figure 6.8 the results of our designed system are now depicted for all the applicable datasets available from BCI Competition IV.

The first thing one can remark when looking at the results are the big differences in performance between subjects. On the one hand we have very good performance, up to 80%, for subject 1E and 1G. Next we have subject '1d' for which an accuracy of about 65% is obtained. For real life applications this might not be useful, but it shows that there is definitely some information in the EEG signals available that can be used to classify between different states. By using the system the subject might train himself and so further increase the performance. For subjects '1b' and '1c' on the other hand (especially 1c), the accuracy hardly exceeds 50%. This means that classification seems to happen randomly.



**Figure 6.8:** The final accuracy results for both training and testing set for all BCI Competition IV

The cause for these differences in between subjects can be sought at different levels:

- **Optimization based on dataset 1E and 1G:** First of all we optimized the system solely based on the 1E and 1G datasets, so it's not that surprising that these systems outperform the other datasets. There is a very likely possibility that the optimal parameters for the other datasets are different than for the datasets we have used. This will most likely be the case for parameters regarding the data itself, like the optimal amount of grid points used for classification. Parameters such as the time window of interest, or the amount of trials used for training the filter are less likely to be strongly dependent on the dataset.
- **Differences among subjects:** The performance of the subjects themselves will also differ. It is quite likely that one person is better in performing motor imagery tasks than another person. This will also be dependent on how much training this person already has had. So the differences in accuracy might also be a consequence of differences among subjects, rather than the consequence of a lack of robustness. To see whether or not this is a valid assumption, we consulted the website of the BCI Competition IV. Here are besides the datasets, also the results of the contest per participant and per subject made available. We see the same trend reappear here, namely the best results are obtained for subject 1E and 1G, followed by dataset '1d' and the lowest accuracies are obtained for dataset '1b' and '1c'. These are results for the evaluation datasets, as we will use in the next section. But we assume that since performance is a characteristic of the subject itself, we can apply the same reasoning for the calibration datasets.
- **Changes in ROI:** The region of interest, i.e. the region where the ERDs are best noticeable, can also differ a lot among subjects. Therefore it might be interesting to use another selection method than 'grid points nearest to the electrodes'.

Another conclusion we can draw based on Figure 6.8 is that the system doesn't seem to be prone to overfitting. We can deduce this from the differences between the accuracy of the training set and the testing set. These are for all datasets practically the same. Furthermore we see that the standard deviation with respect to the average is quite low for all datasets except '1c'. This again indicates the system to be robust.

For dataset '1c' we notice the standard deviation to be much larger than the other values. This again implies a complete random classification. The fact that the training error is also about 50% further implies that there are no discriminating features distinguished.

## 6.4 Evaluation of the evaluation datasets

### 6.4.1 Specifics of the evalauation

Besides evaluating the system based on the result for the synchronous calibration datasets, we also want to evaluate the system based on the results for the asynchronous evaluation

datasets provided by the BCI Competition for each subject. Since the subjects of the evaluation datasets correspond with the subjects of the calibration datasets, we can use the parameters as determined by calibration.

For evaluating the calibration dataset it was necessary to separate the dataset in both a training set and a test set. For the evaluation dataset this is of course not the case. Therefore we train the Fisher LDA classifier with all the trials available (i.e. all trials without the ones with artefacts). The result of this training, i.e.  $B$  and  $c$ , is then saved, and used for the classification of the evaluation dataset.

These datasets are asynchronous, so we won't include time information whatsoever. What we will do is use a sliding window. By using a sliding window we take the information of the previous samples into account. The size of this sliding window will be one of the determinant factors of the final accuracy. The applied method for classification was discussed in detail in 5.4.3. As a final result we will have a certain amount of classification results per sample. The amount is equal to the step size. The final value assigned to a sample will be a number between -1 and 1. -1 indicating left hand MI and 1 indicating right hand MI. To make a final classification of a sample we will have to decide on a threshold that separates the two MI classes. This threshold will also have an influence on the accuracy. Not the complete range of -1 to 1 has to correspond with a certain state. One can assume that results with a value near to zero most likely correspond with the idle state, since they are supposed to be classified randomly left or right. We will discuss this assumption in detail later on.

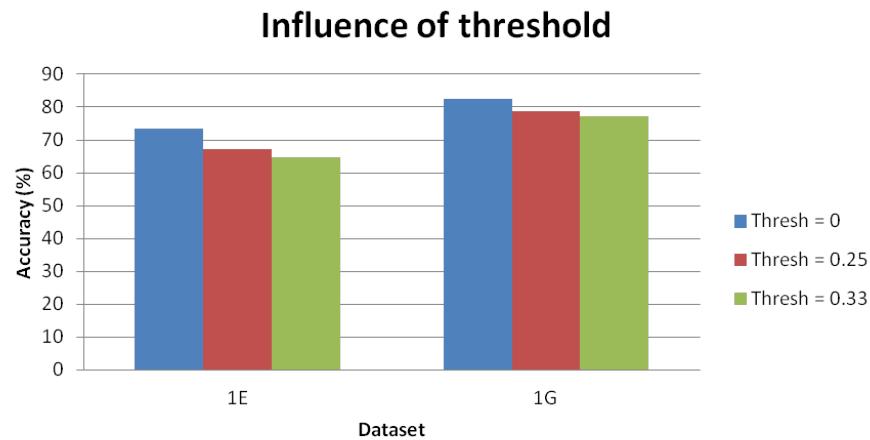
#### 6.4.2 Influence of threshold

The threshold determines which values, and thus samples, are classified as right or as left MI, or not classified at all. The classification accuracy for three different thresholds and both dataset 1E and 1G are shown in Figure 6.9.

