Feasibility of using Error-related potentials as an appropriate method for adaptation in a brain-computer interface

Project group: 17gr10405
Johannes Thorling Hadsund & Kasper Kunz Leerskov

Aalborg University
10. semester Biomedical Engineering and Informatics
School of Medicine and Health

AALBORG UNIVERSITET
STUDENTERRAPPORT

AALBORG UNIVERSITY
STUDENT REPORT

School of medicine and health
Biomedical Engineering and Informatics
Niels Jernes Vej 12, A5
9220 Aalborg Øst
Telefon 9940 8752
http://www.smh.aau.dk/

Title:
Feasibility of using Error-related potentials as an appropriate method for adaptation in a brain-computer interface

Theme:
Master Thesis

Project period:
6th of February 2017 -
7th June 2017

Project group:
17gr10405

Participants:

Johannes Thorling Hadsund

Kasper Kunz Leerskov

Supervisor:
Mads Jochumsen

Pages: 59

Concluded: 07/06/17

Synopsis:

Introduction: Error related potentials are evoked in response to erroneous prediction and actuation by a BCI. Detection of error related potentials provides prediction-independent means for guiding adaptation, and potentially enables more accurate adaptation and means for re-separation of partially confused classes. In this work it was investigated if error related potentials can be detected independently of the misclassified user intention. *Method:* Error and non-error related potentials were provoked in 12 subjects during three different paradigms: Overt motor, covert motor and a game. EEG was recorded from 16 electrodes and divided into epochs containing one second of data post onset of the visual feedback. For each group a classifier was trained with spectral and temporal features to detect error and non-error related potentials. Furthermore each group was classified by a classifier trained using data for the two other groups. *Results:* For the three groups the area under the ROC was 0.92 % for the covert motor, 0.86 % for the overt motor, and 0.68 % for the game. When classifying the groups using the two other groups for training, accuracy changed from to 72% to 74%, from 81% to 77% and from 56 to 60% for the overt, the covert and the game paradigm, respectively, compared to the classification of the single groups. *Discussion:* This study demonstrated that error related potentials evoked during three different paradigms of mental tasks could be comparably accurately classified after training the classifier with either data from the same paradigm as the test data or data from two other paradigms. This strongly suggests that some properties of error related potentials are independent to the misclassified user intention.

Preface

Reading guide

References were made according to the Harvard-method: [Author's surname, year of publication]. I.e. [Jensen2011]. References after a full stop accounts for the previous text section, and references placed before the full stop, accounts only for the previous sentence. Frequent words and phrases are abbreviated after first use. Abbreviations is indicated in parenthesis upon first use of the phrase.

The report is structured in accordance with IMRAD; introduction, methods, results and discussion. However, traditionally the introduction of AAU projects is split into a short introduction of the context of the project, and a longer analysis of a specific problem within this context. The problem analysis serves to specify the problem, and to identify potential solutions. The problem analysis is concluded with a problem statement or a hypotheses.

Table of contents

Dansk resume

Introduktion

Patienter med stærkt nedsat kontrol over bevægeapparatet er afhængige af hjælpemidler for at kunne færdes og begå sig socialt. Electroencefalografi (EEG) er en realtivt billig og udbredt målemetode der gør det muligt at designe grænseflader mellem en patients hjerne og en computer (brain computer interface, BCI). Ved hjælp af et BCI kan brugeren styre eksemplevis en kørestol eller en computer, og dermed genvinde tabt førlighed. Ved brug af mentale strategier, såsom at tænke på bevægelser eller ord, kan brugeren manipulere med EEG'et og dermed sende instruktioner til en computer. Efter opsamling af EEG'et behandles signalet digitalt for at frasortere støj og for at identificere signalegenskaber(features) der sætter systemet i stand til at klassificere brugerens EEG som en given knotrolklasse blandt et sæt af kontrolklasser. EEG'et påvirkes imidlertid ikke kun af de mentale strategier brugeren bevist gennemfører, men også af hjernens øvrige aktivitet og brugerens omverden. For at imødegå skift i EEG'ets karakteristika bør et BCI kunne tilpasse sig EEG'et løbende. Hvis systemet tilpasses ved naivt at træne systemt med udgangspunkt i klassifikationen af EEG'et, er der risiko for at klassifikationsfejl akkumuleres og på sigt hæmmer muligheden for at klassificere nyeligt målt EEG korrekt. Det er blevet påvist at en BCI-brugers reaktion på en fejlklassifikation af en mental strategi, kan spores i EEG'et. Reaktionen betegnes typisk som et fejlpotential (Error related potential, ErrP). Detektion af fejlpotentialer vil gøre det muligt kun at tilpasse systemet til ændringer i EEG'et når brugeren ikke betragter systemets klassifikation af EEG'et som fejlagtig. Dermed bliver systemtilpasningen mere robust overfor fejlklassifikationer. Imidlertid forudsætter fejlpotentialedrevet tilpasning af systemet, at fejlpotentialer kan genkendes uafhængigt af hvilken klassifikationsfejl systemet begår.

Formålet med dette projet var derfor at undersøge om fejlpotentialer kan skelnes fra ikke-fejl uafhængigt af hvilken kontrolklasse der fejlklassificers.

Metode

Der blev opsamlet EEG fra 12 raske forsøgespersoner, der hver deltog i tre forskellige forsøgsparadigmer, der alle havde til formål at provokere fejlpotentialer hos brugerne efter fingeret fejlklassifikation af tre forskellige mentale strategier. De tre mentale strategier var coverte eller overte håndbevægelser og styring af en slange i spillet "snake". I alle tre paradigmer førte forsøgspersonernes mentale strategier til enten forventelig eller uforventelig respons på en computerskærm. EEG'et blev opdelt i epoker af et sekunds varighed. Hver epoke startede samtidigt med at der på computerskræmen blev vist respons på en af forsøgspersonernes mentale strategier. Fra hver epoke blev der udtrukket et sæt af tidslige og spektrale signalegenskaber til brug for træning af en LDA og en Random forest. For hvert af de tre paradigmer blev EEG'et klassificeret efter træning med data fra samme paradigme, og efterfølgende træning med data fra de to øvrige paradigmer.

Resultater

De mest betydende signalegenskaber blev fundet over den centrale del af cortex, i tidsrummet mellem 400 og 800 ms efter epokens start og de væsenligste signalegenskaber viste sig at være korrelation af en epoke til en model af energien i theta båndet. I alle tilfælde gav LDA baseret klassifikation bedre resultater end Random forest, hvorfor alle resultater blev basseret på LDA. Arealet under ROC-kurven for klassifikation af data fra de tre paradigmer var henholdsvis 0.92%, 0.86 % og 0.68% for det coverte paradigme, det overte paradigme og spilparadigmet. Klassifikationsnøjagtigheden var i spilparadigmet signifikant ($p = 0.05$) højere end tilfældighed, men væsenligt lavere end for det overte og det coverte paradigme. Ved træning og test med data fra samme paradigme var klassifikationsnøjagtigheden henholdsvis 72%, 81% og 56% for henholdsvis det overte paradigme, det coverteparadigme og spilparadigme. Efter træning med data fra de to øvrige paradigmer blev der opnået en nøjagtighed på 74% og 77% og 60% for henholdsvis det overteparadigme, det coverte paradigme og spilpradigmet. .

Diskussion

Klassifikationsnøjagtigheden for det coverte og det overte paradigme blev ikke væsenligt påvirket af om der blev anvendt træningesdata fra samme paradigme som testdataen eller træning og test data kom fra forskellige paradigmer. Klassifikation af fejlpotentiler afhænger derfor ikke af at det vides hvilken kontrolklasse der er blevet fejlklassificerret forud for fejlpotentialet.

Chapter 1

Introduction

People living with severe motor impairments, due to disease or trauma have very limited access to the outside world [Bi et al., 2013]. Brain-Computer-Interfaces (BCI) provides means for these patients to interact with assistive technology, giving them the possibility of communicating and controlling effectors such as wheelchairs, keyboards or computer cursors andthus regaining lost control and independence [Minguillon et al., 2017, Bi et al., 2013]. The BCI based human-machine interaction is achieved through monitoring and automated interpretation of the users brain activity, either through blood-oxygen based monitoring techniques such as fMRI or NIRS, or electrical monitoring techniques such as ECoG or EEG [Bi et al., 2013]. Due to low cost and convenience of use, EEG has been the most studied monitoring technique for use in BCI [Bi et al., 2013], and was therefore the focus in this study.

A generic BCI can be separated into three individual processes: Signal acquisition, Signal processing and Application (see figure 1.1) [Minguillon et al., 2017, Bi et al., 2013].

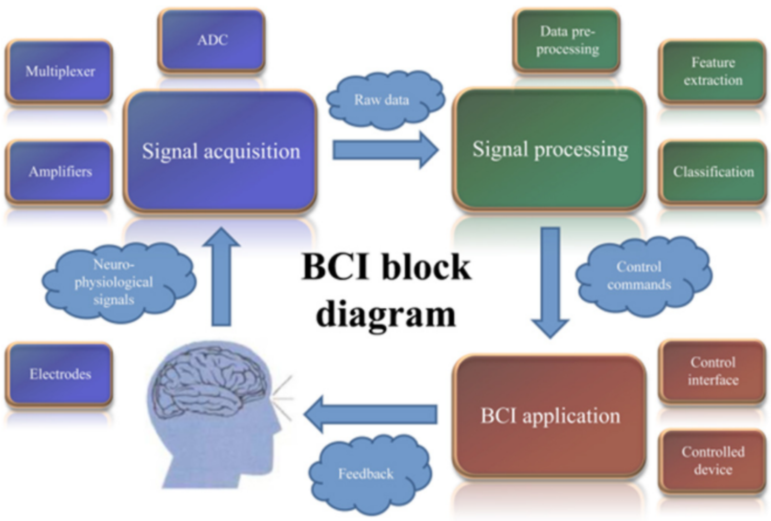


Figure 1.1: A BCI system consists mainly of three parts: Signal acquisition, Signal processing and BCI application. Each of these consists of several elements, all contributing to translating the cortical activity of the user into some action of the effector. From [Minguillon et al., 2017].

Signal acquisition

In EEG based BCI systems, EEG is recorded from electrodes, amplified and finally the electrical signal is converted to discreet values. Conventional BCIs use wet electrodes that need to be prepared and undergo daily maintenance with electrically conductive gel, to ensure sufficient impedance levels between electrode and the users scalp, for appropriate signal acquisition. The electrodes are usually connected to amplifiers through cables, which generally makes the system poorly suited for portable use. This is impractical when considering real-life applications, as many every day tasks requires mobility. Furthermore, preparation time of wet electrodes is time consuming, and usually requires support from technical staff. [Minguillon et al., 2017]

Signal processing

Signal processing in a BCI system concerns preprocessing of the EEG to remove noise and artifacts in the signal, feature extraction to characterize the recorded EEG and classification of the features to determine the intentions of the user, corresponding to actions in an assistive device [Minguillon et al., 2017]. A lot of different preprocessing methods, feature extractors and classifiers exist, with different pros and cons for specific applications. However, generally all elements of the signal processing should be able to run online, which sets demand for high computational efficiency [Minguillon et al., 2017]. During preprocessing, usually different types of filters are used to remove artifacts. However, conventional filters risks removing valuable information from the EEG. Independent Component Analysis (ICA), is another way of removing noise, by extracting independent components from the signal, such as artifacts. However ICA is difficult to use, complex and computationally heavy, which is why a lot of BCIs still use filters. [Bi et al., 2013] Extensive training of BCI users, feature extractors and classification algorithms, is necessary in order to reliably detect the users intentions. A way to minimize initial training is, to apply a pretrained subject-independent BCI [Vidaurre et al., 2011b]. However non-stationary changes in the EEG may cause the users mental actions to be poorly recognizable to the classification algorithms, ultimately compromising the classification accuracy. For that reason, BCI systems need to be regularly re-calibrated, which prevents independent long-term use of BCIs. [Bi et al., 2013]

BCI application

Finally, the BCI application is the actual translation of mental commands into actions by a given effector, such as moving forward in a wheelchair. [Minguillon et al., 2017] Different paradigms for controlling a BCI exists, such as P300, steady-state visually evoked potentials (SSVEP), event-related desynchronization/synchronization (ERD/-S) control [Bi et al., 2013] and slow cortical potentials (SCP) [Thomas et al., 2011].

P300 spellers detects the prominences of either auditory or visually evoked potential 300 ms after stimulations. The potentials are most prominent if the subjects attends the source of the stimulation. Thereby the user can control the system by directing her or his attention towards different sources of stimuli. [Mak et al., 2011, Lule et al., 2013] In another attention based paradigm the user is exposed to multiple sources delivering a constant visual stimulation at different frequencies. When attending one of the sources a steady state visual potential is evoked (SSVEP). The frequency of the signal can be determined from the EEG and as for the P300 paradigms the direction of the users attention can be determined. [Lee, 2012, Combaz et al., 2013].

However both of these paradigms need external stimuli, which is unsuitable for transportation or control type BCIs, as the users need to attend the stimulation from the BCI rather than the task at hand, and is synchronous. ERD/-S and SCP driven BCIs utilize components of EEG, which are modulated when the user performs different mental tasks and is therefore more suitable for a natural control paradigm. [Bi et al., 2013] Such self-paced BCIs will be the primary focus of this study.

Real-life BCI challenges

A problem exists generally for control BCIs, as the achievable information transfer rate (ITR) is insufficient to provide enough BCI commands to control advanced rehabilitation technologies. ITR can be optimized by lowering classification time, enhancing classification accuracy or enabling more classes in the BCI. [Tonet et al., 2008] Additionally, a very scarce effort has been made to test BCIs in a real-life setting, thus very little is known about BCI performance in real-life context [Minguillon et al., 2017].

BCI technology need a range of improvements before real-life applications are realizable [Bi et al., 2013]. This report will focus on enabling the signal processing part of the BCIs to be more robust during unassisted long-term use.

Mainly, the problem within this context, is changes in the EEG signal, which can be related to two things: Changes in artifacts present in the EEG and non-stationarity of EEG.

Artifacts in the EEG, are distortions of the signal caused by events or factors non-related to the users intentions, such as orofacial muscle activity (blinks, teeth clenching etc.), cardiac muscle activity and powerline noise. These artifacts needs appropriate handling as they are often of much higher amplitude than the relevant EEG signal. Furthermore, EEG is often studied in a controlled setting such as a lab, but a variety of noise sources exists in the “outside world” that have yet to be appropriately handled. This emphasizes the relevance of proper artifact rejection or correction. [Minguillon et al., 2017]

Non-stationarity can be defined, as a process that are neither strong nor wide-sense stationary, meaning that its statistical moments are dependent of time [Barkat, 2005]. During classification, events in the EEG are recognized as user intentions based on a model created from training data previously recorded. However non-stationarity implies that the model of intentions eventually will be insufficient for classifying newly recorded events, possibly even hours after recording the training set. [Li and Zhang, 2010] Eventually retraining of the BCI may be necessary to cope with the changes [Kilicarslan et al., 2016].

In a questionnaire study performed by Huggins et al. [2015] on people with spinal cord injuries, it was found that 77% of participants would accept a BCI performance of at least 80%, while 91% would accept a performance of at least 90% [Huggins et al., 2015]. This leaves very little room for performance decay, if it should not result in users rejecting the BCI. Therefore the focus of this report is to investigate the maintenance of classification accuracy, otherwise decaying due to non-stationarity of the EEG.

1.1 Initiating problem statement

Existing studies have shown that adaptation of BCIs may be utilized as a way to adress the issue of retraining of BCIs, due to non-stationarity of the EEG signal [Bi et al., 2013, Vidaurre et al., 2011a]. Hence it is of interest to investigate:

What adaptation methods exists, and how could adaptation be implemented to help realize a real-life BCI applications?

1.2 Literature search

In order to investigate the research regarding adaptation of selfpaced BCI systems, a literature search was conducted on the 9th of february, and a follow-up on error-related potentials the 6th of march. The literature search was done in PubMed and Embase using the search strings specified in table 1.1.

Table 1.1: Search strings used in the literature review, and the resulting hits.

Database	Search string	Hits
Embase 9/02/2017	(online OR live OR 'real time') AND (BCI OR 'brain computer interface' OR HMI OR HCI OR 'human machine interface' OR human computer interface OR 'brain machine interface') AND (EEG OR 'electroencephalography' OR 'electroencephalogram') AND (Adaptive OR update OR heuristic OR nonholonomic OR time-variant OR 'time variant' OR 'self calibrating')	99
Embase 6/3/2017	("error potential" OR Errp) and (EEG OR MEG OR fMRI OR ECOG OR MRI OR BCI)	25
Pubmed 9/02/2017	(online OR live OR 'real time') AND (BCI OR 'brain computer interface' OR HMI OR HCI OR 'human machine interface' OR human computer interface OR 'brain machine interface') AND (EEG OR 'electroencephalography' OR 'electroencephalogram') AND (Adaptive OR update OR heuristic OR nonholonomic OR time-variant OR 'time variant' OR 'self calibrating')	88
Pubmed 6/3/2017	("error potential" OR Errp) and (EEG OR MEG OR fMRI OR ECOG OR MRI OR BCI)	24

After removal of duplicates, 236 articles were identified. The identified literature was screened for relevance to the study in three rounds: Firstly based on title, secondly on abstract and finally on full text review (see figure 1.2 for overview of excluded articles). Exclusion of an article was based on the following exclusion criterias:

- Studies not based on EEG or ECoG
- Did not concern adaptation of BCI systems or error-related potentials
- Animal studies

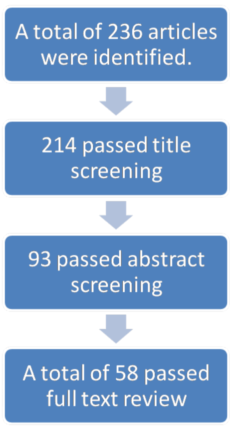


Figure 1.2: Progression of exclusion of identified literature.

Following the review of literature, 58 articles remained relevant to the study, and was further analyzed.

Chapter 2

Problem analysis

In the following, an analysis of the existing problems regarding non-stationarity of EEG, and existing efforts to adapt BCIs accordingly, will be presented. Secondly, knowledge of ErrPs and the feasibility of ErrP guided BCI adaptation will be assesed.

2.1 Non-stationarity

Non-stationarity is a potential problem in all BCIs assuming stationarity between classes [Krauledat and Naturwissenschaften, 2008], which is often the case in conventional BCIs [Hsu et al., 2016]. Changes in electrode impedances, electrode positions, amplifier noise and environmental noise, add to the non-stationarity of EEG. [Sun and Zhang, 2006, Mohammadi et al., 2013]

Shenoy et al. [2006] investigated the importance of a series of reasons for non-stationarity in EEG. They found that the single most important reason for non-stationarity is the differences between calibration sessions and online sessions. Here a difference in background activity was observed, such as the difference in demands for visual processing, in a calibration session and online session with feedback. Additional noise sources such as participant fatigue and alertness, which may fluctuate over time, also contributed to the nonstationarity. The differences in the feature space that they observed, could largely be corrected by displacing the decision plane of an initially trained classifier. An additional rotation of the decision plane further increased the classification accuracy, although relatively less compared to the translation. Despite the simplicity of the correction needed to maintain good classification performance between calibration and online sessions, the work of Shenoy et al. [2006] emphasizes the need for adaptation of online BCI systems.

