Denis Delisle Rodríguez

Brain-Computer Interface Based on Unsupervised Methods to Recognize Motor Intention for Command of a Robotic Exoskeleton

Vitoria - Brazil

December 2017

Denis Delisle Rodríguez

Brain-Computer Interface Based on Unsupervised Methods to Recognize Motor Intention for Command of a Robotic Exoskeleton

Thesis submitted for the Postgraduate Program in Electrical Engineering (PPGEE), Federal University of Espirito Santo (UFES) as a preliminar requirement to obtain the PhD degree in Electrical Engineering.

Federal University of Espirito Santo - UFES, Brazil

Postgraduate Program in Electrical Engineering

Supervisor: Dr. Teodiano Freire Bastos

Co-supervisor: Dr. Anselmo Frizera Neto

Vitoria - Brazil

December 2017

Denis Delisle Rodríguez

Brain-Computer Interface Based on Unsupervised Methods to Recognize Motor Intention for Command of a Robotic Exoskeleton

Thesis submitted for the Postgraduate Program in Electrical Engineering (PPGEE), Federal University of Espirito Santo (UFES) as a preliminar requirement to obtain the PhD degree in Electrical Engineering.

Examined by:

Prof. Dr. Teodiano Freire Bastos (PPGEE/UFES/Brazil)

Supervisor

Prof. Dr. Anselmo Frizera Neto (PPGEE/UFES/Brazil)

Co-supervisor

Prof. Dr. Patrick Marques Ciarelli (PPGEE/UFES/Brazil)

Jury

Prof. Dr. Eduardo Rocon (CAR/CSIC/Spain)

Jury

Prof. Dr. Antonio Padilha Lanari Bó (UnB/Brazil)

Jury

Vitoria, Brazil

"Accept the proposition that humans are not disabled. A person can never be broken. Our built environment, our technologies are broken and disabled. We the people need not accept our limitations, but can transcend disability through technological innovation".

Hugh Herr

"There is no greater disability in society, than the inability to see a person as more".

Robert M. Hensel

I dedicate this doctoral thesis to all my family and friends, especially to persons that supported me to receive each day better and to have focus on my profession, such as: my parents Luciano Delisle Puente and Luz Maria Rodriguez Correa, my sister and brother in law Liumila Delisle Rodriguez and Enry Vaillant, and my wife Elines da Penha Mazzani Delisle.

Acknowledgements

This thesis has been possible due to different institutions and persons. Then, I would like to thank to:

- My supervisor Dr. Teodiano Freire Bastos Filho and my co-supervisor Dr. Anselmo Frizera Neto. They provided me the opportunity to work in this wonderful, and challenging project. In addition, they helped me to grow as a person and professional.

- Project CAPES-MES (1223SC902-006), specially to Dr. Alberto López Delis, Dr. Roberto Sagaró, and Dr. Teodiano Freire Bastos Filho, who guided my research in NTA group at UFES as PhD student.

- Federal University of Espirito Santo (UFES), and Postgraduate Program in Electrical Engineering (PPGEE) at UFES, which accepted my PhD proposal, providing all the institutional support to finish this PhD thesis successfully.

- All professors of the PPGEE/UFES. They shared their knowledge along different courses.

- CAPES and CNPq that supported me these years to complete my PhD thesis successfully.

- University of Oriente (UO) and Center of Medical Biophysic (CBM) of the UO from Cuba, which allowed my PhD studies during 4 years here in Brazil, providing all necessary support. Additionally, both institutions are responsible for my professional formation.

- Emerging Leaders in the America Program (ELAP, Canada), Ryerson University (RU/ Toronto/ Canada), and my Canadian supervisor Dr. Sridhar Krishnan, who provided me a wonderful academic exchange at the SAR (Signal Analysis Research) Group of RU.

- All colleagues and classmates of the Department of Bioengineering from UO and CBM/Cuba, NTA group from UFES/Brazil, and SAR Group/Canada, who shared a lot experience with me in different ways.

- All my family and friends, who are very important in my life, wishing me all the best always, as well as providing all necessary support.

- All persons that contributed to improve the quality of this work.

Now, I would like to thank specially to:

- My supervisor Dr. Teodiano Freire Bastos Filho, who provided me all support and attention as professional and friend, in order to help me to achieve a fast adaptation to Brazil, and to have better days along these 4 years. Thanks for dedicate unconditionally a lot of time to our project, giving me their patience and confidence, as well as helping me with the offer of several opportunities to make me a competitive professional and a better person.

- Dr. Sridhar Krishnan, who provided me all support and attention on both professional and personal formation. I appreciate his patience and confidence, and his guiding, which helped me to be a competitive professional and a better person.

- All my family, especially to that persons that supported me to receive each day better, and to have focus on my profession, such as: my wife Elines da Penha Mazzani Delisle, my parents Luciano Delisle Puente and Luz Maria Rodriguez Correa, and my sister Liumila Delisle Rodriguez and her husband, Enry Vaillant.

- Dr. Alberto López Delis, my great friend, who provided me all support and attention as professional and friend. Thanks to your wonderful family too.

- John Jairo Villarejo Mayor, my great friend, who provided me my first reception in Brazil, doing easy my adaptation. Thanks for dedicating lot of your time in meeting related to our PhD projects.

- Javier Ferney Castillo Garcia and Anibal Cotrina, who introduced me with electroencephalography signals, sharing me lot of their knowledge.

- Manuel Ricardo Alfonso, Sandra Guzman and John Freddy, who provided me an amazing reception in Brazil, doing easy my adaptation.

- Ana Cecilia Villa Parra, my great friend, who provided me all support and attention as professional and friend. Thanks for dedicating lot of your time in meeting related to our PhD projects, and for sharing me your wonderful presence.

- Laura Susana Vargas Valencia, who provided me all support and attention as professional and friend. Thanks for dedicating your wonderful time to discuss about our PhD projects, and sharing me your amazing presence.

- Alan Floriano, who provided me all support and attention as professional and friend. Thanks for helping me with the Portuguese language in your short and important time. Always, I am very glad by our academic discussion.

- Dolores Pinheiro de Souza, my great friend, who provided me all support and attention as professional and friend. Thanks for sharing me your wonderful presence.

- Juan Carlos Garcia, Manuel Lores, Alexander Sóñora, Yamirka Alonso, Yulianela, Albys Ferrer, Tahimy Gonzalez, Lariza Portuondo, Yamari Poulot, Eglis Arévalo, Nelsón Gónzales, Reynier Maturell, Eduardo Duconger, who spended their short time to remain regularly in touch from Cuba, helping or visiting my family in Cuba.

- Mariana Cardoso Melo and Karen Bender, great and amazing persons, who helped me a lot with my adaptation during my visit at Ryerson University (Toronto, Canada), and who are still sharing their short time with me.

- Alice Rueda, Dharmendra Gurve, Jeevan Pan, Sharadha Kolappan, Michael Zara, Giordano Arezza, Abdelrahman Abdou who shared their short time with me, facilitating my adaptation during my visit in Toronto, being very patiente.

- Nicolas Valencia Jimenez and Mario Fernando Jiménez Hernádez, who shared their short time with me, giving their unconditional help when it was necessary.

Abstract

Stroke and road traffic injuries may severely affect movements of lower-limbs in humans, and consequently the locomotion, which plays an important role in daily activities, and the quality of life. Robotic exoskeleton is an emerging alternative, which may be used on patients with motor deficits in the lower extremities to provide motor rehabilitation and gait assistance. However, the effectiveness of robotic exoskeletons may be reduced by the autonomous ability of the robot to complete the movement without the patient involvement. Then, electroencephalography signals (EEG) have been addressed to design brain-computer interfaces (BCIs), in order to provide a communication pathway for patients perform a direct control on the exoskeleton using the motor intention, and thus increase their participation during the rehabilitation. Specially, activations related to motor planning may help to improve the close loop between user and exoskeleton, enhancing the cortical neuroplasticity. Motor planning begins before movement onset, thus, the training stage of BCIs may be affected by the intuitive labeling process, as it is not possible to use reference signals, such as goniometer or footswitch, to select those time periods really related to motor planning. Therefore, the gait planning recognition is a challenge, due to the high uncertainty of selected patterns, However, few BCIs based on unsupervised methods to recognize gait planning/stopping have been explored.

This Doctoral Thesis presents unsupervised methods to improve the performance of BCIs during gait planning/ stopping recognition. At this context, an adaptive spatial filter for on-line processing based on the Concordance Correlation Coefficient (CCC) was addressed to preserve the useful information on EEG signals, while rejecting neighbor electrodes around the electrode of interest. Here, two methods for electrode selection were proposed. First, both standard deviation and CCC between target electrodes and their correspondent neighbor electrodes are analyzed on sliding windows to select those neighbors that are highly correlated. Second, Zscore analysis is performed to reject those neighbor electrodes whose amplitude values presented significant difference in relation to other neighbors.

Furthermore, another method that uses the representation entropy and the maximal information compression index (MICI) was proposed for feature selection, which may be robust to select patterns, as only it depends on cluster distribution. In addition, a statistical analysis was introduced here to adjust, in the training stage of BCIs, regularized classifiers, such as support vector machine (SVM) and regularized discriminant analysis (RDA).

Six subjects were adopted to evaluate the performance of different BCIs based on the proposed

methods, during gait planning/stopping recognition.

The unsupervised approach for feature selection showed similar performance to other methods based on linear discriminant analysis (LDA), when it was applied in a BCI based on the traditional Weighted Average to recognize gait planning. Additionally, the proposed adaptive filter improved the performance of BCIs based on traditional spatial filters, such as Local Average Reference (LAR) and WAR, as well as others BCIs based on powerful methods, such as Common Spatial Pattern (CSP), Filter Bank Common Spatial Pattern (FBCSP) and Riemannian kernel (RK). RK presented the best performance in comparison to CSP and FBCSP, which agrees with the hypothesis that unsupervised methods may be more appropriate to analyze clusters of high uncertainty, as those formed by motor planning.

BCIs using adaptive filter based on Zscore analysis, with an unsupervised approach for feature selection and RDA showed promising results to recognize both gait planning and gait stopping, achieving for three subjects, good values of true positive rate (>70%) and false positive (<16%). Thus, the proposed methods may be used to obtain an optimized BCI that preserves the useful information, enhancing the gait planning/stopping recognition. In addition, the method for feature selection has low computational cost, which may be suitable for applications that demand short time of training, such as clinical application time.

Keywords: artifact reduction, brain-computer interface, EEG, EOG, Laplacian, spatial filter, feature selection, gait planning, gait stopping, SSVEP.

List of Tables

List of Figures

Contents

Glossary

A Attenuation.

ACC Accuracy.

Ad LAR Adaptive Local Average Reference.

Ad WAR Adaptive Weighted Average Reference.

ALLOR Advanced Lower-Limb Orthosis for Rehabilitation.

ANC Adaptive Noise Cancelling.

AR Autoregressive.

BCI Brain-Computer Interface.

BeP Bereitschaftspotential.

BF Biceps Femoris.

BP Band Power.

CCA Canonical Correlation Analysis.

CCC Concordance Correlation Coefficient.

CSP Common Spatial Pattern.

EEG Electroencephalographic Signal.

EOG ocular movements.

ERD Event-Related Synchronization.

ERD/ERS Event-Related Desynchronization/Synchronization.

ERP Event-Related Potentials.

ERS Event-Related Desynchronization.

ES Erector Spinae.

F1 F1 value.

FBCSP Filter-Bank Common Spatial Pattern.

FD Fractal Dimension.

FDH Fractal Dimension from Higuchi.

FDSH Fractal Dimension from Sevcik's and Higuchi.

FFT Fast Fourier Transform.

FPR False Positive Rate.

G Gastrocnemius.

GA Genetic Algorithm.

HAL Hybrid Assistive Leg.

ICA Independent Component Analysis.

K-E Knee Extension.

K-F Knee Flexion.

LAR Local Average Reference.

LDA Linear Discriminant Analysis.

MAV Mean Absolute Value.

MFC medial frontocentral.

MIBIF Mutual Information-based Best Individual Feature.

MICI Maximal Information Compression Index.

MRCP Motor Related Cortical Potential.

PCA Principal Component Analysis.

PPALO Pneumatic Power Active Lower Limb Orthosis.

PSD power spectral density.

QDA Quadratic Discriminant Analysis.

RBF Radial Basic Function.

RCF Rectus Femoris.

RDA Regularized Discriminant Analysis.

RE Representation Entropy.

RF Reference-Free.

RK Riemmaninan Geometry based Kernel.

S Semitendinosus.

SCI Spinal Cord Injury.

SCM Sample Covariance Matrix.

SCP Slow Cortical Potential.

SD standard deviation.

sEMG surface Electromyographic Signal.

Si-R Sit Rest.

SMA sensorimotor areas.

SNR Signal to Noise Ratio.

SPD Symmetric Positive Definite.

SSVEP steady-state visual evoked potentials.

SVM Support Vector Machine.

TPR True Positive Rate.

VD Virtual Distance.

VL Vastus Lateralis.

WAR Weighted Average Reference.

WL Wavelength.

WSE Walking Supporting Exoskeleto.

Chapter 1

Introduction

1.1 Motivation

Nowadays, overweight, obesity and ageing affect many low- and middle-income and risk countries worldwide (WHO, 2017a, 2016, 2015b, 2015a). Overweight and obesity are defined as abnormal or excessive fat accumulation that may cause a risk to health. A person with a body mass index (BMI) from 25 to 30 is considered overweight, and obese for BMI higher than 30 (WHO, 2015c). In 2014, approximately a total of 1.9 billion adults (18 years old, and older) were overweight, while a total of 600 million were obese. In 2015, overweight population in Brazil represented a total of 67% and 74% of men and women, respectively (WHO, 2015c).
On the other hand, people that overcome the 60 years old will rise from 900 million to 2 billion between 2015 and 2050, which represent proportions from 12 to 22% of the world population. Additionally, older people of low- and middle-income countries are more risks to acquire diseases of morbidity than those in the rich world (WHO, 2016, 2015b, 2015a), such as neural disorders, osteoarthritis, back and neck pain.
Overweight, obesity and ageing have high risk to develop cardiovascular diseases such as stroke, which may provide motor damages or impairs (WHO, 2015c, 2014, 2015b). Cardiovascular diseases are classified as the main cause of death and disability in the world. As a result, all years 17.3 million cardiovascular deaths are reported. Additionally, heart attacks and strokes are

responsible for 7.3 million and 6.2 million, respectively (WHO, 2014).

A stroke is caused by the interruption of the blood fluid to the brain, usually when a blood vessel bursts, or it is blocked by a clot. Thus, these events may cut off the supply of oxygen and nutrients, causing damage to the brain tissue (WHO, 2014). The effects of a stroke depend on which part of the brain is injured and how severely it is affected (WHO, 2014). In 2005, a prevalence of 32% of Brazilian population suffered cardiovascular disease (WHO, 2005). Similar values were prognosticated over the next 10 years (WHO, 2005).

Moreover, around 1.25 million people died from road traffic injuries in 2013, where another group of people (from 20 to 50 million) incurring non-fatal injuries as a result of road traffic collisions or crashes. Road traffic represent the ninth leading cause of death globally (WHO, 2017b), where it together to fall and violence may cause spinal cord injury (SCI). SCI may totally affect the lower limb mobility (WHO, 2013a, 2013b), representing a prevalence per year of several new cases per million population.

Several strategies can be adopted on people post-stroke and with SCI, in order to reduce dependency and traumatic incidence, and avoid secondary complications. For example, barriers removing and adequate supports may be used to minimize the poverty and social exclusion associated to the injury. Additionally, devices for motor rehabilitation and assistance to reduce dependence can also be carried out (PONS, 2008). This way, several technologies as robotic exoskeletons have been developed, which may operate alongside human limbs to assist during walking (PONS, 2008; HORTAL et al., 2016; VILLA-PARRA et al., 2015).

Robotic training has several advantages, such as the reduction of the effort of physical therapists per patient and the possibility to objectively quantify rehabilitation parameters (JIANG et al., 2015). For this purpose, at UFES/Brazil, it is developing an active knee exoskeleton named ALLOR “Advanced Lower-Limb Orthosis”, which may be used on patients with motor deficits in the lower extremities for knee rehabilitation and gait assistance. However, the effectiveness of robotic exoskeletons may be reduced by the autonomous ability of the robot to complete the movement without the patient involvement (JIANG et al., 2015). It has been demonstrated that the active participation of the patient may be crucial to improve the outcome of the rehabilitation (JIANG et al., 2015). As a result, several brain-computer interfaces (BCIs)

based on electroencephalography (EEG) signals have been proposed for conveying control commands directly from the brain toward robotic exoskeletons, in order to reproduce more natural movements, increase the engagement and motivation of the patient, as well as to enhance neuroplasticity (JIANG et al., 2015; XU et al., 2014; HASHIMOTO; USHIBA, 2013; GALLEGO et al., 2012; HORTAL et al., 2016; SBURLEA; MONTESANO; MINGUEZ, 2015; SBURLEA et al., 2015). BCIs are communication systems that do not depend on the brain's normal output pathways of peripheral nerves and muscles (WOLPAW et al., 2000; PFURTSCHELLER; FLOTZINGER; KALCHER, 1993; WOLPAW; MCFARLAND, 1994). These BCIs may measure specific components from EEG, in order to be used on users with severe motor disability to control end-applications, such as cursors (WOLPAW et al., 1991; MCFARLAND et al., 1997), televisions (ALSHBATAT et al., 2014), wheelchairs (MÜLLER; BASTOS; FILHO, 2013; DIEZ et al., 2013), autonomous car, word processing, electrical stimulator (XU et al., 2014), robotic prosthesis and exoskeletons (XU et al., 2014; ROCON et al., 2010).

Thus, a suitable BCII may be proposed to conveying control command for our exoskeleton ALLOR, in order to produce plasticity on people post-stroke or SCI, and to improve its activities of daily living. The next sections address scientific issues around the state-of-art of BCIs to recognize gait intention, as well as the hypothesis, objectives and justification of this thesis.

1.2 Scientific Issues

Several BCIs have been proposed to conveying control commands for lower-limb robotic exoskeletons through a closed-loop (JIANG et al., 2015; HORTAL et al., 2016; XU et al., 2014; HASHIMOTO; USHIBA, 2013; SBURLEA; MONTESANO; MINGUEZ, 2015; SBURLEA et al., 2015; VELU; SA, 2013), in order to provide support to people with neural injury, such as spinal cord injury, stroke, among others. For this purpose, BCIs measure patterns related to motor intention, such as event-related desynchronization/synchronization (ERD/ERS)(HORTAL et al., 2016) and motor related cortical potentials (MRCPs) have been used to anticipate movements (SHIBASAKI; HALLETT, 2006), providing a direct control of robotic exoskeletons with more natural movements (JIANG et al., 2015; XU et al., 2014; HASHIMOTO; USHIBA, 2013;

HORTAL et al., 2016). For example, previous works have demonstrated that the EEG power decreases on the mu-ERD band (from 8 to 12 Hz) before the movement onset (from -2.0 to 0 s) (PFURTSCHELLER; SILVA, 1999), while the EEG power increases on the beta-ERS band (from 26 to 30 Hz) around +1.5 s after the movement offset (PFURTSCHELLER; SILVA, 1999). However, these potentials may be easily blurred due to the typical poor signal to noise ratio (SNR) of EEG signals. SNR of EEG may be affected by common interferences and artifacts, such as electrode and eyes movements, eye blinks, myoelectric and cardiac activities, non-μ-rhythm of EEG components as visual alpha rhythm, among others (MCFARLAND et al., 1997; KILICARSLAN; GROSSMAN; CONTRERAS-VIDAL, 2016). Physiological artifacts, such as blinks and ocular movements (EOG) are present in most EEG recording, which due to the volume conduction, corrupt all EEG electrode measurements, changing profiles and amplitude distributions (KILICARSLAN; GROSSMAN; CONTRERAS-VIDAL, 2016).