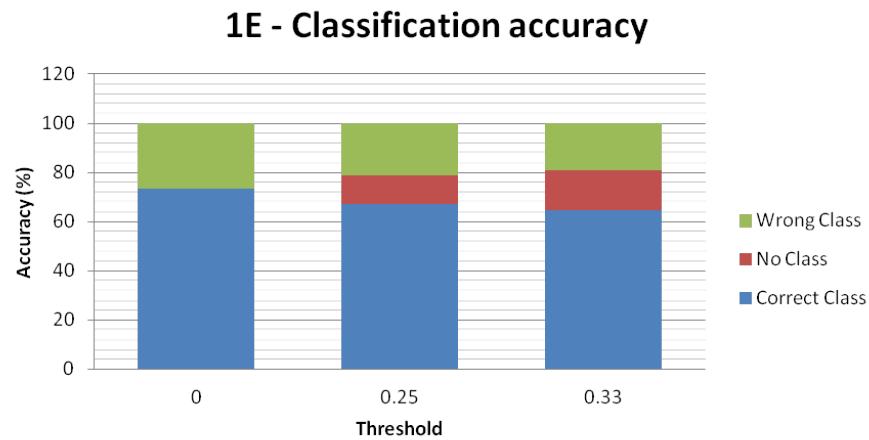
The first thing we notice is that there is a decrease in accuracy if the threshold increases. As we explained earlier, in the case of a threshold equal to 0, all samples are classified. In the case of a threshold equal to 0.25, the values from -1 to -0.25 are classified as left MI, and the values from 0.25 to 1 are classified as right MI. Analogue for a threshold of 0.33. So the decreasing accuracy is the direct consequence of the fact that just less samples are classified.

From this Figure 6.9 we can't deduce what the effect is of a change in threshold on the amount of actual wrong classified samples. Therefore we also plotted Figure 6.10. This figure makes a distinction between the amount of wrongly classified samples, the amount of samples that are not classified at all, and the correctly classified samples for different thresholds. Here we see that the amount of samples that are no longer classified in Figure 6.9 were equally often

correctly and wrongly classified. This is in contradiction what one had hoped, namely that by changing the threshold the amount of wrongly classified samples would decrease. Why would one want to offer overall classification accuracy for a lower rate of wrongly classified samples? In some cases it might be worse to make a wrong decision than no decision at all. Choosing a higher threshold would then be equal to lowering the amount of false positives or false negatives. But apparently this is not the case here. But after all it is not that surprising, since we didn't build a classifier for the idle state.



**Figure 6.9:** Accuracy of evaluation dataset for different thresholds

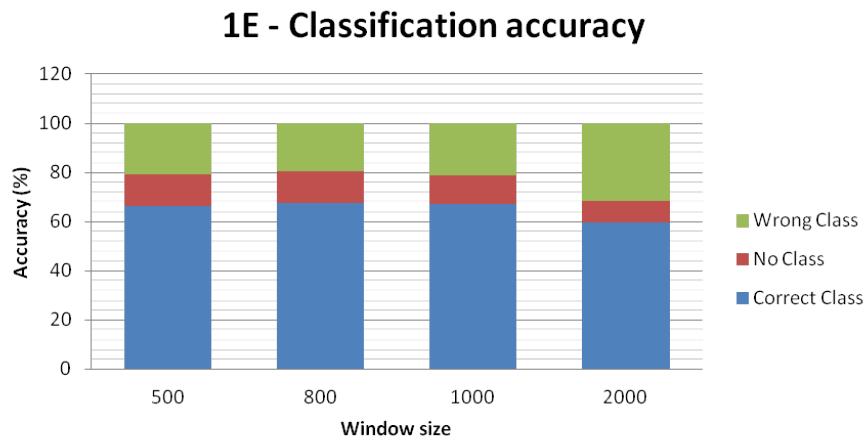


**Figure 6.10:** Accuracy of evaluation dataset for different thresholds

### 6.4.3 Influence of the window size

The window size is one of the parameters we have to determine when dealing with asynchronous BCIs. They will have an impact on the speed of the system: the shorter the window, the faster the BCI can make a decision regarding the class. So it will influence the ITR.[1] We vary the window size from 500 until 2000 samples, at a sampling frequency of 1000 Hz, this corresponds with a delay of 0.5 to 2 seconds between the sampling and the final task execution (not taking into account any further delays due to machinery or other software). Of course, the shorter the delay, the better. This is why this parameter is so important.

In Figure 6.11 the accuracy for dataset 1E is depicted using different window sizes. We see that this window size doesn't have much influence on the accuracy for the first 3 values they all are about 67%. However for a window size of 2000 samples, there is a remarkable decrease in the accuracy. So despite the higher amount of information the system has, it still classifies worse. This can be explained by considering the fact that we're dealing with an asynchronous BCI. When classifying a windowed EEG-signal, we're no longer sure the complete window corresponds to 1 single state. The larger the window, the bigger the chance is that there are samples of different states in the EEG-signal. Of course it is then impossible to classify the complete signal correctly. Even if the second present state is the idle state, one will lose a lot of information for classification. Therefore the accuracy will decrease with an increasing window size.