2.2 Adaptation strategies

A way to deal with the non-stationarities in EEG is the application of adaptivity in the BCI system [Galan et al., 2008, Li and Zhang, 2010]. Adaptation can be done either in the preprocessing, feature extraction or in the classification of data [Hsu et al., 2016, Mohammadi et al., 2013, Li and Zhang, 2010]. Continuous adaptation enables the BCI system to track the EEG state of the subject and make appropriate corrections to maintain good classification accuracy [Li and Zhang, 2010].

Three main approaches for implementing adaptivity in BCIs have been identified: 1. Robustification of the features and/or classifiers, 2. adaptive filtering and 3. adaptivity of features and/or classifiers.

2.2.1 Robust features and classifiers

Assuming that the EEG, though non-stationary, varies within limits, it might be possible to either find features less sensitive to drift in the EEG or construct classifiers general enough to model the time-dependent variance in the feature space caused by non-stationarity. Here A few studies investigating robust signal processing are mentioned. A few studies using robust signal processing are mentioned here (for more studies identified using adaptive filtering see table 2.1).

Shin et al. [2015] proposed a sparsely representation classification (SRC) as a robust classifier, and compared it to support vector machines. The SRC consists of a library of feature vectors each representing one epoch of training data from a specific class. The classification of a new test epoch is then made, by representing the test epoch as a linear combination of all training epochs, but minimizing the total scalars needed to represent the test using the training epoch library. Finding the train epoch which contributes most to the minimal representation of the test epoch, identifies the class of the test, as the same class of the training epoch. Their comparison showed that SRC had a higher mean accuracy across 20 subjects ($p<0.01$), and that the SRC were consistently and progressively better with higher signal-to-noise ratio as test data were gradually contaminated with gaussian noise. [Shin et al., 2015]

Galan et al. [2008] proposed to identify features that consistently performed best across different conditions. To the extent that the variance across scenarios covers the variance that can be caused by non-stationarity the selected features will be as robust as possible. Unfortunately it is not reported how this feature selection strategy compared to traditional feature selection.

Table 2.1: Studies utilizing robustification for overcoming non-stationarities. Performance is stated as obtained accuracy, with increase to compared method in brackets, unless stated otherwise. Not all studies showed grand averages, in these cases classification accuracies are approximated from reported results. Classes denote the numbers of classes in the classifier and the used actions in the respective study. H = healthy subjects.

Robust Features

Studies	Performance (%)	Classes	Subjects
[Chamanzar et al., 2015]	86.5[a]	2: MI	n = 6 H
[Fernandez-Vargas et al., 2013]	ITR[b] 11.9 (4.9) bit/m	4: SSVEP	n = 18 H
[Galan et al., 2008]	43.1	3: MI/abstract	n = 2 H
[Kam et al., 2013]	60 (2)	4: MI	n = 9[c] H

Robust Classifier

Studies	Performance (%)	Classes	Subjects
[Li et al., 2010]	95.46 (11.17)%	3: MI	n = 1 H
[Shin et al., 2015]	73.6 (8.65)%	2: MI	n = 20 H

a: -608 ms mean latency of onset. ***b***: Information transfer rate, bits pr. minute. ***c***: BCI Competition IV set IIa. ***c***: BCI Competition III set IVc

2.2.2 Adaptive filtering

Adaptive filtering is used as a description of preprocessing adapting to varying noise sources, to provide a more stable EEG for feature extraction. A few studies using adaptive filtering are mentioned here (for more studies identified using adaptive filtering see table 2.2).

Hsu et al. [2016] have developed an online version of ICA, called online recursive independent component analysis (ORICA). The general idea of ORICA, is to provide a computational cheap ICA usable in online scenarios which quickly converge to a solution, thus when changes in the sources occurs, the preprocessing is quickly adjusted accordingly. The ORICA and the RUNICA function in EEGLab (used in the study as “golden standard”), produced comparable sources, however the ORICA was less computational heavy, suggesting its feasibility over RUNICA, for online implementation. [Hsu et al., 2016]

Woehrle et al. [2015] implemented a variation of the xDAWN algorithm, for adaptive filtering. The purpose of xDAWN is to find a transformation that optimizes the discriminability between signal and noise, by comparing the EEG to a model of the noise and a model of the signal of interest. The modified version of the algorithm was able to update with every new data point, and obtained the best performance when compared to the static xDAWN. [Woehrle et al., 2015]

Table 2.2: Studies utilizing adaptive filtering for overcoming non-stationarity. Performance is stated as obtained accuracy, with increase to compared method in brackets, unless stated otherwise. Not all studies showed grand averages, in these cases classification accuracies are approximated from reported results. Classes denote the numbers of classes in the classifier and the used actions in the respective study. MI = motor imagery, H = healthy subjects.

Studies	Performance (%)	Classes	Subjects
[Hsu et al., 2016]	Less computation time than RUNICA	Eriksen flanker test	n = 1 H
[Hsu et al., 2015]	Succesful adjustment	Eyes closed w. music & eyes open typing	Simulation + n = 1 H
[Wang et al., 2011]	87.41	4: MI	n = 3 H
[Wang et al., 2013]	Improved spectral estimation	4: MI	n = 9[b] H
[Woehrle et al., 2015]**	85 (4)[a]	Oddball	n = 6 H

** : Studies which also utilized adaptive classifiers
[a]: Training in one session, testing in another
[b]: BCI Competition IV

2.2.3 Adaptive features and classifiers

While Adaptive filtering is utilized to account for non-stationary noise-sources, the use of adaptive features and classifiers aims at maintaining high classification accuracy through continuous retraining. A few studies using adaptive features and classifiers are mentioned here (for more studies identified using adaptive features and classifiers see table 2.3 and 2.4).

Classifier adaptation

Thomas et al. [2011] used an adaptive classification of motor imagery classes, by utilizing an updating weight of discriminative frequency components. The weights were based on Fisher ratio, calculated for a series of frequency bands. Following and initial calibration, the discriminative frequency bands were continuously re-evaluated. The adaptive algorithm, using true labels to update, provided a mean accuracy of 89.58 and 88.33% respectively for two online sessions seperated by five weeks, in comparison to 85.00 and 83.75% for a static classifier. [Thomas et al., 2011]

Scherer et al. [2015] used a similar approach to Thomas et al. [2011], updating the most discriminative combination of a pair of mental actions (of four possible), channel and frequency band, based on Fisher ratio. In contrast to Thomas et al. [2011], Scherer et al. [2015] retrained their classifier by estimating the new best frequency band, every time 5 new artifact free EEG trials were available. They were able to reach a classification accuracy of 69.14% in patients suffering from severe motor impairment of various causes. [Scherer et al., 2015]

Hazrati and Erfanian [2008] used an adaptive probabilistic neural network (PNN) in a virtual reality environment, where the objective was to grab a falling ball by hand grasp motor imagery (MI). The kernel nodes of the PNN was initially trained, but updated with every classification following training. The adapting neural network obtained an accuracy of 89% compared to a static version obtaining 81%. [Hazrati and Erfanian, 2008]

Table 2.3: Studies utilizing adaptive classifiers for overcoming non-stationarities. Performance is stated as obtained accuracy, with increase to compared method in brackets, unless stated otherwise. Not all studies showed grand averages, in these cases classification accuracies are approximated from reported results. Classes denote the numbers of classes in the classifier and the used actions in the respective study. MI = motor imagery, H = healthy subjects, I = impaired subjects.

Studies	Performance (%)	Classes	Subjects
[Bamdadian et al., 2013]	71.81 (4.21)	2: Idle & MI	n = 12 H
[Bryan et al., 2013]	83.7 (0.82) ***a***	5: SSVEP	n = 10 H
[Faller et al., 2012]	76	2: MI	n = 12 H
[Faller et al., 2014]	68.6	3: Idle & MI	n = 22 I
[Hazrati and Erfanian, 2008]	89	2: Idle & MI	n = 6 H
[Hazrati and Erfanian, 2010]	75.4***b***	2: Idle & MI	n = 10 H
[Li et al., 2006]	92.3 (22.15)	6 x 6 speller	n = 10 H
[Li and Guan, 2006a]*	87.66	2: MI	n = 3 H
[Li and Guan, 2006b]*	93	2: MI	n = 3 H
[Li and Zhang, 2010]	86 (9.01)	2: MI	n = 6 H
[Lu et al., 2009]	98.2***d***	6 x 6 speller	n = 10 H
[Marathe et al., 2016]	93***e***	2: Target & non-target	n = 15 H
[Mondini et al., 2016]*	73.86	2: MI	n = 10 H
[Müller-Putz et al., 2014]	69.7	2: MI	n = 12 H
[Nicolas-Alonso et al., 2015]	77 (4)	4: MI	n = 9***f*** H
[Perdikis et al., 2016]	88	2: MI	n = 10 H
[Sannelli et al., 2012]*	80***g***	2: MI	n = 3 H
[Sannelli et al., 2016]*	75.9	2: MI	n = 20 H
[Scherer et al., 2015]	69.14	2: abstract & MI	n = 7 I
[Shenoy et al., 2006]	81	3: MI	n = 5 H
[Song and Yoon, 2015]*	74.66 (14.7)	4: MI	n = 3 H & 5 H***h*** and n = 9 H & 9 H***i***
[Sykacek et al., 2004]	A: 68.25 (5)	A: 2: Idle, abstract & MI	A: n = 8 H
	B: 79.5 (1.25)	B: 3: abstract & MI	B: n = 10 H
[Thomas et al., 2011]	75 (4)	2: MI	n = 3 H
[Vidaurre et al., 2005]	77.32	2: MI	n = 12 H
[Vidaurre et al., 2006]	73.36	2: MI	n = 12 H
[Vidaurre et al., 2007]	84	2: MI	n = 18 H
[Vidaurre et al., 2011a]	80	3: MI	n = 11 H
[Vidaurre et al., 2011b]	85.4	2: MI	n = 11 H
[Vidaurre et al., 2011c]	86	2: MI	n = 14 H

*: Studies which also utilized adaptive feature extractors. ***a***: and a 1.2 sec decrease in decision time. ***b***: First run in a series of runs. ***c***: BCI Competition III set IVa. ***d***: After 400 rounds of flashing. ***e***: By using 25% of data, comparable to 10 cross-validation using 90% of data. ***f***: BCI Competition IV set IIa ***g***: In best run. ***h***: BCI Competition III set IIIa & IVa. ***i***: BCI Competition IV set IIa & IIb.

Table 2.3: Continued...

Studies	Performance (%)	Classes	Subjects
[Wang et al., 2007]	79.48	3: MI	n = 5 H
[Wang et al., 2014]	70.6	2: No seizure / seizure	n = 5 E
[Wu and Ge, 2013]	85.55 (11)	2: M	n = 2 H
[Yoon et al., 2009]	$\text{AUC}^{\boldsymbol{d}}$: 0.7891	2: Idle & M	n = 2 H
[Yoon et al., 2011]	$\text{AUC}^{\boldsymbol{d}}$: 0.9380	3: MI	n = 5 H
[Yuxiao Yang et al., 2014]	$\text{MDPE}^{\boldsymbol{e}}$ 0.508 (8.98)	Drug concentration	Simulation
[Woehrle et al., 2015]**	85 $(4)^{\boldsymbol{f}}$	Oddball detection	n = 6 H

*: Studies which also utilized adaptive feature extractors. ** : Studies which also utilized adaptive filtering. $\boldsymbol{d}$: Area under receiver operating characteristic curve. $\boldsymbol{e}$: Median prediciton error decrease. $\boldsymbol{f}$: Training in one session, testing in another

Feature adaptation

Mohammadi et al. [2013] used covariate shift minimization (CSM) to adapt the training features to better represent test features. The CSM was implemented by fitting a polynomial function to features' values over time. From this function, future feature distributions were estimated, and the shift in future and present feature means were added to the means of training data, and the classifier was updated. Using a Linear Discriminant Analysis (LDA), the CSM adaptation improved classification accuracy from 88.4 - 94.1 %. [Mohammadi et al., 2013]

Sun and Zhang [2006] used an extension of the feature extractor common spatial pattern (CSP), called adaptive CSP (ACSP). The basic idea was to update the CSP derived covariance matrix, at the arrival of new data, using previous estimated covariance matrix. Additionally they utilized multiclass CSP in a manner of one-versus-rest, in a three class paradigm: left hand MI, right hand MI and word generation. They found that ACSP improved the classification accuracy from 58.25-65.12 % compared to conventional CSP. [Sun and Zhang, 2006]

Table 2.4: Studies utilizing adaptive feature extractors for overcoming non-stationarities. Performance is stated as obtained accuracy, with increase to compared method in brackets, unless stated otherwise. Not all studies showed grand averages, in these cases classification accuracies are approximated from reported results. Classes denote the numbers of classes in the classifier and the used actions in the respective study. M = movement, MI = motor imagery, H = healthy subjects, I = impaired subjects.

Studies	Performance (%)	Classes	Subjects
[Ang and Guan, 2016]	77 (8)	2: Idle & MI	n = 6 I
[Jeyabalan et al., 2010]	94.23	2: M	n = 1 H[a] & n = 3[b]
[Li and Guan, 2006a]*	87.66	2: MI	n = 3 H [c]
[Li and Guan, 2006b]*	93	2: MI	n = 3 H [c]
[Mohammadi et al., 2013]	TPR[d]: 97.1 (4.3)	2: Idle & M	n = 7 H
[Mondini et al., 2016]*	73.86	2: MI	n = 10 H
[Sannelli et al., 2012]*	80[e]	2: MI	n = 3 H
[Sannelli et al., 2016]*	75.9	2: MI	n = 20 H
[Song and Yoon, 2015]*	74.66 (14.7)	4: MI	n = 3 H & 5 H[f] and n = 9 H & 9 H[h]
[Sun and Zhang, 2006]	65.12 (5.45)	3: Abstract & MI	n = 3 H
[Wang et al., 2011]	87.4	2: MI	n = 3 H

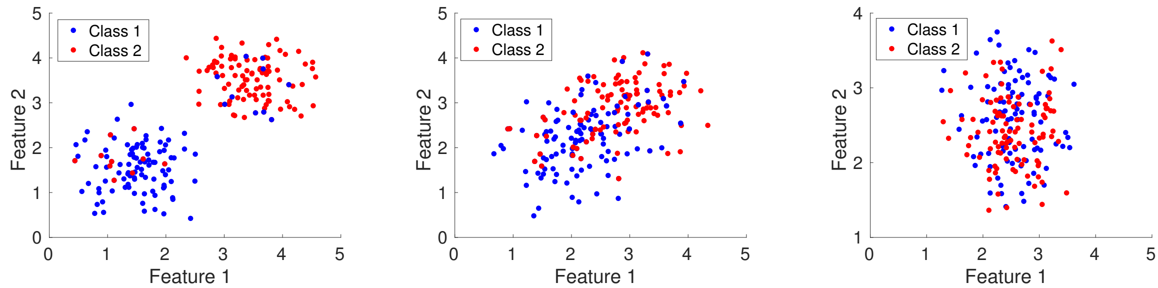
*: Studies which also utilized adaptive classifiers. [a]: BCI Competition III. [b]: BCI7 dataset. [c]: BCI Competition III set IVa. [d]: True positive rate. [e]: In best run. [f]: BCI Competition III set IIIa & IVa. [h]: BCI Competition IV set IIa & IIb

2.2.4 Summary

As is shown evidently through table 2.1 to 2.4, adaptation methods for coping with non-stationarities have been extensively studied in recent years. However, calibrations are a noisance and impractical for real-life BCI applications, [Marathe et al., 2016], and robust design of either features or classifiers is speculated to be insufficient for long term classification of EEG, without eventual recalibration. The periods between calibration may be extended compared to non-robust features and classifiers, but optimally calibration should be done only initially, to make the systems easy to use, and gain as high independencies of health care professionals.

Adaptive filtering may be a good implementation choice for real-life BCIs, as it aims to enhance signal-to-noise ratio, by adapting to non-stationary noise sources [Woehrle et al., 2015]. However, it is speculated that in the absence of adapting features or classifiers in addition to adaptive filtering, the underlying tendencies which are truly informative about the state of the BCI user, may change enough to be unrecognizable to the classifier, regardless of the signal-to-noise ratio. This implies that adaptive filtering is most appropriately implemented in conjunction with adaptive feature extractors or classifiers, as in Woehrle et al. [2015].

Feature and classifier adaptation in the mentioned studies can be divided into two groups: Either the adaptation is done using true labels, which are never available in real-life BCIs [Galan et al., 2008], or predicted labels are obtained from the classification, and naively used to adapt the features and classifier. The latter can be a great estimation and achieve remarkable classification accuracies (see figure 2.1a for a fictitious initial separated BCI training set). However, eventually mistakes will be made due to fatigue, changes in noise sources etc., which will affect future classification accuracies [Galan et al., 2008]. In case an epoch is misclassified and naively pooled with training data from a wrong class, that class will drift towards the true class of the epoch, thus the two classes will be less separable and more likely to be confused in future classifications (see figure 2.1b for an



(a) Two fictitious initial training sets of Class 1 and Class 2, with slight impurities.

(b) Naive adaptation may lead to blending of the classes, making the classes less seperable.

(c) The furthest consequence of naive adaptation may be inseperably confusing the classes.

Figure 2.1: Fictious example of worst-case scenario of naive adaptation.

increasingly mixed training set). As more and more epochs are misclassified and used for updating the wrong pools of training data, the two pools might converge, or one of the pools might grow to the extend that most of the epochs will be assigned to that pool. When two training pools converge they can hardly be separated again as new epochs of both classes is effectively randomly assigned to the two pools (see figure 2.1b for an inseparably mixed training set).

2.3 Error related potentials

To cope with the problem of possibly mixing training data from different classes, there is a need for an estimate of when adaptation is preferable or "safe", using predicted labels. Amongst others, Milekovic et al. [2012] demonstrated that a potential could be identified in the ECoG when a user got feedback which did not match the users intention. This potential is called an error related negativity or error related potential (ErrP). The ErrP appears shortly after an effector (e.g. a virtual spaceship) has initiated an action or event unintended by the user of the system (e.g. moving left instead of right) [Milekovic et al., 2012]. Thus classification of the users response to erroneous feedback (e.g. the left move of the spaceship), enables the system to evaluate whether the classification of the intention and the resulting action is considered correct by the user. Therefore, tracking the users responses to the feedback of the BCI system, might make a great basis for deciding when a predicted label should *not* be used to update a classifier, dependent on the presence of an ErrP.

2.3.1 Types of Error related potentials

The term ErrP broadly covers potentials elicited in response to errors in human-machine-interface (HMI) control. ErrPs have been identified across several modalities, such as fMRI [Diedrichsen et al., 2006], MEG [Llera et al., 2011], ECoG [Milekovic et al., 2013, 2012] and EEG [Blankertz et al., 2003, Kreilinger et al., 2016]. Across modalities several distinguishable kinds of ErrPs have been identified:

Target error potentials are elicited when the overall goal of an action changes during or before the action is intended, i.e. the subjects intend to reach a target, but the target unexpectedly moves [Milekovic et al., 2012].

Outcome error potentials are elicited when the system correctly detects the users intention and performs the corresponding action, but the outcome of the action is not as intended by the user (see figure 2.2 for an example of an outcome ErrP). For example outcome ErrPs occurs if the user despite full control of an avatar in a video game, misguide the avatar and unintendedly makes it collide with an obstacle. [Milekovic et al., 2012].

Execution error potentials are elicited when the HMI does not perform the action intended by the user. [Milekovic et al., 2012]. This type of ErrP is very relevant when considering BCI applications, as the detection of an execution ErrP indicates misclassification of the user intention.