In order to improve the SNR of EEG signals, and consequently the motor intention recognition, several authors have proposed BCIs based on independent component analysis (ICA) (JIANG et al., 2015; VELU; SA, 2013), fastICA, Common Spatial Patterns (CSP) (BLANKERTZ et al., 2008; MÜLLER-GERKING; PFURTSCHELLER; FLYVBJERG, 1999), Filter Bank - Common Spatial Patterns (FBCSP) (ANG et al., 2012a), Riemannian based kernel of covariance matrices (RK) (BARACHANT et al., 2013), and filters based on Euclidean distances such as Laplacian, Local Average Reference (LAR) and Weighted Average Reference (WAR) (HORTAL et al., 2016; SBURLEA; MONTESANO; MINGUEZ, 2015; MCFARLAND et al., 1997).

Regarding ICA (JUNG et al., 1998a), although offline acquired mixing/demixing can be applied for real time applications, its performance over some large periods of time may be affected, due to the fact that it is practically impossible to obtain a mixing matrix from a training database that contains all possible physiological and non-physiological artifacts, as well as several internal states of the user due to changes in cognitive-motor-affective behaviors, medication and health status (KILICARSLAN; GROSSMAN; CONTRERAS-VIDAL, 2016). Other filters such as Laplacian, LAR and glswar do not depend on underlying data, therefore, they could contribute to add on the electrode of interest undesirable artifacts from neighbor electrodes (PFURTSCHELLER; NEUPER; BERGER, 1994; WOLPAW; MCFARLAND, 1994).

Additionally, the combination of bandpass filters with methods for feature extraction, such as CSP (BLANKERTZ et al., 2008; MÜLLER-GERKING; PFURTSCHELLER; FLYVBJERG, 1999), FBCSP (ANG et al., 2012a), and RK (BARACHANT et al., 2013) have been applied in BCIs with promissing results during on-line processing, to recognize imagery motor and mental tasks (BLANKERTZ et al., 2008; MÜLLER-GERKING; PFURTSCHELLER; FLYVBJERG, 1999; BARACHANT et al., 2013; ANG et al., 2012a). Anyways, BCIs based on these approaches have been little explored to control exoskeletons, and their challenging purposes, such as gait rehabilitation or assistance (JIANG et al., 2015; HORTAL et al., 2016; VELU; SA, 2013; SBURLEA; MONTESANO; MINGUEZ, 2015). In contrast, several BCIs based on traditional Laplacian filters have been proposed to recognize the gait onset (JIANG et al., 2015; HORTAL et al., 2016; VELU; SA, 2013).

Different from the movement onset recognition, the motor planning recognition (around 2 s before the movement onset) may allow to detect the motor intention with very short delay (some hundred ms), so that the resulting control can be perceived as a closed loop (XU et al., 2014). Nevertheless, little BCIs have been proposed to recognize both gait planning/stopping (HORTAL et al., 2016; SBURLEA; MONTESANO; MINGUEZ, 2015; SBURLEA et al., 2015), which is very important to reproduce a more natural movement after a voluntary intention.

Motor anticipation recognition may be affected by several factors, such as time variability linked to the intra-subject reaction, EEG variability, and others unknown criteria as event duration. Additionally, patterns related to motor planning should be labeled before any muscular activation (from -2.0 to 0 s). Thus, it is not possible to use on this time period some external reference (goniometer, footswitch, among others) to locate when these events occur. Then, the labeled process is a task with high uncertainty. Therefore, BCIs based on unsupervised methods may be suitable to reduce effects of these issues. However, unsupervised methods have been little explored in the learning stage of a BCI, during the gait intention recognition (HORTAL et al., 2016; SBURLEA; MONTESANO; MINGUEZ, 2015; SBURLEA et al., 2015; VELU; SA, 2013; JIANG et al., 2015; DO et al., 2013; SBURLEA; MONTESANO; MINGUEZ, 2017).

Some researchers have used methods for feature selection , such as genetic algorithms (GAs) to improve the BCI speed (JIANG et al., 2015; BULEA et al., 2014; BAI et al., 2011; GARRETT

et al., 2003), however, their performance are highly correlated to the classifier complexity, increasing the time delay of the training stage.

In summary, few BCIs have been addressed to recognize gait planning/stopping, although it has been demonstrated their relevant role to enhance neuroplasticity on post-stroke people (SBURLEA; MONTESANO; MINGUEZ, 2015; SBURLEA et al., 2015; HORTAL et al., 2016). Additionally, BCIs based on unsupervised methods have been little explored for preprocessing and training stages, although they could improve the BCI's performance during static and dynamic movements, specially walking.

1.3 Hypothesis

It is hypothesized that a BCI based on an adaptive filter using Concordance Correlation Coefficient (CCC) to preserve the information of interest, and unsupervised methods, such as feature selection using the representation entropy (RE) index on the training stage, may be robust to uncertain patterns, which could improve the gait planning/stopping recognition.

1.4 Objectives

The main objective of this research is to propose a BCI based on unsupervised methods of low computational cost for conveying control commands to a robotic exoskeleton, in order to provide support to patients with severe motor disability during the gait starting.

The following specific objectives are proposed to obtain the BCI.

1. A method based on the Concordance Correlation Coefficient to provide an adaptive behavior to traditional spatial filters, such as LAR and WAR, in order to reduce common interferences throughout EEG, while preserving the information of interest.

2. An unsupervised method of low computational cost for feature selection, robust enough to uncertain patterns, in order to improve the BCI speed.

3. A statistical analysis to select the classifier that provides the best performance during the training stage of the BCI.

4. A BCI based on the combination of the proposed methods, in order to preserve the information of interest on EEG signals, providing a better calibration stage to improve the pattern recognition.

5. An evaluation on the information preservation of the proposed adaptive filter throughout an SSVEP database.

6. An experimental protocol to study motor patterns related to gait planning/stopping.

7. An evaluation of several classifiers, such as Linear Discriminant Analysis (LDA), Regularized Discriminant Analysis (RDA) and Support Vector Machine (SVM) with kernel lineal, coming to obtain the best BCI to recognize gait planning/stopping.

1.5 Justification

Lower-limbs robotic exoskeletons may be used to assist people with neural injury, such as spinal cord injury, stroke, among others (PONS, 2008). Additionally, a BCI can be used to convey control commands from the brain for a robotic exoskeleton to reproduce more natural movements after a voluntary intention, aiming to produce neuroplastic changes in the cortical pathways on these subjects (JIANG et al., 2015; HORTAL et al., 2016; VELU; SA, 2013; SBURLEA; MONTESANO; MINGUEZ, 2015). For this purpose, patterns related to pre-movement actions or motor planning (from -2 s before the movement onset), such as ERD/ERS (PFURTSCHELLER; SILVA, 1999) and MRCP (SHIBASAKI; HALLETT, 2006) potentials, may be used to recognize motor intention with very short delay, so that the resulting control can be perceived as a closed loop (XU et al., 2014).

Events related to motor intention, such as ERD/ERS, can be blurred by the poor SNR of EEG, due to several sources of interferences and artifacts, such as electrode and eyes movements, eyes blink, myoelectric and cardiac activities, non-μ-rhythm of EEG components as

visual alpha rhythm, among others (MCFARLAND et al., 1997; KILICARSLAN; GROSSMAN; CONTRERAS-VIDAL, 2016). Moreover, the motor anticipation recognition may be affected by process of high uncertainty, such as pattern labeling, as these events are located before any muscular activation (from -2.0 to 0s), where it is not possible to use goniometer and footswitch, in order to select correctly these patterns. Therefore, BCIs based on unsupervised methods may be suitable to reduce effects of these issues.

Several BCIs have been proposed to recognize gait planning (SBURLEA; MONTESANO; MINGUEZ, 2015; SBURLEA et al., 2015; SBURLEA; MONTESANO; MINGUEZ, 2017; DELISLE-RODRIGUEZ et al., 2017), onset (JIANG et al., 2015; HORTAL et al., 2016; VELU; SA, 2013), and stopping (HORTAL et al., 2016; DELISLE-RODRIGUEZ et al., 2017). BCIs for gait onset recognition have been more explored, for example, an MRCP template extracted from several trials have been used in a BCI based on the Matched filter to detect gait onset (JIANG et al., 2015). Here, ICA and Laplacian filters were also applied in the pre-processing stage to improve the SNR of EEG signals, improving the BCI performance with true positive rate (TPR) and false positive rate (FPR) of $76.9 \pm 8.97\%$ and 2.93 ± 1.09 per minute, respectively. MRCPs occur with small amplitudes (typically lies between 5 and 30 μV) at low frequency (from 0.1 to 4 Hz) (SHIBASAKI; HALLETT, 2006; JIANG et al., 2015), therefore, it may be masked by potentials of higher frequency band. Thus, BCIs based on MRCPs need several trials in the calibration stage to obtain an averaged MRCP, in order to remove the background noise of each trial (SBURLEA; MONTESANO; MINGUEZ, 2015).

In addition, a BCI based on the MRCP phase representation was proposed to decrease the time-consuming in the calibration stage, enhancing the gait intention recognition (SBURLEA; MONTESANO; MINGUEZ, 2017). As a result, a comparison of three BCIs based on MRCP amplitude, MRCP instantaneous phase, and a combination of both representation was carried out, where the combination of both MRCP features showed the best performance (accuracy of 66.5% on healthy subjects, and 63.3% on stroke patients).

Other BCIs that compute features on mu and beta frequency bands using the power spectral density method have been proposed to recognize gait intention (DO et al., 2013; HORTAL et al., 2016), with accuracy (ACC) of 68.6% and 86.3%, respectively. Both researches were con-

ducted on subjects with spinal cord injury, but three classes (resting, starting, and stopping) were considered in (HORTAL et al., 2016).

Although patterns related to motor planning may improve the closed-loop during the user-exoskeleton interactions, few BCIs based on motor planning have been proposed (SBURLEA; MONTESANO; MINGUEZ, 2015; SBURLEA et al., 2015). For example, a combination of both MRCP and ERD/ERS was evaluated on nine post-stroke patients, in order to recognize pre-movement states (ACC of 64%, in a range from 18 to 85.2%) by a BCI during walking (SBURLEA et al., 2015). It is worth noting that patterns related to motor planning on gamma-ERS band were used (PFURTSCHELLER et al., 2003; SEEBER et al., 2015; PFURTSCHELLER; SILVA, 1999). Although, this band has been little explored to recognize gait intention, researches have recently demonstrated that high and low gamma-ERS bands may be used for gait analysis (SEEBER et al., 2015; COSTA et al., 2016).

Additionally, several works have been conducted during walking to recognize action types and directions (VELU; SA, 2013), as well as to detect unexpected obstacles (SALAZAR-VARAS et al., 2015). For these objectives, coefficients from 9 levels of discrete wavelet decompositions using Daubechies Wavelet 4 as a mother wave were computed (VELU; SA, 2013), where coefficients obtained at low frequency bands were highly weighted. This evidences an important role of low frequency components, also when using slow cortical potentials (SCP) or MRCPs, to recognize motor intention, which agrees with (JIANG et al., 2015; SBURLEA; MONTESANO; MINGUEZ, 2015; LEW et al., 2012; SHIBASAKI; HALLETT, 2006).

Regarding BCIs specially designed for gait intention recognition, several authors have commonly used classifiers, such as Linear Discriminant Analysis (LDA) (JIANG et al., 2015), Regularized Linear Discriminant Analysis (RDA) (VELU; SA, 2013) and Support Vector Machine (SVM) (HORTAL et al., 2016). Furthermore, pre-processing methods, such as ICA, and Laplacian filters have been adopted to improve the SNR of EEG signals (HORTAL et al., 2016; JIANG et al., 2015; VELU; SA, 2013). Moreover, since EEG signals are highly variable between subjects and recording sessions, typically BCIs are calibrated at the beginning of each session. This calibration stage should be short, especially when patients that suffering locomotor disabilities cannot maintain a bipedal position for a long time. To improve this issue,

a BCI to execute continuous recognition of gait onsets without session-to-session recalibration was proposed (SBURLEA; MONTESANO; MINGUEZ, 2015). On the other hand, BCIs based on unsupervised learning to recognize gait planning/stops have been little explored.

BCIs with acceptable speed are suitable to obtain closed-loop control that produce more natural movements after a voluntary intention, which provide a neuroplastic change in the cortical pathways on people with neural injury. Previous works have demonstrated that the BCI speed and accuracy can be affected by the SNR achieved through spatial and temporal filtering methods (MCFARLAND et al., 1997). In the literature, few adaptive methods have been reported for BCIs, in order to improve the motor patterns recognition during dynamic movements, as walking. Nowadays, almost all BCIs developed on system for gait rehabilitation/assistance, are based on ICA and spatial filters based on Euclidean distance (Laplacian, LAR and WAR) (JIANG et al., 2015; HORTAL et al., 2016; VELU; SA, 2013). ICA can be also applied for real time applications but the performance of real-time data over long period of time may be affected, as it is practically impossible to obtain a mixing matrix from a training database that contains all possible physiological and non-physiological artifacts (KILICARSLAN; GROSSMAN; CONTRERAS-VIDAL, 2016). Furthermore, the operation of Laplacian, LAR and WAR filters does not depend on underlying data, which may contribute to add on the electrode of interest undesirable artifacts.

Other adaptive filters have been proposed to remove artifacts on EEG signals (KILICARSLAN; GROSSMAN; CONTRERAS-VIDAL, 2016; KHATUN; MAHAJAN; MORSHED, 2016), but they have not been applied in BCIs for gait intention recognition. For example, principal component analysis (PCA) have been used to remove artifacts in offline analysis (JUNG et al., 1998b). These BCIs based on ICA and PCA have demonstrated a good performance in motor intention recognition. However, additional stages, such as artifact annotation by trained specialist and new classifiers must be included to identify possible patterns correlated to artifacts throughout EEG.

In particular (MULLEN et al., 2013), a method based on PCA, denoted as artifact subspace reconstruction, was proposed to remove artifacts on a limited time window of EEG, which yields promising results for real-time applications. Moreover, adaptive noise cancelling

(ANC) methods have been developed for signal extraction from noise corrupted measurements, in which several adaptive estimators as recursive least squares and Kalman filter have been used (VASEGHI, 2008; SWEENEY et al., 2012; SWEENEY; MCLOONE; WARD, 2013; GUERRERO-MOSQUERA; NAVIA-VÁZQUEZ, 2012). Many of these methods do not adopt the non-stationary behavior of EEG signals, which is considered as time-varying frequency characteristics in general. This condition may not exist in some cases, however, it is likely to exist in EEG recordings over duration of an experimental session, and across sessions. These fluctuations create time-varying characteristics over the course of an experimental session, which require adaptive and robust filters to reject these abnormal events.

Additionally, several studies have used standard electrode locations close to the ocular source, such as frontal electrodes FP1, FP2, FT9, and FT10 to provide a reference input with EOG components for methods as ANC to remove artifacts (PUTHUSSERYPADY; RATNARAJAH, 2005). Other authors have used H^{∞} to obtain an adaptive filter using EOG components acquired around the eyes as reference input, which considers full or partial superposition of artifacts, such as eye blink, eye movement, and signal drift throughout EEG (KILICARSLAN; GROSSMAN; CONTRERAS-VIDAL, 2016). Furthermore, another method has been proposed for a single channel scenery, using the wavelet transform decomposition (KHATUN; MAHAJAN; MORSHED, 2016). These aforementioned methods have demonstrated a good performance to remove EOG artifacts on EEG, preserving the neural information on the reconstructed segments.

In summary, one of the most important challenges for BCIs to command robotic exoskeletons is to provide a closed-loop control with more natural movements, as well as reduce time-consuming related to session-by-session re-calibration (or training). Thus, methods of low computational cost, robust to artifacts and uncertainty due to pattern labeling must be explored, in order to obtain a better BCI performance. In spite of the large number of stroke incidents affecting the locomotion, few attention has been devoted to study the pre-movement recognition during walking.

1.6 Organization

This document is composed of four chapters. Chapter 1 introduces the motivation, scientific issue and the justification of this PhD proposal.

Chapter 2 presents several technologies for lower-limb rehabilitation, including the role of the brain to improve the close-loop control, between human and end-applications, such as exoskeletons. Additionally, brain areas and cortical rhythms related to motor intention are described. Furthermore, this chapter introduces several methods used in BCIs, to recognize mainly gait intention recognition.

Chapter 3 describes proposed methods of low computational cost to improve the BCI performance during gait planning/stopping recognition. This way, experimental protocol and methodology of evaluation are presented, in order to validate proposed designs of BCIs to recognize gait planning/stopping.

Chapter 4 presents obtained results and the discussion, including a comparison with methods of state-of-art. Finally, the Conclusion and Future Works are presented.

Chapter 2

Theoretical Background

This chapter presents several existing technologies for lower-limbs rehabilitation/assistance, such as lower-limb exoskeletons, and active ortheses (see Section 2.1). Additionally, this chapter discusses how biofeedback information as brain signals can increase the people involvement with these technologies during motor rehabilitation, enhancing neuroplasticity. Here, the role of the brain to command motor tasks is introduced, and oscillations throughout raw EEG related to real and imaginary motor movements are described. EEG signals have small amplitudes and poor SNR, therefore, the information of interest can be easily blurred. For this reason, different factors that affect EEG signals, and consequently the aforementioned brain oscillations are discussed here. Finally, a summary of previous methods proposed for BCIs to recognize gait planning/stopping are presented in Section 2.3.

2.1 Robotic orthoses to support lower-limb

Wearable robots are persons-oriented robots, which can be used to support the function of a limb, or replace it completely (PONS, 2008). These robotic devices named, exoskeleton or orthosis, may be used alongside of human limbs. An orthosis is a mechanical structure designed to map the anatomy of the human limb, in order to restore lost or weak functions of subjects with neural disease, or dysfunction. The robotic counterparts of orthoses are robotic exoskele-

tons (PONS, 2008). In this case, the objective of exoskeletons is to provide support for the correspondent human limb complete a defined movement, restoring the handicapped function. Exoskeletons can be defined, according to the application, as augmentative, alternative, ambulatory, functional, and for rehabilitation. Additionally, these robotic devices can be classified according to the section of the body assisted, for instance, upper limb, lower limb, and full body. Other definition can be carried out referencing degrees of freedom, type of sensors that may provide feedback using biosignals, functional electric stimulation, among others (VITECKOVA; KUTILEK; JIRINA, 2013; HUSSAIN; XIE; LIU, 2011a; PONS, 2008). On the other hand, also several orthoses have been built to support people during walking or for knee rehabilitation (VITECKOVA; KUTILEK; JIRINA, 2013; HUSSAIN; XIE; LIU, 2011a). The Hybrid Assistive Leg (HAL) was the first exoskeleton used on a hospital for rehabilitation training (HUSSAIN; XIE; LIU, 2011b), while the Self-Adjusting Knee Exoskeleton was designed to assist people with knee injuries (ERGIN; PATOGLU, 2011).

Figure 2.1 shows other examples of orthoses, such as the Vanderbilt lower limb exoskeleton (FARRIS; QUINTERO; GOLDFARB, 2011), Pneumatic Power Active Lower Limb Orthosis (PPALO) (HUANG; CHEN, 2013), Robotic Knee Exoskeleton (GAMS et al., 2013), and Walking Supporting Exoskeleton (WSE) (ONEN et al., 2014). The first one was built to provide mobility to paraplegics, while PPALO was developed to aid patients with spinal cord injuries or poliomyelitis, and the WSE was proposed to support patients during walking, sitting and standing.