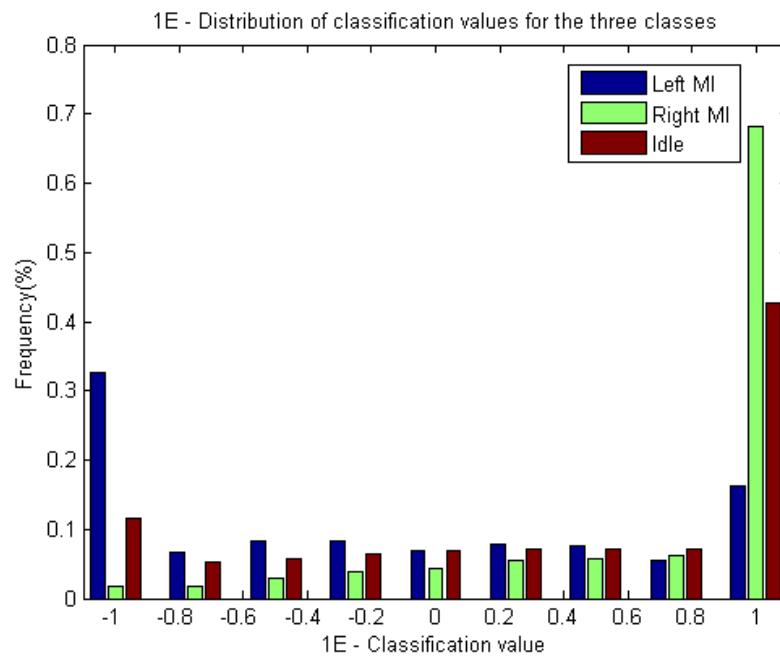


**Figure 6.11:** Accuracy of evaluation dataset for different window sizes

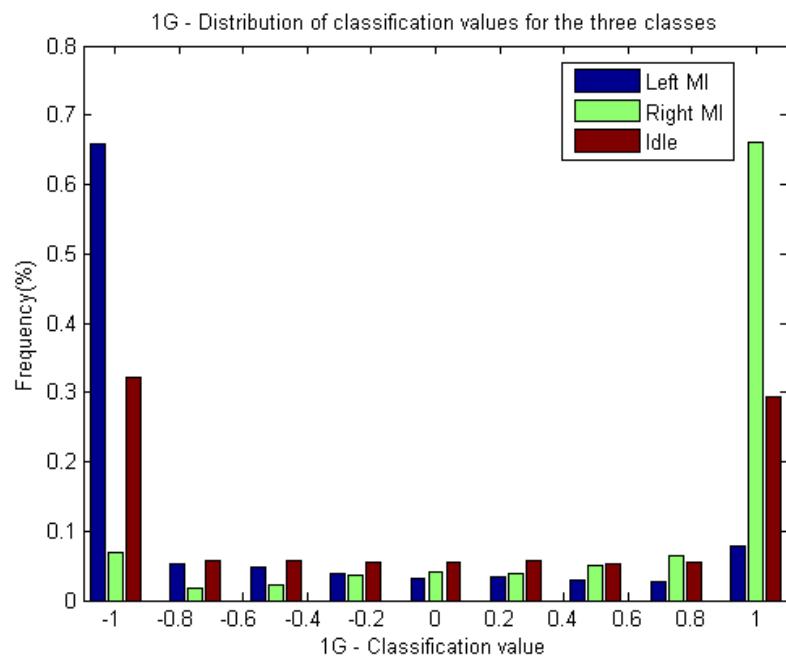
### 6.4.4 Idle state

Another reason for the use of thresholds was the idea that samples with a classification value of about 0 most probably would belong to the idle state. Since these samples should be

classified randomly, the classification values should average each other out. To see if this is a valid assumption, we plotted the histogram of the values corresponding with each class, see Figure 6.12 and 6.13. We don't see an obvious peak for the Idle state around the value of 0, so supposedly this is not a valid assumption. What we do see, especially for dataset 1G, is that the idle state is as often classified as left as as right. So there is a random classification of the idle state, but probably not at the level of the sample, but more likely at the level of the complete trial. This teaches us that if we want to be able to classify the idle state in the future, we'll have to deal with this as a separate class and not as a state 'between' left and right.



**Figure 6.12:** Histogram of distribution of classification values for right MI, left MI and idle state for dataset 1E



**Figure 6.13:** Histogram of distribution of classification values for right MI, left MI and idle states for dataset 1G

# Chapter 7

## Results: Influence of the head model

### 7.1 Introduction

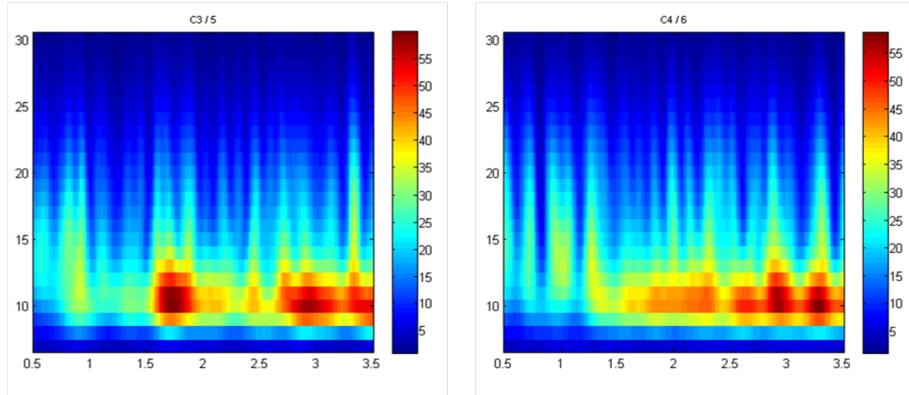
In the previous chapter we tried to optimize the system by investigating the influence of the different parameters of the system on the performance. We tested this system on the evaluation dataset and noticed that the accuracy didn't decrease. We can thus be quite sure we have developed a good working system.

We now want to use this system to investigate what the influence of the head model is on the performance. Therefore we're going to use the previous system and the dataset as defined in Section 4.1.2. The head models have also been described in Section 4.2. The models we use are a single shell spherical head model, a 3-shell spherical head model, a template BEM-model and a personalized BEM-model, based on MRI recordings. Finally we're going to combine the personalized BEM-model with a personalized ROI, based on fMRI fingertapping experiments. Every model has an increased degree of complexity. The question now is if the accuracy benefits this increasing complexity.