In figure 2.2 morphological properties of execution ErrPs is illustrated compared to outcome ErrPs. In the spectral domain the ErrP contained most energy in the gamma and theta band 2.3. Milekovic et al. [2012] additionally investigated the evocation latency of execution ErrPs, and found that these most frequenctly peaked 600 ms following the feedback triggering the ErrP. Furthermore they found that 90% of elicited execution ErrPs peaked within the time window 100-800 ms following the triggering event.

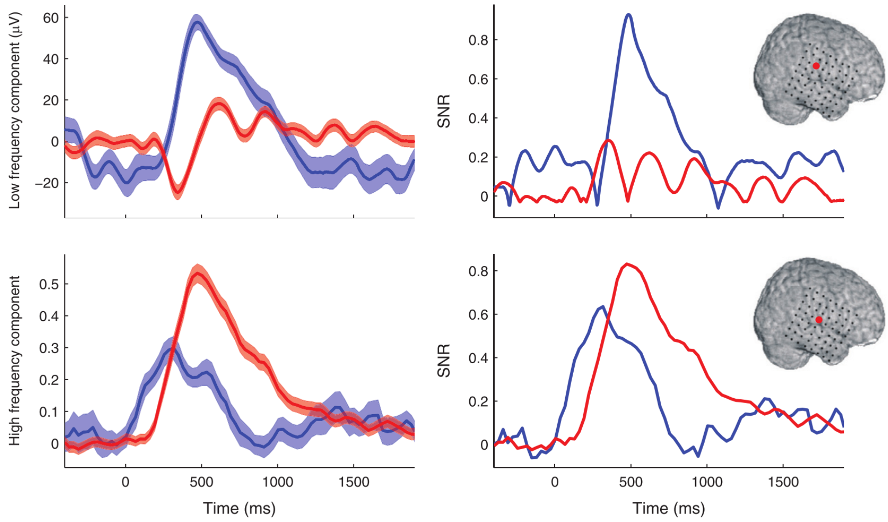


Figure 2.2: Temporal representation of outcome (blue) and execution (red) ECoG ErrPs following feedback (time 0 ms), recorded from the right parietal region (red dot). Left side: Morphology of outcome and execution ErrPs. Right side: Difference in signal to noise ratio between ErrPs and baseline. Upper and lower half: Representation of low (0 to 7.85 Hz) and high (60 to 128 hz) frequency components, respectively. Modified from [Milekovic et al., 2012].

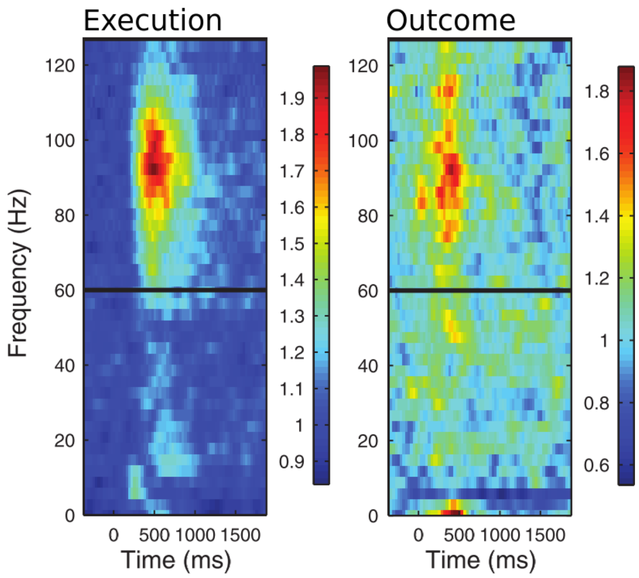


Figure 2.3: Joint time-frequency representation of Execution (left) and Outcome (right) ECoG ErrPs following feedback (time 0 ms), recorded from the right temporal region. Activity is found in the gamma band and for outcome ErrPs also in the delta and theta band. Modified from [Milekovic et al., 2012].

2.3.2 Studies utilizing ErrP in BCIs

Multiple studies have already studied the impact of utilizing ErrPs in BCIs. It has been found that the characteristics of the ErrP is dependent on the experimental paradigm [Milekovic et al., 2012, Diedrichsen et al., 2006]. In Milekovic et al. [2012] it was demonstrated that outcome and execution ErrPs could be discriminated with and accuracy of 83 % or more. Thus characteristics of ErrPs depends on the paradigm in which they are elicited.

The identified studies solely investigated the detection of ErrPs or the use of ErrPs for correcting or stopping misclassified intentions (see an overview of the studies in table 2.5).

Table 2.5: Overview of used features, time segments and classifiers in the identified studies investigating ErrPs in BCI context, and the respective classification accuracies of the ErrPs. M denotes motor tasks.

Study (subjects #)	Feature & Time	Classifier	Result	Classes	Feedback
[Ferrez and del R Millan, 2008] n = 5	Temporal 1 - 10 Hz (Broad region) 150-650 ms	Gaussian	Sensitivity 79% Specificity 84%	2:M	1 sec
[Buttfield et al., 2006] n = 3	Temporal < 128 Hz (Cz, Fz) 150-650 ms	Gaussian	Sensitivity 78% Specificity 82%	2:M	0 ms
[Parra et al., 2002] n=1 (outcome)	Temporal < 125 Hz (Broad region) 0-100 ms	Gaussian	auc-ROC 0.84%	2:M	0 ms
[Blankertz et al., 2003] n = 8	Temporal < 10 Hz (Broad region) 20-40 & 220-260 ms	LDA	Sensitivity 85% FP 2%	2:M	0 ms
[Kreilinger et al., 2016] n = 8	Temporal < 100 Hz Spectral 8 to 30 Hz (C3,Cz, C4) 0-1000 ms	LDA	Sensitivity 69 %	*$\boldsymbol{\mu\beta}$	indeterminable
[Parra et al., 2003] n = 7 (outcome)	Temporal < 125 Hz (Broad region) 0-100 ms	LDA	Not reported	2:M	0 ms
[Schalk et al., 2000] n = 4	Spectral 8 - 30 Hz (Cz) 0 -200 ms	LDA	Not reported	4:MI	80 ms

*$\boldsymbol{\mu\beta}$ = power in the mu and beta band. **M** = Overt motor action. **MI** = motor imaginary.

2.3.3 Cortical localization of ErrPs

It has been proposed that ErrPs are elicited mainly from the anterior cingulate cortex as a step in the reinforcement learning process upon the experience of an error [Holroyd and Coles, 2002]. The importance of the central region of the cortex is emphasize by Kreilinger et al. [2016], Buttfield et al. [2006] and [Schalk et al., 2000] in which ErrPs have been detected solely using EEG recorded in the central region. Furthermore, Blankertz et al. [2003], and [Ferrez and del R Millan, 2008] confirmed that ErrPs were prominent in this region.

Time

In Schalk et al. [2000], Blankertz et al. [2003], Parra et al. [2003], and [Parra et al., 2002] it was found that the ErrP can be detected within the first 260 ms of the EEG post feedback. In contrast it was found that using ECoG, the ErrP is most prominent after 600 ms, and most information about the ErrP can be found within the first 800 ms. Similarly Ferrez and del R Millan [2008] found that ErrP peaks occurs within the first 800 ms post stimuli.

Feature types

Ferrez and del R Millan [2008], Buttfield et al. [2006], Parra et al. [2002], Blankertz et al. [2003], Kreilinger et al. [2016], and [Parra et al., 2003] all used simple temporal features to detect ErrPs. Either single samples or means of small time segments were used as features. Ferrez and del R Millan [2008], and [Blankertz et al., 2003] did only use frequency components below 10 Hz for calculating the temporal features, whereas Buttfield et al. [2006], Parra et al. [2002, 2003], and [Kreilinger et al., 2016] utilized frequency components up to between 100 and 128 Hz. Futhermore Kreilinger et al. [2016], and [Schalk et al., 2000] utilized the power of spectral bands between 8 and 30 Hz (mu, alpha, and beta band).

Classifier performance

As the main focus of the reported studies were intention correction sensitivity and specificity is not reported for all studies. Ferrez and del R Millan [2008], Buttfield et al. [2006], Blankertz et al. [2003], and [Kreilinger et al., 2016] reported ErrP sensitivities between 69 % and 85 %. In Ferrez and del R Millan [2008], and [Buttfield et al., 2006] the specificity was reported to be between 82% and 84 %. Finally a review from 2014 concludes that accuracy in ErrP detection is commonly about 80 % [Chavarriaga et al., 2014].

Summary

Several studies have detected ErrPs in the EEG and demonstrated feasibility of using ErrPs to correct erroneous classification of control classes. However the potential of using ErrP for continuously adaptation of BCIs using active classes is uncovered. Due to the presence of an execution ErrP in the case of a erroneous BCI command, it seems feasible to use execution ErrP (henceforth referred to as only ErrP) for determining whether predicted labels can be used for adaptation of features and classifiers in BCIs.

2.4 Problem statement

The aim of this project was to determine to what extend EEG based ErrP detection is feasible for adaptation, to enhance the naive adaptation of classifiers. The feasibility of ErrP driven adaptation relies on a set of assumptions related to the detection of ErrPs. In this study, three criterias for feasible ErrP driven adaptation was established:

1. Highly sensitive ErrP detection

As adaption based on predicted labels may be inappropriate, as non-stationarity can lead to mixing of classes, it is important that ErrPs are correctly classified. Specifically it is essential that the sensitivity of the ErrP detection is high, leaving the chance of false positives less important. This is true as online use of a BCIs will generate a lot of epochs belonging to the control class, thus a sufficient amount of epochs usable for adaptation will remain, although not all control epochs without ErrPs are correctly identified. Should this criteria not be met, a BCI would not benefit from ErrP detection, as too many ErrPs would be included in the re-training data. In this work the quality of the system is evaluated as the ratio of non-errps in the adaptation training set versus errps.

2. The detection of ErrPs should be intention independent

The properties of an ErrP might depend on both the control class, and the machines response to that control signal. If so, potentionally one ErrP and non-ErrP subclass exists for each combination of control class and machine response. As it cannot be assumed that the classification of the intention is correct, information about what intention has been performed prior to classification of the response to feedback, cannot be available during the classification of response. Thus all ErrP subclasses must be distinguishable from all non-ErrP subclasses. Effectively it must be possible to pool all ErrP subclasses and all non-ErrP subclasses in the ErrP classifier, and obtain sufficient classification accuracy. Should this criteria not be met, the BCI would either be limited in the number of classes possible to implement and benefit from ErrP detection, or the BCI would not benefit at all.

3. Naive adaptation of the ErrP-classifier should be possible

In an adaptive BCI, it has been proven necessary to adapt the control classes, thus it is only natural to presume that adaptation of a ErrP classifier is needed as well. The adaptation of the control classes can be guided by ErrP classification, but the adaptation of the ErrP classes can only be done naively. Thus it is necessary to prove that ErrP and non-ErrP classes can be successfully adapted naively. Given that the second criteria is met, classification of RTF remains a two class problem, no matter the number of intention classes utilized in the BCI. In Ferrez and del R Millan [2008] a classifier was trained to discriminate between ErrPs and non-ErrP and obtained an accuracy of 83.8 %. When tested on data recorded up to three months before the training data the classifier obtained a similar accuracy (81.4 %). This indicates that long term detection of ErrPs might be possible using either robust classifiers or naive adaptation.

Though not exhaustively, previous work provides evidence for the first and third criteria. Especially longterm detection of ErrPs should be further investigated. However the primary aim in the current study was to investigate the second criteria. Furthermore, as mainly temporal features has been investigated in EEG literature, feasibility of both temporal and spectral features for ErrP detection was investigated. This gives the following scientific questions:

Can ErrPs reliably be elicited and classified for different control classes, and can ErrPs be detected independently of the control-class prior to the ErrP?

Chapter 3

Methods

In the following, all methodological considerations and procedures will be described. Firstly the design of the error-potential elicitation experiment will be described. Secondly the dataanalysis following the experiment will be clarified.

3.1 Subjects and Experimental setup

12 subjects (6 females and 2 left-handed) of mean age 24.3 ± 2.3, participated in the study. None had any known neurological disorders.

A NuAmp EEG-recorder was used to record the EEG from 16 electrodes positioned according to the international 10/20 standard in position: Fp1, Fp2, F3, Fz, F4, FT7, FT8, C3, Cz, C4, TP7, TP8, P3, P4, O1 and O2, in order to cover all cortical regions. Additionally an EMG electrode was placed above the left eyebrow to act as an EOG electrode. All electrodes were grounded to the left earlope, and referred to the right, although in cases where subjects had small earlopes, the ground and the reference electrodes were placed behind the ears on the mastoid process. Recording was done using Neuroscan 4.5 software, using a sampling frequency of 500 Hz.

3.2 Experiment

To asses the feasibility of ErrP detection for adaptation in BCIs the paradigms for elecitation of the ErrPs should resemble a realistic use case of a BCI. Thus ideally the ErrPs should be recorded during use of an online BCI. Implementation of a such a system was not possible within the time frame of this project. Instead three paradigms were designed inspired by studies like Blankertz et al. [2003] and Milekovic et al. [2012], whom showed that simple gamified tasks could elicit execution ErrPs, when the system responded unexpectedly to control commands from the subjects.

All subjects completed three separate paradigms: Overt arrow paradigm (OAP), covert arrow paradigm (CAP) and the game paradigm (GP). The order in which the subjects completed the paradigms followed a cross-design. Herein each paradigm were equally often completed as the fist, second and third. With 12 subjects, each order of the three paradigms were completed by two subjects. After completion of one session of each paradigm, all three paradigms were repeated. The sessions were repeated to ensure a sufficient amount of data, while limiting the duration of each session, to avoid inducing habituation and fatigue. During all recordings the subjects were seated in a comfortable chair facing a monitor and instructed to minimize blinking and unneccessary movements.

The overt arrow paradigm (OAP)

In the OAP, subjects were presented with a series of big black arrows, indicating either left or right direction. Subjects were instructed to press a button with either left or right hand in correspondence with the direction of the arrow. Following the keystroke, a smaller red arrow appeared in the center of the black arrow. Subjects were informed that the orientation of the red arrow, corresponded solely to their mental activity during the keystroke. Subjects were encouraged to focus on the movement, to increase the performance of the system, and were told that errors were caused by poor or wrong focus.

However, the orientation of the red arrow was independent of the subjects mental activity. In 85 % of the trials the orientation of the red arrow corresponded to the orientation of the black arrow, giving the subjects feedback corresponding to a correctly translated intention. In the last 15 % of the trials, the red arrow pointed in the opposite direction of the black, giving the subjects feedback corresponding to an erroneous translation of the intention. The subjects' mental response to the correct and erroneous feedback was used for non-ErrP and ErrP detection.

Each trial consisted of three phases: A preparation phase of random duration between 1 and 3 seconds, a cue phase of subject controlled duration, and a feedback phase of 1 second. The cue phase lasted until the subject hit a key with either right or left hand. Trials were completed in blocks of 20 repetitions. Between blocks subjects' had a small break which lasted 1 min or until the subject felt ready to continue. A total of 60 erroneous trials and 340 correct trials were recorded. The timeline of the OAP phases are depicted in figure 3.1.

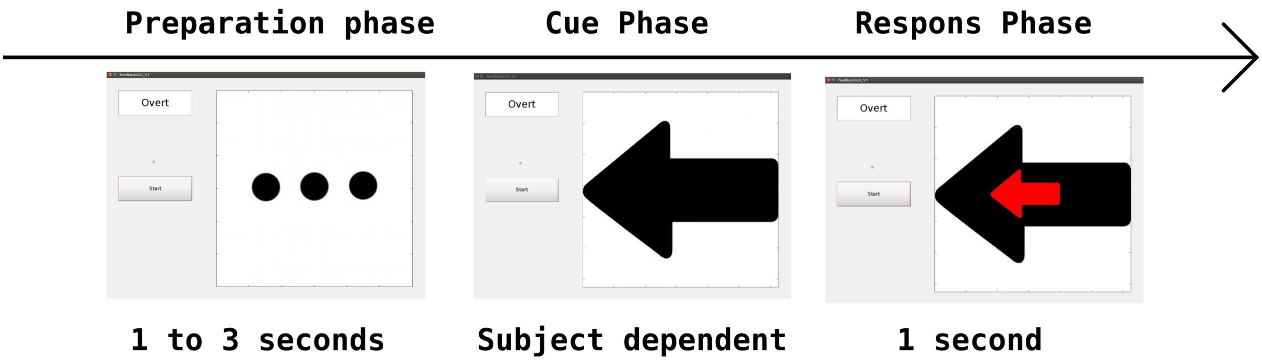


Figure 3.1: User interface and timeline of one trial in the OAP. The first frame indicates that a trial will soon begin. The second frame cues the subject to perform a keystroke. This cue lasts until the the keystroke is completed. The third frame is presented 40 ms after the keystroke and indicates the systems response to the users' mental action, and lasts for 1 s. The EEG recorded during this frame constitutes one epoch of data, used for ErrP detection.

The covert arrow paradigm (CAP)

The CAP was conducted similarly to the OAP, although subjects were instructed to only use mental strategies to control the red arrow, without any overt movement. Subjects were instructed to use MI and/or sensations in the right or left hand/arm corresponding to the direction of the black arrow. Prior to recording, subjects were trained in performing MI, until they felt comfortable in their performance.

The CAP showed a countdown following the presentation of the black arrow. The countdown showed the numbers "3, 2, 1", one at a time with one second delay. Subjects were instructed to perform the mental activity during the countdown, which ended at the time "0" could be expected, resulting in the presentation of the red feedback arrow.

The CAP consisted of the same phases as the OAP: A preparation phase of random duration between 1 and 3 seconds, a cue phase of 4 seconds, and a feedback phase of 1 second. Trials were completed in blocks of 10 repetitions as the CAP trials were longer than the OAP. Between blocks subjects had a small break which lasted 1 min or until the subject felt ready to continue. A total of 60 erroneous trials and 340 correct trials were recorded. The timeline of the CAP phases are depicted in figure 3.2.

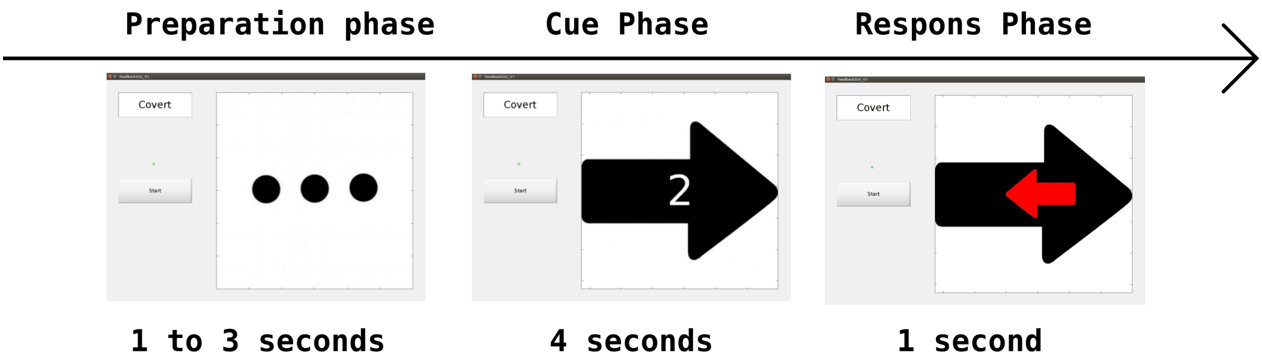


Figure 3.2: User interface and timeline of one trial in the CAP. The first frame indicates that a trial will soon begin. The second frame cues the subject to perform a MI task. The cue lasts for 4 seconds. During the 4 seconds a count down from 3 to 0 seconds is presented in the center of the black arrow. The third frame is presented 40 ms after the end of the countdown and indicates the systems response to the users' MI. The third frame lasts for 1 s and the EEG recorded during this frame constitutes one epoch of data, used for ErrP detection.