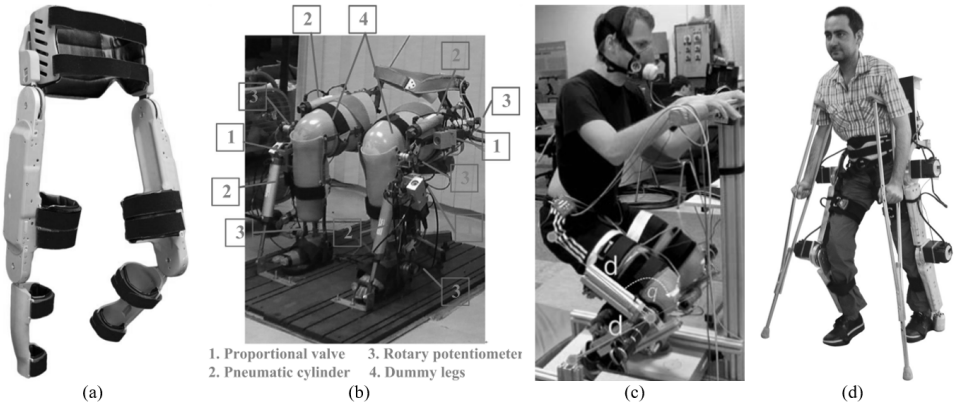

Figure 2.1: Examples of lower-limb exoskeletons: (a) Valderbilt; (b) PPALO lower limb orthosis; (c) Robotic Knee Exoskeleton; (d) Walking Supporting Exoskeleton (WSE).

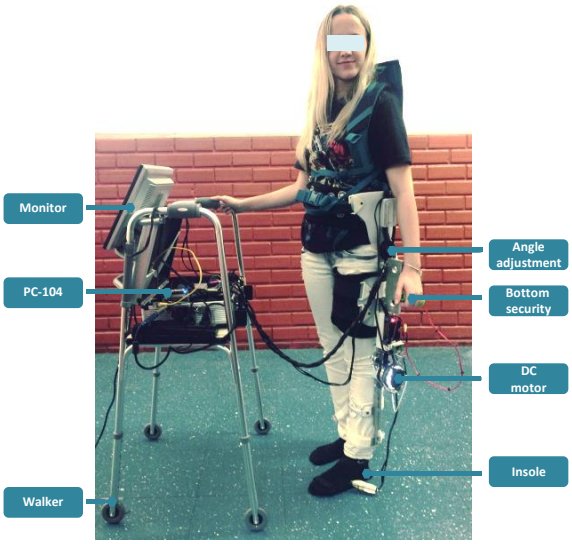

Figure 2.2: Advanced Lower-limb Orthersis for Rehabilitation (ALLOR) of UFES/Brazil.

Figure 2.2 shows the active knee exoskeleton developed at the Universidade Federal de Espirito Santo (UFES)/ Brazil, termed ALLOR (Advanced Lower-Limb Orthosis for Rehabilitation), which can be applied for knee and gait rehabilitation/assistance. ALLOR provides mechanical power at the left knee joint and torque for subject-robot interaction. Position and velocity controllers are used at ALLOR to allow passive training, tracking trajectories defined by a physiotherapist, during knee or gait rehabilitation/assistance.

Robotic training has several advantages, such as the reduction of the effort of physical therapists per patient and the possibility to objectively quantify rehabilitation parameters (JIANG et al., 2015). Also, the effectiveness of the rehabilitation may be reduced by the autonomous ability of the robot to complete the movement without the patient involvement (JIANG et al., 2015). In addition, it has been demonstrated that the active participation of the patient may be crucial to improve the outcome of the rehabilitation (JIANG et al., 2015). As a result, electroencephalography (EEG) signals have been proposed for conveying control commands directly from the brain toward robotic exoskeletons, in order to reproduce more natural movements, increase the engagement and motivation of the patient, as well as to enhance neuroplasticity (JIANG et al., 2015; XU et al., 2014; HASHIMOTO; USHIBA, 2013; GALLEGO et al., 2012; HORTAL et al., 2016; SBURLEA; MONTESANO; MINGUEZ, 2015; SBURLEA et al., 2015). Brain-computer interfaces (BCIs) are communication systems that do not depend on the brain's normal output pathways of peripheral nerves and muscles (WOLPAW et al., 2000;

PFURTSCHELLER; FLOTZINGER; KALCHER, 1993; WOLPAW; MCFARLAND, 1994). The next section provides general details related to the brain and the role of BCIs to command motor task.

2.2 Motor rhythms on EEG, and its role for motor rehabilitation

Figure 2.3 shows brain regions, such as lobes frontal, parietal, occipital, and temporal. Furthermore, regions linked to motor tasks are presented. Also, Figure 2.3b shows that upper limbs, eyes, and other body parts are controlled by the brain, using their respective contra-lateral regions. Except for lower-limbs, brain signals linked to motor functions of several body parts, such as upper limbs, can be more accessible to study. For this reason, researches have shown more progress in studies of upper limbs (LEW et al., 2012; BLANKERTZ et al., 2008; BARACHANT et al., 2013; PFURTSCHELLER; SILVA, 1999; PFURTSCHELLER et al., 2003).

The brain potentials linked to motor activities can be understood in real time through raw

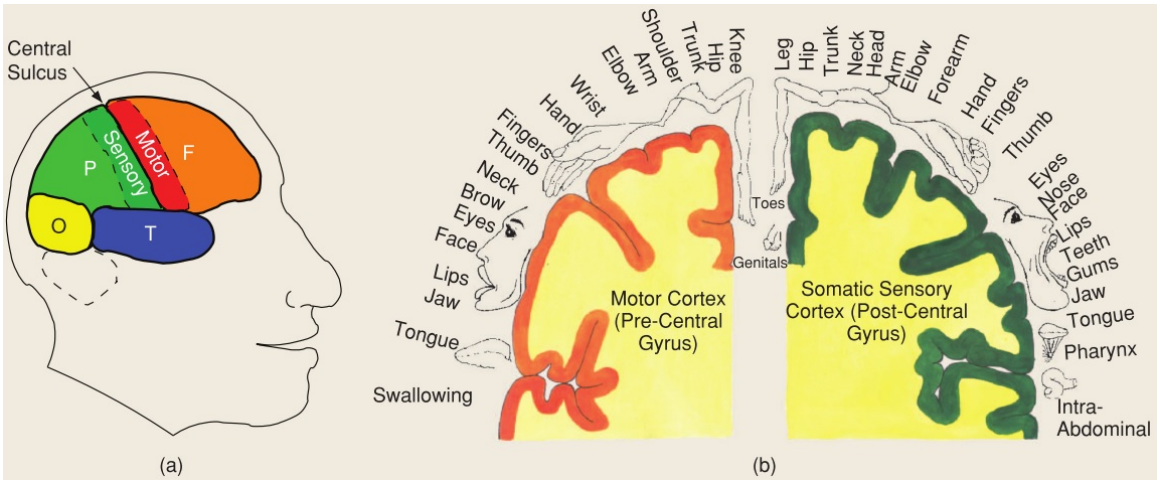

Figure 2.3: Representation of brain region linked with motor tasks. (a) Lobes of the brain, such as frontal, parietal, occipital, and temporal. The central sulcus separates the frontal and parietal lobe; (b) Geometric mapping between body parts, and motor/somatosensory cortex. Here, the motor cortex and the somatosensory cortex are shown at the left and right sides, respectively. Notice that in each hemisphere there is one motor area (frontal to the central sulcus) and one sensori area (posterior to the central sulcus). From (BLANKERTZ et al., 2008).

EEG, which has been used to predict the next voluntary motor task (JIANG et al., 2015; PFURTSCHELLER; SILVA, 1999). For this aim, previous researches have been conducted with promissing results to recognize the motor intention of upper and lower limbs, using brain rhythms and event-related potentials (ERPs) (HASHIMOTO; USHIBA, 2013; SBURLEA et al., 2015; PFURTSCHELLER; SILVA, 1999; PFURTSCHELLER et al., 2003; HORTAL et al., 2016). For example, it has been demonstrated that EEG of persons physically impaired may be used in robotic exoskeletons to improve the closed-loop control through motor planning recognition, which enhances the neuroplasticity during the motor rehabilitation (XU et al., 2014). ERPs are related to time-locked changes in the activity of neuronal populations (PFURTSCHELLER; SILVA, 1999), while the evoked activity, or useful signal, has a more or less fixed time-delay to the stimulus, turning the EEG signal blurred by additive noise. Then, averaging techniques may be applied on raw EEG to enhance the SNR, detecting ERPs. However, a simple processing technique cannot be adopted, as certain events can block or desynchronize the ongoing alpha rhythm (from 8 to 12 Hz) activity. In addition, ERPs may be affected by additional uncorrelated noise, such as visual stimuli, which can reduce the EEG amplitude (PFURTSCHELLER; SILVA, 1999).

ERD/ERS are phenomena that represent frequency-specific changes during the EEG activity, which decrease/increase the power in their correspondent frequency band (PFURTSCHELLER; SILVA, 1999), respectively. ERD/ERS reflect changes in the activity of local interactions between main neurons and interneurons that control the frequency components of the ongoing EEG. In general, the frequency of brain oscillations is negatively correlated to their amplitude (PFURTSCHELLER; SILVA, 1999), thus, components of high frequencies may be more affected by additive noise .

Figures 2.4a - 2.4c show the generation of ERDs around the finger tapping onset. ERD potentials may be found on the frequency band of alpha rhythms (8 to 12 Hz) during real and imaginary movement tasks, in which high amplitudes are obtained when the task complexity or attention increase (PFURTSCHELLER; SILVA, 1999; HASHIMOTO; USHIBA, 2013; COSTA et al., 2016). Alpha band desynchronization is not an unitary phenomenon, therefore, two distinct patterns related to alpha desynchronization may be observed. For example, lower

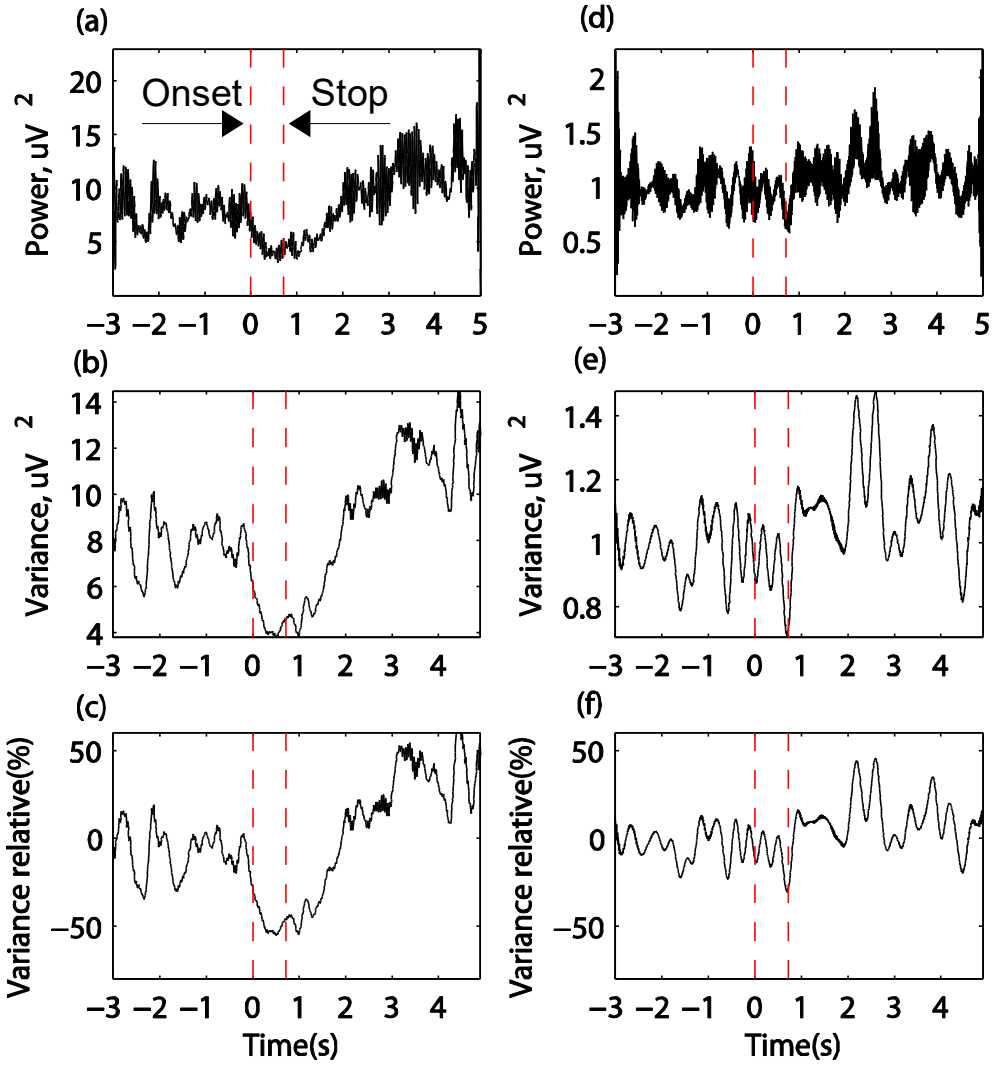

Figure 2.4: Representation of the ERD and ERS measured from the C3-FZ electrodes, using the Emotiv equipment. The vertical dotted lines represent the movement onset and offset. a) Power spectrum smoothed of the ERD; b) Variance inter-trial smoothed of the ERD; c) Variance relative of the ERD; d) Power spectrum smoothed of the ERS; e) Variance inter-trial smoothed of the ERS; f) Variance relative of the ERS.

.

alpha desynchronization may be obtained at the frequency band from 7 to 10 Hz, in response to almost any type of task. It may occur on wide areas of the scalp and probably reflects general task demands and attentional processes (PFURTSCHELLER; SILVA, 1999). In addition, upper alpha desynchronization or mu-rhythm is very often topographically restricted, which is generated at the frequency band from 10 to 12 Hz, during the processing of sensory-semantic information above parieto-occipital areas (PFURTSCHELLER; SILVA, 1999).

Voluntary movements result in a circumscribed desynchronization in the upper alpha and lower beta bands located close to sensorimotor areas (PFURTSCHELLER; SILVA, 1999). This desynchronization starts around 2 s prior to the movement onset over the contralateral Rolandic region and becomes bilaterally symmetrical immediately before the movement execution (GRAIMANN et al., 2002; PFURTSCHELLER; SILVA, 1999; PFURTSCHELLER et al., 2003; HASHIMOTO; USHIBA, 2013). It is referred as pre-movement or before the movement, whose time periods do not present muscle movement or unrelated if it occurs, but the subject is fully familiar with the action that he/she is going to perform in the near future. In this time interval (from 0.5 to 2 s prior to the movement onset), the cortex is adapted for the implementation of action (SHIBASAKI; HALLETT, 2006; LEW et al., 2012; NESHIGE; LÜDERS; SHIBASAKI, 1988).

Oscillations with good SNR are obtained on the human scalp, which are defined as post-movement beta ERS (PFURTSCHELLER; SILVA, 1999; HORTAL et al., 2016). These oscillations may be found around +1 s after completing a voluntary movement (PFURTSCHELLER; SILVA, 1999; PFURTSCHELLER et al., 2003).

In addition to oscillations in alpha and lower beta bands, induced ERS potentials have been found in the frequency band around 40 Hz (lower gamma band) and 60 Hz (upper gamma band), during visual stimulation and movement (PFURTSCHELLER; SILVA, 1999; HORTAL et al., 2016; PFURTSCHELLER et al., 2003; SEEBER et al., 2015). These potentials reach a peak around hundred millisecond before movement onset, and during the movement execution (PFURTSCHELLER; SILVA, 1999; PFURTSCHELLER et al., 2003).

The execution of a real or imaginary motor task in humans, measured over the primary motor cortex, is preceded by a slow potential that decreases the EEG amplitude (within at least 500 ms), which is called slow cortical potential (SCP) or motor related cortical potential (MRCP) (SHIBASAKI; HALLETT, 2006; LEW et al., 2012; JIANG et al., 2015; SBURLEA; MONTESANO; MINGUEZ, 2015). Researches suggest that medial frontocentral (MFC) and sensorimotor areas (SMA) are probable generators of MRCP (TOMA et al., 2002), which may be recorded on subcortical structures, such as basal ganglia and thalamus (SHIBASAKI; HALLETT, 2006). On MRCP, SCPs related to motor planning may be located around 1.5 to 0.5

s prior to begin a voluntary movement (LEW et al., 2012). These potentials, called Bereitschaftspotential (BeP), may be produced by both sensorimotor areas (NESHIGE; LÜDERS; SHIBASAKI, 1988), and with consistent potentials on both ipsilateral and contralateral supplementary motor areas (IKEDA et al., 1992).

In summary, ERD/ERS potentials and MRCPs may be measured over the motor cortex during a real or imaginary motor task. Therefore, these potentials are useful for motor rehabilitation of patients with severe motor disability, but still with ability to attempt and imagine a movement. However, these potentials can be blurred by interferences and artifacts that cause a poor SNR on EEG signals, which are described in the next section.

2.2.1 Factors influencing EEG signals

Motor intention recognition is a challenging task, as related patterns on EEG can be blurred by several sources of common interferences and artifacts. In our study, common interferences are considered those signals or undesirable physiological events broadcasting on the scalp or around the electrode of interest, with the same phase and similar amplitude (LAPLANTE, 2000), such as interferences caused by power line, eye movements, eye blinks, myoelectric and cardiac activities, non-μ-rhythm of EEG components as visual alpha rhythm, among others (MCFARLAND et al., 1997; KILICARSLAN; GROSSMAN; CONTRERAS-VIDAL, 2016). Physiological artifacts, such as blink and ocular movements (EOG), are present in most EEG recording, which, due to the volume conduction, corrupt all electrode measurements in changing profiles and

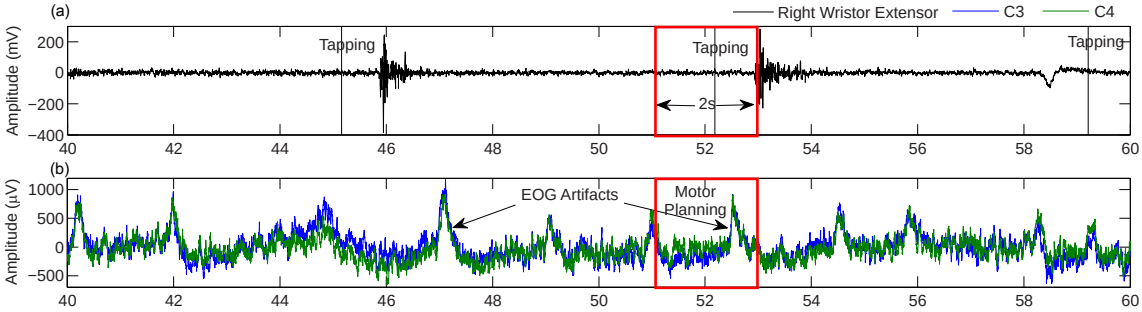

Figure 2.5: Representation of EOG artifacts throughout raw EEG, captured from the acquisition equipment BrainNet BNT during right finger tapping. a) myoelectric activity on the wrist extensor muscle; b) raw EEG on the brain hand region.

amplitude distributions (KILICARSLAN; GROSSMAN; CONTRERAS-VIDAL, 2016). These are maximal at low frequency ($\leq$ 0.5 Hz), and affect the EEG measured on the brain frontal region (MCFARLAND et al., 1997; KILICARSLAN; GROSSMAN; CONTRERAS-VIDAL, 2016). In particular, EOG artifacts may present high amplitudes, as shown in Figure 2.5, and a similar shape that an MRCP (JIANG et al., 2015). Additionally, the visual alpha rhythm can be extended to central scalp regions, but it is most predominant over the parieto-occipital cortex (PFURTSCHELLER; SILVA, 1999). These contaminations are also subject-specific (due to slower and faster blinks) and may also dynamically vary in amplitude across sessions, as well as in experimental conditions, changing environments or excessive electrode impedance differences (KILICARSLAN; GROSSMAN; CONTRERAS-VIDAL, 2016). Anyways, these unwanted effects contribute to reduce the SNR of EEG signals.