The main objective from the previous chapter was to define the optimal parameters for the system. However, some parameters are subject or model dependent and will have to be adjusted:

- **Frequency band:** We know that the frequencies in which the MI is most prominent present differs from person to person. Therefore we already decided from the beginning that this parameter would be subject-specific. To determine the bandpass frequencies we take all the left and right trials in the new dataset and perform a time-frequency analysis. The result can be seen in Figure 7.1. We'll limit the bandpass filter to 9 to 13 Hz.

- **Amount of grid points:** In the previous chapter we chose a fixed amount of grid points for all the subjects. Now we're going to adapt this amount to the model used. The reason for this is that as discussed in Section 4.2 the amount of grid points and the grid point spacing differs from model to model. So the same amount of grid points corresponds with a different area. Since we use one dataset, the active area should be the same over the different head models, and thus the amount of grid points has to be variable. But we do have to remark that even if the grid points cover the same area, choosing a smaller grid spacing might influence the accuracy of the solution. There is namely a trade-off between the ability to resolve sources that are close together and the possibility of introducing spurious sources. In other words, if the grid spacing is too big, separate sources will be considered one source.



**Figure 7.1:** Time-frequency analysis between 7 and 30 Hz, left: for left trials at C4, right: for right trials at C3

## 7.2 Description

In this section we keep the data constant, the only thing that changes is the head model and the amount of grid points used for feature extraction. The head models used differ in shape and complexity. For the personalized BEM model we can add some extra features. Once we add a predefined dipole orientation. Here one assumes the dipole orientations are perpendicular at the cortex (see Section 1.2.1). We also have fMRI data of the same subject, see Section 4.2.3. This can be used to locate more specific ROIs.

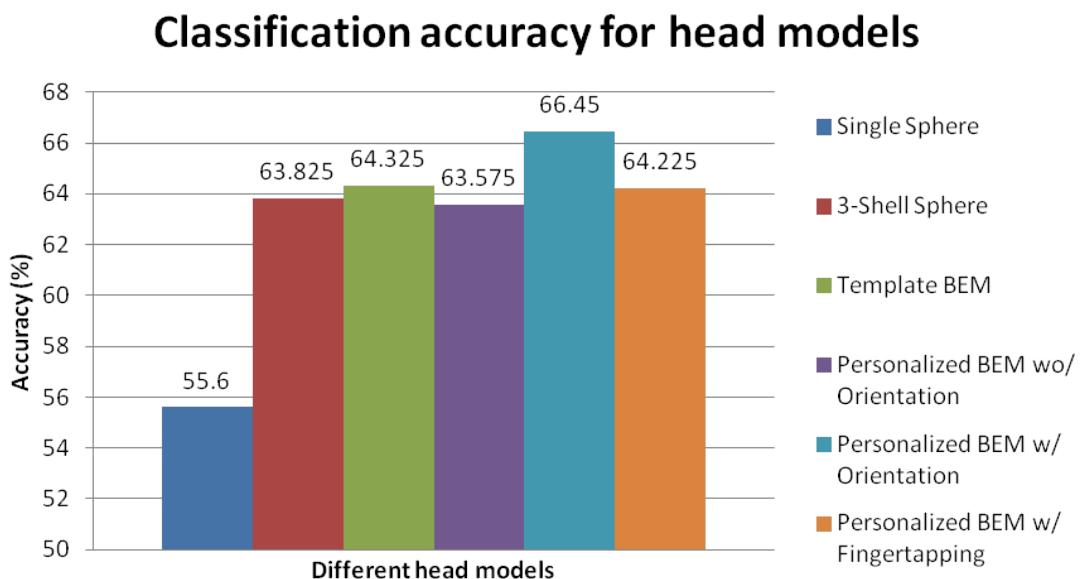
## 7.3 Results

In Figure 7.2 we see the classification accuracy for the different head models. The general trend here is an increasing accuracy with an increasing complexity of the model. But we can also notice some more specific characteristics.

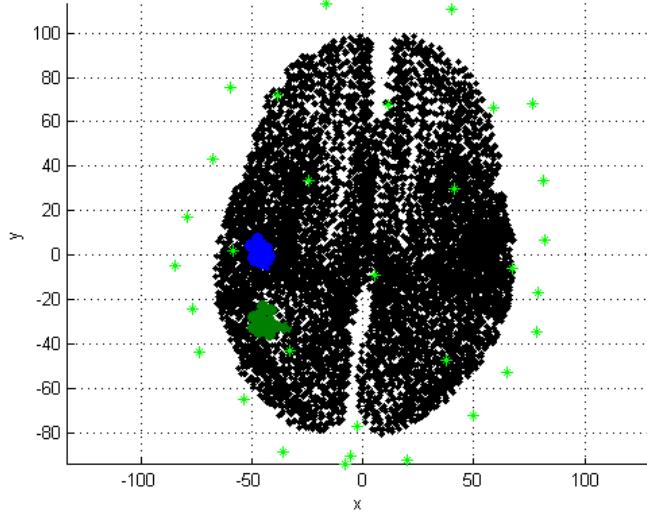
- **Single-shell spherical head model:** For the single shell spherical model, we see that the accuracy is about 10% lower than for the other methods. This can be explained by the fact that this model is extremely simple. It only takes into account one conductivity, thereby neglecting the much lower conductivity of the skull.
- **Three-shell spherical head model:** The three-shell spherical head model already performs much better, and (almost) equally performs as the realistic shaped head models. The main conclusion we thus can draw is that it is far more important to have a realistic model of the conductivities than of the head shape itself. This is due to two factors. The assumed conductivity of the different shells will determine how one assumes the signals will attenuate and spread towards the scalp. It is thus impossible to do realistic forward calculations, without more or less correct assumptions regarding the different layers of the head. The second factor is that the brain may not be spherical, but still some parts of the brain come closer to a spherical shape than others. The sensorimotor cortex is a part which can quite good be approximated by a sphere. Therefore the influence of the difference in shape between the template BEM model and the spherical model will be rather limited.
- **Template BEM model:** This model performs better than the 3-shell spherical head model, and even outperforms the personalized head model without predefined dipole orientation. It is quite logic that it is better than the spherical head model, since it is more sophisticated. And even if it is an averaged model, it probably comes closer to the real model than the spherical head model.
- **Personalized BEM model without predefined dipole orientation:** One would expect this model to perform better than the not personalized BEM models. Especially since these also don't take specific dipole orientation into account. The cause of the slightly lesser performance might be found in the grid model which is less dense for the realistic BEM model than for the template BEM model. This might cause the source reconstruction to be less accurate. The difference in accuracy between this model and the 3-shell spherical model is almost negligible. A direct explanation why they are so close together is not available.
- **Personalized BEM model with predefined dipole orientation:** This model performs the best of all models. It takes both the real shape of the head and the real direction of the dipoles into account. This makes for a very accurate lead field matrix, which is obviously reflected in the final performance.
- **BEM model with fMRI information:** In the other models we choose the ROI based on the fact that regions active during MI lay generally in the close proximity of the C3 and C4 electrodes. However as we stated during the problem description, we're