The game paradigm (GP)

The GP paradigm was inspired by Milekovic et al. [2012, 2013] whom elicited ErrPs by disturbing the subjects' control-inputs in a simple video game. In the GP the subjects played an open-source implementation of the classical Snake game (see [Gillespie, 2016]) using the arrow keys on a keyboard, while EEG were recorded. The time of each key stroke were registered, and used to sync epochs in the EEG to the specific event. For 80% of the key strokes the snake changed direction accordingly, but for the last 20% a script reversed the user input, so that the snake changed to the opposite direction of what the subject intended. The order of correct and reverse response was randomized, to prevent the subjects predicting the response. Unlike the CAP and the OAP, and the game played in Milekovic et al. [2012, 2013], the GP did not consist of phases, instead the subjects played continuously, to better mimic a real-life BCI scenario.

In each session the subjects played the game ten times of one minute duration. The game finished when the snake hit either it self or the edge of the game-platform. Upon game-over, the game was restarted and the subjects continued playing until the end of the current trial. As the frequency of keystroke were controlled by subjects, the number of ErrP and non-ErrP trials varied across subjects.

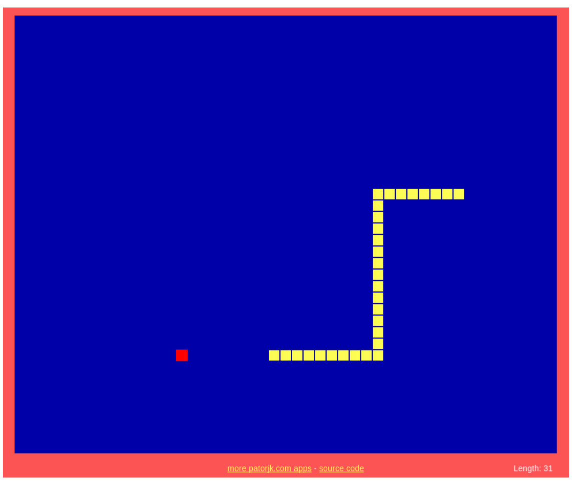


Figure 3.3: The interface of the game paradigm as seen by the subjects. the game is an unmodified copy of work by [Gillespie, 2016]. However the user commands (up, down, left, right) is not passed to the game directly, but through an interface, that randomly reversed the commands.

3.3 Dataanalysis

The dataanalysis was designed to explore the established criterias for ErrP usability. Firstly to investigate settings for optimal ErrP detection including relevance of features, channels and time segment, secondly to investigate the separability of ErrP and non-ErrP classes respectively. The separability test sought to investigate the similarity between ErrPs and non-ErrPs evoked under different paradigms.

3.3.1 Temporal filtering

All filtering were performed with an IIR implementation of chebyshev type 2 filters, to gain steep attenuation without passband ripples. Filters were of order 10 and had a stopband attenuation of 20 dB, however, the filtering was employed in MATLAB using the function filtfilt, effectively making the filters of order 20 and providing a stopband attenuation of 40 dB. Filtfilt was preferable as it preserves phase information. The procedure implies that the filtering of EEG cannot be performed immediately after recording, as the filter needs data both before and after an epoch to stabilize. This is however not a problem as ErrP driven adaptation, even when running online, does not need to occur immediately. All bandpass filters were implemented as one high and one low pass filter. Preprocessing of the EEG was done by bandpass filtering between 0.05 and 100 Hz. Following filtering, EEG data was segmented into epochs, which consisted of 1 s. of data, following feedback in the OAP and CAP, and 1 s. without further keystrokes in the GP.

3.3.2 Spatial noise reduction

Online ErrP detection prevents manual epoch selection. To reduce noise from EOG, movement and other non-cortical sources, three common noise rejection approaches were implemented and tested. Following preprocessing, noise was removed using common average reference (CAR) or independent component analysis (ICA) or the combination of the two.

CAR was implemented simply by subtracting a mean of all EEG channels, with respect to time, from each individual channel. By removing a mean of all channels the effect of noise sources affecting all the channels is reduced.

ICA was implemented in three steps: The EEG was transformed into the ICA domain using the MATLAB rica function. EOG artifacts was visually prominent in two ICA dimensions. The energy of the most contaminated dimension or both dimensions was removed (henceforth referred to as ICA-1 and ICA-2), and the residual EEG was transformed back to the ordinary spatial domain.

3.3.3 Feature extraction

Features in the present study were kept computationally simple, due to the prospect that feature calculation should be cheap computation-wise, in order ensure the usability of the features in online implementation. The features consisted of simple means, FFT components and template matching.

Temporal means
Multiple studies have found that temporal features could discriminate ErrPs from non-ErrPs in different time windows, spanning to 1 sec following feedback (Blankertz2003, Buttfield2006, Kreilinger2016, Parra2003, Parra2002,Milekovic2012). Following these suggestions simple means of EEG-data were calculated in steps of 50 ms per channel, resulting in a total of 320 simple mean features.

FFT components
Milekovic2013 & Milekovic2012 found spectral features capable of discriminating ErrPs from non-ErrPs. Therefore FFT components matching familiar EEG-bands, delta, theta, alpha, beta, gamma and high gamma, were used as features. The frequency bands are defined as in by Sanei et al. [2007], with the addition of "high gamma" and can be seen in table 3.1.

Frequency band	Sanei 2007	This study
Delta	0.5 - 4 Hz	1-3 and 2.5-5 Hz
Theta	4 - 7.5 Hz	5 - 7.5 Hz
Alpha	8 - 13 Hz	5 - 15 Hz
Beta	14 - 26 Hz	15 - 30 Hz
Gamma	30 - 45 Hz	30 - 45 Hz
High Gamma	-	45 - 100 Hz

Table 3.1: Normal EEG frequency bands as stated by Sanei et al. [2007], and as defined in this study.

For alpha, beta, gamma and high gamma frequency band, spectral features were calculated for five time windows of 200 ms with no overlap, from 0 to 1 s. after onset of the system feedback. Theta power was derived for three windows of 400 ms, with 100 ms overlap between 0 and 1 s. Finally the delta band was split into a high and a low subband. The high subband covered 2.5 to 5 Hz, and were analyzed in the same windows as the theta band. The low delta comprised 0.5 to 2.5 Hz and covered the entire epoch.

Template matching

Blankertz et al. [2003], Buttfield et al. [2006], Kreilinger et al. [2016], Parra et al. [2002, 2003] and Ferrez and del R Millan [2008] all found temporal properties of ErrPs, by using single data points of filtered or down-sampled EEG as features. The simple mean feature in this study might capture some of this information, but does not necessarily capture all morphological information. Morphological information is sometimes quantified using template matching [Jochumsen et al., 2013] and Aliakbaryhosseinabadi et al. [2015]. Instead of using all data points of an epoch as features, the correlation between the epoch and a template is used as a feature. Thereby the number of features is reduced. The templates were derived for each channel individually, as a mean of training epochs. For each channel the epochs were correlated to two templates, one derived as a mean of all ErrP training epochs, and one as a mean of all non-ErrP training epochs. Two kinds of templates were calculated: temporal templates, and spectral power templates (see examples of different templates in figure 3.4).

Temporal templates were derived from lowpass filtered epochs, as Jochumsen et al. [2013] and Aliakbaryhosseinabadi et al. [2015], demonstrated morphological information is prominent in the delta and theta band.

Spectral power templates contained information about amplitude modulation within a series of spectral bands. As demonstrated by Pfurtscheller and Lopes [1999] the power of a frequency band might have a distinct morphological profile during a mental action. Therefore the power of the frequency bands from theta to high gamma (see table 3.1) were calculated. First signal components within the bands were isolated by temporal filtering. Then power were calculated as the square of each data point. Finally the signal was filtered with a moving average with a window size of the same length as one period of the lowest frequency within the band. As for the temporal templates one template for ErrPs and one for non-ErrPs was derived for each channel, and each epoch was correlated to both templates.

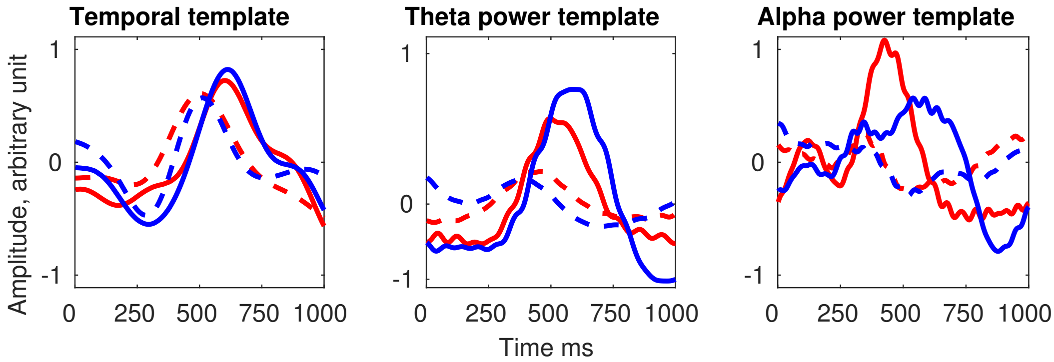


Figure 3.4: One subjects temporal, theta and alpha templates for ErrPs (full line) and non-ErrPs (dashed line), for both the OAP (blue) and the CAP (red).

Feature importance and selection

In order to make a comparison between temporal and spectral features' predictability of ErrP versus non-ErrP, feature importance was evaluated. Several approaches for evaluating features' importance. In this study three feature importance assessment methods were evaluated on performance in the classifiers: LDA single feature performance, Random Forest (RF) feature importance, and a Principal Component Analysis (PCA) dimension variance. Following evaluation of feature importance, the classifiers were trained, using the most important features. For all three feature selection approaches a within training data 10 fold cross validation were used to determine the best number of features to use.

Single feature importance was based on single feature prediction accuracy in a LDA. Each calculated feature were evaluated on prediction accuracy of training data on its own. The features selected for classification, were the features providing the highest accuracy in a single feature classification.

RF feature importance was based on the feature importance calculated in the MATLAB function "Treebagger", used to generate the RF. The importance of each feature is evaluated by comparing the accuracy between a classification using all features, and a classification where the values of the current feature is randomly interchanged between epochs. If the interchange impaired the classification it must mean that the feature contributed with important information, the more impaired classification, the more important the feature.

PCA feature importance was based on PCA transforming the feature matrix. Importance of features were derived as variance following the PCA transformation.

3.3.4 Classification

In this study two types of classifiers were used, LDA and a RF classifier. These two classifiers were chosen as they meet two different potential challenges in the classification of the data. LDA is a fairly simple, yet broadly used classifier [Vidaurre et al., 2011a, Kreilinger et al., 2016, Mohammadi et al., 2013]. For limited amounts of training data, as is the case in the present study, LDA approximates means of classes for identifying prominent tendencies in the data with a limited risk of over fitting to the training data. However for complex patterns, LDA might not provide a good model of the data, due to its simplistic decision plane. Random Forest is an ensemble of decision trees, which classifies based on votes from each tree [Breiman, 2001]. Unlike LDA the shape of the discrimination plan is not constricted and thus provides a more complex model that might better fit the data. The RF was trained using 2000 decision trees, in order to evaluate feature importance [Breiman, 2002], but only used 128 when used for classification [Oshiro and Perez, 2012]. Both classifiers were trained and evaluated in a 10-fold cross validation. To ensure fair comparison between the classifiers, the same randomization pattern was used for separating training and test data for both classifiers.

Cost

Both classifiers were trained with an unequal cost of misclassification of ErrPs and non-ErrPs. The cost was implemented to bias the classification towards higher sensitivity to ErrPs, as this is highly important in ErrP based adaption. It was attempted to optimize a cost function for the classifiers, to bias the classification in order to gain higher ErrP sensitivity (at the cost of lower specificity, see section 3.3.5). The optimization was unsuccessful, thus the cost of classifying non-ErrPs wrong was set as 1, and the cost of classifying ErrPs wrong was defined as the ratio between the number of non-ErrPs and ErrPs in the training set. Thereby equalling the cost of misclassifying all ErrPs and all non-ErrPs.

Performance measure: Correctness ratio

As sensitivity to ErrP detection is of greatest importance when utilizing ErrPs for choosing epochs for adaptation, plain accuracy was not a sufficient measure of classification results. Instead, we here introduce the *correctness ratio*. The correctness ratio is a measure of the ratio between non-ErrPs and ErrPs sent to adaptation of a BCI when an equal number of ErrPs and non-ErrPs are presented to the ErrP-classifier. The correctness ratio is calculated as:

$$\boldsymbol{C.ratio = \frac{specificity}{specificity + (1 - sensitivity)}}$$

and thus describes the relationship between true non-ErrPs and true non-ErrPs plus false non-ErrPs. In a naive adaptation setting, this would result in a ratio of 1/2, when presented with an equal number of ErrPs and non-ErrPs.

The correctness ratio has a downside in that it is vulnerable to very low specificity, which would result in a high ratio, without any substantial amount of non-ErrPs being available for adaptation. However, if data is sufficient, and a feasible ROC-curve for classification of ErrPs and non-ErrPs exists, then the correctness ratio would be optimal with a decent number of non-ErrPs being detected.

3.3.5 Optimization of templates, cost and number of features

In order to maximize the detection of ErrPs, efforts were made to optimize the template feature, the cost of misclassification and the number of features used in classification. Each of these optimizations were made strictly on training data, to avoid overfitting to the test data.

For template features, it was investigated which window-size for the moving average filter for each respective template, yielded the best results. It was found that the optimum window-size in all cases were a single period of the lowest frequency, in the respective template. In some cases the same result was obtained for several window-sizes, however, for simplicity it was chosen to fix the window size to the lowest frequency, in all the respective template.

For cost optimization, it was investigated which cost should be chosen for misclassifying an ErrP, in order to obtain high ErrP sensitivity, but still have a feasible specificity. It was found that the optimal cost varied greatly across the different training sets within each paradigm for each subject, and thus it was concluded that the amount of training data was not sufficiently big to estimate an optimal cost. Therefore, when the best classification, preprocessing and remaning optimizations had been found, a ROC-curve was made by training a classifier with increasing costs, to investigate the impact of cost on the classifications. Finally the impact of different numbers of features used in the classifiers were investigated, in order to find an optimum. In all optimization the aim were to gain as high a correctness ratio as possible.

When classifying a limited dataset promising results might occur randomly without representing any true tendency in the investigated phenomenon. In Müller-putz et al. [2008] minimum accuracy needed to statistically classify better than random given the sample size and a confidence level is calculated and presented. In the present work, a confidence level of $\boldsymbol{\alpha} = 0.05$, and individual subjects generated roughly 60 ErrPs, and when pooled a total of 720 ErrPs were available. In order to not underestimate the chance level, the individual chance level was based on 40 epochs (only 40 and 80 epochs are presented in Müller-putz et al. [2008]) of each class, and the pooled on 160 (upper limit of presented chance levels in Müller-putz et al. [2008]), giving a true chance level of 60 % and 55.6 % respectively. These were the chance levels considered when presenting results.

Feature, channel and time importance

The individual feature importances were found using the LDA single feature importance, anad was used to evaluate the importance of single channels, time segment and feature types by averaging the importance of all features.

While importance does give a general indication of the time and location of information enabling detection of ErrPs, it does not enable direct comparison of importance between channels or time segments, as it does not tell if all features within i.e. a channel provides the same or different information.

Time segment importance was calculated as the mean importance of features within or partly within time segments in steps of 50 ms. Template features and the spectral feature Low delta power were not included as they contained information in the entire epoch. Channel importance was calculated as a mean of all features derived from the same channel. Finally features were divided into four types, temporal means, spectral power, ErrP-templates and non-ErrP templates to assess the importance of each type of feature.

Chapter 4

Results

The results are divided into three parts. Firstly the different performances across subjects for each configuration of spatial filtering, feature selection and classification is compared. Secondly, classification and feature importance within each paradigm is described. Finally classification across paradigms is investigated, to asses if ErrP detection complies with the second criteria for feasibility of ErrP guided adaptation of BCIs.

4.1 ErrP classification accuracy

In order to show the obtainable sensitivity, specificity and correctness ratio of ErrPs, the results of the different tests and optimization efforts are presented here.

4.1.1 Classifier performances

A LDA classifier was compared to a RF in order to see whether choice of classifier impacts the detection accuracy of ErrPs. In addition each classifier was evaluated using different preprocessing techniques (see table 4.1 for classification accuracies). As the GP were subject to epoch rejection, the results presented for the GP are based on the number of epochs specified in tabel 4.2.

Table 4.1: Mean correctness ratio of the two classifiers, in the three paradigms, under different spatial filtering methods. The best combination preprocessing, feature reduction method and classifier for each paradigm is presented in bold.

	Reduction method	*Linear Discriminant Analysis*					
		None	CAR	ICA-1	ICA-2	CAR + ICA-1	CAR + ICA-2
	Single feature performance	0.77	0.76	0.80	0.74	0.78	0.73
OAP:	RF feature importance	**0.83**	0.78	0.80	0.78	0.78	0.77
	PCA Variance	0.72	0.70	0.74	0.73	0.81	0.74
	Single feature performance	0.79	0.79	0.82	0.81	0.82	0.75
CAP:	RF feature importance	**0.88**	0.85	0.85	0.85	0.84	0.80
	PCA Variance	0.80	0.78	0.79	0.76	0.83	0.79
	Single feature performance	0.57	0.62	0.55	0.55	0.56	0.56
GP:	RF feature importance	0.63	0.62	0.64	0.60	0.61	0.59
	PCA Variance	0.63	**0.70**	0.60	0.61	0.67	0.68
			Random Forest				
	Single feature performance	0.70	0.67	0.71	0.66	0.71	0.66
OAP:	RF feature importance	0.73	0.71	0.73	0.71	0.71	0.71
	PCA Variance	0.58	0.58	0.60	0.59	0.62	0.57
	Single feature performance	0.71	0.74	0.76	0.72	0.74	0.70
CAP:	RF feature importance	0.75	0.77	0.79	0.77	0.79	0.71
	PCA Variance	0.61	0.62	0.62	0.57	0.65	0.62
	Single feature performance	0.53	0.56	0.53	0.54	0.55	0.55
GP:	RF feature importance	0.57	0.55	0.55	0.55	0.57	0.55
	PCA Variance	0.55	0.57	0.55	0.55	0.56	0.59

Table 4.2: Number of ErrP and non-ErrP epochs in the GP.

Subject:	1	2	3	4	5	6	7	8	9	10	11	12
ErrPs	76	70	60	49	66	73	34	59	61	69	67	64
Non-ErrPs	294	241	251	273	282	221	150	269	276	302	279	275

With no exceptions, the LDA achieved a higher correctness ratio than the random forrest, in all three paradigms. Additionally the CAP achieved the greatest correctness ratio between the paradigms and GP the lowest. Lastly the RF feature reduction method generally worked best for the OAP and the CAP, whereas the PCA variance reduction method worked better for the GP.

For both the OAP and the CAP the best configuration was found to be without spatial filtering, using RF feature reduction and a LDA classifier. The CAP achieved the biggest correctness ratio of the three paradigms. The GPs best configuration was using CAR filtering and PCA variance feature reduction and a LDA classifier. See the individual subjects' sensitivity, specificity and correctness ratio under these configurations for each paradigm, in table 4.3.