In addition, components of MRCP can be produced by various factors, such as preparatory state, level of intention, movement selection, pace of movement repetition, speed and precision of movement, praxis movement, perceived effort, force exerted, discreteness and complexity of movement, learning and skill acquisition, and pathological injuries of various brain structures (SHIBASAKI; HALLETT, 2006; SHAKEEL et al., 2015).

2.3 Brain-computer interfaces

Figure 2.6 shows the block diragram with main stages that can be adopted on BCIs, such as preprocessing, feature extraction, feature selection, and classification. The aim of BCIs is to take signals produced by the brain, and translate them into useful commands with no intervention on muscles (MCFARLAND et al., 2006). To achieve this goal, either regression (MCFARLAND; WOLPAW, 2005) or classification (JIANG et al., 2015; HORTAL et al., 2016) algorithms can be used. BCI-based on classification algorithms are the most popular approaches, where these algorithms are used to identify patterns of brain activity (LOTTE et al., 2007; HORTAL et al., 2016; BLANKERTZ et al., 2008). For instance, LDA, RDA, and SVM (with kernal linear and Guassian) classifiers have been widely used in BCIs to recognize gait intention.

Classification mistakes may increase by several factors that affect the EEG signal, as shown

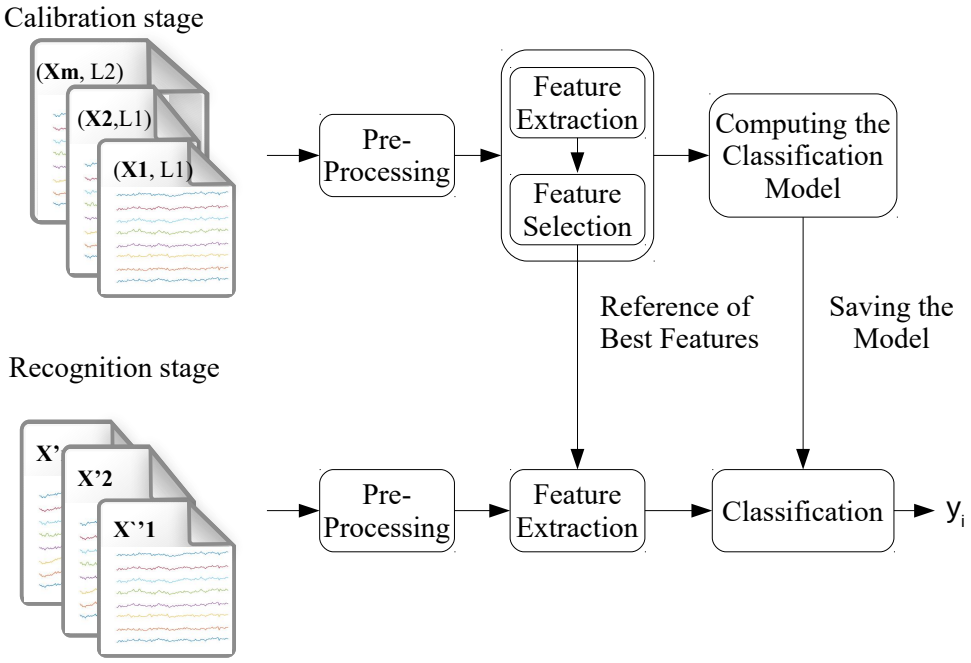

Figure 2.6: Block diagram of a brain-computer interface.

in Section 2.2.1. As a result, other methods for pre-processing (PFURTSCHELLER; NEUPER; BERGER, 1994; WOLPAW; MCFARLAND, 1994), feature extraction (MÜLLER-GERKING; PFURTSCHELLER; FLYVBJERG, 1999; ANG et al., 2012a; MCFARLAND et al., 2006; BARACHANT et al., 2013), and feature selection (DELISLE-RODRIGUEZ et al., 2017; LOTTE et al., 2007; GARRETT et al., 2003) have been proposed to improve the BCI performance.

This section describes some proposed methods, which were focused on improving the BCIs during gait intention recognition.

2.3.1 Pre-processing

Several methods, such as Principal Component Analysis (PCA) (JUNG et al., 1998a; MULLEN et al., 2013), Independent Component Analysis (ICA) (JIANG et al., 2015; VELU; SA, 2013), fastICA, and Adaptive Noise Cancelling (ANC) have been proposed to remove EOG artifacts,

showing promissing results. However, these methods may also remove the information of interest on raw EEG.

In particular, the combination of ICA and Laplacian filters in BCIs, to recognize the gait intention have been more explored (JIANG et al., 2015; HORTAL et al., 2016; VELU; SA, 2013). Spatial filters, such as LAR, and WAR have been widely used in BCIs to reduce common interferences on EEG signals, computing the average voltage from a set of surrounding electrodes with respect to a central electrode for LAR, taking into account the distance between electrodes (PFURTSCHELLER; NEUPER; BERGER, 1994; HORTAL et al., 2016). The common interferences may be reduced from raw EEG, through the following equations:

$$V_i^{LAR} = V_i^{CR} - \sum_{j \in S_i} g_{ij} V_j^{CR}, \; g_{ij} = \frac{1}{d_{ij}} \sum_{j \in S_i} \frac{1}{d_{ij}} \tag{2.1}$$

$$V_i^{WAR} = V_i^{CR} - \sum_{j=1}^{N} g_{ij} V_j^{CR}, \; g_{ij} = \frac{1}{d_{ij}} \sum_{j=1;\; j \neq i}^{N} \frac{1}{d_{ij}}, \tag{2.2}$$

where V_i^{CR} is the potential between the electrode of interest i and the reference electrode, S_i is the set of surrounding electrodes or neighbors, d_{ij} is the distance between the electrode of interest i and the neighbor electrode j, N is the number of electrodes, and g_{ij} is the weight index.

2.3.2 Feature extraction

A great variety of features from EEG have been used to design BCIs, such as amplitude values of signals, band powers (BP), power spectral density (PSD), autoregressive (AR) model, among others (HORTAL et al., 2016; MCFARLAND et al., 2006). During the BCI design, it is important to understand the feature properties, and how they are used, in order to select the most appropriate classifier.

Some critical properties of BCI features can be considered (LOTTE et al., 2007). For example, BCI features can be noisy or contain outliers, as EEG signals have a poor SNR. Additionally, feature vectors are often of high dimensionality, as they may be generally extracted from sev-

eral channels and with different time segments before being concatenated into a single feature vector. In addition, BCI features must contain time information as brain activity patterns, which are generally related to specific time variations of EEG. Moreover, EEG signals are non-stationary, since they may rapidly vary over time, and more especially over sessions (LOTTE et al., 2007). Furthermore, the training sets are normally small, since the training process is time consuming, and demanding for the subjects (LOTTE et al., 2007).

Thus, most brain activity patterns used to drive BCIs are related to particular time variations of EEG, possibly in specific frequency bands (PFURTSCHELLER; SILVA, 1999; HORTAL et al., 2016; JIANG et al., 2015; PFURTSCHELLER et al., 2003). Therefore, the time course of EEG signals should be taken into account during feature extraction. For this reason, concatenation of features from different time segments have been widely used, which consists of extracting features from several time segments, and concatenating them into a single feature vector.

Features in time, frequency and spatial domains, used in this work to recognize gait planning/stopping, are presented in next subsections.

Time domain

Mean absolute value

Mean absolute value (MAV) is widely used on sEMG and EEG signals analysis (PHINYOMARK; PHUKPATTARANONT; LIMSAKUL, 2012). MAV feature is an average of absolute values, computed from EEG amplitudes x in a segment, which can be defined as:

$$MAV = \frac{1}{N}\sum_{i=1}^{N}|x_i| \ . \tag{2.3}$$

Wavelength

Wavelength (WL) may be used to measure the EEG complexity (PHINYOMARK; PHUKPATTARANONT; LIMSAKUL, 2012). It is defined as cumulative length of the EEG waveform over

the time segment, and can be calculated by:

$$WL = \sum_{i=1}^{N} |x_{i+1} - x_i|. \tag{2.4}$$

Fractal dimension from Higuchi

Fractal dimension (FD) is a measurement of data complexity, generally evaluated in phase space by mean of correlation dimension. Fractal dimension from Higuchi (FDH) was proposed to estimate the FD (HIGUCHI, 1988), directly in time domain without reconstructing the strange attractor (PHOTHISONOTHAI; WATANABE, 2013). This feature may be used on short time segments to improve the estimation of the FD. Considering that the FDH algorithm is based on a given finite time series $y = \{y(1), y(2), ..., y(N)\}$, a new time series y_m^k can be constructed using FDH by the following equation:

$$y_m^k = \left\{ y(m), y(m+k), y(m+2k), ..., y\left(m + \frac{N-m}{k}\right) \cdot k \right\} \ , \ m = 1, 2, 3, ..., k \tag{2.5}$$

where m is an integer that indicates the initial time, k is an integers that means the time interval.

Then, the length $L_m(k)$ can be computed for each curve using the following equation:

$$L_m(k) = \frac{1}{k} \left\{ \left(\sum_{i=1}^{\left[\frac{N-m}{k}\right]} |y(m+ik) - y(m+(i-1)k)| \right) \cdot \frac{N-1}{\left[\frac{N-m}{k}\right] \cdot k} \right\} \ , \ i = 0, 1, 2, .., N \tag{2.6}$$

The length of a curve $L_m(k)$, for the time interval k, can computed as the average of m curves. Then, FDH, defined as D_{FDH}, can be calculated following the relationship $L_m(k) \propto k^{-D_{FDH}}$ (HIGUCHI, 1988), (PHOTHISONOTHAI; WATANABE, 2013).

Fractal dimension from Sevcik

In Sevcik's method, the data are first normalized to be within a unit square by rescaling the abscissa (time axis) and the ordinate (EEG signal) of the data space:

$$i^n = \frac{i}{N},\ s^n(i^n) = \frac{s(i) - s_{\min}}{s_{\max} - s_{\min}}, \tag{2.7}$$

where $s(i)$ and $s^n(i^n)$ are the original and normalized EEG signals of the ith data point; $s_{\max}$ and $s_{\min}$ are the maximal and minimal values of the signal; $i = 1, 2, 3, ..., N$ is the serial number of the data points and i^n is the normalized one.
Then, FD can be calculated by:

$$FD = 1 + \frac{\log(L)}{\log(2(N-1))}, \tag{2.8}$$

where L is the total length of the data segment in the normalized coordinate system.

Frequency domain

Band power

The power of specific frequency bands is widely used in on-line applications (HORTAL et al., 2016) to measure oscillatory activities on EEG, especially those associated with motor activity. Band power (BP), defined in Equation (2.9), can be extracted by both bandpass filters (in time domain), and power spectral density (PSD) from the fast Fourier transform (FFT). In relation to PSD, BP can be computed using absolute values in a specific frequency band, such as mu-ERD and/or beta-ERD (HASHIMOTO; USHIBA, 2013; HORTAL et al., 2016; SBURLEA; MONTESANO; MINGUEZ, 2017).

$$BP = \sum_{f=1}^{N} \left| y_f^2 \right| \ , \tag{2.9}$$

where y is the power value at the frequency f_i, and N is the size of the analysis window. Although the phase information can be obtained from FFT (SBURLEA; MONTESANO; MINGUEZ, 2017), it was not considered here, in order to reduce computational cost, thus avoiding large dimensional feature vectors.

Spatial domain

In addition to the traditional spatial filters LAR and WAR (PFURTSCHELLER; NEUPER; BERGER, 1994; WOLPAW; MCFARLAND, 1994), described in Section 2.3.1, other methods, such as Common Spatial Pattern (CSP) (MÜLLER-GERKING; PFURTSCHELLER; FLYVBJERG, 1999; BLANKERTZ et al., 2008), Filter-Bank CSP (ANG et al., 2012a), and Riemannian covariance matrices (RK) (BARACHANT et al., 2013) can be used as features in spatial domain. These last aforementioned methods have been combined on BCIs, with bandpass filters in a frequency range from 8 to 30 Hz, in order to recognize mental and imagery motor tasks (BLANKERTZ et al., 2008; BARACHANT et al., 2013). A brief description of these methods are given in next subsections.

Common Spatial Pattern

This section describes a mathematical concept that computes relevant features, called Common Spatial Pattern (CSP), which captures the class-invariant characteristic from some trials using multi-channels analysis (MÜLLER-GERKING; PFURTSCHELLER; FLYVBJERG, 1999; BLANKERTZ et al., 2008), during the calibration stage. Figure 2.7 shows a block diagram of this powerful tool.

Let $\mathbf{X} \in \mathbb{R}^{Ch \times T}$ be a short time segment of EEG signal in a calibration or training set, which corresponds to a trial of real or imagery movement, Ch is the number of EEG channels, and T is the number of samples in this trial. In addition, let two classes $\mathbf{A}$ and $\mathbf{B}$ that can be described by $\mathbf{X}_i$. The normalized covariance matrices for $\mathbf{A}$, and its corresponds mean value

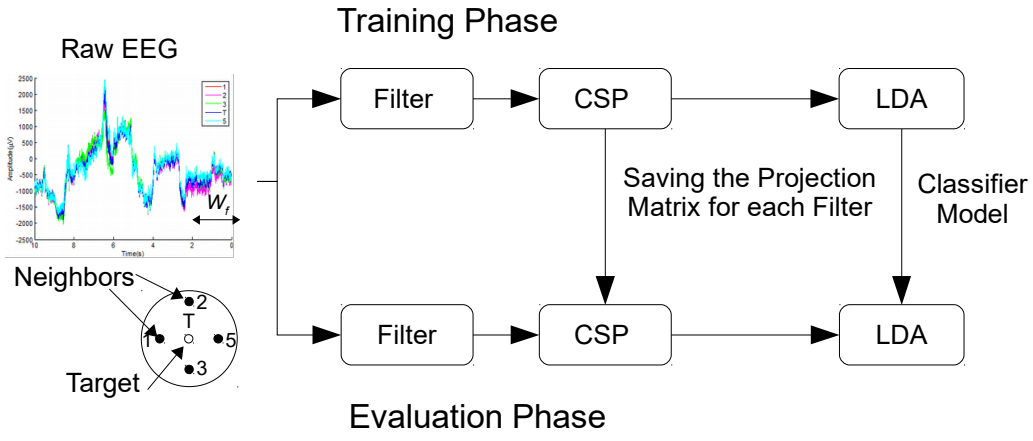

Figure 2.7: Block diagram of the Common Spatial Pattern algorithm (CSP).

can be written as:

$$\mathbf{R}_i^{\mathbf{A}} = \frac{\mathbf{X}_i^{\mathbf{A}} \mathbf{X}_i^{\mathbf{A}\dagger}}{\mathrm{tr}\left(\mathbf{X}_i^{\mathbf{A}} \mathbf{X}_i^{\mathbf{A}\dagger}\right)} , \tag{2.10}$$

$$\mathbf{R}^{\mathbf{A}} = \frac{1}{M} \sum_{i=1}^{M} \mathbf{R}_i^{\mathbf{A}} , \tag{2.11}$$

Similarly, the normalized covariance matrices $\mathbf{R}_i^{\mathbf{B}}$ of $\mathbf{B}$, and its corresponds mean value $\mathbf{R}^{\mathbf{B}}$, can be computed in the same way followed for $\mathbf{B}$. Here, $\dagger$ denotes the transpose operator.

Once computed $\mathbf{R}^{\mathbf{A}}$ and $\mathbf{R}^{\mathbf{B}}$, their additive value ($\mathbf{R}^{\mathbf{C}} = \mathbf{R}^{\mathbf{A}} + \mathbf{R}^{\mathbf{B}}$) can be factored into its eigenvectors by the following expression:

$$\mathbf{R}^{\mathbf{C}} = \mathbf{B}_{\mathbf{C}} \lambda \mathbf{B}_{\mathbf{C}}^{\dagger} , \tag{2.12}$$

where $\mathbf{B}_{\mathbf{C}}$ is an $Ch \times Ch$ matrix of normalized eigenvectors, $\mathbf{B}_{\mathbf{C}} \times \mathbf{B}_{\mathbf{C}}^{\dagger} = 1_{Ch \times Ch}$, and λ is the corresponding diagonal eigenvalues with dimension $Ch \times Ch$.

Then, whitening transformation $\mathbf{W} = \lambda^{-1/2} \mathbf{B}_{\mathbf{C}}^{\dagger}$ can be used to transform $\mathbf{R}^{\mathbf{A}}$ and $\mathbf{R}^{\mathbf{B}}$, using Equation 2.13. $\mathbf{W}$ equalizes the variances in the space spanned by the eigenvectors in $\mathbf{B}_{\mathbf{C}}$. Then, $\mathbf{S}_{\mathbf{A}}$ and $\mathbf{S}_{\mathbf{B}}$ share the same eigenvectors, since $\mathbf{S}_{\mathbf{A}} + \mathbf{S}_{\mathbf{B}} = \mathbf{W} R^{\mathbf{C}} \mathbf{W}^{\dagger} = 1_{Ch \times Ch}$. Therefore, the eigen decomposition of $\mathbf{S}_{\mathbf{A}}$ and $\mathbf{S}_{\mathbf{B}}$ can be read as $\mathbf{U} \psi^{\mathbf{A}} \mathbf{U}^{\dagger}$ and $\mathbf{U} \psi^{\mathbf{B}} \mathbf{U}^{\dagger}$, respectively, being $\mathbf{U}$ orthonormal, and $\psi^{\mathbf{A}} + \psi^{\mathbf{b}} = \mathbf{I}$.

$$S_{\mathbf{A}} = \mathbf{W} R^{\mathbf{A}} \mathbf{W}^{\dagger} \; and \; S_{\mathbf{B}} = \mathbf{W} R^{\mathbf{B}} \mathbf{W}^{\dagger}, \tag{2.13}$$

In consequence, the projection of whitened EEG epochs on $\mathbf{U}$ may provide feature vectors, which are optimal in the least squares to discriminate both classes $\mathbf{A}$ and $\mathbf{B}$. Then, the decomposition of each trial can be computed by Equation (2.15) Then, raw EEG can be represented as a linear combination of time-invariant modes by the following expression:

$$\mathbf{P}^{\dagger} = \mathbf{U}^{\dagger}\mathbf{W} \tag{2.14}$$

$$\mathbf{Z}_i = \mathbf{P}^{\dagger}\mathbf{X}_i \ , \tag{2.15}$$

where $\mathbf{X}_i$ is the raw EEG, $\mathbf{P}^{\dagger}$ is the projection matrix with dimension $Ch \times Ch$, and $\mathbf{Z}_i$ are expansion coefficients with the same dimension of $\mathbf{X}_i$.

The method of CSPs looks a way to obtain at the same time maximal variance on the first rows of $\mathbf{Z}_i$ for trials of the class $\mathbf{A}$, and minimal variance on the last rows of $\mathbf{Z}_i$ for trials of class $\mathbf{B}$. This property of decomposition of CSPs ensures that the variances of the first and last rows of $\mathbf{Z}_i$ contain the most relevant information to discriminate both classes $\mathbf{A}$ and $\mathbf{B}$. Notice that the variance of a row in $\mathbf{Z}_i$ can be high only if the spatial distribution of the raw EEG amplitudes is similar to the corresponding patterns.