dealing with people and thus we always have to deal with inter-subject differences. By now including person specific information about the MI active regions, we endeavour to respond to these differences. So one expects to have a lower classification error if such information is included. However here we see a decrease in accuracy. This can be due to an electrode cap which is placed slightly incorrect. Hereby we address immediately a non-negligible problem: namely that the fMRI data and the EEG data are acquired with different methods and mostly at different moments. This complicates the alignment between the different datasets and compromises the advantages. Another possible cause could be that the region activated during actual finger tapping might not be exactly be the same to the region active during motor imagery. The difference between the ROI used for the other BEM models and for this model is depicted in Figure 7.3.

We see that even for the best performing system the accuracy doesn't exceed 67%. One can argue this is quite low. This is indeed not that high. However we have to remark here that first of all this is a dataset from an untrained subject. The dataset was also already analysed in another thesis. Here they obtained similar accuracy values for the classification between left and right hand motor imagery. Another problem is that we only have 50 trials per class. This then has to be divided in a training and a testing set, leaving us with a rather small training set. This definitely won't be in favour of the classification accuracy.



**Figure 7.2:** Classification accuracies for different head models



**Figure 7.3:** Black: Grid points of personalized head model, Blue: standard chosen ROI (closest to C3), Green: ROI based on fingertapping information

## 7.4 Remarks: Amount of electrodes and position of the electrodes

As discussed in Chapter 3, the forward model is partly also determined by the electrodes. This means that these will also have an influence on the accuracy. The more precise one knows the position of the electrodes, the better the forward model will resemble the reality and the better the classification will be. There already have been experiments where EEG and fMRI are performed simultaneously.[24] This has the advantage that one can know the exact position of the electrodes with respect to the head. If this is performed simultaneously with fMRI the alignment between the fMRI data and the electrode positions can be measured exactly. This will most likely increase the accuracy.

Another factor we want to address is the amount of electrodes used. We want to point out that the dataset used in this section was only recorded from 31 electrode positions, whereas the BCI Competition IV data was recorded from 59 electrode positions. Laarne et. al. investigated the effect of EEG electrode density on dipole localization accuracy and concluded that an increasing number of recording electrodes seems to improve the localization accuracy in the presence of noise. [15]. So the rather low amount of electrodes used here, might also partly explain the low classification accuracy.

# Chapter 8

## Conclusion

At the beginning of this dissertation, we listed three main objectives. First of all we wanted to design a robust BCI system which could correctly classify between two different states, namely right and left hand MI and which uses beamformers as spatial filters. We believe we herein succeeded. We investigated the influence of different parameters on the system and found some interesting results:

- During the preprocessing step it is interesting to perform baseline correction to take into account changes in noise level. However in order to do this, we need for each trial an estimate of the noise level at that time. Theretofore pretrial data is used. In the case of asynchronous data we don't have this kind of information, and thus baseline correction is impossible to apply.
- Based on what we saw for 2 subjects, it seems to be also beneficiary for the system to completely reject trials contaminated with artefacts. Despite the loss in data, the performance will increase.
- The final important conclusion is that the amount of grid points (and thus the amount of features) used for classification is one of the most determining parameters for the accuracy.

After optimization of the different parameters we tested the system on some new datasets. The accuracies found for these new datasets varied from pretty low ( $\pm 52\%$ ) to quite good ( $\pm 70\%$ ). So even for the best performing dataset, the classification accuracy was still about 10% lower than for the datasets used for the optimization process. These differences in classification accuracy can have many causes, such as the fact that the worst performing dataset was an artificial dataset. But most likely this difference is at least partly due to the fact that the system wasn't optimized specifically for these datasets. Therefore it seems appropriate to optimize the complete system per subject.

In the second part of this discourse we tested the system on an asynchronous dataset. The results were dependent on the threshold and on the window size, but in general we could state that the system performed really good. The most interesting part here was that there was only a limited decrease in classification accuracy for decreasing window sizes. This is extremely important since the window size is the most determining factor for the delay between the user's thoughts and the activation of the external device. We can thus conclude that beamformers are a good choice for real-life applications.

Our last objective was to investigate the influence of the use of different head models and the incorporation of personalized fMRI data on the final accuracy of the system. This is extremely interesting, since inter-subject differences are one of the main problems in BCI research. With respect to the head models we can conclude that the more complex the head model becomes, the better the performance. Personalized head models indeed outperform the template head model. The best results were obtained for a personalized head model which assumed dipoles orientated perpendicular to the cortex. Including personalized fMRI data about the active regions during motor tasks, didn't add much to the performance. This could be caused by deviations in alignment between the different modalities, or by a deviation between the active region for finger tapping and the region activated during hand MI.