Table 4.3: Subjects' individual sensitivity, specificity and correctness ratio under the optimal configurations for each paradigm. Means for each paradigm is calculated using 0, in case of NaN. Better than random performance, as defined by Müller-putz et al. [2008], is marked with bold.

		OAP			CAP			GP	
Sub:	Sens.	Spec.	C. ratio	Sens.	Spec.	C. ratio	Sens.	Spec.	C. ratio
1	**0.86**	**0.57**	**0.80**	**0.98**	**0.25**	**0.93**	0.88	0.27	0.70
2	**0.88**	**0.82**	**0.87**	**0.92**	**0.81**	**0.91**	1.00	0.00	NaN
3	**0.95**	**0.51**	**0.91**	**0.90**	**0.93**	**0.91**	0.98	0.09	0.83
4	**0.82**	**0.93**	**0.84**	**0.87**	**0.97**	**0.88**	0.88	0.18	0.60
5	**0.96**	**0.95**	**0.96**	0.74	0.28	0.52	**0.68**	**0.88**	**0.73**
6	**0.84**	**0.95**	**0.86**	**0.86**	**0.89**	**0.86**	0.91	0.12	0.57
7	1.00	0.01	1.00	**0.94**	**0.95**	**0.94**	1.00	0.00	NaN
8	**0.98**	**0.81**	**0.98**	**0.90**	**0.60**	**0.86**	0.97	0.01	0.32
9	**0.82**	**0.55**	**0.75**	0.90	0.29	0.75	0.99	0.03	0.74
10	0.53	0.52	0.53	**0.90**	**0.90**	**0.90**	0.97	0.09	0.74
11	0.91	0.10	0.52	**1.00**	**0.96**	**1.00**	1.00	0.00	NaN
12	1.00	0.08	1.00	**0.87**	**0.83**	**0.86**	**0.64**	**0.84**	**0.70**
Mean:	**0.88 ± 0.13**	**0.57 ± 0.35**	**0.84 ± 0.17**	**0.90 ± 0.07**	**0.72 ± 0.29**	**0.86 ± 0.12**	**0.91 ± 0.13**	**0.21 ± 0.32**	**0.49 ± 0.32**

Compared to the chance level, as defined by Müller-putz et al. [2008] to be 60% accuracy for each subject in the setting of this study, not all subjects achieved better than random classification accuracies. Subject 7, 10, 11 and 12 did not achieve above chance level in the OAP, subject 5 and 9 did not achieve above chance level in the CAP and only subject 5 and 12 achieved above chance level for the GP. When pooling epochs, the chance level is 55.6%, which all paradigms achieved, although the GP is just barely above chance level.

Additionally it was found that great intersubject variance existed, evident by the generally high standard deviations. The lowest standard deviation however, was in the CAP. Thus it seems that subjects where more consistent in their performance during the CAP, which is also suggested by the ROC-curves in figures 4.1-4.3.

Based on the findings presented in table 4.1 the following results for the three paradigms will be based on the configurations: RF feature importance without spatial filtering for OAP and CAP, and PCA variance and CAR filtering for GP.

4.1.2 Features reduction methods

The influence of the number of features on the correctness ratio was investigated. This was done in an effort to minimize overfitting to training data. The included features were chosen based on three feature evaluation methods. The correctness ratio achieved with different numbers of features for the optimal configuration of the three paradigms, is presented in table 4.4.

Table 4.4: Mean correctness ratio, with respect to number of included features, for the optimal configurations for each paradigm. The best obtained correctness ratio for each paradigm is marked in bold.

Features (n)	OAP	CAP	GP
2	68.5 ± 12.7	72.2 ± 8.1	54.4 ± 5.1
4	69.1 ± 11.2	74.2 ± 9.2	54.7 ± 4.5
8	71.7 ± 12.0	75.9 ± 9.3	54.7 ± 3.8
16	73.4 ± 12.9	81.0 ± 9.8	**57.6 ± 4.7**
32	75.3 ± 12.0	81.6 ± 9.1	57.4 ± 6.8
64	**76.2 ± 12.7**	**81.9 ± 8.4**	56.4 ± 5.3
128	75.6 ± 12.5	79.8 ± 6.1	53.1 ± 1.3
256	72.6 ± 12.5	75.0± 7.5	48.2 ± 15.2

The optimal number of features were lowest for the GP (16) and the same for the OAP and CAP having the optimal at 64 features. Adding more or less features to the classification of the respective paradigms resulted in a decrease in correctness ratio. This is believed to be the result of either overfitting to, or underrepresenting the training data. Generally the standard deviation on the correctness ratio was biggest for the OAP and lowest for the GP, suggesting greatest variability in subjects performance in the OAP, and less variability in GP. However, taken the correctness ratios of the GP into consideration, the low standard deviation may just be representative for the difficulty of achieving better results in the GP. Since subjects' individual optimal number of features differed, calculations were done using the individual optimal number of features.

4.1.3 Influence of cost - ROC curves

Following the identification of the best configuration and the optimal number of features, for each respective paradigm and subject, a ROC curve was produced for each paradigm, to find the influence of cost on the specific configurations. The mean ROC-curve of each paradigm can be seen in figure 4.1-4.3 respectively.

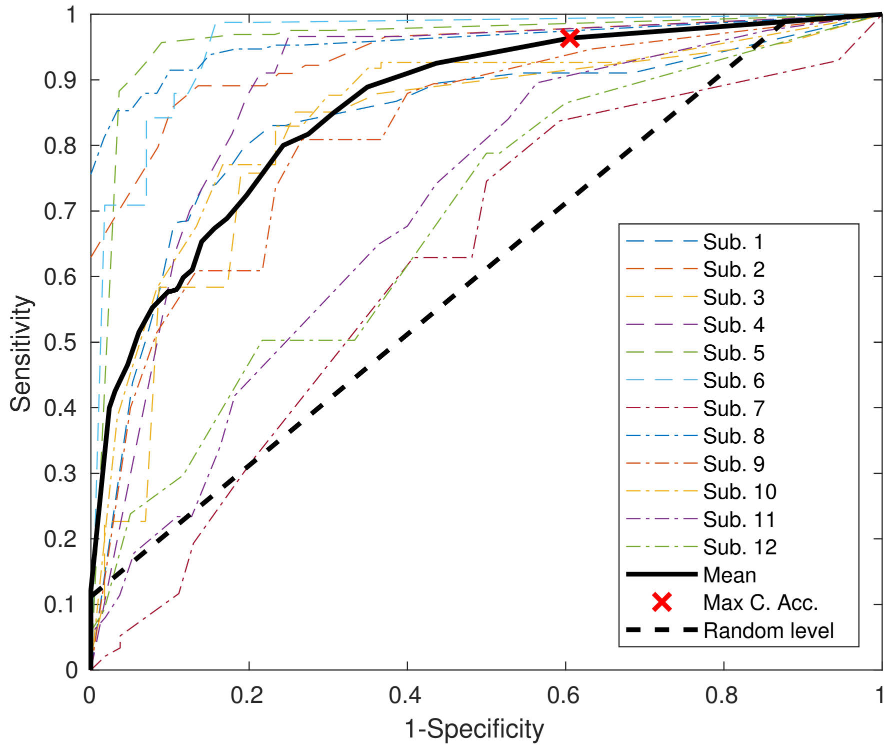


Figure 4.1: Mean ROC-curve across subjects for the best pre-processing classification configuration in the OAP. The red x marks the position on the average ROC-curve, where the correctness ratio was the highest. In this point, the correctness ratio was 0.9156. The mean ROC curve has an area under curve of 0.8557

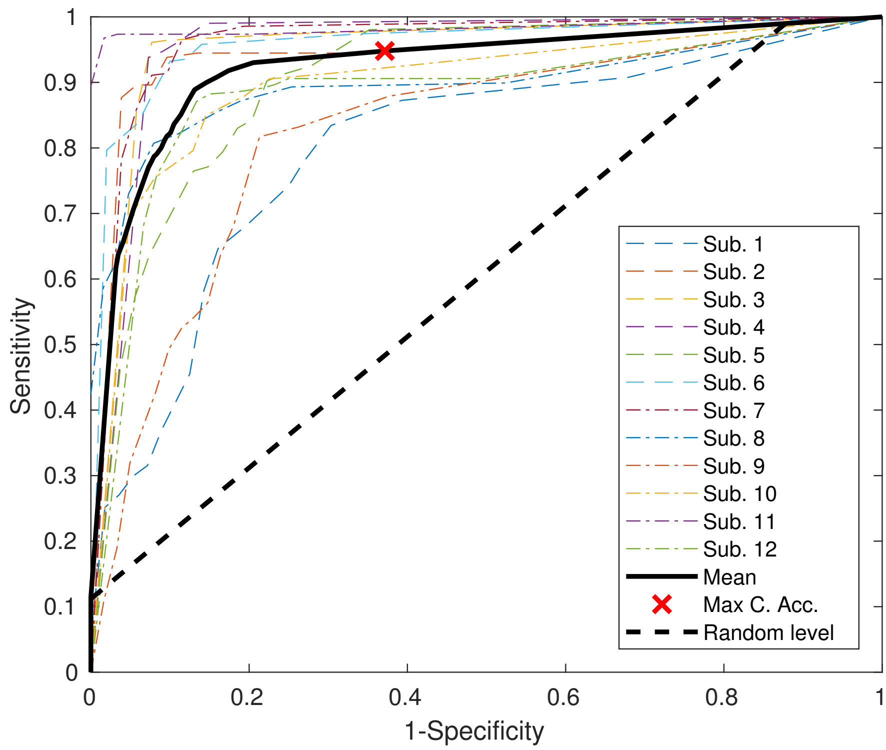


Figure 4.2: Mean ROC-curve across subjects for the best pre-processing classification configuration in the CAP. The red x marks the position on the average ROC-curve, where the correctness ratio was the highest. In this point, the correctness ratio was 0.9238. The mean ROC curve has an area under curve of 0.9207

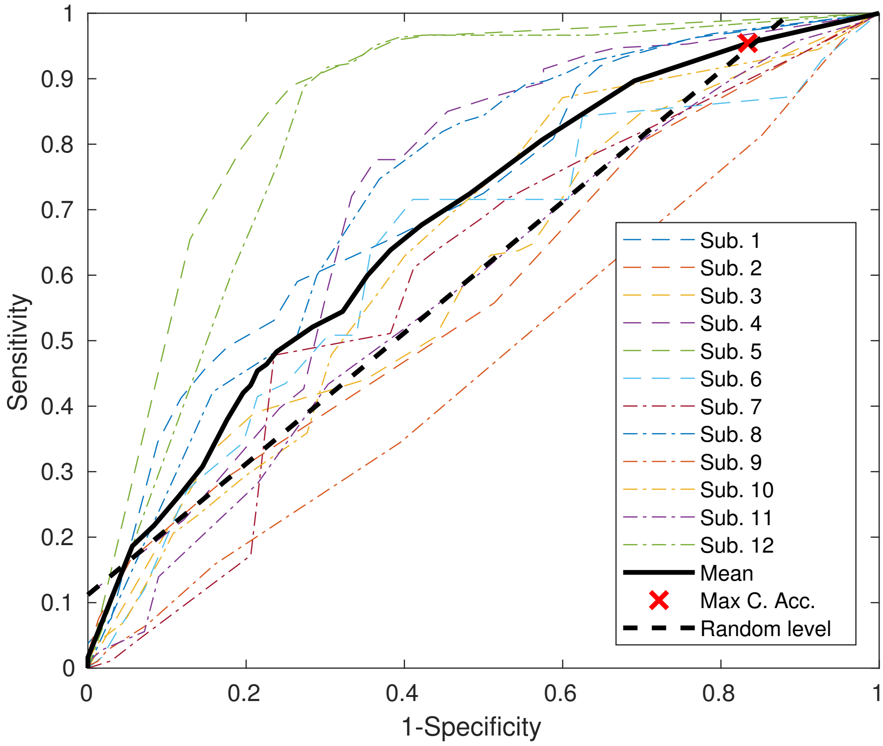


Figure 4.3: Mean ROC-curve across subjects for the best pre-processing classification configuration in the GP. The red x marks the position on the average ROC-curve, where the correctness ratio was the highest. In this point, the correctness ratio was 0.7640. The mean ROC curve has an area under curve of 0.6755

It was found that the optimal cost for each subject could not reliably be identified using the data obtained in the present study. However, for OAP and CAP, there is a clear influence on the achievable classification accuracy, by varying the cost. For the GP only subject 5 and 12 showed feasible detection accuracies of ErrPs, whereas subjects 1, 4, 6, 8 and 10 perform with limited feasibility, keeping in mind that sensitivity is of the essence, with a smaller emphasis on specificity. The remaining subjects did not perform very well in the GP.

The random level as defined by Müller-putz et al. [2008] for the pooled data (0.556), is plotted in the ROC-curves. As evident by figure 4.2 and 4.1, the mean ROC curve for the OAP and CAP are well above the random classification level. The mean ROC curve for the GP is above the random classification level as well, but did not show the same sensitivity to cost as the OAP and CAP.

Both the OAP and the CAP achieved +0.9 correctness ratio at the best points on the average ROC-curve, indicating that, at this point of the ROC curve, 9/10 epochs used for adaptation, would be truly corresponding to the underlying intention of the user, given an equal amount of correct and misclassified control epochs. The GP reached 0.76 corrected accuracy at the best point, indicating that 3/4 of epochs used for classification, given an equal amount of correct and misclassified control epochs, would truly correspond to the underlying intention of the user.

The three mean ROC-curves, achieved a AUC of 0.8557, 0.9207 and 0.6755, for the OAP, CAP and the GP respectively.

4.1.4 Important features in the optimal configurations

In the optimal configuration for each paradigm it was investigated which features had contributed to the classification accuracy. Additionally, feature type, channel and time segment importance was investigated. All feature importances reported here, were calculated prior to feature selection, and thus constitute the importances on which the selection of features was based, and not the importance of features chosen for classification. The GP achieved the best correctness ratio using PCA variance feature reduction. However the features importance calculated using this method is radically different compared to RF, which was the best feature reduction method for the OAP and

CAP. Thus for comparison purposes, the feature importances for the GP will be reported using the best performing RF feature reduction configuration, which utilized ICA-1 filtering.

Feature Type Importance

Three types of features were used for ErrP detection.

1. Simple temporal features derived as a mean of small time segments of lowpass filtered EEG
2. Series of spectral power within a set of frequenzy bands
3. Template matching based on:
    a) ErrP temporal and frequency power templates matched to both ErrPs and non-ErrPs
    b) non-ErrP temporal and frequency power templates matched to both ErrPs and non-ErrPs

In figure 4.4 the importance for temporal features, spectral features, ErrP template features, and non-ErrP template features is illustrated. For all three paradigms the ErrP-template matching features were the most important features. For the GP, also temporal and spectral features and non-ErrP templates provided information. non-ErrP templates seemed more important for the CAP than for the OAP.

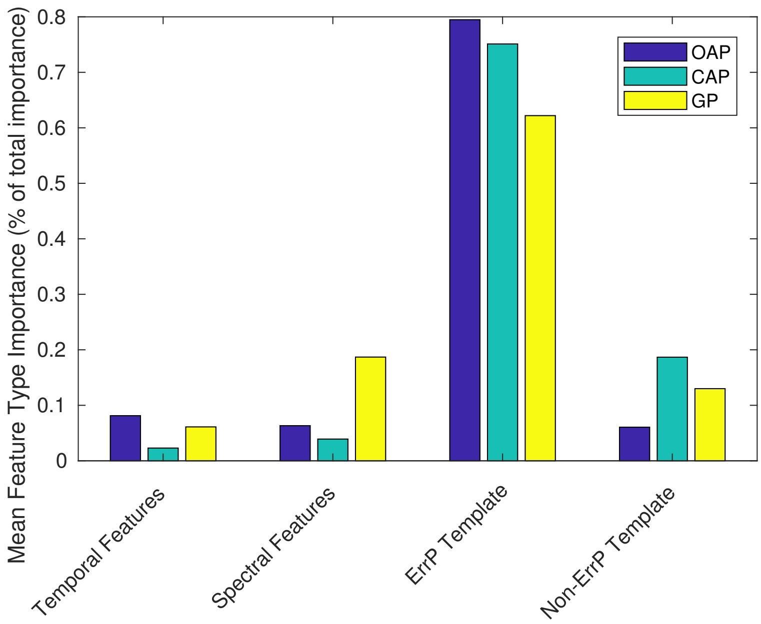


Figure 4.4: Barplot of importance of each feature type utilized in the study across subjects, for all paradigms.

Channel Importance

The importance of each channel was assessed as the total importance of features derived from the respective channel. The importance of each channel for the three paradigms is illustrated in figure 4.5. Cz was generally an important channel, being the most important for the OAP and GP, and the second most important for the CAP. For the CAP the most important channel was Fz. Generally there was important information in all electrodes, although a few electrodes constituded the most importance, with the two most important electrodes of each paradigm comprising Cz, Fz and Fp2. In addition, TP7 for the OAP and O1 for the GP had considerable high importance. The electrodes of least importance was FT7 and FT8.

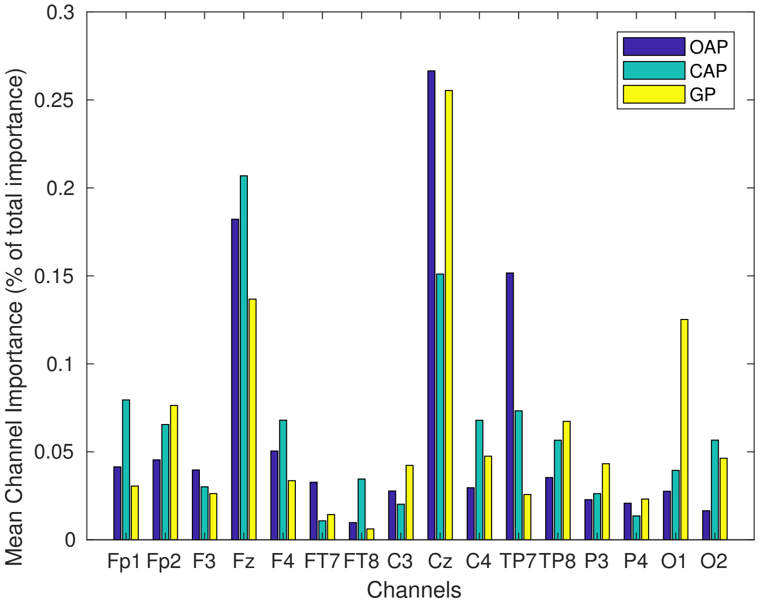


Figure 4.5: Barplot of channel importance across subjects, for all paradigms.

Time Segment Importance

The simple temporal and spectral features were calculated for small time segments within each epoch. Though not representing all discriminative information used for classification, the average importance of the features from each segment provided some information about what parts of the epochs that contained most discriminative information. The importance of each time segment for the three paradigms is illustrated in figure .

For both the OAP and CAP the most discriminative information appeared between 400 and 800 ms, with peak importance between 400 and 600 ms. The most discriminative information appeared between 600 and 1000 ms with peak importance between 800 and 1000 ms for GP. For the GP, importance was much more equally distributed across time segments. Unlike for the OAP and the CAP, the data does not indicate that all discriminative information has appeared within the first second following an event for the GP.