By now, steps have been presented to obtain the projection matrix $\mathbf{P}^{\dagger}$, but it is necessary to find the first and last m rows of $\mathbf{P}^{\dagger}$ that allow obtaining the expansion coefficients $\mathbf{Z}_i$, which discriminates both classes $\mathbf{A}$ and $\mathbf{B}$ with good performance. Then, the full training set is analyzed in a cross-validation ($k = 10$), to be divided in two new training and testing sets. The training set, obtained for each k-fold, is used to compute the projection matrix $\mathbf{P}^{\dagger}$. Once computed the projection matrix $\mathbf{P}^{\dagger}$, and using a Linear/Quadratic Discriminant Classifier (LDA/QDA), it is possible to find the first and last m rows that provide the best expansion coefficients $\mathbf{Z}_i$, improving the discrimination of both classes $\mathbf{A}$ and $\mathbf{B}$. The features ons $\mathbf{Z}_i$ must be normalized through Equation (2.16).

$$f_p^i = log\left(\frac{var_p^i}{\sum_{p=1}^{2m} var_p^i}\right) \ , \tag{2.16}$$

where var_p^i is the variance of the p-th row of $\mathbf{Z}_i$, p is running from 1 to m, and from $N-m+1$ to N, and f_p^i is the normalized feature.

Filter Bank-Common Spatial Pattern

Filter Bank-Common Spatial Pattern (FBCSP) has been used in BCIs to recognize imagery motor tasks (ANG et al., 2012a), which is an extension of CSP, but including filter banks as shown in Figure 2.8. For example, some authors have used 9 bandpass filters in a frequency range of 4 to 40 Hz, such as 4-8 Hz,8-12 Hz,...,36-40 Hz. Then, after applying bandpass filters, CSP must be individually applied for each filtered output to detect ERD/ERS. Details of CSP

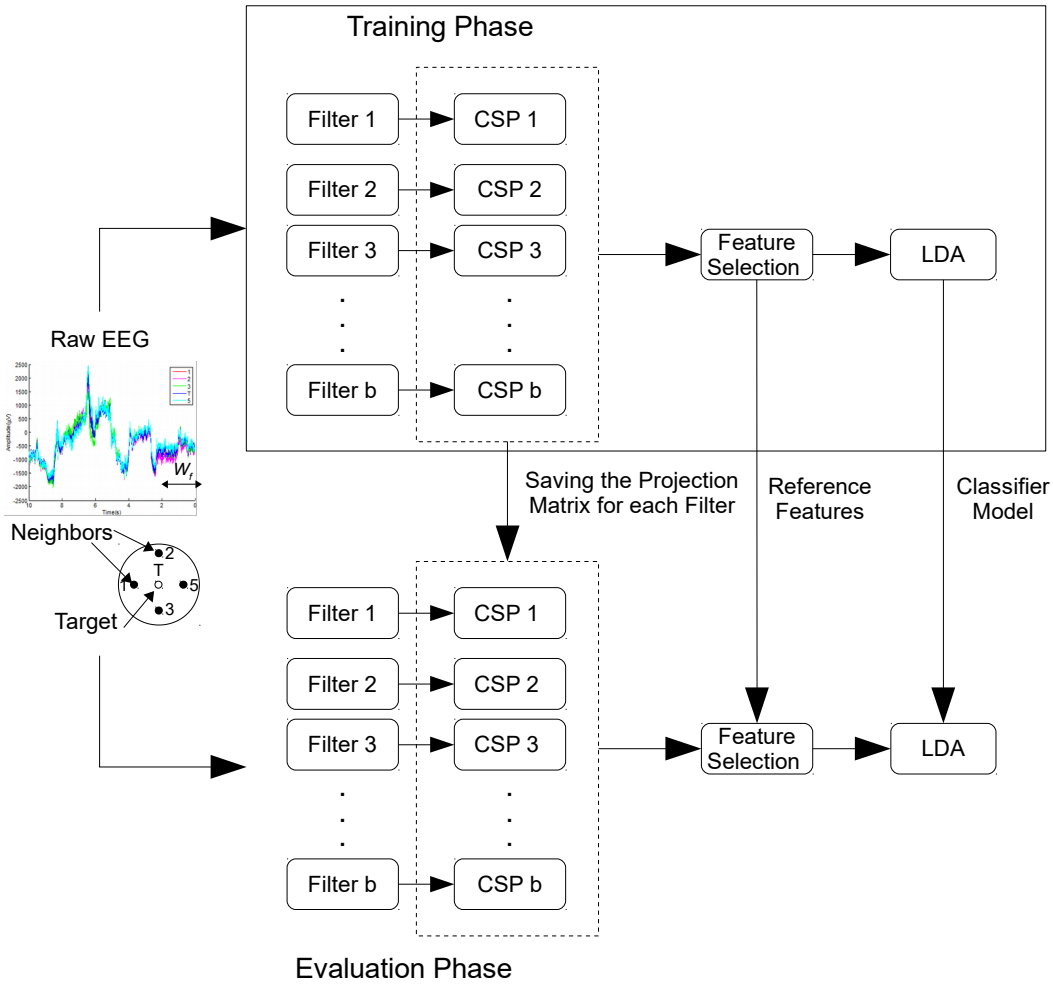

Figure 2.8: Block diagram of the Filter-Bank Common Spatial Pattern algorithm.

are given in previous subsection.

$$\mathbf{Z}_i = \mathbf{W}_b^{\dagger}\mathbf{E}_{b,i} \ , \tag{2.17}$$

where $\mathbf{E}_{b,i} \in \mathbb{R}^{Ch\times T}$ denotes the single trial of EEG measurement from the bth bandpass filter of the ith trial, $\mathbf{Z}_i \in \mathbb{R}^{Ch\times T}$ denotes $\mathbf{E}_{b,i}$ after spatial filtering, $\mathbf{W}_b \in \mathbb{R}^{Ch\times Ch}$ is the CSP projection matrix, Ch is the number of EEG channels, T is the number of EEG samples per channel, and † denotes the transpose operator.

CSP computes the transformation matrix $\mathbf{W}_b$ to provide features whose variances are optimal to discriminate two classes, solving the eigenvalue decomposition as:

$$\Sigma_{b,1}\mathbf{W}_b = (\Sigma_{b,1} + \sigma_{b,2})\,\mathbf{W}_b\mathbf{D}_b \ , \tag{2.18}$$

where $\Sigma_{b,1}$ and $\Sigma_{b,2}$ are covariance matrices of both classes 1 and 2, computed from bth bandpass filtered EEG, $\mathbf{D}_b$ is the diagonal matrix that contains the eigenvalues of $\Sigma_{b,1}$. Technically, $\mathbf{W}_b$ can be computed in Matlab using the command $W = eig(S1, S1 + S2)$, where $W,\ S1,\ and\ S2$ represent here $\mathbf{W}_b$, $\Sigma_{b,1}$ and $\Sigma_{b,2}$, respectively.

Then, the spatial filtered signal $\mathbf{Z}_i$ can maximize the difference between variances of both classes 1 and 2 through the projection matrix $\mathbf{W}_b$ given by Equation (2.18), as shown in Equation (2.17).

$$\mathbf{v}_{b,i} = \log\left[\frac{\overline{\mathbf{W}}_b^{\dagger}\mathbf{E}_{b,i}\mathbf{E}_{b,i}^{\dagger}\overline{\mathbf{W}}_b}{\mathrm{tr}\left(\overline{\mathbf{W}}_b^{\dagger}\mathbf{E}_{b,i}\mathbf{E}_{b,i}^{\dagger}\overline{\mathbf{W}}_b\right)}\right] \ , \tag{2.19}$$

where $\mathbf{v}_{b,i} \in \mathbb{R}^{2m}$; $\overline{\mathbf{W}}_b$ represents the first and the last m rows of $\mathbf{W}_b^{\dagger}$; $\mathrm{diag}(\cdot)$ returns the diagonal elements of the square matrix; and $\mathrm{tr}(\cdot)$ returns the sum of the diagonal elements in the square matrix.

The FBCSP feature vector of 9 bandpass filters, obtained from ith trial, can be formed as $\mathbf{v}_i = [\mathbf{v}_{1,i}, \mathbf{v}_{2,i}, \ldots, \mathbf{v}_{9,i}]$, where $\mathbf{v}_i \in \mathbb{R}^{9*2m}$, $i = 1, 2, \ldots, n$, n denotes the total number of trials. Then, the training data that comprised the extracted features, and its correspondent true class

labels, can be respectively denoted, as:

$$\overline{\mathbf{V}} = \begin{bmatrix} \overline{\mathbf{V}}_{\mathbf{1}} \\ \overline{\mathbf{V}}_{\mathbf{2}} \\ \vdots \\ \overline{\mathbf{V}}_{\mathbf{n}} \end{bmatrix}, \tag{2.20}$$

$$\overline{\mathbf{y}} = \begin{bmatrix} \overline{\mathbf{y}}_{\mathbf{1}} \\ \overline{\mathbf{y}}_{\mathbf{2}} \\ \vdots \\ \overline{\mathbf{y}}_{\mathbf{n}} \end{bmatrix}, \tag{2.21}$$

where $\overline{\mathbf{V}} \in \mathbb{R}^{n \times 9*2m}$; $\overline{\mathbf{y}} \in \mathbb{R}^{n \times 1}$; $\overline{\mathbf{v}}_i$ and $\overline{\mathbf{y}}_i$ denote the feature vector and true class label from the ith training trial, $i = 1, 2, \ldots, n$; and n denotes the total number of trials in the training data.

Finally, a feature selection algorithm on $\overline{\mathbf{V}}$ is applied to obtain CSP features that increase the discrimination between both classes 1 and 2. For this purpose, a comparison between two methods for feature selection, such as Mutual Information-based Individual Feature (MIBIF) and Mutual Information-based Rough Set Reduction has been carried out (ANG et al., 2012a), obtaining for MIBIF the best performance. Therefore, MIBIF is commonly used to select the best CSP features. Details on MIBIF are given in Appendix A.2.1.

Similar to CSP, a cross-validation stage with the training set must be carried out.

Theory of Riemannian geometry

A general theory of Riemmanian geometry is presented in this section, but more details can be consulted in (BARACHANT et al., 2013, 2012).

Let $\mathbf{X} \in \mathbb{R}^{Ch \times T}$ be a trial, with Ch being the number of EEG channels, and T the epoch duration, which is expressed in number of samples. It is assumed that the raw EEG has been bandpass filtered, in order to take into account frequency bands, such as mu-ERD (8-12 Hz) and beta-ERD (18-24 Hz), which are related to the motor planning action (PFURTSCHELLER;

SILVA, 1999; SBURLEA; MONTESANO; MINGUEZ, 2015). For each trial $\mathbf{X}_p$ of known class $y_p \in \{-1, 1\}$ may be used to estimate the spatial covariance matrix of the EEG random signal by the $Ch \times Ch$ sample covariance matrix (SCM). SCM can be computed by the following expression:

$$\mathbf{C}_p = \frac{1}{T-1}\mathbf{X}_p\mathbf{X}_p^{\dagger} \, . \tag{2.22}$$

The space of symmetric positive definite (SPD) matrices with dimension $Ch \times Ch$ will be denoted $P(Ch)$ in the rest of this study. This space of SPD forms a differentiable manifold $\mathcal{M}$ of dimension $Ch = Ch(Ch+1)$. Thus, the suitable manipulation of these matrices relies on a special branch of differential geometry, named as Riemannian geometry (BARACHANT et al., 2013). This way, at each point $\mathbf{C}$ of the mainfold $\mathcal{M}$, a scalar product can be defined in the associated tangent space $\mathcal{T}_{\mathbf{C}}\mathcal{M}$ (BARACHANT et al., 2013).

Let $\mathbf{S}_1$ and $\mathbf{S}_2$ be two tangent vectors, where its scalar product on the tangent space at $\mathbf{C}$ may be defined through Equation (2.23). Furthermore, the logarithm map in Equation (2.24) projects locally all covariance matrices $\{\mathbf{C}_p\}_{p=1}^{P}$ on the tangent plane.

$$\langle \mathbf{S_1}, \mathbf{S_2} \rangle_{\mathbf{C}} = \mathrm{tr}\left(\mathbf{S}_1\mathbf{C}^{-1}\mathbf{S}_2\mathbf{C}^{-1}\right) \tag{2.23}$$

$$\begin{aligned} \mathbf{S_p} &= \log(\mathbf{C}_p) \\ &= \mathbf{C}^{1/2}\,\mathrm{logm}\left(\mathbf{C}^{-1/2}\mathbf{C}_p\mathbf{C}^{-1/2}\right)\mathbf{C}^{1/2} \, , \end{aligned} \tag{2.24}$$

where $\mathrm{logm}(\cdot)$ denotes the logarithm of a matrix. The logarithm of a diagonalizable matrix $\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}$ is defined as $\mathrm{logm}(\mathbf{A}) = \mathbf{V}\mathbf{D}^{\dagger}\mathbf{V}^{-1}$, where the diagonal elements are represented as $d_{i,i}^{T} = \log(d_{i,i})$ (BARACHANT et al., 2013, 2012). In addition, the exponential map allows projecting an element of the tangent space back to mainfold $\mathcal{M}$, through an inverse operation as shown in Equation (2.25).

$$\begin{aligned} \mathbf{C_p} &= \mathrm{Exp}_{\mathbf{C}}(\mathbf{S}_p)) \\ &= \mathbf{C}^{1/2}\,\mathrm{expm}\left(\mathbf{C}^{-1/2}\mathbf{S}_p\mathbf{C}^{-1/2}\right)\mathbf{C}^{1/2} \, , \end{aligned} \tag{2.25}$$

where $\mathrm{expm}(\cdot)$ denotes the exponential of a matrix.

On other hand, given an scalar product that is defined in this tangent plane at $\mathbf{C}_{ref}$, as shown

in Equation (2.24), the Riemannian-based kernel is given by the following expression:

$$\begin{aligned} k_R(\text{vect}(\mathbf{C}_i), \text{vect}(\mathbf{C}_j); \mathbf{C}_{ref}) &= \langle \phi\mathbf{C}_i, \phi\mathbf{C}_j \rangle_{\mathbf{C}_{ref}} \\ &= \langle \phi\mathbf{S}_i, \phi\mathbf{S}_j \rangle_{\mathbf{F}} \\ &= \text{vect}(\mathbf{S}_i)^{\dagger} \, \text{vect}(\mathbf{S}_j) \end{aligned} \tag{2.26}$$

$$\phi\mathbf{C} = \log_{\mathbf{C}_{ref}}(\mathbf{C}) \tag{2.27}$$

$$\begin{aligned} \widetilde{\mathbf{S}}_i &= \mathbf{C}_{ref}^{-1/2} \log_{\mathbf{C}_{ref}}(\mathbf{C}_i) \mathbf{C}_{ref}^{-1/2} \\ &= \text{logm}\left(\mathbf{C}_{ref}^{-1/2} \mathbf{C}_i \mathbf{C}_{ref}^{-1/2} \right) \ , \end{aligned} \tag{2.28}$$

where $\widetilde{\mathbf{S}}_i$ and $\widetilde{\mathbf{S}}_j$ are symmetric matrices.

$\mathbf{C}_{ref}$ is a free parameter that defines the point in the mainfold $\mathcal{M}$ where the tangent plane is computed (BARACHANT et al., 2013). Here, $\mathbf{C}_{ref}$ is the average of the whole set of labeled covariance matrices, which can be computed by the following equation:

$$\mathfrak{A}(\mathbf{C}_1, ..., \mathbf{C}_p) = \frac{1}{P} \sum_{p=1}^{P} \mathbf{C}_p \ , \tag{2.29}$$

where p is the number of labels. The Riemannian distance between SPD matrices can be computed by the Equation (2.31) through the real eigenvalues $\{\lambda_i\}_{i=1}^{E}$ of $\mathbf{C}_1^{-1}\mathbf{C}_2$.

$$\mathfrak{G}(\mathbf{C}_1, ..., \mathbf{C}_p) = \text{argmin}_{\mathbf{C}} \sum_{p=1}^{P} \delta_R^2(\mathbf{C}, \mathbf{C}_p) \tag{2.30}$$

$$\delta_R^2(\mathbf{C}_1, \mathbf{C}_2) = \left|\left| \log(\mathbf{C}_1^{-1}, \mathbf{C}_2) \right|\right|_F = \left[\sum_{i=1}^{E} \log^2 \lambda_i \right]^{1/2} \tag{2.31}$$

2.3.3 Feature Selection

Feature selection and dimensionality reduction may be performed on a feature set to eliminate any redundant features, reduce computational complexity, prevent over-fitting during classifier training, and increase classification performance (LEW et al., 2012).

High-resolution analysis of spatial, temporal, and spectral aspects of the data may conduct to a

very high-dimensional feature space (GARRETT et al., 2003). Leveraging a higher percentage of potential features in the measured data requires powerful methods for signal processing, and classification capabilities. Then, the selection of a features subset that provides good accuracy on new data is most useful for classification problem. For example, the mutual information, such as MIBIF has been used to select optimal-temporal patterns, after applying FBCSP (ANG et al., 2012a, 2012b). Moreover, Canonical Variant Analysis (commonly known as Multivariate Discriminant Analysis) have used in BCIs based on MRCPs to select those channels more correlated, comparing movement preparation period to the non-movement related period (LEW et al., 2012). In addition, some authors have used Wavelets that provid information of high-resolution frequency, to obtain feature set of high-dimension (for example n channels $\times$ 358 coefficients per trial) (VELU; SA, 2013). As a result, PCA has been applied by these authors to reduce all feature set at 10 principal components (VELU; SA, 2013).

Unsupervised techniques for linear dimensionality reduction, such as PCA, may be affected by the low SNR of EEG. For instance, PCA reduces dimensionality by maximizing data variance in the projected subspace using a linear transformation. The transformation, dictated by the eigenvectors that correspond to the largest eigenvalues of the data covariance matrix, may discard the useful information for classification, which is contained in lower energy dimensions of raw EEG.

In BCIs, genetic algorithms (GAs) based on classifiers, such LDA, local Fisher's discriminant analysis, and SVM have been successfully to reduce a feature set (GARRETT et al., 2003; BULEA et al., 2014). These methods are evolutionary algorithms inspired from biological evolution to solve a defined task, with capacity of finding hidden relations in data. Figure 2.9 shows a block diagram of genetic algorithms based on classifiers. Several BCIs that use methods, such as GA for feature selection, search the best feature set while its performance is increasing continuously towards an unknown optimum value (GARRETT et al., 2003). However, this strategy could be unsuitable due to the EEG variability and its low SNR. Thus, a methodology to compute a reference value linked to the evaluation index (for example the accuracy) may be introduced in BCIs to obtain an optimal feature set.