If we now compare our results with the existing literature about the influence of head models on the performance of BCI systems based on beamformers, we see that the conclusions drawn are similar. With respect to the influence of head models on EEG simulations Ramon et al. state that the complexity of head models strongly influences the scalp potentials and the inverse source localizations. A more complex head model performs better in inverse source localizations as compared to a model with lesser tissue surfaces. [23] This is confirmed by our results, which show an increasing accuracy if we go from a single shell spherical head model to a realistic template BEM-model. Research of the influence of head models specifically in combination with beamforming based BCIs was already conducted by Murzin et al. They also concluded that anatomically constrained beamforming procedures along with realistic geometry and conductivity-based forward solutions provides a viable and quite accurate approach to the inverse problem in EEG. [19] The main difference between their research and this master's thesis is that they didn't use realistic datasets in their dissertation, but used simulated signals from point sources at well-defined locations. The combination however of beamforming techniques with personalized datasets and personalized head models was to date not found in the literature by the author, and the final conclusions therefore can't be compared.

Concerning the classification accuracies it is hard to find literature that really can be compared with what we did in our dissertation. For the results of the BCI Competition IV dataset

we can address ourselves to the results from participants of the competition. However these are for the evaluation datasets and expressed in a value for the mean squared error, instead of classification accuracy. But what we can deduce from these results is that in general subjects 1e and 1g performed best, and subject 1c was the worst subject.[6] So this indicates that the difference in accuracy we obtained between subjects 1e and 1g, and the other subjects was not solely because of the optimization process. The data recorded at Ghent University was already used by Strobbe et al. They obtained with the application of Common Spatial Pattern (CSP) filters an accuracy of maximum 65% for this dataset. [28] So the personalized BEM head model with predefined dipole orientations in combination with beamforming (66.5%) slightly outperforms CSP filtering. Grosse-Wentrup et al. also compared beamforming with CSP filters and laplacian spatial filters. The accuracies over the different subjects ranged for beamforming from 53% to 93%. They concluded that the optimal filter method was subject dependent. Except for if the EEG data is extremely noisy, then beamformers will substantially outperform CSP filtering. [9]

# Chapter 9

# Discussion and Future Work

## 9.1 Introduction

Since it was the first time for me to work with BCI systems, I had a lot to discover. My first goal was to reach the objectives stated in Chapter 2. Throughout this dissertation we proved that we indeed reached those objectives. We built a good working BCI system, which also seems to be applicable in real-time. Also the possibilities of using personalized head models were researched. However during the development of the system, often problems were encountered which couldn't be resolved immediately. But over time I thought about or read about a solution for some of them. In the next paragraphs, I list some of the problems and their possible solutions.

## 9.2 Building an accurate BCI System

In the first part of this dissertation we wanted to build an accurate system. The resultant system was successful but quite simple. So here we propose some ideas which might make the BCI even more robust and more accurate. We discuss the possible improvements per step in the translation process.

### Preprocessing

In the current system we use a bandpass filter to restrict the data to the frequency band active during MI. We determine the bandpass frequencies by visual inspection of the time frequency analysis of the average over all trials of a state. However the optimal frequency band might vary over time. Therefore it might be interesting to use adaptive band pass filtering. The signal will then always be filtered at the best possible frequency band. Especially for real-time purposes this might be interesting. [14]

### Feature extraction

The parameter with the biggest influence on the classification performance is the number of grid points used as input for the feature vector. An optimal selection of grid points is therefore necessary. We propose two methods which may possibly improve the grid point selection.

First of all, if no fMRI data available from the subject, one could perform an R-squared analysis on the training data. This is a statistical technique which quantifies how significant the difference is between the average band powers of the two states. This can be calculated per electrode. By visualizing the results, one gets a better image of where the active brain regions appear. If they appear to be moved away from electrodes C3 and C4, it might be interesting to also adjust the ROIs.[28, 18]

Another way to obtain a better feature vector is to make a difference between the number of grid points and the number of features used for classification. Based on the signals at the grid points in the ROI, one can construct a first feature vector. Next, one can search for an optimal subset of features. There exist different methods to select such a subset, such as t-tests, wrapping methods,... [25, 7]

### Classification

At this point we use quite a straightforward classification algorithm in the system, namely Fisher LDA. We pointed out earlier why we think it is interesting to use this classifier. However one could also apply Regularized Fisher LDA. This is almost the same classification procedure, but it introduces a regularization parameter that can allow or penalize classification errors on the training set. This makes the algorithm better equipped to deal with outliers, which are quite common in EEG. [17]

## 9.3 Real-time classification

The second part of this dissertation was addressed to real-time classification and its problems. We noticed that the BCI system performed really good in real-time, even if only a relative small part of the signal was used (about 0.5-0.8 sec). Since we know that the system seems to work for an asynchronous dataset, the next step could be to apply it in real life. Then feedback to the user can be applied, which should further increase the performance of the system.

For real-time classification it might also be interesting to recognize a third state, namely the idle state. We saw that we cannot distinguish the idle state in a dataset based on the filters for right and left hand MI. Therefore the third class has to be defined explicitly by its own filters and own classification parameters.

## 9.4 Head models

In the last part of this dissertation we tried to address the inter-subject dependencies of BCI systems by using more sophisticated and personalized head models, even in combination with personal neurophysiological data.

The personalized head model was a BEM model, based on three different layers. However it might be more interesting to use an FDM model in the future. This is a numerical method for approximating the solutions of differential equations using finite difference equations to approximate the derivatives. In previous research it was already shown that this is a solid method to use in forward modelling. The use of FDM might be preferred, because it is known to give a better image of the cortical structures than BEM. It can also include more than three layers, unlike the BEM-models created in FieldTrip. By incorporating more layers, one makes the model more realistic and thus the forward solution more accurate. This should also lead to an increased accuracy of the system. Ramon et al. state that the CSF layer plays an important role in modifying the scalp potentials and also influences the inverse source localizations. [23]. Since this layer was thus far neglected in our models, FDM models are most probably a solid choice.