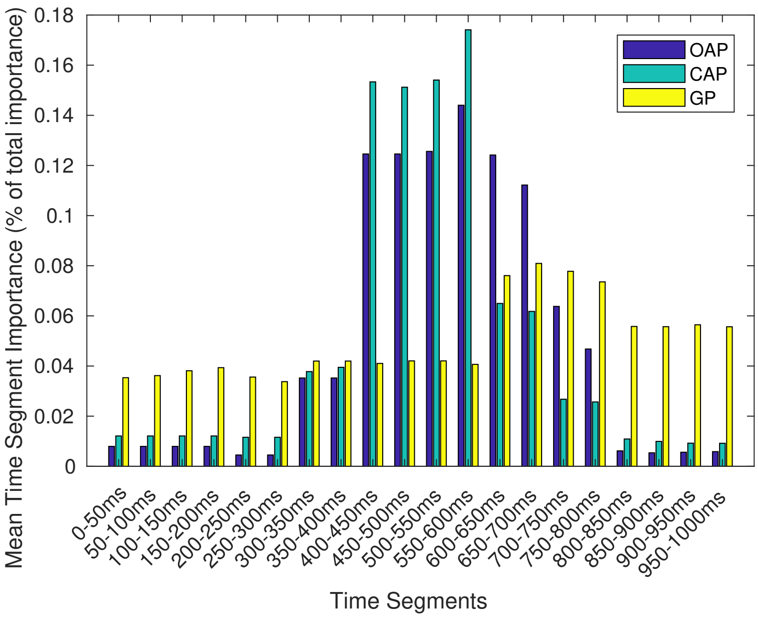


Figure 4.6: Barplot of time segment importance across subjects, for all paradigms.

4.1.5 Feature importance across subjects

In order to investigate the single features' relative importance compared to each other, a 3-dimensional plot was constructed to illustrate the importance of each features with respect to channels. See the 3-dimensional feature importance plots for each paradigm in figures 4.7-4.9.

For the OAP the most important features were ErrP theta template-matching, especially in Cz and Fz. Among the spectral feature the most important features were located in the gamma and theta band in the central and frontal region. The most important simple temporal features were located in the central frontal, central and central parietal region. Amongst the temporal and spectral features the most important features represented activity between 550 and 800 ms post feedback onset. Importance of all features for the OAP is illustrated in figure 4.7.

As for the OAP the single most important feature for the CAP was ErrP theta template matching, but unlike the OAP, also non-ErrP theta template matching was an important feature. As for the OAP the most important channel was Cz, but more channels were important to the classification of ErrPs in the CAP. For both ErrP and non-ErrP theta template matching the most important channels were Cz, C3, C4, FT8, FZ, F4. simple spectral and temporal features were less important for the CAP than for the OAP. The most important spectral features were gamma, delta and theta power in the central, parietal, tempo parietal and occipital regions. The most important spectral and temporal features represented information between 400 and 750 ms post feedback onset. Importance of all features for the CAP2 is illustrated in figure 4.8.

Unlike the OAP and the CAP the most important template features was the high gamma and gamma ErrP templates. The most important channel for the high gamma and the gamma template was O1. Unlike OAP and CAP the occipital region were more important than the central region for the template matching features. The simple temporal features did not contain as much information as for the OAP and the CAP. The most important spectral features were delta and theta features in the frontal parietal, and the frontal region. Importance of all features for the GP is illustratet in figure 4.9.

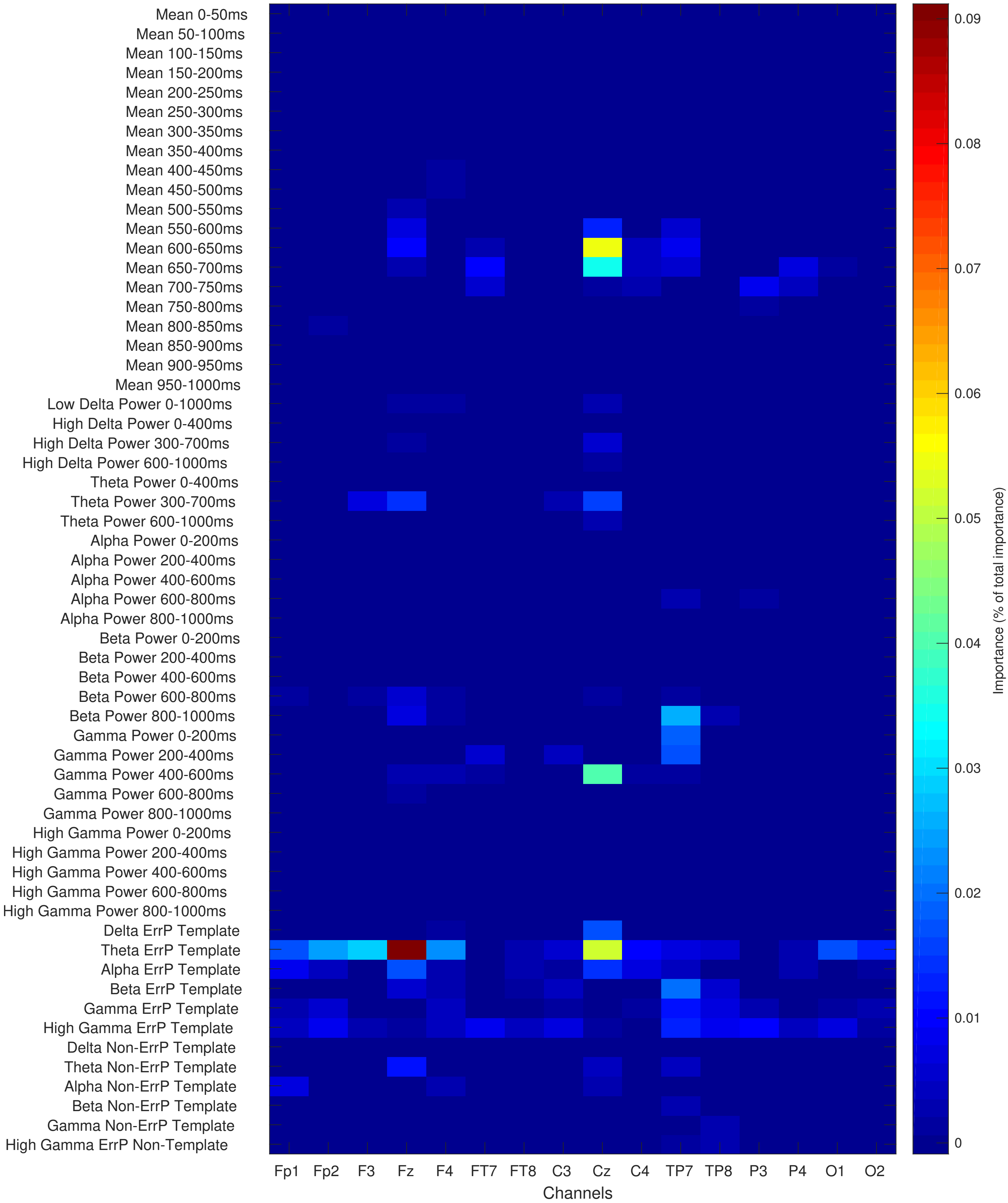


Figure 4.7: Individual mean importance of all features across all subjects for the OAP. EEG channels are listed along the x-axis, and individual features along the y-axis. The color of each feature indicates its importance. Dark blue indicates low importance (random classification) and dark red indicates high importance (more accurate classification).

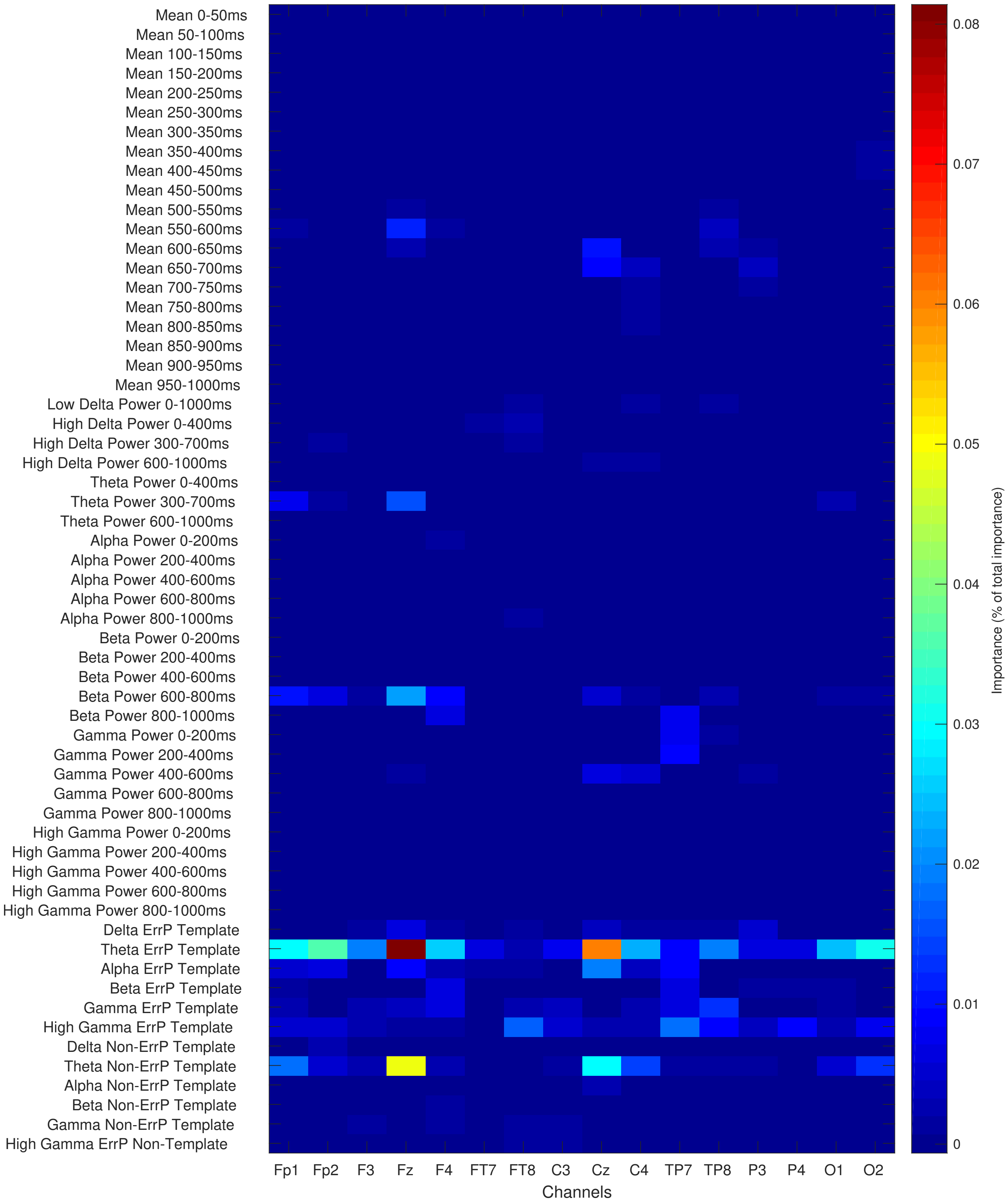


Figure 4.8: individual mean importance of all features across all subjects for the CAP. EEG channels are listed along the x-axis, and individual features along the y-axis. The color of each feature indicates its importance. Dark blue indicates low importance (random classification) and dark red indicates high importance (more accurate classification).

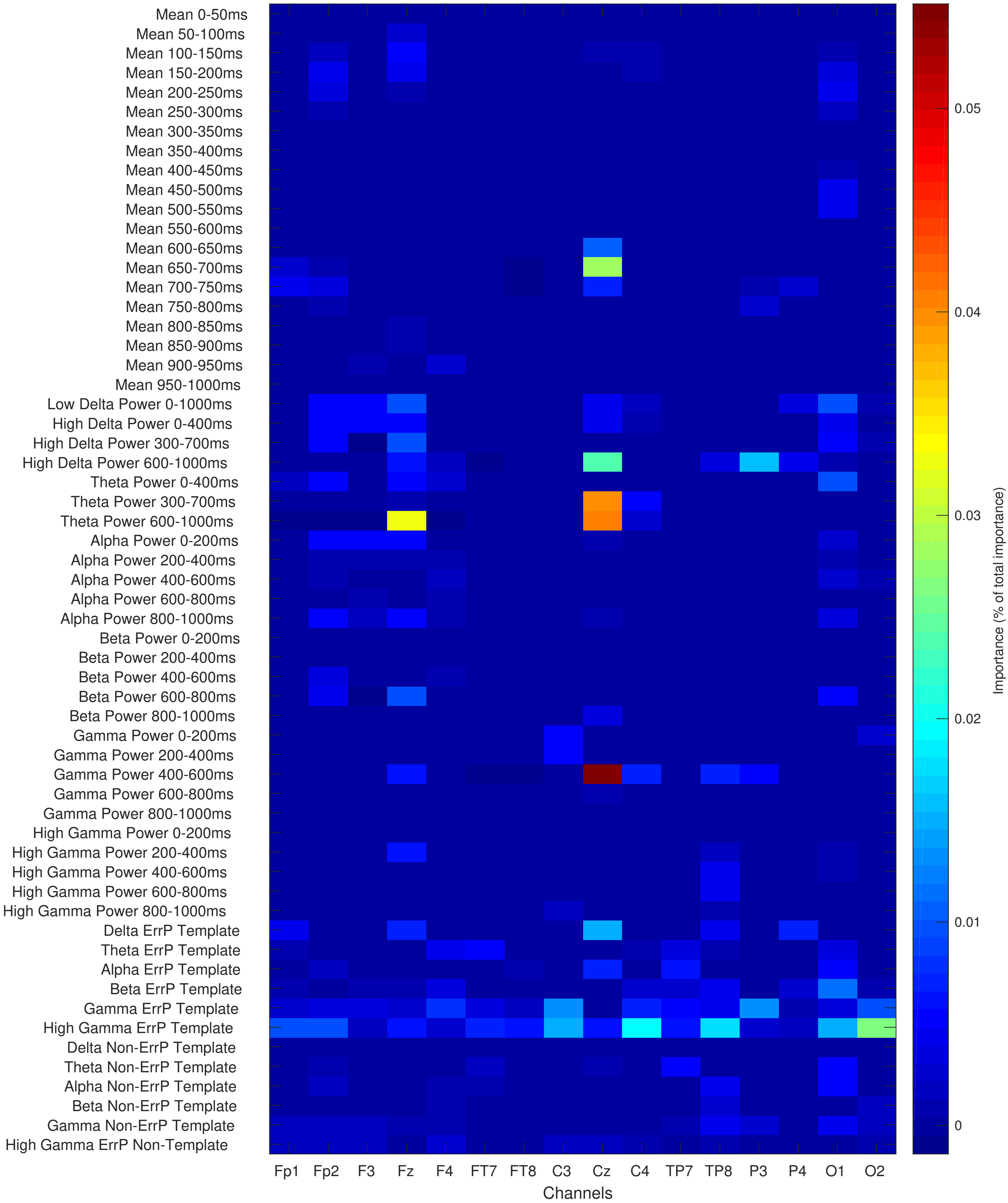


Figure 4.9: Individual mean importance of all features across all subjects for the GP.EEG channels are listed along the x-axis, and individual features along the y-axis. The color of each feature indicates its importance. Dark blue indicates low importance (random classification) and dark red indicates high importance (more accurate classification).

4.2 ErrP and non-ErrP separability across paradigms

The second criteria for successful utilization of ErrPs in adaptive BCIs, demand the ErrPs to be separable from all non-ErrP responses in a given BCI. In order to investigate whether this was realizable in the current study. A LDA and RF were trained using data from one or two paradigms and tested on a second or third. The results of the test is presented in table 4.5.

Table 4.5: Results obtained using the best configuration, trained on two paradigms and tested on the third. The sensitivity, specificity and the correctness ratio obtained is listed for each subject under the paradigm that constituted the test data. Results not better than random are marked in italic.

	LDA								
	OAP			CAP			GP		
	Sens.	Spec.	C. ratio	Sens.	Spec.	C. ratio	Sens.	Spec.	C. ratio
Sub 1	0.60	0.82	0.67	0.88	0.70	0.85	0.44	0.80	0.59
Sub 2	0.85	0.77	0.84	0.94	0.73	0.93	*0.03*	*0.95*	*0.50*
Sub 3	0.64	0.90	0.71	0.52	0.80	0.62	*0.26*	*0.75*	*0.50*
Sub 4	0.79	0.82	0.79	0.76	0.92	0.80	0.52	0.91	0.65
Sub 5	0.64	0.96	0.73	0.63	0.97	0.72	0.68	0.76	0.70
Sub 6	0.95	0.81	0.94	0.88	0.81	0.87	*0.25*	*0.93*	*0.55*
Sub 7	*0.17*	*0.81*	*0.49*	0.85	0.66	0.81	*0.18*	*0.90*	*0.52*
Sub 8	0.93	0.84	0.93	0.57	0.90	0.68	0.55	0.79	0.64
Sub 9	0.77	0.63	0.73	0.74	0.63	0.71	*0.35*	*0.84*	*0.56*
Sub 10	0.67	0.76	0.69	0.92	0.47	0.85	*0.57*	*0.61*	*0.59*
Sub 11	0.51	0.84	0.63	0.97	0.84	0.96	*0.45*	*0.68*	*0.55*
Sub 12	0.55	0.78	0.64	0.57	0.81	0.65	*0.64*	*0.53*	*0.59*
Mean:	0.67 ± 0.21	0.81 ± 0.08	0.73 ± 0.13	0.77 ± 0.16	0.77 ± 0.14	0.79 ± 0.11	0.41 ± 0.20	0.79 ± 0.13	0.58 ± 0.06

Generally the LDA outperformed the RF in this classification scheme. As previously observed within single paradigms It was found that all paradigms achieved better than random classification using the LDA, trained on the other paradigms. Of the three paradigms, the CAP was best classified and achieved the biggest correctness ratio while The GP was the worst classified and achieved the smallest correctness ratio. The correctness ratio achieved by the CAP was considerable (78.8).

Chapter 5

Discussion

In the present study, error related potentials were elicited during three different paradigms, the overt arrow paradigm (OAP), the covert arrow paradigm (CAP) and the game paradigm (GP). It was found that the highest classification accuracy and correctness ratio could be achieved in the CAP using no spatial filtering, selecting important features based on Random Forest feature importance, and using a LDA classifier. This configuration resulted in a sensitivity of 0.90 $\pm$ 0.07, specificity of 0.72 $\pm$ 0.29 and a correctness ratio of 0.86 $\pm$ 0.12. Most discriminative information was found in central region of cortex between 400 and 800 ms after feedback onset. The most important features was templates representing modulation of the spectral power of the theta band. For the three groups the area under the ROC was 0.92 %, for the CAP, 0.86 % for the OAP and 0.68 % for the GP.

Additionally it was found that ErrPs and non-ErrPs could be classified better than random when training epochs from two paradigms were used to classify the third. Using the CAP as test data achieved the best results, giving a sensitivity of 0.77 $\pm$ 0.16, a specificity of 0.77 $\pm$ 0.14 and a correctness ratio of 0.79 $\pm$ 0.11.

For the CAP when trained and classified alone, a mean correctness ratio of 86 % was achieved. If used for guiding adaptation of a naive system with an accuracy of 50 %, 86 % of the control epochs used for classification would be correctly labeled. However for a system with an accuracy of e.g. 85 %, the percent of correctly labeled epochs used for adaptation would be substantially higher, equalling 97 %.