Analysis of feature set with high dimension is commonly used in BCIs. Thus, indices of low

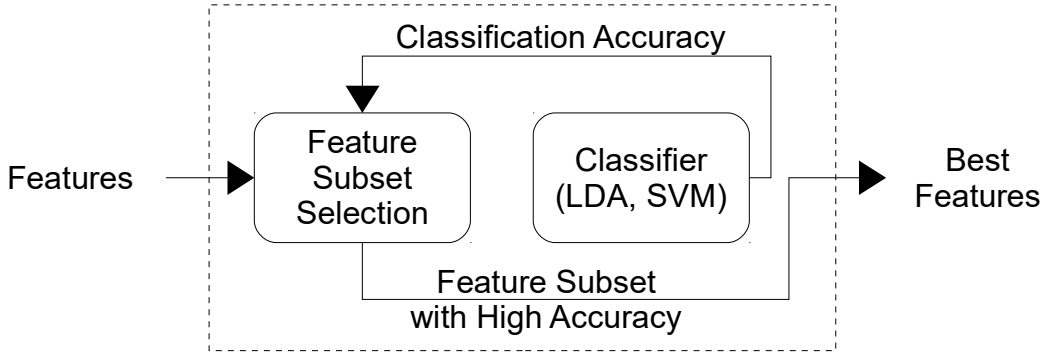

Figure 2.9: Block diagram of genetic algorithms.

computational cost, such as the Representation Entropy (RE) (MITRA; MURTHY; PAL, 2002), may be more suitable than classifiers, in order to obtain a smallest time during this process. Additionally, patterns related to motor anticipation occur before the movement onset, therefore, the labeling process of these patterns by specialists has a high uncertainty. Then, unsupervised methods, such as RK (BARACHANT et al., 2013) for feature selection, as well as metrics as RE (MITRA; MURTHY; PAL, 2002), in which their performance only depend on cluster distribution, could be appropriated in BCIs to recognize gait planning. Details of RE and RK may be consulted in Appendices A.4 and A.3, respectively.

2.3.4 Classification

In order to choose the most appropriate classifier for a given set of features, the properties of the available classifiers must be known. For instance, several definitions are commonly used to describe different kinds of available classifiers, such as:

Generative-discriminative: Generative (also known as informative) classifiers, such as Bayes quadratic, learn the class models. To classify a feature vector, generative classifiers compute the likelihood of each class and choose the most likely. Discriminative ones, as Support Vector Machines (SVM) (HORTAL et al., 2016; GARRETT et al., 2003), only learn the way to discriminate classes in order to classify a feature vector directly (LOTTE et al., 2007).

Static-dynamic: Static classifiers, for example, MultiLayer Perceptrons, cannot take into ac-

count temporal information during classification as they classify a single feature vector. In contrast, dynamic classifiers, such as Hidden Markov Model, can classify a sequence of feature vectors, and consequently, catch temporal dynamics (LOTTE et al., 2007).

Stable-unstable: Stable classifiers, such as Linear Discriminant Analysis (LDA), have low complexity. They are said stable as small variations in the training set does not affect considerably their performance (FRIEDMAN, 1989; GARRETT et al., 2003). However, unstable classifiers, such as MultiLayer Perceptron have high complexity, and small variations in the training set may lead to important changes in performances (LOTTE et al., 2007).

Regularized: Regularization consists of carefully controlling the complexity of a classifier in order to prevent overtraining. A regularized classifier has good generalization performances and is more robust with respect to outliers (LOTTE et al., 2007; FRIEDMAN, 1989; GARRETT et al., 2003).

On the other hand, while performing a pattern recognition task, classifiers may be facing several issues related to feature properties, such as outliers, overtraining, among others. In addition, two main problems need to be underlined for BCI applications: feature vector dimensionality and the bias-Variance tradeoff.

In relation to the feature vector dimensionality, several classes may be properly separated, increasing exponentially the training data and the size of feature vectors (LOTTE et al., 2007). Then, the classification accuracy for smallest training data with high feature dimension may be very low. Therefore, it is recommended to use a proportion from 5 to 10 times between the training set and the feature vector dimension (LOTTE et al., 2007). Unfortunately, this rule cannot be applied in BCIs for several conditions, such as rehabilitation of patients with severe motor disabilities.

Moreover, in relation to the bias-variance tradeoff, classification consists of finding the true label y of a feature vector $\mathbf{X}$ using a mapping $\mathbf{f}$. This mapping is obtained from a training set, while the best mapping $\mathbf{f}$ that can be generated from labels is unknown.

Classification error can be influenced by noise, bias, and variance. Error related to the noise within the system is an irreducible error. Bias represents the divergence between the estimated mapping, and the best mapping. Therefore, the classifier $f(linear, quadratic, \ldots)$ selection

plays an important role to reduce the bias. Finally, variance reflects the sensitivity to the training set used.

Notice that to achieve the lowest classification error, both bias and variance must be low. However, there is a natural bias-variance tradeoff. Then, stable classifiers have high bias and low variance, and vice-versa, for unstable classifiers. This can explain why simple classifiers sometimes outperform more complex ones.

Several techniques, known as stabilization techniques, can be used to reduce the variance. Among them, it can be cited combination of classifiers (LIN et al., 2014) and regularization (GARRETT et al., 2003; FRIEDMAN, 1989). EEG signals are known to be non-stationary, thus, training sets coming from different sessions are likely to be relatively different. Then, a low variance can be a solution to cope with the variability problem in BCIs.

Regularized classifiers such as RDA and SVM are described in the next subsections.

Regularized Discriminant Analysis

Linear Discriminant Analysis (LDA has been widely used in BCIs to recognize real and imagery motor tasks (LEW et al., 2012; GARRETT et al., 2003; XU et al., 2014). It has demonstrated that regularized classifiers, such as Regularized Discriminant Analysis (RDA), can provide higher performance than LDA, but curiously it has been little used in BCIs to recognize gait intention (VELU; SA, 2013). RDA formulation is given by the following expression:

$$\Sigma_k(\lambda) = \frac{(1-\lambda)\mathbf{S}_k + \lambda\mathbf{S}}{(1-\lambda)\mathbf{W}_k + \lambda\mathbf{W}} \tag{2.32}$$

$$\Sigma_k(\lambda,\gamma) = (1-\gamma)\Sigma_k(\lambda) + \frac{\gamma}{p}\operatorname{tr}(\Sigma_k(\lambda))\mathbf{I} \tag{2.33}$$

$$d_k(\mathbf{X}) = (\mathbf{X}-\overline{\mathbf{X}}_k)^{\dagger}\mathbf{\Sigma}_k^{-1}(\lambda,\gamma)(\mathbf{X}-\overline{\mathbf{x}}_k) + \log|\mathbf{\Sigma}_k(\lambda,\gamma)| - 2\log(\pi_k)\ , \tag{2.34}$$

where $\mathbf{X}$ is a known object (describe by p features), $\overline{\mathbf{X}}_k$ is the average vector of the class k, $\mathbf{S_k}$ is the covariance matrix computed from all objects that correspond to the class k, $\mathbf{S}$ is the mean covariance matrix $\mathbf{S} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{S}_k$, $\mathbf{W_k}$ is the assigned weight for each object k, $\mathbf{W}$ is $\sum_{k=1}^{K} W_k$the assigned weight for each object k, π_k is the prior probability of each object k, Σ is

the covariance matrix, λ and γ are regularization parameters, being $0 \leq \lambda \leq 1$ and , $0 \leq \gamma \leq 1$. Both λ and γ are chosen, looking in the training stage a classification model with the lowest bias, which provides low misclassification of future patterns.

RDA formulation may be used to obtain several setups of classifiers, which includes well-known classifiers, such as LDA and Quadratic Discriminant Analysis (QDA). For example, four corners can be defined by λ and γ. The lower left corner ($\lambda = 0$, $\gamma = 0$) represents QDA, and the lower right ($\lambda = 1$, $\gamma = 0$) represents LDA. The upper-right corner ($\lambda = 1$, $\gamma = 1$) corresponds to the nearest means classifier well-known in pattern recognition, where an observation is assigned to the class with the closest (Euclidean distance) mean. The upper-left corner of the plane represents a weighted nearest means classifier with the class weights being inversely proportional to the average variance of the measurement variables within the class. Moreover, holding γ fixed at zero and varying λ produces models in between QDA and LDA. Furthermore, holding λ fixed at zero and increasing γ attempts to unbias the sample based on eigenvalue estimates. Finally, holding λ fixed at one and increasing λ gives rise to a ridge regression analogue for LDA.

Support Vector Machine

Support Vector Machine (SVM) has been used in BCIs with characteristic linear, polynomial and Gaussian (kernel Radial Basic Function (RBF), being SVM linear and RBF more used with promissing results (HORTAL et al., 2016; BARACHANT et al., 2013; GARRETT et al., 2003).

Regarding SVM linear, it uses a discriminant hyperplane to identify classes (LOTTE et al., 2007; HORTAL et al., 2016; GARRETT et al., 2003) through linear decision boundaries, where margins (or distances) between nearest training points of different classes are maximized selecting suitable hyperplanes. Then, the generalization capabilities of classification models may increase, maximizing margins (LOTTE et al., 2007). This way, SVM uses a regularization parameter C that enables accommodation to outliers, but allows errors during the training stage. This classifier has been applied with success for several BCIs (BARACHANT et al.,

2013).

Different to SVM linear, SVM RBF can be used to create nonlinear decision boundaries, increasing slightly the classifier complexity. It is based on a kernel function $K(x, y)$, presented in Equation (2.35), which allows mapping the data to another space, generally with higher dimensionality. This classifier has also showed promising results in BCIs to recognize gait starting and gait stopping (HORTAL et al., 2016).

$$K(x, y) = \exp\left[\frac{-||x - y||^2}{2\sigma^2}\right] , \tag{2.35}$$

SVM classifiers have few hyperparameters that need to be defined by hand, namely, the regularization parameter C and width σ (only for RBF, as shown in Equation (2.35)). It is worth noting that BCIs based on SVM have low speed of execution. Additionally, few authors have published suitables values of C and σ of BCIs being applied to recognize gait planning. Moreover, C and σ values have been manually defined by several authors, while an automatic regularized process may help to reduce the bias.

2.3.5 Evaluation

The aim of the BCIs proposed on this research is to recognize user's intention through raw EEG, providing a communication pathway to control end-devices and applications. A confusion matrix $\mathbf{C}$ is a solution to evaluate the different BCIs, where $\mathbf{C} = \{c_{ij}\}$, i is the row index, j is the column index, $i, j \in \{1, 2, \ldots, l\}$, l is the total number of classes.

It is possible to make on $\mathbf{C}$ the following observations:

1. $\sum_{j=1}^{l} c_{ij}(f) = c_i(f)$ denotes the total number of examples of class i in the test set.

2. $\sum_{i=1}^{l} c_{ij}(f) = c_j(f)$ denotes the total number of examples assigned to class j by the classifier f.

3. All the diagonal entries c_{ii} denote the correctly classified examples for class i. Hence,

$\sum_{i=1}^{l} c_{ii}(f)$ denotes the total number of examples classified correctly by the classifier f.

4. All the non-diagonal entries denote misclassifications. Hence $\sum_{i,j:i\neq j}^{l} c_{ij}(f)$ denotes the total number of examples assigned to wrong classes by classifier f.

Then, two analysis from $\mathbf{C}$, such as multi-class and single-class, can be carried out. Multi-class analysis has been widely used in BCI through accuracy (ACC) and Kappa indices (JIANG et al., 2015; HASHIMOTO; USHIBA, 2013), while single-class analysis has been performed using true positive rate (TPR), false positive rate (TPR), and F measure.

ACC measures the fraction of correctly classified instances in the test set, which can be used to summarize the overall performance, taking into account all data classes.

$$ACC(f) = \frac{\sum_{i=1}^{l} c_{ii}(f)}{\sum_{i,j=1}^{l} c_{ij}(f)}\ , \tag{2.36}$$

TPR measures the proportion of examples that correspond to class i, correctly assigned to this class by the learning algorithm. In contrast, FPR quantifies the instances assigned to this class i of interest that actually do not belong to this class. Both TPR and FPR can be calculated, respectively, through the following expressions:

$$TPR_i(f) = \frac{c_{ii}(f)}{\sum_{j=1}^{l} c_{ij}(f)}\ , \tag{2.37}$$

$$FPR_i(f) = \frac{\sum_{j:j\neq i}^{l} c_{ji}(f)}{\sum_{j,k:j\neq i}^{l} c_{jk}(f)}\ . \tag{2.38}$$

F measure combines precision and recall in a single metric. More specifically, F measure is a weighted harmonic mean of precision and recall (JAPKOWICZ; SHAH, 2011), as shown in

Equation (2.40). Several variations of F measure can be obtained. For instance, the balanced F (for $\alpha = 1$) measure weights the recall ($Rec = TPR$) and precision ($Prec$) of the classifier evenly. $Prec_i$ measures the proportion of correctly assigned i examples. In other words, this metric measures how precise the algorithm is when identifying examples of a given class i.

$$Prec_i(f) = \frac{c_{ii}}{\sum_{j=1}^{l} c_{ji}} , \tag{2.39}$$

$$F\alpha(f) = \frac{(1+\alpha)[Prec(f) \times Rec(f)]}{[\alpha \times Prec(f)] + Rec(f)} , \tag{2.40}$$

In addition, PNM measures the combined misclassification in the prediction of each class (resting from stand position, gait planning, and gait stopping) (CASTRO; ARJUNAN; KUMAR, 2015), which can be computed by the following equation:

$$PNM_i = \frac{\left(C_{ii} - \sum_{j=1, j\neq i}^{K} C_{ij}\right) + \left(C_{ii} - \sum_{i'=1, j=i, i'\neq j}^{K} C_{i'j}\right)}{\sum_{j=1}^{K} C_{ij} + \sum_{i'=1, j=i}^{K} C_{i'j}} , \tag{2.41}$$

where C_{ij} is an element of the confusion matrix that corresponds to the row i and column j, and K is the total number of classes. PNM takes values from 1, if all predictions are correct, to -1 if all predictions are wrong.

2.4 Conclusion

This chapter 2 introduced several technologies for lower-limb rehabilitation, such as exoskeleton and ortheses, including the role of the brain using cortical motor patterns, to design BCIs that allow improving the close-loop control between human and end-applications. It was also discussed how EEG may be blurred by several artifacts and interferences, which may affect the performance of BCIs. Consequently, several methods for BCIs focused on recognizing the motor intention, specially, the gait intention, were presented.

The next chapter presents the proposed methods in this thesis to improve the performance of BCIs during gait intention recognition, which will be applied to the exoskeleton ALLOR of UFES/Brazil.

Chapter 3

Materials and Methods

Several BCIs have been proposed to recognize motor intention of upper and lower limbs (JIANG et al., 2015; LEW et al., 2012; XU et al., 2014), such as shown in Section 2.3. Pre-processing methods play an important role to preserve the information of interest, during the reduction of common interferences.

Traditional spatial filters, such as LAR and WAR have been commonly used on the pre-processing stage of BCIs for gait intention recognition. These filters use weight indices based on inter-electrode distances, which do not depend on underlying data. Therefore, these methods may add on the target electrode undesirable artifacts from neighbor electrodes, and may remove patterns related to gait intention.

Such as aforementioned, events related to gait planning may be generated 2 s before the movement onset, while gait stopping associated events may be found +1 s after finishing the movement. However, it is a hard task to assign labels on time periods in which these actions occur, therefore, these patterns are labeled with uncertainty. For this reason, unsupervised methods may be suitable to process these patterns.

This chapter presents our proposed BCI based on adaptive spatial filter to preserve the useful information on EEG signals, in addition to other unsupervised method for feature selection. Additionally, a statistical analysis to obtain a classifier that improve the BCI performance is also introduced.

3.1 Proposed Brain-Computer Interface

3.1.1 Adaptive spatial filter

Figure 3.1 shows the simplified process to reduce, from neighbor electrodes (1, 2, 3 and 5), common interferences that can be present on the electrode of interest (T). This method provides an adaptive behavior for the traditional spatial filters LAR and WAR, which are termed Ad LAR and Ad WAR, respectively. Similar to LAR (PFURTSCHELLER; NEUPER; BERGER, 1994), Ad LAR aims to reduce common interferences on the electrode of interest (or target electrode), using the nearest neighbor electrodes (Ad LAR-small) or next nearest neighbors (Ad LAR-large). The method is based on a two-stage approach, in which the first stage, termed "Neighbor Selection", is responsible for the selection of neighbor electrodes, while the second stage, termed "Virtual Distance & Weights Computation", computes the virtual distances between electrodes. The objective of this approach is to select those neighbor electrodes around the electrode of interest that provide the best estimation of the common interferences. Lowest

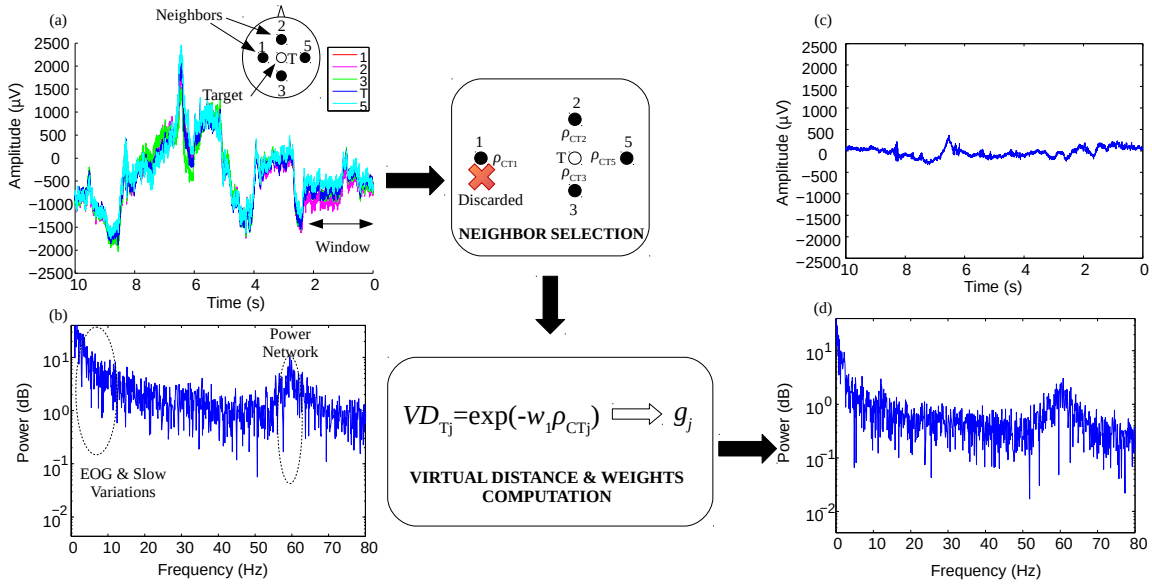

Figure 3.1: Block diagram of the proposed adaptive filter for EEG signals. Signals acquired around Cz (target electrode, T) during a cycle of gait, with processing through Ad LAR-small setup. a) raw EEG signals; b) power spectrum on T location using Fast Fourier Transform (FFT); c) filtered signal from the location T; d) power spectrum of the filtered signal using FFT showing an attenuation on components of low frequency ($\leq$ 10 Hz) and 60 Hz (power line).

values of weight are assigned to electrodes that contain information highly correlated to the electrode of interest, in order to preserve the useful information. This way, the concordance correlation coefficient (CCC) (LAWRENCE; LIN, 1989) is introduced here to analyze the interchangeability between the electrode of interest and neighbor electrodes. CCC values, denoted as ρ_c in Equation (3.1), are computed on a window W_f, shifted for each sample throughout the time, where higher values are related to a good interchangeability (LAWRENCE; LIN, 1989). This index takes values between [-1, 1] and provides accuracy and precision criteria between observations, which may be used as rejection criterion for neighbor electrodes that can affect the SNR of the target electrodes during the reduction of common interferences. For n independent pairs of samples, ρ_c can be computed by the following equation:

$$\rho_{cij} = \frac{2\frac{1}{n}\sum_{k=1}^{n}\left(Y_{ki}-\overline{Y_i}\right)\left(Y_{kj}-\overline{Y_j}\right)}{\left[\frac{1}{n}\sum_{k=1}^{n}(Y_{ki}-\overline{Y_i})^2\right]^2+\left[\frac{1}{n}\sum_{k=1}^{n}(Y_{kj}-\overline{Y_j})^2\right]^2+\left(\overline{Y_i}+\overline{Y_j}\right)^2}, \tag{3.1}$$

where Y is the current segment of raw EEG, n is the size of window W_f, i is the electrode of interest, and j is the neighbor electrode.