While discussing the head models in Chapter 4, we always mentioned the number of grid points used in this model. The number of grid points used for the calculation of the filter ranged from about 1700 for the spherical model to more than 20000 for the template BEM model. This difference in amount of grid points might induce a bias in the performance of the systems. For the spherical model the grid spacing will be rather large. The larger the grid spacing is, the less accurate the source reconstruction will be. However if the grid spacing is really small, the source reconstruction might be more prone to overfitting. It could be interesting to investigate what the exact influence is of the number of grid points used on the final classification accuracy. The influence of the grid points was also investigated by Fuchs et al. They concluded that more than 500 nodes per compartment are needed for reliable BEM models. [8]

# Bibliography

- [1] *Toward Brain-Computer Interfacing*. Massachusetts Institute of Technology, 2007.
- [2] *Brain Signal Analysis : Advances in Neuroelectric and Neuromagnetic methods*, chapter A Practical Guide to Beamformer Source Reconstruction for EEG. MIT Press, 2009.
- [3] *Brain-Computer Interfaces: Revolutionizing Human-Computer Interaction*, chapter An introduction to brain - computer interface systems. Springer Publishing, 2011.
- [4] B.Z. Allison, R. Leeb, C. Brunner, G.R. Mller-Putz, G. Bauernfeind, J.W. Kelly, and C. Neuper. Toward smarter bcis: extending BCIs through hybridization and intelligent control. *Journal of Neural Engineering*, 2012.
- [5] P. Belardinelli, E. Ortiz, and C. Braun. Source activity correlation effects on lcmv beamformers in a realistic measurement environment. *Computational and Mathematical Methods in Medicine*, 2012.
- [6] B. Blankertz, G. Dornhege, M. Krauledat, K.R. Mller, and G. Curio. The non-invasive berlin brain-computer interface: Fast acquisition of effective performance in untrained subjects. *NeuroImage*, 2007.
- [7] J. Fruitet and M. Clerc. Reconstruction of cortical sources activities for online classification of electroencephalographic signals. *Conf Proc IEEE Eng Med Biol Soc*, 2010:6317–6320, 2010.
- [8] M. Fuchs, M. Wagner, and J. Kastner. Boundary element method volume conductor models for EEG source reconstruction. *Clin Neurophysiol*, 112(8):1400–1407, Aug 2001.
- [9] M. Grosse-Wentrup, C. Liefhold, K. Gramann, and M. Buss. Beamforming in noninvasive brain-computer interfaces. *IEEE Trans Biomed Eng*, 56(4):1209–1219, Apr 2009.
- [10] <http://brainscraper.com/>.
- [11] <http://en.wikipedia.org/wiki/Electroencephalography>.
- [12] <http://www.bem.fi/book/13/13x/1302x.htm>.

- [13] <http://www.ias.informatik.tu-darmstadt.de/Research/BCIRoboticsRehabilitation>.
- [14] V. Jeyabalan, A. Samraj, and Kiong L.C. Motor imaginary signal classification using adaptive recursive bandpass filter and adaptive autoregressive models for brain machine interface designs. *International Journal of Biological and Life Sciences*, 2007.
- [15] P. H. Laarne, M. L. Tenhunen-Eskelin, J. K. Hyttinen, and H. J. Eskola. Effect of EEG electrode density on dipole localization accuracy using two realistically shaped skull resistivity models. *Brain Topogr*, 12(4):249–254, 2000.
- [16] S. Lemm, K. R. Muller, and G. Curio. A generalized framework for quantifying the dynamics of EEG event-related desynchronization. *PLoS Comput. Biol.*, 5(8):e1000453, Aug 2009.
- [17] F. Lotte, M. Congedo, A. Lecuyer, F. Lamarche, and B. Arnaldi. A review of classification algorithms for EEG-based brain-computer interfaces. *J Neural Eng*, 4(2):R1–R13, Jun 2007.
- [18] K.-R. Müller, M. Krauledat, G. Dornhege, G. Curio, and B. Blankertz. Machine learning techniques for brain-computer interfaces. *BIOMEDICAL ENGINEERING*, pages 11–22, 2004.
- [19] V. Murzin, A. Fuchs, and J. A. Kelso. Anatomically constrained minimum variance beamforming applied to EEG. *Exp Brain Res*, 214(4):515–528, Oct 2011.
- [20] L. F. Nicolas-Alonso and J. Gomez-Gil. Brain computer interfaces, a review. *Sensors (Basel)*, 12(2):1211–1279, 2012.
- [21] R. Oostenveld, P. Fries, E. Maris, and J. M. Schoffelen. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci*, 2011:156869, 2011.
- [22] G. Pfurtscheller, C. Guger, G. Müller, G. Krausz, and C. Neuper. Brain oscillations control hand orthosis in a tetraplegic. *Neuroscience Letters*, 292(3):211 – 214, 2000.
- [23] C. Ramon, P. H. Schimpf, and J. Haueisen. Influence of head models on EEG simulations and inverse source localizations. *Biomed Eng Online*, 5:10, 2006.
- [24] P. Ritter and A. Villringer. Simultaneous EEG-fMRI. *Neurosci Biobehav Rev*, 30(6):823–838, 2006.
- [25] Y. Saeys, I. Inza, and P. Larrañaga. A review of feature selection techniques in bioinformatics. *Bioinformatics*, 23(19):2507–2517, Oct 2007.
- [26] S. J. Smith. EEG in neurological conditions other than epilepsy: when does it help, what does it add? *J. Neurol. Neurosurg. Psychiatr.*, 76 Suppl 2:8–12, Jun 2005.