ErrP detection

The best configurations for classification of ErrPs versus non-ErrPs, was in this study found to be: no spatial filtering, RF feature selection and LDA classification for both the OAP and the CAP, while CAR filtering and PCA variance feature selection and LDA classification worked optimally for the GP. Based on these findings it is assumed that spatial filtering distorted the characteristics of the ErrP epochs sufficiently to alter classification accuracy for the OAP and CAP.

Unlike during the OAP and the CAP it was not possible to ask the subjects to minimize blinking as they played continuously, and thus did not have breaks for blinking. It is therefore likely that the epochs in the GP were more contaminated by blinks. Additionally, it is speculated that the complexity of the GP resulted in generally low classification accuracies. These limitations on the GP make room for spatial filtering to improve the signal quality, for detection. This could be investigated by recalculating the system performance after removing epochs visually contaminated by blinks. In a ErrP based adaptation paradigm it would be adequate to always remove EOG contaminated epochs to insure that the artifacts deos not impair adaptation.

Subject variance

Within each of the three paradigms the performance of the subjects varied greatly. In the OAP, the performance of subject 7, 10, 11 and 12 was not above random classification, while subject 2, 4, 5, 6 and 8 were classified with a specificity and a sensitivity above 80 %. In a study by Blankertz et al. [2010] it was found that the strength of a subjects resting state sensory motor rhythm (SMR) was correlated to the accuracy the subject could obtain using an SMR-modulation based BCI. It was therefore speculated that some subjects might be neurophysiologically disposed for beeing BCI illiterate. However, in the present study, subjects how did not performing better than chance level in one paradigm, either the OAP or the CAP, performed better than chance level in the other.

Hence poor subject performance can not fully be explained by general BCI illiteracy. If the sum of sensitivity and specificity equals 1, the classification is random. For subjects performing close to randomly, the sensitivity tends to be close to 1 and the specificity close to 0. In these case the correctness ratio was very high and indicated good subject performance. Although this is a bad indicator of the individual subjects performance, it does not influence the calculation of the mean correctness ratio across subjects, as the ratio was calculated from the mean sensitivity and the mean specificity instead of a mean of each subjects correctness ratio, thereby the ratio was not falsely enhanced by poorly performing subjects.

Timing of important features

In line with work by Milekovic et al. [2012] it was found that most of the discriminative features for ErrP detection appears within the first 800 ms post system feedback. However, Schalk et al. [2000], Blankertz et al. [2003], Parra et al. [2003] and Parra et al. [2002] suggests that the ErrP appears in the EEG within the first 260 ms, whereas the present work suggests that information feasible for ErrP detection is mostly present after 300 ms post feedback. In both the present work and the mentioned studies, the subjects constantly faced a monitor that immediately following users actions, delivered the feedback that elicited the ErrP. Thus the experimental designs does not suggest that the high importance of information post 300 ms in the present work is caused by a delay in the experiments deliverance of feedback. It was not investigated whether it was possible to detect ErrPs solely using data from the first 300 ms post feedback, and Schalk et al. [2000], Blankertz et al. [2003], Parra et al. [2003] and Parra et al. [2002] did not analyse data post 260 ms after feedback, thus the studies are not completely comparable. However, analyzing features spanning over an entire second following feedback, the present study suggests higher importance of features within the window 300-800 ms post feedback, compared to pre 260 ms following feedback.

The most important time segments in the classification of epochs from the GP, appears later than for the other paradigms. Additionally there is no indication that all discriminative information is found within one second post feedback, as all time segments contained important features. The later part of the epochs in the GP might be contaminated with preparatory activity such as Bereitschaftspotentials, as the game was played continuously. It is also possible that cortical activity involved in correcting the previous mistake in the GP, is a benefactor to importance of late time segments, as this correcting action would only be present following an ErrP. However, the analysis of the GP here presented, is insufficient to answer whether cortical action-correction processes were present in the late ErrP epochs. However, it remains important that ErrPs can be accurately detected under different cortical loads, as real life BCI use might involve scenarios, where the user has to consider more than only BCI control, and thus inevitably contaminate the control signal with other mental activity. Another reason why the performance of the GP was lower than for the OAP and CAP might be that subjects were aware that control errors in the game was not related to their performance. This is worth investigating as it would affects how future ErrP studies should be designed.

Localization of ErrPs

Across the three investigated paradigms OAP, CAP and GP, Cz and Fz were the most important channels. This is in accordance with Ferrez and del R Millan [2008], Buttfield et al. [2006], Kreilinger et al. [2016] and Schalk et al. [2000] that also found that ErrPs can be detected in frontal to central regions and supports the assumption that ErrPs are elicited from the anterior cingulate cortex [Holroyd and Coles, 2002]. With the channels used in this study, the spatial resolution obtainable suggests that ErrPs appears to be well localized, or at the least most quantifiable, at channels Cz and Fz.

Milekovic et al. [2012, 2013] also found that the ErrPs could be found in the parietial and temporal-parietial region. In the present work this is to some extend supported by the relative high importance of TP7, however only during the OAP.

Important features in ErrP detection

In previous work, classification was based either on down sampled EEG or means of small segments [Blankertz et al., 2003, Buttfield et al., 2006, Kreilinger et al., 2016, Ferrez and del R Millan, 2008, Parra et al., 2002] and [Parra et al., 2003] or power within the alpha and beta band [Schalk et al., 2000]. In this study, theta band-power template-matching consistently provided a more important feature, compared to simple temporal features. Some of the information found in the theta band, might have been captured in previous studies using temporal features derived from low pass filtered EEG, but the present work suggests that more information is available if the band power of the signal is considered. In the GP, information from the gamma band was more important than information from the theta band. To the best of our knowledge this is inconsistent with previous findings in EEG studies of ErrPs elicited post movement or movement intention, or at the least uninvestigated. However discriminative information in this frequency band has been found using ECoG Milekovic et al. [2012] for outcome errors but not for execution errors. This suggests that the present study might have elicited outcome ErrPs. However, the design of the experimental paradigms, were tuned to elicit execution errors, while being inspired by other studies finding execution errors, thus the presence of important theta frequencies does not imply the elicitation of outcome errors.

Cost Estimation

For subject 5 and 9 the ROC curve indicates that by using the right cost, the performance would be considerably better than what is reported in table 4.3. This emphasizes the importance of adjusting the cost to the individual subjects, and optimization in general. It was unsuccessfully attempted to optimize the cost using only training data. This indicates that for the suggested optimization approach, the training data was not sufficiently big to fully model the ErrP and non-ErrP classes. Further work is needed to determine if more or another within-training-data optimization approach can reliably estimate proper subject specific misclassification cost. The possibility of insufficiently big sets of training data is further emphasized by the performance of subject 10 during the OAP. Using the general cost function used across subjects, both sensitivity and specificity of subject 10 is comparable to random classification, but the ROC curve for subject 10 indicates that remarkably better performance could be obtain for that subject.

Generalizability of ErrPs across paradigms

In order to investigate the generalizability of ErrPs across paradigms, a leave-one-out approach was used, to see how a classifier would perform if trained on all data from two paradigms, and tested on the third. This was done for all paradigms.

It was found that classification accuracies for the different paradigms were generally high, considering that no optimization efforts were made, to the specific classification scheme. The results imply that ErrP detection in this study, complies with the second criteria of usage of ErrPs for BCI adaptation. As it was possible to better than randomly classify ErrPs and non-ErrPs for a paradigm, when trained on others, there must be some general separability of ErrPs and non-ErrPs, independently on the control classes performed, before ErrPs where elicited.

The similarity of ErrPs across the three paradigms is promising. Further work shoud asses whether ErrPs also can be detected across motor paradigms, abstract paradigms and speech paradigms. Finally ErrP guided adaptation should be tested online, in a long-term scenario. Two questions remains: how does ErrP guided adaptation compare to naive adaptation of control classes in practice, and can the ErrP and non-ErrP class be naively adapted without losing accuracy?

5.1 Conclusion

Error related potentials are evoked in response to erroneous predection and actuation by a BCI. Detection of error related potentials provides prediction-independent means for guiding adaptation, and potentially enables more accurate adaptation of a BCI. In this work it was investigated if error related potentials could be reliably classified and detected independently of the user intention. This study demonstrated that error potentials evoked during three different paradigms of intention classes, could be comparably accurately classified after training the classifier with either data from the same paradigm as the test data or data from two other paradigms. This strongly suggests that some properties of error related potentials are independent of the misclassified user intention, and thus error related potentials can be used for nominating epochs usable for adaptation in a BCI.

References

Aliakbaryhosseinabadi, Jiang, Petrini, Farina, and Dremstrup, **2015**. S Aliakbaryhosseinabadi, N Jiang, L Petrini, D Farina, and K Dremstrup. *Robustness of Movement Detection Techniques from Motor Execution : single trial movement related cortical potential.* Neural Engineering, 7th Intern, 13–16, 2015. doi: 10.1109/NER.2015.7146548.

Ang and Guan, **dec 2016**. Kai Keng Ang and Cuntai Guan. *EEG-based strategies to detect motor imagery for control and rehabilitation.* IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society, 2016. ISSN 1558-0210 (Electronic). doi: 10.1109/TNSRE.2016.2646763.

Bamdadian, Guan, Ang, and Xu, **2013**. A Bamdadian, C Guan, K K Ang, and J Xu. *Improving session-to-session transfer performance of motor imagery-based BCI using Adaptive Extreme Learning Machine.* Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 2013, 2188–2191, 2013. ISSN 1557-170X. doi: 10.1109/EMBC.2013.6609969.

Barkat, **2005**. Mourad Barkat. *Signal Detection and Estimation*, volume 1. Artech House, 2 edition, 2005. ISBN 9788578110796. doi: 10.1017/CBO9781107415324.004.

Bi, Fan, and Liu, **2013**. Luzheng Bi, Xin-an Fan, and Yili Liu. *EEG-Based Brain-Controlled Mobile Robots : A Survey.* IEEE Transactions on Human-Machine Systems, 43(2), 161–176, 2013. ISSN 21682291. doi: 10.1109/TSMCC.2012.2219046.

Blankertz, Dornhege, Schafer, Krepki, Kohlmorgen, Muller, Kunzmann, Losch, Curio, Schäfer, Krepki, Kohlmorgen, Müller, Kunzmann, Losch, and Curio, **jun 2003**. Benjamin Blankertz, Guido Dornhege, Christin Schafer, Roman Krepki, Jens Kohlmorgen, Klaus-Robert Muller, Volker Kunzmann, Florian Losch, Gabriel Curio, C Schäfer, Roman Krepki, Jens Kohlmorgen, K.-R. Müller, Volker Kunzmann, Florian Losch, and Gabriel Curio. *Boosting bit rates and error detection for the classification of fast-paced motor commands based on single-trial EEG analysis.* IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society, 11(2), 127–131, 2003. ISSN 1534-4320 (Print). doi: 10.1109/TNSRE.2003.814456.

Blankertz, Sannelli, Halder, Hammer, Kübler, Müller, Curio, and Dickhaus, **2010**. Benjamin Blankertz, Claudia Sannelli, Sebastian Halder, Eva M. Hammer, Andrea Kübler, Klaus Robert Müller, Gabriel Curio, and Thorsten Dickhaus. *Neurophysiological predictor of SMR-based BCI performance.* NeuroImage, 51(4), 1303–1309, 2010. ISSN 10538119. doi: 10.1016/j.neuroimage.2010.03.022.

Breiman, **2001**. Leo Breiman. *Random Forests.* pages 1–33, 2001. ISSN 1098-6596. doi: 10.1017/CBO9781107415324.004.

Breiman, **2002**. Leo Breiman. *Manual on setting up, using, and understanding random forests v3.1.* page 29, 2002. doi: 10.2776/85168.

Bryan, Martin, Cheung, and Rao, **dec 2013**. Matthew J Bryan, Stefan A Martin, Willy Cheung, and Rajesh P N Rao. *Probabilistic co-adaptive brain-computer interfacing.* Journal of neural engineering, 10(6), 66008, 2013. ISSN 1741-2552 (Electronic). doi: 10.1088/1741-2560/10/6/066008.

Buttfield, Ferrez, and Millán, **jun 2006**. Anna Buttfield, Pierre W. Ferrez, and José Del R Millán. *Towards a robust BCI: Error potentials and online learning.* IEEE Transactions on

Neural Systems and Rehabilitation Engineering, 14(2), 164–168, 2006. ISSN 15344320. doi: 10.1109/TNSRE.2006.875555.

Chamanzar, Malekmohammadi, Bahrani, and Shabany, **2015**. A Chamanzar, A Malekmohammadi, M Bahrani, and M Shabany. *Accurate single-trial detection of movement intention made possible using adaptive wavelet transform.* Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 2015, 1914–1917, 2015. ISSN 1557-170X. doi: 10.1109/EMBC.2015.7318757.

Chavarriaga, Sobolewski, and Millan, **2014**. Ricardo Chavarriaga, Aleksander Sobolewski, and Jose Del R Millan. *Errare machinale est: the use of error-related potentials in brain-machine interfaces.* Frontiers in neuroscience, 8, 208, 2014. ISSN 1662-4548 (Print). doi: 10.3389/fnins.2014.00208.

Combaz, Chatelle, Robben, Vanhoof, Goeleven, Thijs, Van Hulle, and Laureys, **2013**. Adrien Combaz, Camille Chatelle, Arne Robben, Gertie Vanhoof, Ann Goeleven, Vincent Thijs, Marc M. Van Hulle, and Steven Laureys. *A Comparison of Two Spelling Brain-Computer Interfaces Based on Visual P3 and SSVEP in Locked-In Syndrome.* PLoS ONE, 8(9), 2013. ISSN 19326203. doi: 10.1371/journal.pone.0073691.

Diedrichsen, Hashambhoy, Rane, and Shadmehr, **2006**. Jörn Diedrichsen, Yasmin Hashambhoy, Tushar Rane, and Reza Shadmehr. *Neural Correlates of Reach Errors.* 25(43), 9919–9931, 2006.

Faller, Vidaurre, Solis-Escalante, Neuper, and Scherer, **2012**. J Faller, C Vidaurre, T Solis-Escalante, C Neuper, and R Scherer. *Autocalibration and recurrent adaptation: towards a plug and play online ERD-BCI.* IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society, 20(3), 313–319, 2012. ISSN 1558-0210. doi: 10.1109/TNSRE.2012.2189584.

Faller, Scherer, Costa, Opisso, Medina, and Müller-Putz, **2014**. J Faller, R Scherer, U Costa, E Opisso, J Medina, and G R Müller-Putz. *A co-adaptive brain-computer interface for end users with severe motor impairment.* PLoS ONE, 9(7), 2014. ISSN 1932-6203. doi: 10.1371/journal.pone.0101168.

Fernandez-Vargas, Pfaff, Rodríguez, and Varona, **2013**. J Fernandez-Vargas, H U Pfaff, F B Rodríguez, and P Varona. *Assisted closed-loop optimization of SSVEP-BCI efficiency.* Frontiers in Neural Circuits, (FEBRUARY 2013), 2013. ISSN 1662-5110. doi: 10.3389/fncir.2013.00027.

Ferrez and del R Millan, **mar 2008**. Pierre W Ferrez and Jose del R Millan. *Error-related EEG potentials generated during simulated brain-computer interaction.* IEEE transactions on bio-medical engineering, 55(3), 923–929, 2008. ISSN 0018-9294 (Print). doi: 10.1109/TBME.2007.908083.

Galan, Nuttin, Lew, Ferrez, Vanacker, Philips, and Millán, **2008**. F Galan, M Nuttin, E Lew, P W Ferrez, G Vanacker, J Philips, and J d. R Millán. *A brain-actuated wheelchair: Asynchronous and non-invasive Brain-computer interfaces for continuous control of robots.* Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology, 119, 2159–2169, 2008. doi: 10.1016/j.clinph.2008.06.001[doi].

Gillespie, **2016**. Patrick Gillespie. *JavaScript Snake*, 2016. URL https://github.com/patorjk/JavaScript-Snake.

Hazrati and Erfanian, **2008**. M K Hazrati and A Erfanian. *An on-line BCI for control of hand grasp sequence and holding using adaptive probabilistic neural network.* Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society.

IEEE Engineering in Medicine and Biology Society. Conference, 2008, 1009–1012, 2008. ISSN 1557-170X.

Hazrati and Erfanian, 2010. M K Hazrati and A Erfanian. *An online EEG-based brain-computer interface for controlling hand grasp using an adaptive probabilistic neural network.* Medical Engineering and Physics, 32(7), 730–739, 2010. ISSN 1350-4533. doi: 10.1016/j.medengphy.2010.04.016.

Holroyd and Coles, 2002. Clay B. Holroyd and Michael G.H. Coles. *The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity.* Psychological Review, 109(4), 679–709, 2002. ISSN 1939-1471. doi: 10.1037/0033-295X.109.4.679.

Hsu, Pion-Tonachini, Jung, and Cauwenberghs, 2015. Sheng-Hsiou Hsu, Luca Pion-Tonachini, Tzyy-Ping Jung, and Gert Cauwenberghs. *Tracking non-stationary EEG sources using adaptive online recursive independent component analysis.* Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 2015, 4106–4109, 2015. ISSN 1557-170X (Print). doi: 10.1109/EMBC.2015.7319297.

Hsu, Mullen, Jung, and Cauwenberghs, mar 2016. Sheng-Hsiou Hsu, Tim R Mullen, Tzyy-Ping Jung, and Gert Cauwenberghs. *Real-Time Adaptive EEG Source Separation Using Online Recursive Independent Component Analysis.* IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society, 24 (3), 309–319, 2016. ISSN 1558-0210 (Electronic). doi: 10.1109/TNSRE.2015.2508759.

Huggins, Moinuddin, Chiodo, and Wren, 2015. Jane E. Huggins, Aisha A. Moinuddin, Anthony E. Chiodo, and Patricia A. Wren. *What would brain-computer interface users want: opinions and priorities of potential users with spinal cord injury.* Archives of physical medicine and rehabilitation, 96(3), e1–e5, 2015. ISSN 1532821X. doi: 10.1016/j.apmr.2014.05.028.

Jeyabalan, Samraj, and Loo, 2010. V Jeyabalan, A Samraj, and C K Loo. *Motor imaginary-based brain-machine interface design using programmable logic controllers for the disabled.* Computer methods in biomechanics and biomedical engineering, 13(5), 617–623, 2010. doi: 10.1080/10255840903405678[doi].

Jochumsen, Niazi, Mrachacz-Kersting, Farina, and Dremstrup, 2013. Mads Jochumsen, Imran Khan Niazi, Natalie Mrachacz-Kersting, Dario Farina, and Kim Dremstrup. *Detection and classification of movement-related cortical potentials associated with task force and speed.* Journal of neural engineering, 10(5), 2013. ISSN 1741-2552. doi: 10.1088/1741-2560/10/5/056015.

Kam, Suk, and Lee, 2013. Tae Eui Kam, Heung Il Suk, and Seong Whan Lee. *Non-homogeneous spatial filter optimization for ElectroEncephaloGram (EEG)-based motor imagery classification.* Neurocomputing, 108, 58–68, 2013. ISSN 09252312. doi: 10.1016/j.neucom.2012.12.002.

Kilicarslan, Grossman, and Contreras-Vidal, 2016. A Kilicarslan, R G Grossman, and J L Contreras-Vidal. *A robust adaptive denoising framework for real-time artifact removal in scalp EEG measurements.* Journal of Neural Engineering, 13(2), 2016. ISSN 1741-2552. doi: 10.1088/1741-2560/13/2/026013.