Neighbor Selection

Based on CCC and standard deviation

The estimation of common interferences may be obtained from all neighbor electrodes. However, due to possible electrode and cable movements, some electrodes may present high values of amplitude throughout raw EEG, providing a low performance during the interference rejection. Thus, a first stage is proposed here to select neighbor electrodes that are more suitable to reduce the interferences and artifacts on the electrode of interest.

According to the filter setup (LAR-small, LAR-large and WAR), ρ_{cij} and the standard deviation (SD) of amplitude Y_{std}^{ik} values are computed combining the electrode of interest ith with the neighbor electrodes jth. Y_{std}^{ik} is computed on the current sample from the neighbor electrodes

jth respect to the electrode of interest ith, using Equation (3.2). From these values, threshold values such as Y^{i}_{stdTH} and ρ_{cTH} are calculated using Equations (3.3) and (3.4), respectively, in order to regulate the selectivity of this process.

$$Y^{ik}_{std} = \sqrt{\frac{\sum_{j=1}^{N} \left(Y_{jk} - Y_{ik}\right)^2}{N-1}} \tag{3.2}$$

$$Y^{i}_{stdTH} = 2Y^{i}_{std\,\min} + \left| (Y^{i}_{std\,\max} - Y^{i}_{std\,\min}) \frac{Y^{i}_{std\,\max} - Y^{ik}_{std}}{Y^{i}_{std\,\max} - 2Y^{i}_{std\,\min}} \right| \tag{3.3}$$

$$\rho_{cTH} = \begin{cases} \rho_{cM} - M(\rho_{cM} - \rho_{cij}) - 0.05, & M(\rho_{cM} - \rho_{cij}) \leq 0.05 \\ \rho_{cM} - M(\rho_{cM} - \rho_{cij}) \ , & otherwise \end{cases} \tag{3.4}$$

where $Y_{std\,\min}$ and $Y_{std\,\max}$ are the minimum and maximum values respectively, obtained respect to the electrode i throughout the time, ρ_{cM} is the median value calculated on the ρ_{cij} values, i is the current target electrode, j is the current neighbor electrode, k is the current sample, N is the total of channels, and $M(\cdot)$ is the median operator.

A combination of maximum and minimum values of the standard deviation is carried out to establish an adaptive threshold through Equation (3.3), which allows deciding when applying the electrode selection. Notice that values of Y^{i}_{std} close to $Y_{std\,\min}$ may increase the threshold value around $Y_{std\,\max}$, and vice-versa. In addition, SD value may present high values due to artifacts not located on all channels. Thus, the condition $Y_{std\,\max} \leq 10Y_{std\,\min}$ was introduced in Equation (3.3), which was adjusted empirically, using a database composed of raw EEG acquired during steady-state visual evoked potentials (SSVEP) (CHEN et al., 2015). Appendix A.6 describes details about this database.

On the other hand, Equation (3.4) is used to consider all neighbor electrodes during the computation of the common interferences in case of their ρ_{cij} values be very close, as shown at the first condition of this formulation; else some electrodes may not be taken into account to improve the estimation.

All neighbor electrodes are considered as suitable candidates to obtain the common reference if the current Y^{ik}_{std} value is lower than Y^{i}_{stdTH}; else, only neighbor electrodes with ρ_{cij} value higher than ρ_{cTH} are selected.

Based on Zscore analysis

Figure 3.2 shows a second approach based on Zscore analysis, which is used to reject those neighbor electrodes that may be not suitable to compute the common interference in the current sample k. For this purpose, on the current sample k amplitude values from all neighbor electrodes around the electrode of interest are analyzed, using the Zscore definition. Zscore is a useful statistic index that computes the probability of a score occurring within a normal distribution. Thus, it may be used to detect outliers on EEG, during offline processing (NOLAN; WHELAN; REILLY, 2010). This formulation is introduced here to reject neighbor electrodes during on-line processing, which may be not suitable to reduce common interferences, and preserve the useful information. Then, a logarithm transform is applied on raw EEG as a first stage to reduce effects that EEG variability produces on the mean and standard deviation values. For this purpose, Equations (3.5) and (3.6) are used to compute the logarithm on the positive PS and negative NS amplitudes, respectively. After, the cumulative value for PS and NS is stored at PA and NA, when it is higher than PA and NA, respectively. Here, PA and NA values are used according to Equation (3.7) to follow the direction of the predominant amplitudes, which may be associated to artifacts. Thus, the electrode rejection is only applied

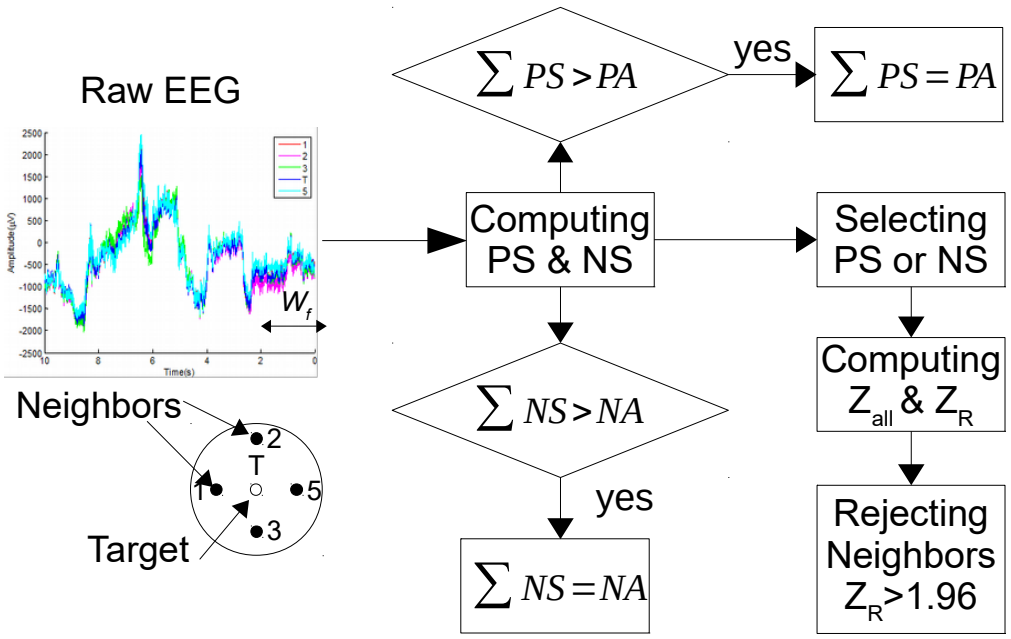

Figure 3.2: Representation of the proposed method based on Zscore to reject neighbor electrodes non-suitables.

on the way (positive or negative amplitudes) of possible artifacts.

$$\mathbf{PS}_k = \log_2(\mathbf{Y}_k)\ ,\ y_i = 1\ \forall\ y_i < 1,\ y \in \mathbf{Y}_k \tag{3.5}$$

$$\mathbf{NS}_k = \log_2(-\mathbf{Y}_k)\ ,\ y_i = 1\ \forall\ y_i < 1,\ y \in -\mathbf{Y}_k\ , \tag{3.6}$$

$$\mathbf{YN}_k = \begin{cases} \mathbf{PS_k}\ ,\ \forall\ PA_k \geq 3.3NA_k \vee (3.3PA_k \geq NA_k \wedge PA_k \geq NA_k) \\ \mathbf{NS_k}\ ,\ otherwise \end{cases} \tag{3.7}$$

where $\mathbf{Y}_k$ are the amplitude channels, $\mathbf{YN}_k$ are logarithm amplitudes, and k is the current sample. Notice that a difference of 3.3 times (see Equation (3.7)) between logarithm values computed with base 2 means a difference higher than 10 times between their non-transformed data. Once obtained $\mathbf{YN}_k$, both Zscore values $\mathbf{Z_{all}}$ and $\mathbf{Z_R}$ are computed on $\mathbf{YN}_k$, in order to reject those neighbor electrodes with $\mathbf{Z_R} > 1.96$, being at the same time the correspondent $\mathbf{Z_{all}}$ value of the electrode of interest (T) less than 1.65.

$$\mathbf{Z_{all}}_k = \frac{(\mathbf{YN}_k - \overline{\mathbf{YN}_k})}{\sigma_{\mathbf{YN}_k}} \tag{3.8}$$

$$\mathbf{Z_R}_k = \begin{cases} \frac{(\mathbf{YN}_k - \mathbf{E})}{\sigma_{saved}}\ ,\ \forall\ \sigma_{saved} > 0 \\ \frac{(\mathbf{YN}_k - \mathbf{E})}{\sigma_{\mathbf{YN}_k}}\ ,\ otherwise\ , \end{cases} \tag{3.9}$$

where $\mathbf{Z_{all}}_k$ is the Zscore value computed from mean and standard deviation values of $\mathbf{YN}_k$, E is the logarithmic amplitude of the neighbor electrode that presents its $\mathbf{Z_{all}}$ value more close to 1.65 (≤ 1.65), σ_{saved} is the highest value obtained from $\sigma_{\mathbf{YN}_k}\ \forall\ \mathbf{Z_{all}} \leq 1.65$, and $\mathbf{Z_R}_k$ is an adaptive Zscore value.

Virtual Distance and Weights Computation

The second stage is used to compute virtual distances (VD) between the electrode of interest and the selected neighbor electrodes, which are based on the similarity indice ρ_c. VD depends on ρ_c index, such as shown in Equation (3.10). This curve searches the minimum value of VD

for high similarity ($\rho_c \to 1$) between electrodes.

$$VD_{ij} = \exp\left[-w_1 \rho_{cij}\right] \ , \tag{3.10}$$

where i is the electrode of interest, j is the selected neighbor electrode, ρ_{cij} is the concordance correlation coefficient, and w_1 is a adjusted coefficient through EEG signals with known neural information (as SSVEP), such as shown in Section 3.1.4.

Finally, a weighted index is obtained for each neighbor electrode from VD, in order to estimate the common interferences. As result, for Ad LAR and Ad WAR filters, Equations (3.11) and (3.12) are applied, respectively, to provide highest values for neighbor electrodes with low similarity respect to the target electrode.

$$V_i^{Ad\ LAR} = V_i^{CR} - \sum_{j \in S_i} g_{ij} V_j^{CR}, \ g_{ij} = \frac{VD_{ij}}{\sum\limits_{j \in S_i} VD_{ij}} \tag{3.11}$$

$$V_i^{Ad\ WAR} = V_i^{CR} - \sum_{j=1}^{N} g_{ij} V_j^{CR}, \ g_{ij} = \frac{VD_{ij}}{\sum\limits_{j=1;\ j \neq i}^{N} VD_{ij}} \ , \tag{3.12}$$

where V_{CR}^i is the measured potential between electrode of interest i and reference electrode, S_i is the set of surrounding electrodes or neighbor electrodes to the electrode i, VD_{ij} is the virtual distance between the electrode i and the selected neighbor j, N is the number of electrodes, and g_{ij} is the weight index. The proposed formulation for the Ad LAR and Ad WAR filters may contribute to reduce, on the target electrode, undesirable artifacts and noise, because the weighted indices are related to the similarity between the electrode of interest and the selected neighbor electrodes. It is possible to see that this new method provides higher weight g_{ij} to neighbor electrodes with low similarity to the target electrode, in order to preserve the useful neural information. The next section introduces two-approaches for feature selection based on representation entropy (RE).

3.1.2 Feature Selection

Based on information compression and entropy

An unsupervised method of low computational cost based on the Representation Entropy (RE) is proposed here for feature selection (DELISLE-RODRIGUEZ et al., 2017), aiming to improve the BCI performance during the analysis of clusters with uncertainty. Figure 3.3 shows a block diagram of the proposed method for feature selection, which is composed of two processes.

The first process, highlighted with dotted lines, allows computing threshold values related to RE (MITRA; MURTHY; PAL, 2002). For this purpose, three stages were proposed: 1) Clusters of different sizes (10 to 80% of the total features, with increment of 10%) are formed to compute the RE value of several feature combinations from the original set, in which high value of RE is related to low redundancy. More details are given in Section A.4. 2) RE values are arranged in a descending order, and analyzed through Z-test, in order to obtain RE values that do not present a significant difference ($p < 0.05$) respect to the maximum value. 3) The Mahalanobis distance is computed on these last values to select an RE value as threshold, which is the value that presents the minimum distance respect to all values.

The second process is used for feature selection through the maximal information compression index (MICI) and RE, which is composed of the following three stages: 4) The updated feature set is analyzed through MICI to obtain feature combinations with lowest value or high

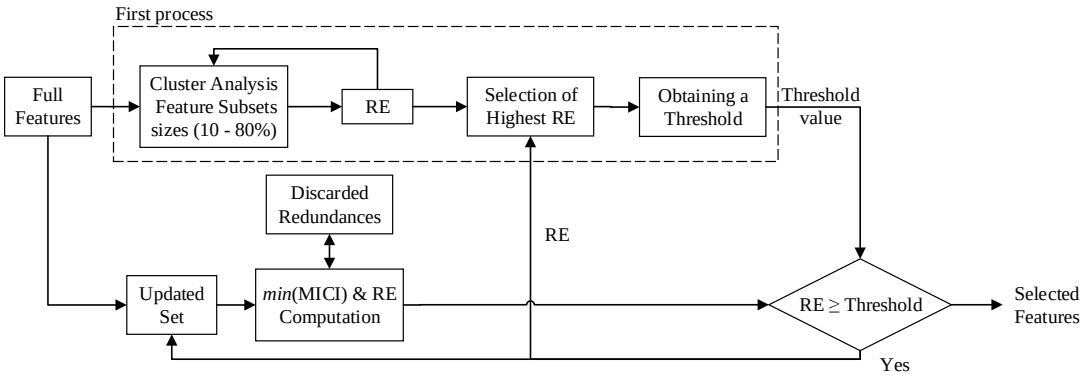

Figure 3.3: Representation of the proposed method for feature selection the maximal information compression index (MICI) and the representation of entropy (RE).

redundancy (MITRA; MURTHY; PAL, 2002). More details of MICI are given in Section A.5. 5) These new redundancies compete with rejected features, in order to obtain an updated set formed by the feature that provides the highest RE value during the combination with non-redundant features. 6) This current value of RE is analyzed together with the highest RE values following again the steps 2 and 3 (see first process), in order to update the threshold value. The process is repeated until the RE value of the updated set is higher than the threshold value.

Based on entropy and LDA

A supervised method based on the RE index (DELISLE-RODRIGUEZ et al., 2017) and classifiers (LOTTE et al., 2007) is introduced here for feature selection, aiming also to improve the BCI performance during the analysis of clusters with uncertainty. Figure 3.4 shows the block diagram of the proposed method for feature selection, which is composed of two processes.
The first process, highlighted with red lines, allows computing threshold values of ACC (JAPKOWICZ; SHAH, 2011). For this purpose, five stages were proposed: 1) Clusters of different sizes (10 to 80% of the total features, with increment of 10%) are formed to compute the RE value of several feature combinations from the original set, which high value of RE is related to low redundancy. 2) RE values are arranged in a descending order, and analyzed through Z-test, in order to obtain a preliminary size of feature sets that provide highest values of RE without

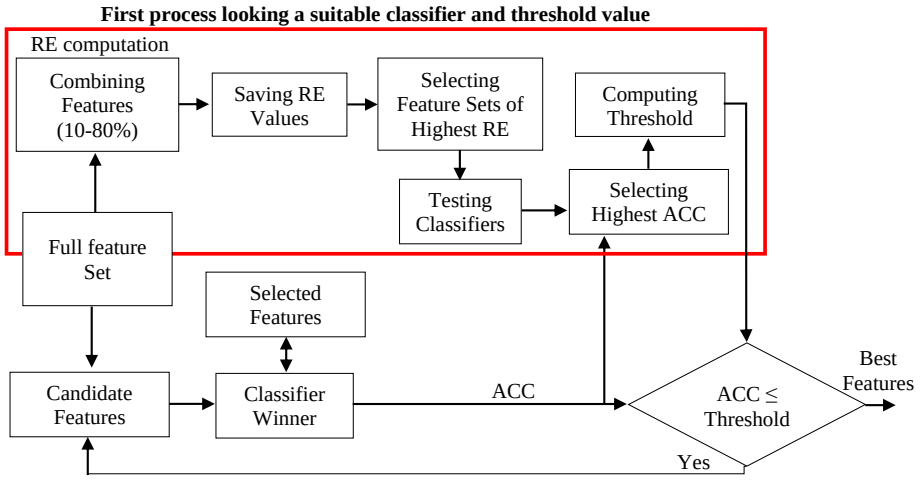

Figure 3.4: Block diagram of the proposed method to obtain a feature set using supervised feature selection.

significant differences ($p < 0.05$) respect to the maximum value. 3) Feature combinations with the preliminary size are used to test several classifiers through ACC value, where the classifier of best performance is selected as winner. 4) ACC values are arranged in a descending order, and analyzed through Z-test, in order to obtain ACC values that do not present a significant difference ($p < 0.05$) respect to the maximum value. 5) The Mahalanobis distance is computed on these last values to select an ACC value as threshold, which is the value that presents the minimum distance respect to all values.

The second process is a genetic algorithm (GA) for feature selection, using the winner classifier and its ACC for each input feature set. This algorithm is composed of the following three stages: 6) Each candidate feature is analyzed through all possible combinations with the selected features, which begin as a null set, using the winner classifier. During this process each candidate feature competes with all selected features, thus, it is possible that features selected previously return again as candidates. 7) The combination of highest ACC is taken as a new updated feature set, and its ACC is analyzed together with the other highest ACC values following again the steps 4 and 5 such as shown in the first process, in order to update the threshold value. The process is repeated until the ACC value of the updated set is higher than the threshold value.

3.1.3 Classification: Training stage

Figure 3.5 shows a block diagram of the proposed method to search automatically from a training set, a suitable setup of the classifier, and a classification model to recognize patterns of interest, in this case the rest state, gait planning, and gait stopping. First, the full training set is divided on both training and testing set, during an inner cross-validation $k1$-fold ($k1$ up to 40). For each fold, both group are applied to several classifiers to obtain their performance using the index Idx, which is derived from the combination of weighted parameters, such as true positive rate (TPR), false positive rate (FPR), Kappa, accuracy (ACC) (JAPKOWICZ;

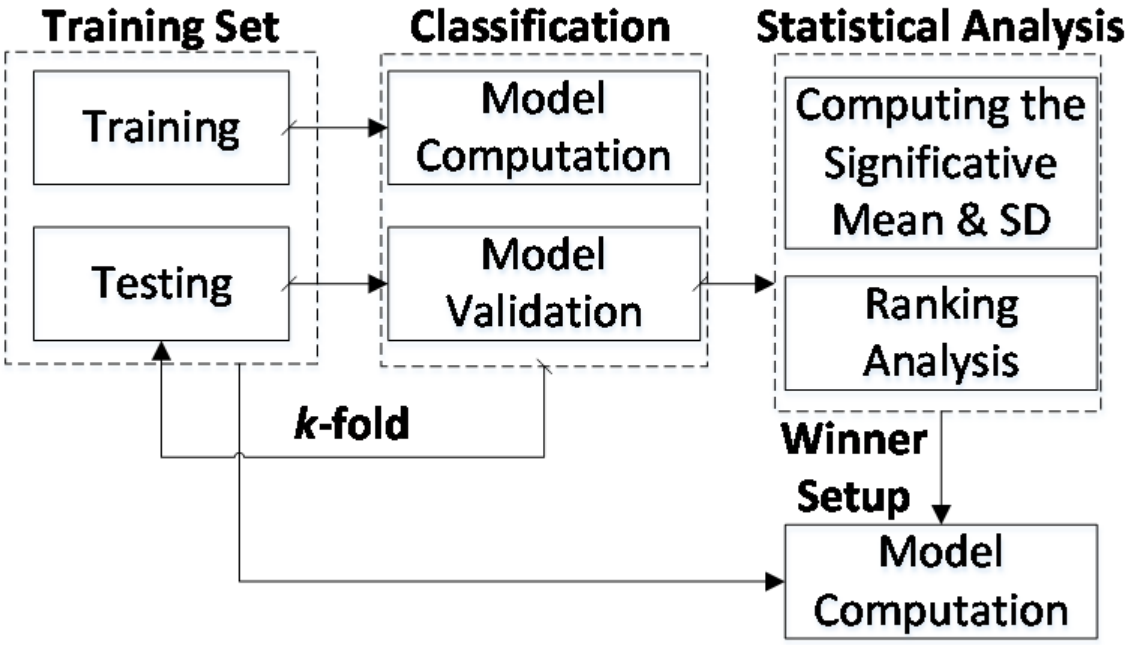

Figure 3.5: Block diagram of the proposed method to obtain a classification model from several setups.