- [27] M. B. Sterman. Physiological origins and functional correlates of EEG rhythmic activities: implications for self-regulation. *Biofeedback Self Regul*, 21(1):3–33, Mar 1996.
- [28] G. Strobbe and R. Van de Vaerd. Geschiktheidsstudie van connectiviteitsmaten in het electroencephalogram voor brain-computer interfaces. Master’s thesis, Ghent University, 2010.
- [29] B. D. Van Veen, W. van Drongelen, M. Yuchtman, and A. Suzuki. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans Biomed Eng*, 44(9):867–880, Sep 1997.
- [30] S. VandenBerghe, H. Hallez, R. Declerck, and J. Cornelis. Biomedical signals and images.

# List of Figures

1.1	Basic block diagram of a BCI system incorporating signal acquisition, processing and deployment [10] . . . . .	2
1.2	Three different ways to detect the brain's electrical activity: EEG, ECoG, and intracranial recordings . . . . .	3
1.3	Left: Functional areas of the brain: the cortical area particularly important for motor imagery based BCIs are the motor are and the somatosensory cortex, Right: Pyramidal cell . . . . .	5
1.4	Electrode placement as defined by the International 10-20 system. The names of the electrodes are based on the area of the cerebral cortex corresponding with the electrode position, i.e. Frontal, Central, Parietal and Occipital. Electrodes on the left hemispheres are given odd numbers, whereas electrodes on the right hemisphere are given even numbers. . . . .	5
1.5	ERD and ERS . . . . .	7
1.6	General preprocessing steps . . . . .	8
1.7	Feature Extraction steps . . . . .	9
1.8	Signal Classification steps . . . . .	10
1.9	Relation between forward and inverse problem of EEG source localization . .	11
1.10	Relationship between BCI application areas, BCI information transfer rates and user capabilities . . . . .	13
1.11	BCI and virtual reality . . . . .	14
1.12	BCI as interface for hand orthosis . . . . .	14
2.1	Left and right hand motor imagery . . . . .	18
3.1	Elements determining lead field matrix for forward problem . . . . .	22
3.2	Effect of noise on beamformer source reconstruction. (a) The average ERP, timefrequency plot, and beamformer output for a pair of bilateral sources in auditory cortex for data with a high signal-to-noise ratio. (b) The same signal with increased background noise. Information about the activity of interest is difficult to extract from the background noise, resulting in lower power in the timefrequency domain and spurious sources in the beamformer output . . . . .	26

4.1	Acquisition protocol for BCI competition IV calibration datasets . . . . .	28
4.2	Acquisition protocol for BCI competition IV evaluation datasets . . . . .	29
4.3	Acquisition protocol for own recorded dataset . . . . .	30
4.4	Schematic view of the 3-shell spherical head model . . . . .	31
4.5	BEM model of 3-shell spherical head model . . . . .	32
4.6	Schematic view of the 3-shell realistic head model . . . . .	33
4.7	BEM model of standardized head model, up: faces, down: edges, green cross = electrode position . . . . .	33
4.8	Determination of left ROI based on fMRI data . . . . .	35
4.9	Determination of right ROI based on fMRI data . . . . .	36
5.1	From continuous data to trials separated per condition . . . . .	38
5.2	Time-Frequency analysis between 7 and 30 Hz, left: for right MI trials at C3, right: for left MI trials at C4 for dataset 1g . . . . .	39
5.3	Trial selection for training of beamformer filter . . . . .	41
5.4	Trial selection for training of beamformer filter . . . . .	42
5.5	Right: Combination of averaged signals at 11 nearest grid points at C3 (blue) and C4(red) for right hand MI. Left: Combination of averaged signals at 11 nearest grid points at C3 (blue) and C4(red) for left hand MI. We indeed see a decrease of the signal power at the contra-lateral electrode. . . . .	43
5.6	Example of Fisher LDA for a two dimensional feature vector . . . . .	44
6.1	The influence of baseline correction and artefact removal on the accuracy . .	52
6.2	The influence of the amount of trials for training of the filter on the accuracy	53
6.3	From continuous data to trials separated per condition . . . . .	53
6.4	Influence of the time window on the classification accuracy . . . . .	54
6.5	Influence of the number of grid points / features used on the classification accuracy . . . . .	54
6.6	Influence of the trial size on the classification accuracy . . . . .	55
6.7	The influence of the training percentage on the classification accuracy . . .	56
6.8	The final accuracy results for both training and testing set for all BCI Competition IV . . . . .	57
6.9	Accuracy of evaluation dataset for different thresholds . . . . .	60
6.10	Accuracy of evaluation dataset for different thresholds . . . . .	60
6.11	Accuracy of evaluation dataset for different window sizes . . . . .	61
6.12	Histogram of distribution of classification values for right MI, left MI and idle state for dataset 1E . . . . .	62
6.13	Histogram of distribution of classification values for right MI, left MI and idle states for dataset 1G . . . . .	63

<i>List of Figures</i>	80
------------------------	----

7.1 Time-frequency analysis between 7 and 30 Hz, left: for left trials at C4, right: for right trials at C3 . . . . .	65
7.2 Classification accuracies for different head models . . . . .	67
7.3 Black: Grid points of personalized head model, Blue: standard chosen ROI (closest to C3), Green: ROI based on fingertapping information . . . . .	68

# List of Tables

4.1	Specifications 3-shell spherical model . . . . .	31
5.1	Example of confusion matrix for a two class problem . . . . .	47
6.1	The specifics of the available and used data sets . . . . .	50
6.2	Parameters list: standard values and possible values . . . . .	50
6.3	Confusion matrix - Standard parameters - 1E . . . . .	51
6.4	Confusion matrix - Standard parameters - 1g . . . . .	51
6.5	Parameters list: final values . . . . .	56