Krauledat and Naturwissenschaften, 2008. Matthias Krauledat and Doktor Der Naturwissenschaften. *Analysis of Nonstationarities in EEG Signals for Improving Brain-Computer Interface Performance.* page 139, 2008.

Kreilinger, Hiebel, and Muller-Putz, mar 2016. Alex Kreilinger, Hannah Hiebel, and Gernot R Muller-Putz. *Single Versus Multiple Events Error Potential Detection in a BCI-Controlled Car Game With Continuous and Discrete Feedback.* IEEE transactions on bio-medical engineering, 63(3), 519–529, 2016. ISSN 1558-2531 (Electronic). doi: 10.1109/TBME.2015.2465866.

Lee, **2012**. Kijoon Lee. *Function for estimating the Sample Entropy*, 2012.

Li, Li, and Guan, **2006**. Huiqi Li, Yuanqing Li, and Cuntai Guan. *An effective BCI speller based on semi-supervised learning.* Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 1, 1161–1164, 2006. ISSN 1557-170X (Print). doi: 10.1109/IEMBS.2006.260694.

Li and Zhang, **2010**. J Li and L Zhang. *Bilateral adaptation and neurofeedback for brain computer interface system.* Journal of neuroscience methods, 193, 373–379, 2010. doi: 10.1016/j.jneumeth.2010.09.010[doi].

Li and Guan, **2006a**. Y Li and C Guan. *A semi-supervised SVM learning algorithm for joint feature extraction and classification in brain computer interfaces.* Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference, 1, 2570–2573, 2006. ISSN 1557-170X.

Li and Guan, **2006b**. Y Li and C Guan. *An extended em algorithm for joint feature extraction and classification in brain-computer interfaces.* Neural Computation, 18(11), 2730–2761, 2006. ISSN 0899-7667. doi: 10.1162/neco.2006.18.11.2730.

Li, Kambara, Koike, and Sugiyama, **jun 2010**. Yan Li, Hiroyuki Kambara, Yasuharu Koike, and Masashi Sugiyama. *Application of covariate shift adaptation techniques in brain-computer interfaces.* IEEE transactions on bio-medical engineering, 57(6), 1318–1324, 2010. ISSN 1558-2531 (Electronic). doi: 10.1109/TBME.2009.2039997.

Llera, Gerven, Gómez, Jensen, and Kappen, **2011**. A Llera, M A J van Gerven, V Gómez, O Jensen, and H J Kappen. *On the use of interaction error potentials for adaptive brain computer interfaces.* Neural Networks, 24(10), 1120–1127, 2011. ISSN 0893-6080. doi: 10.1016/j.neunet.2011.05.006.

Lu, Guan, and Zhang, **2009**. S Lu, C Guan, and H Zhang. *Unsupervised brain computer interface based on intersubject information and online adaptation.* IEEE Transactions on Neural Systems and Rehabilitation Engineering, 17(2), 135–145, 2009. ISSN 1534-4320. doi: 10.1109/TNSRE.2009.2015197.

Lule, Noirhomme, Kleih, Chatelle, Halder, Demertzi, Bruno, Gosseries, Vanhaudenhuyse, Schnakers, Thonnard, Soddu, Andrea Kubler, and Laureys, **2013**. Dorothée Lule, Quentin Noirhomme, Sonja C. Kleih, Camille Chatelle, Sebastian Halder, Athena Demertzi, Marie-Aurélie Bruno, Olivia Gosseries, Audrey Vanhaudenhuyse, Caroline Schnakers, Marie Thonnard, Andrea Soddu, Andrea Kubler, and Steven Laureys. *Probing command following in patients with disorders of consciousness using a brain-computer interface.* Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology, 124 (1 PG - 101-106), 101–106, 2013. doi: 10.1016/j.clinph.2012.04.030[doi].

Mak, Arbel, Minett, McCane, Yuksel, Ryan, Thompson, Bianchi, and Erdogmus, **2011**. J N Mak, Y Arbel, J W Minett, L M McCane, B Yuksel, D Ryan, D Thompson, L Bianchi, and D Erdogmus. *Optimizing the P300-based brain-computer interface: current status, limitations and future directions.* Journal of neural engineering, 8(2), 025003, 2011. ISSN 1741-2560. doi: 10.1088/1741-2560/8/2/025003.

Marathe, Lawhern, Wu, Slayback, and Lance, **mar 2016**. Amar R Marathe, Vernon J Lawhern, Dongrui Wu, David Slayback, and Brent J Lance. *Improved Neural Signal Classification in a Rapid Serial Visual Presentation Task using Active Learning.* IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society, 24(3), 333–343, 2016. ISSN 1558-0210 (Electronic). doi: 10.1109/TNSRE.2015.2502323.

Milekovic, Ball, Schulze-Bonhage, Aertsen, and Mehring, **apr 2012**. Tomislav Milekovic, Tonio Ball, Andreas Schulze-Bonhage, Ad Aertsen, and Carsten Mehring. *Error-related electrocorticographic activity in humans during continuous movements.* Journal of neural engineering, 9 (2), 26007, 2012. ISSN 1741-2552 (Electronic). doi: 10.1088/1741-2560/9/2/026007.

Milekovic, Ball, Schulze-Bonhage, Aertsen, and Mehring, **2013**. Tomislav Milekovic, Tonio Ball, Andreas Schulze-Bonhage, Ad Aertsen, and Carsten Mehring. *Detection of error related neuronal responses recorded by electrocorticography in humans during continuous movements.* PloS one, 8(2), e55235, 2013. ISSN 1932-6203 (Electronic). doi: 10.1371/journal.pone.0055235.

Minguillon, Lopez-Gordo, and Pelayo, **2017**. J Minguillon, M A Lopez-Gordo, and F Pelayo. *Trends in EEG-BCI for daily-life: Requirements for artifact removal.* Biomedical Signal Processing and Control, 31, 407–418, 2017. ISSN 1746-8108. doi: 10.1016/j.bspc.2016.09.005.

Mohammadi, Mahloojifar, and Coyle, **2013**. Raheleh Mohammadi, Ali Mahloojifar, and Damien Coyle. *Unsupervised short-term covariate shift minimization for self-paced BCI.* Proceedings of the 2013 IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain, CCMB 2013 - 2013 IEEE Symposium Series on Computational Intelligence, SSCI 2013, (i), 101–106, 2013. doi: 10.1109/CCMB.2013.6609172.

Mondini, Mangia, and Cappello, **2016**. Valeria Mondini, Anna Lisa Mangia, and Angelo Cappello. *EEG-Based BCI System Using Adaptive Features Extraction and Classification Procedures.* Computational intelligence and neuroscience, 2016, 4562601, 2016. ISSN 1687-5273 (Electronic). doi: 10.1155/2016/4562601.

Müller-Putz, Steyrl, Faller, Muller-Putz, Steyrl, Faller, Müller-Putz, Steyrl, and Faller, **2014**. G R Müller-Putz, David Steyrl, Josef Faller, Gernot R Muller-Putz, David Steyrl, Josef Faller, G R Müller-Putz, David Steyrl, and Josef Faller. *Adaptive hybrid brain-computer interaction: ask a trainer for assistance!* Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 2014, 1493–1496, 2014. ISSN 1557-170X. doi: 10.1109/EMBC.2014.6943884.

Müller-putz, Scherer, Brunner, Leeb, and Pfurtscheller, **2008**. Gernot R Müller-putz, Reinhold Scherer, Clemens Brunner, Robert Leeb, and Gert Pfurtscheller. *Better than random ? A closer look on BCI results.* International Jouranl of Bioelectromagnetism, 10(1), 52–55, 2008.

Nicolas-Alonso, Corralejo, Gomez-Pilar, Álvarez, and Hornero, **2015**. L F Nicolas-Alonso, R Corralejo, J Gomez-Pilar, D Álvarez, and R Hornero. *Adaptive semi-supervised classification to reduce intersession non-stationarity in multiclass motor imagery-based brain-computer interfaces.* Neurocomputing, 159(1), 186–196, 2015. ISSN 1872-8286. doi: 10.1016/j.neucom.2015.02.005.

Oshiro and Perez, **2012**. Thais Mayumi Oshiro and Pedro Santoro Perez. *Machine Learning and Data Mining in Pattern Recognition.* 7376(November), 2012. doi: 10.1007/978-3-642-31537-4.

Parra, Spence, Gerson, and Sajda, **2003**. L C Parra, C D Spence, A D Gerson, and P Sajda. *Response error correction - A demonstration of improved human-machine performance using real-time EEG monitoring.* IEEE Transactions on Neural Systems and Rehabilitation Engineering, 11(2), 173–177, 2003. ISSN 1534-4320. doi: 10.1109/TNSRE.2003.814446.

Parra, Alvino, Tang, Pearlmutter, Yeung, Osman, and Sajda, **2002**. Lucas Parra, Chris Alvino, Akaysha Tang, Barak Pearlmutter, Nick Yeung, Allen Osman, and Paul Sajda. *Linear Spatial Integration for Single-Trial Detection in Encephalography.* NeuroImage, 17, 223–230, 2002. ISSN 10538119. doi: 10.1006/nimg.2002.1212.

Perdikis, Leeb, and Millán, **2016**. S Perdikis, R Leeb, and J D R Millán. *Context-aware adaptive spelling in motor imagery BCI.* Journal of Neural Engineering, 13(3), 2016. ISSN 1741-2552. doi: 10.1088/1741-2560/13/3/036018.

Pfurtscheller and Lopes, 1999. G Pfurtscheller and F H Lopes. *Event-related EEG / MEG synchronization and desynchronization : basic principles.* Clinical Neurophysiology, 110, 1842–1857, 1999.

Saeid Sanei, J. A. Chambers, and Karl E. Misulis. Introduction to EEG. In *EEG Signal Processing*, chapter 1, pages 1–34. 2007. ISBN 9780199985906. doi: 10.1093/med/9780199985906.003.0001.

Sannelli, Vidaurre, Müller, and Blankertz, 2012. C Sannelli, C Vidaurre, K R Müller, and B Blankertz. *Common Spatial Pattern Patches: online evaluation on BCI-naive users.* Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference, 2012, 4744–4747, 2012. ISSN 1557-170X.

Sannelli, Vidaurre, Müller, and Blankertz, 2016. C Sannelli, C Vidaurre, K.-R. Müller, and B Blankertz. *Ensembles of adaptive spatial filters increase BCI performance: An online evaluation.* Journal of Neural Engineering, 13(4), 2016. ISSN 1741-2552. doi: 10.1088/1741-2560/13/4/046003.

Schalk, Wolpaw, McFarland, and Pfurtscheller, dec 2000. G Schalk, J R Wolpaw, D J McFarland, and G Pfurtscheller. *EEG-based communication: presence of an error potential.* Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology, 111 (12), 2138–2144, 2000. ISSN 1388-2457 (Print).

Scherer, Faller, Opisso, Costa, Steyrl, and Muller-Putz, 2015. R Scherer, J Faller, E Opisso, U Costa, D Steyrl, and G R Muller-Putz. *Bring mental activity into action! An enhanced online co-adaptive brain-computer interface training protocol.* Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 2015, 2323–2326, 2015. ISSN 1557-170X. doi: 10.1109/EMBC.2015.7318858.

Shenoy, Krauledat, Blankertz, Rao, and Müller, 2006. Pradeep Shenoy, Matthias Krauledat, Benjamin Blankertz, Rajesh P N Rao, and Klaus-Robert Müller. *Towards adaptive classification for BCI.* Journal of neural engineering, 3(1), R13–R23, 2006. ISSN 1741-2560. doi: 10.1088/1741-2560/3/1/R02.

Shin, Lee, Ahn, Cho, Jun, and Lee, 2015. Y Shin, S Lee, M Ahn, H Cho, S C Jun, and H.-N. Lee. *Noise robustness analysis of sparse representation based classification method for non-stationary EEG signal classification.* Biomedical Signal Processing and Control, 21, 8–18, 2015. ISSN 1746-8108. doi: 10.1016/j.bspc.2015.05.007.

Song and Yoon, 2015. X Song and S.-C. Yoon. *Improving brain-computer interface classification using adaptive common spatial patterns.* Computers in Biology and Medicine, 61, 150–160, 2015. ISSN 1879-0534. doi: 10.1016/j.compbiomed.2015.03.023.

Sun and Zhang, 2006. S Sun and C Zhang. *Adaptive feature extraction for EEG signal classification.* Medical and Biological Engineering and Computing, 44(10), 931–935, 2006. ISSN 0140-0118. doi: 10.1007/s11517-006-0107-4.

Sykacek, Roberts, and Stokes, 2004. P Sykacek, S J Roberts, and M Stokes. *Adaptive BCI Based on Variational Bayesian Kalman Filtering: An Empirical Evaluation.* IEEE Transactions on Biomedical Engineering, 51(5), 719–727, 2004. ISSN 0018-9294. doi: 10.1109/TBME.2004.824128.

Thomas, Guan, Lau, Vinod, and Ang, 2011. K P Thomas, C Guan, C T Lau, A P Vinod, and K K Ang. *Adaptive tracking of discriminative frequency components in electroencephalograms for a robust brain-computer interface.* Journal of Neural Engineering, 8(3), 2011. ISSN 1741-2560. doi: 10.1088/1741-2560/8/3/036007.

Tonet, Marinelli, Citi, Rossini, Rossini, Megali, and Dario, **2008**. Oliver Tonet, Martina Marinelli, Luca Citi, Paolo Maria Rossini, Luca Rossini, Giuseppe Megali, and Paolo Dario. *Defining brain-machine interface applications by matching interface performance with device requirements.* Journal of neuroscience methods, 167(1), 91–104, 2008. ISSN 0165-0270. doi: 10.1016/j.jneumeth.2007.03.015.

Vidaurre, Schlögl, Cabeza, Scherer, and Pfurtscheller, **2005**. C Vidaurre, A Schlögl, R Cabeza, R Scherer, and G Pfurtscheller. *Adaptive on-line classification for EEG-based Brain Computer Interfaces with ARR parameters and band power estimates.* Biomedizinische Technik, 50(11), 350–354, 2005. ISSN 0013-5585. doi: 10.1515/BMT.2005.049.

Vidaurre, Schlogl, Cabeza, Scherer, and Pfurtscheller, **jun 2006**. C Vidaurre, A Schlogl, R Cabeza, R Scherer, and G Pfurtscheller. *A fully on-line adaptive BCI.* IEEE transactions on bio-medical engineering, 53(6), 1214–1219, 2006. ISSN 0018-9294 (Print). doi: 10.1109/TBME.2006.873542.

Vidaurre, Schlögl, Cabeza, Scherer, and Pfurtscheller, **2007**. C Vidaurre, A Schlögl, R Cabeza, R Scherer, and G Pfurtscheller. *Study of on-line adaptive discriminant analysis for EEG-based brain computer interfaces.* IEEE Transactions on Biomedical Engineering, 54(3), 550–556, 2007. ISSN 0018-9294. doi: 10.1109/TBME.2006.888836.

Vidaurre, Kawanabe, Von Bünau, Blankertz, and Müller, **2011a**. C Vidaurre, M Kawanabe, P Von Bünau, B Blankertz, and K R Müller. *Toward unsupervised adaptation of LDA for brain-computer interfaces.* IEEE Transactions on Biomedical Engineering, 58(3 PART 1), 587–597, 2011. ISSN 0018-9294. doi: 10.1109/TBME.2010.2093133.

Vidaurre, Sannelli, Müller, and Blankertz, **2011b**. C Vidaurre, C Sannelli, K.-R. Müller, and B Blankertz. *Machine-learning-based coadaptive calibration for brain-computer interfaces.* Neural computation, 23(3), 791–816, 2011. ISSN 1530-888X. doi: 10.1162/NECO_a_00089.

Vidaurre, Sannelli, Müller, and Blankertz, **2011c**. Carmen Vidaurre, Claudia Sannelli, Klaus-Robert Müller, and Benjamin Blankertz. *Co-adaptive calibration to improve BCI efficiency.* Journal of Neural Engineering, 8(2), 025009, 2011. ISSN 1741-2552. doi: 10.1088/1741-2560/8/2/025009.

Wang, Xu, Wang, and Zhang, **2011**. J Wang, G.-z. Xu, L Wang, and H.-y. Zhang. *Features extraction of brain-computer interface based on adaptive autoregressive models.* Journal of Clinical Rehabilitative Tissue Engineering Research, 15(48), 9007–9010, 2011. ISSN 1673-8225. doi: 10.3969/j.issn.1673-8225.2011.48.019.

Wang, Xu, Wang, Yang, and Yan, **jan 2007**. Lei Wang, Guizhi Xu, Jiang Wang, Shuo Yang, and Weili Yan. *Feature extraction of mental task in BCI based on the method of approximate entropy.* Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 2007, 1941–4, 2007. ISSN 1557-170X. doi: 10.1109/IEMBS.2007.4352697.

Wang, Chaovalitwongse, and Wong, **2014**. S Wang, W A Chaovalitwongse, and S Wong. *A gradient-based adaptive learning framework for online seisure prediction.* International journal of data mining and bioinformatics, 10(1), 49–64, 2014. ISSN 1748-5673.

Wang, Veluvolu, and Lee, **2013**. Y Wang, K C Veluvolu, and M Lee. *Time-frequency analysis of band-limited EEG with BMFLC and Kalman filter for BCI applications.* Journal of NeuroEngineering and Rehabilitation, 10(1), 2013. ISSN 1743-0003. doi: 10.1186/1743-0003-10-109.

Woehrle, Krell, Straube, Kim, Kirchner, and Kirchner, **jul 2015**. Hendrik Woehrle, Mario M Krell, Sirko Straube, Su Kyoung Kim, Elsa A Kirchner, and Frank Kirchner. *An Adaptive Spatial Filter for User-Independent Single Trial Detection of Event-Related Potentials.* IEEE transactions on bio-medical engineering, 62(7), 1696–1705, 2015. ISSN 1558-2531 (Electronic). doi: 10.1109/TBME.2015.2402252.

Wu and Ge, **2013**. Y Wu and Y Ge. *A novel method for motor imagery EEG adaptive classification based biomimetic pattern recognition.* Neurocomputing, 116, 280–290, 2013. ISSN 0925-2312. doi: 10.1016/j.neucom.2012.03.030.

Yoon, Roberts, Dyson, and Gan, **2009**. J W Yoon, S J Roberts, M Dyson, and J Q Gan. *Adaptive classification for Brain Computer Interface systems using Sequential Monte Carlo sampling.* Neural Networks, 22(9), 1286–1294, 2009. ISSN 0893-6080. doi: 10.1016/j.neunet.2009.06.005.

Yoon, Roberts, Dyson, and Gan, **sep 2011**. Ji Won Yoon, Stephen J Roberts, Mathew Dyson, and John Q Gan. *Bayesian inference for an adaptive Ordered Probit model: an application to Brain Computer Interfacing.* Neural networks : the official journal of the International Neural Network Society, 24(7), 726–734, 2011. ISSN 1879-2782 (Electronic). doi: 10.1016/j.neunet.2011.03.019.

Yuxiao Yang, Shanechi, Yang, Shanechi, Yuxiao Yang, and Shanechi, **2014**. Yuxiao Yang, Maryam M Shanechi, Yuxiao Yang, Maryam M Shanechi, Yuxiao Yang, and Maryam M Shanechi. *An adaptive brain-machine interface algorithm for control of burst suppression in medical coma.* Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 2014, 1638–1641, 2014. ISSN 1557-170X. doi: 10.1109/EMBC.2014.6943919.