SHAH, 2011), and PNM (CASTRO; ARJUNAN; KUMAR, 2015), as shown in Eq. (3.1):

$$Idx = 1 - \sum_{i=1}^{n} w_i P_i \ , \tag{3.13}$$

where w_i is the index weighted assigned to the parameter i, $w = [0.3, 0.3, 0.15, 0.25]$, and $P = [Kappa, (1 - FPR_{max}), ACC, PNM_{min}]$.

After finishing the $k1$-fold process, a statistical evaluation is carried out through the bootstrap confidence interval and density distribution analysis. First, the bootstrap confidence analysis was applied on the mean and median values throughout 2000 iterations, using the normal criteria. Second, the outputs of the bootstrap samples are analyzed from the density distribution to compute expected values for both mean and median. Finally, a ranking analysis is performed between every classifiers, using the mean and upper values of confidence intervals, obtained for both mean and median values related to classification outputs. The next section presents the methodology used to fit the proposed model (see Equation (3.10)) and the size of the window W_f. In addition, methodologies to evaluate the proposed methods are described.

3.1.4 Statistical Analysis

Model fitting and optimization

In the literature, little information about methods used to fit models of spatial filter was found. Then, the well-known SSVEP was adopted here as a methodology to fit the proposed model, looking for a way to preserve the useful information when rejecting artifacts. For this purpose, ten subjects (SV1 to SV10) from a SSVEP dataset (see Appendix A.6) with 35 subjects (17 females, aged 17-34 years, mean age: 22 years) obtained from <ftp://sccn.ucsd.edu/pub/ssvep_benchmark_dataset> were used to fit the proposed model through several trials. Ten subjects were used here for adjusting the model. In this manner, the rest of the database can be used for evaluation purposes. Then, EEG signals of these ten subjects were analyzed through the power spectrum density, which was computed using the FFT on the locations O1, O2, Oz, PO3, PO4, PO7, PO8, POz, P1, P2, P3, P4 and Pz, such as suggested by (COTRINA et al., 2017). Twelve trials were randomly selected for each subject, and up sampled at 400 Hz. Thus, a total of 120 trials was employed from these 10 subjects to fit the proposed model.

Notice that the proposed model in Equation (3.10) is based on several parameters (W_f, w_1) to compute virtual distances (VD) between electrodes, which should be empirically fitted. For this purpose, several setups of the proposed method were applied on the SSVEP database during an on-line processing, in order to obtain the one that provides output signals with the best values of SNR and attenuation (A). SNR and A values were computed from the FFT using Equations (3.14) and (3.15), respectively (CHEN et al., 2015). This way, several windows W_f with different sizes (20, 50, 100, 150, 200, 250 and 300 ms), as well as different threshold values (5, 10, 15, 20) related to the ratio $V^i_{std\,\max}/V^i_{std\,\min}$ were combined with the coefficient w_1 (from 1 to 100, with increment of 5) of VD (see Equation (3.10)).

$$SNR = 20 \times log_{10}\left(\frac{O(f)}{\frac{1}{10}\sum_{k=1}^{5}[(f-0.2\times k)+O(f+0.2\times k)]}\right) \tag{3.14}$$

$$A = 20 \times log_{10} \left[\frac{O(f)}{I(f)} \right]^{sign[I(f)-O(f)]} , \tag{3.15}$$

where I and O are the magnitude spectrum from the FFT of the input and output signals, respectively, and f are the SSVEP components. Results of model fitting are shown in Section 4.1.

Preservation of SSVEP components

The EEG signals from all the 35 subjects were used from the SSVEP database (see Appendix A.6) to study the capability of the proposed method to preserve the main SSVEP components after applying the adjusted Ad WAR and WAR filters on three EEG electrodes O1, O2 and Oz. Then, indices such as attenuation (A), SNR, and coherence were analyzed (QUIROGA et al., 2002). SNR was computed around the SSVEP components using Equation (3.14), and A value was calculated on the main frequencies through Equation (3.15). The coherence was computed through the following equation (QUIROGA et al., 2002):

$$\Gamma_{xy}(\omega) = \frac{|C_{xy}(\omega)|}{\sqrt{C_{xx}(\omega)}\sqrt{C_{yy}(\omega)}} , \tag{3.16}$$

where x is the EEG segment without processing, y is the output segment after processing, Γ_{xy} is the coherence function that provides a measure of the linear synchronization between x and y as a function of the frequency ω, ω is the discrete frequency, and C_{xx}, C_{yy} and C_{xy} are defined as the Fourier transform of the cross correlation.

Canonical Correlation Analysis (CCA) has been used in BCIs to recognize SSVEP targets (LIN et al., 2007). Here, CCA was applied to recognize the 40-targets from the database, using six harmonics at the locations O1, O2, Oz, PO3, PO4, PO5, PO6, POz, and Pz, such as done by (CHEN et al., 2015). ACC and FPR were adopted to analyze the performance of the BCI (JAPKOWICZ; SHAH, 2011), and the non-parametric Wilcoxon rank sum test was used to obtain a statistical comparison between both Ad WAR and WAR filters for different p-values (0.05, 0.01, 0.001, and 0.0001).

Application in a BCI for gait planning/stopping recognition

This work is a first stage of a robotic system based on a lower-limb robotic exoskeleton, built at UFES/Brazil (VILLA-PARRA et al., 2015), which must be commanded by a BCI during walking. For this reason, the proposed methods were evaluated in BCIs to recognize the gait planning/stopping, which were composed of the following four-stages: pre-processing, feature extraction, feature selection, and classification. Section 3.1.5 describes an experimental protocol, which was carried out with six volunteers (S1 to S6) to evaluate the performance of the BCI based on Ad LAR and Ad WAR in the pre-processing stage, during the gait planning recognition.

A total of 24 repetitions for gait planning (-1.5 to 0 s from the gait onset) and gait stopping (-0.5 to +1.0 s from the gait offset), and 36 repetitions of rest state from the stand position (from 0 to +2.0 s, using the stimulus reference) were collected on each subject during six sessions. The feature vector was obtained using sliding windows of 250 ms in length for each sample. Thus, from each segment of 1.5 s (gait planning and stopping) and 2.0 s (resting from stand position) of duration, a total of 1250 (separately for each one) and 1750 trials, respectively, were obtained.

Different BCIs were developed here to recognize the gait planning/stopping, which are based on Ad LAR, LAR, Ad WAR, and WAR filters, with the aims of comparing their performance. These filters were used to obtain the reference-free (RF) EEG on Cz, CP1, and CP2. The electrode locations FC3, FC1, C3, C4, and Pz surrounding Cz, CP1, and CP2 were adopted in a similar way as done by (JIANG et al., 2015; XU et al., 2014; HORTAL et al., 2016), when applying WAR filters. In contrast, for LAR the following locations FCz, C1, C2, CPz, CP3, and CP4 were adopted.

For these BCIs based on the aforementioned spatial filters, several features were explored on Cz, CP1, and CP2 locations, in order to obtain feature vectors in time and frequency domains. The analyzed features were the RF EEG (MCFARLAND et al., 2006; QIN; DING; HE, 2004; LEMM et al., 2005), mean absolute value (MAV), wavelength (WL), fractal dimension from Higuchi (FDH) (PHOTHISONOTHAI; WATANABE, 2013; LIU et al., 2005), fractal dimension from Sevcik's and Higuchi (FDSH) (LIU et al., 2005), and band power (BP) (HASHIMOTO;

USHIBA, 2013; HORTAL et al., 2016). Seven features related to BP were computed through FFT on the following frequency bands: 0.1 to 4 Hz, 8 to 12 Hz, 13 to 17 Hz, 18 to 24 Hz, 26 to 30 Hz, 30 to 50 Hz, and 50 to 70 Hz (HASHIMOTO; USHIBA, 2013; HORTAL et al., 2016). All features were obtained through a sliding window of 250 ms in length for each sample, from the planning interval (-1.5 to 0 s from the onset reference), stopping interval (-0.5 to +1.0 s from the offset reference) and resting interval, from the stand position (0 to +2.0 s from the stimulus reference). In addition, BCIs based on bandpass filter (from 8 to 30 Hz), combined with methods for feature extraction such as CSP (BLANKERTZ et al., 2008), FBCSP (MÜLLER-GERKING; PFURTSCHELLER; FLYVBJERG, 1999), and RK (BARACHANT et al., 2013) were also analyzed here, in order to establish a comparison with the BCIs based on proposed filters. Specifically, for FBCSP, a filter bank (5th order, Butterworth) was applied in the following bands: 0.1-4Hz, 8-12 Hz, 13-17 Hz, 18-24 Hz, 26-30 Hz, 30-50 Hz, and 50-70 Hz. CSP, FBCSP, and RK methods were applied on the same electrode locations of WAR and Ad WAR filters, using the same sliding windows of 250 ms in length for each sample. The appendices A.3, A.1, and A.2 show details about CSP, FBCSP, and RK methods.

In order to compare these BCIs, a cross-validation technique ($k2 = 3$) was used to obtain the training and validation sets, both formed by independent sessions. Here, the training set was used only to adjust the classifiers SVM, regularized discriminant analysis (RDA), and linear discriminant analysis (LDA), as well as to obtain the classification model. This way, the values of box constraint C (0.01, 0.05, 0.1, 1, 5, 10) of the SVM classifier were scanned in a inner cross-validation ($k1 = 2$), as shown at the training stage (Section 3.1.3). Similarly, the RDA classifier was adjusted, scanning for both λ and γ parameters, values from 0 to 1 with increment of 0.1 (VELU; SA, 2013; FRIEDMAN, 1989). Additionally, in order to obtain the best performance of the (LDA) classifier, three covariance matrices, such as diagonal, full, and spherical were evaluated (VELU; SA, 2013).

The indices ACC, TPR, FPR, and F1 (JAPKOWICZ; SHAH, 2011) were adopted to evaluate the performance of each BCI. Furthermore, the latency and continuous recognition throughout the planning intervals were also analyzed, and the non-parametric Wilcoxon rank sum test was used to obtain a statistical comparison between the BCIs for different p-values (0.05, 0.01,

0.001, and 0.0001).

3.1.5 Protocol for gait planning recognition

An experimental protocol was implemented to evaluate the proposed methods in a BCI for gait planning/stopping recognition. For this purpose, six healthy subjects (males, 31.0 ± 5.08 years old, height 1.75 ± 0.07 m, and weight 78.35 ± 12.72 kg) without lower limb injury or locomotion deficits were selected to participate in this study, which was approved by the Ethics Committee of UFES (Protocol number: 47024214.5.0000.5060). All subjects (S1 to S6) provided written informed consent prior to the data collection, and a background of this study was explained. Figure 3.6 shows the electrode locations and experimental setup.

In this experiment, the equipment BrainNet BNT 36 (EMSA, Brazil) with 20 EEG channels was used to acquire all the signals. A cap with 64 electrodes (Ag-AgCl) was used to acquire the brain activity around the primary and supplementary motor cortex, at locations FC3, FC1, FCz, FC2, FC4, C5, C3, C1, Cz, C2, C4, C6, CP5, CP3, CP1, CPz, CP2, CP4, CP6 and Pz, according to the international 10-20 system. The reference electrodes were located on earlobes A1 and A2, and the ground electrode was located between the eyebrows. These electrode lo-

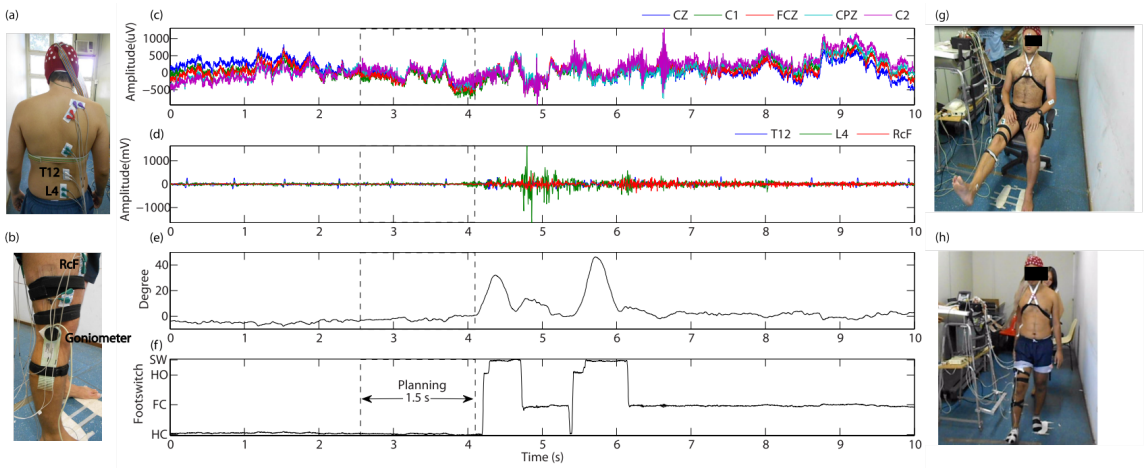

Figure 3.6: Representation of the experimental setup, used at the protocol. Some EEG and sEMG channels acquired during two-steps are displayed. a) sEMG electrodes placed on the erector spinae muscle at five levels; b) sEMG electrodes and goniometer sensor placed on the right leg; c) EEG signals; d) sEMG signals; e) knee angle; f) signal related to the foot contacts on the floor during two-steps; g) knee extension; h) walking.

cations have been previously used by other authors to study motor intention (JIANG et al., 2015; XU et al., 2014; HORTAL et al., 2016; SBURLEA; MONTESANO; MINGUEZ, 2015). In addition, electromyography signals (sEMG) were acquired on the following muscles: rectus femoris (RCF), vastus lateralis (VL), biceps femoris (BF), semitendinosus (S) and gastrocnemius (G), erector spinae (ES) at levels C7, T3, T7, T12 and L4. During the protocol, gel was used to improve the skin impedance. Figures 3.6a-3.6b show the location of the sEMG sensors. These sEMG signals were used here as reference to ensure no muscular contractions, during the annotation of time periods related to the gait planning action.

The signal acquisition equipment was setup with a band-pass filter from 0.1 to 100 Hz, notch filter in 60 Hz, and sampling rate of 400 Hz. Additionally, a goniometer and footswitch sensors, located on the right leg, were used to measure the knee angle and gait phase, respectively, using a frequency range from DC to 70 Hz. A software was developed to guide the subjects during the experiment through visual and sound cues, and an Arduino board was used to generate a synchronous signal. The acquisition system was attached to a mobile platform in order to follow the subjects during walking.

Several experiments were conducted to evaluate the capability of the spatial filters to reduce common interferences, but preserving the neural information related to gait planning.

Two sessions of sequential and random tasks were proposed to this study. The first session was conducted with 10 sub-sessions, using the following tasks: (1) knee extension (K-E)/knee flexion (K-F)/sit rest (Si-R). The second session was conducted with 6 sub-sessions for the following random tasks: (2) K-F, K-E, sit-to-stand, stand-to-sit, stand rest, sit rest and walking (two normal steps). All sub-sessions had 10 repetitions of 7 s of motor task, with 3 min of rest. Furthermore, all subjects were asked to avoid using their arms as extra support, and talk during each session. Figures 3.6g-3.6h show a subject executing the knee extension and two-steps, respectively.

To study the gait planning/stopping, the classes gait planning, gait stopping and rest from stand position were considered. The walking task had 4 repetitions per session, 7 s of duration. Figures 3.6d-3.6f show, respectively, acquired signals from sEMG (on the location T12, L4 and RCF), goniometer and footswitch. These signals were used to annotate manually the gait onset

and offset, in order to locate segments (free of muscular contractions) associated with actions of gait planning (-1.5 s to 0 s, from gait onset or pre-swing) and gait stopping (-0.5 s to +1.0 s, from gait offset). The sEMG channels were used to guarantee all time periods annotated free of myoelectric activities related to walk. The rest state was selected from 0 s to +2.0 s, using the stimulus signal as reference during stand rest position.

3.2 Conclusion

This chapter introduced methods that may be used in BCIs to preserve the neural information of EEG, and to improve the training stage of BCIs. For instance, a pre-processing method based on similarity analysis was introduced here, looking those neighbor electrodes suitable to reduce common interferences, while preserving the useful information. This way, two methods were proposed to select neighbor electrodes during the EEG pre-processing. Additionally, unsupervised methods were proposed here for the training stage of BCIs to increase the robustness to uncertainty patterns, such as patterns related to motor planning actions.

The experimental protocol, and the statistical analysis used to evaluate the proposed methods in designed BCIs to recognize gait planning/stopping were presented.

Chapter 4

Results and Discussion

4.1 Model fitting based on SSVEP

SSVEP patterns that arise in reaction to flickering stimuli can be detected mainly around occipital location from EEG channels (COTRINA et al., 2017). Thus, EEG signals of ten subjects (SV1 to SV10), obtained during SSVEP stimuli (see Section A.6), were analyzed with electrodes firstly located on the occipital region (O1, O2 and Oz), in order to fit the proposed model (see Equation (3.10)).

After analyzing the data, a unique model, described by $VD_{ij} = \exp[-5\rho_{cij}]$, was obtained with window sizes of 100 ms. Thus, for pre-processing signals of target electrodes, this adjusted model was indistinctly applied for the Ad LAR, Ad WAR, and Ad WARb filters. Here, Ad LAR and Ad WAR nomenclatures refer to combinations of the proposed model in Equation (3.10) with the proposed method based on CCC and standard deviation to select neighbor electrodes, while Ad WARb is used when the neighbor selection is based on Zscore.

4.1.1 Analysis of SSVEP components preservation

EEG signals of 35 healthy subjects (from SV1 to SV35), obtained from the SSVEP database (see Section A.6), were used to study the preservation of the main components related to each

stimulus, after applying Ad WAR and WAR filters on O1, O2 and Oz locations.

Figures 4.1a-4.1b show that the frequencies related to the stimuli were significantly ($p < 0.0001$) less attenuated when using the adaptive filter. Figure 4.1a shows the mean values of the main components, where both methods presented differences around -4 dB throughout all stimuli. Figure 4.1b shows that Ad WAR caused the highest attenuation on O2, but this was still a significant improvement over the WAR filter. Figures 4.1c-4.1d show that similar values of signal to noise ratio (SNR) between main components and their correspondent neighbor frequencies were obtained for both methods, with better values on Oz. It is worth mentioning that the WAR filter improved significantly ($p < 0.0001$) the SNR, as shown in Figures 4.1c-4.1d. Figures 4.1e-4.1f show that highest values of coherence were obtained throughout all stimuli on Oz, using the WAR filter. However, these values on channels O1 and O2 were slightly improved using the adaptive filter.

Another comparison was carried out between both WAR and Ad WAR filters, for the BCI

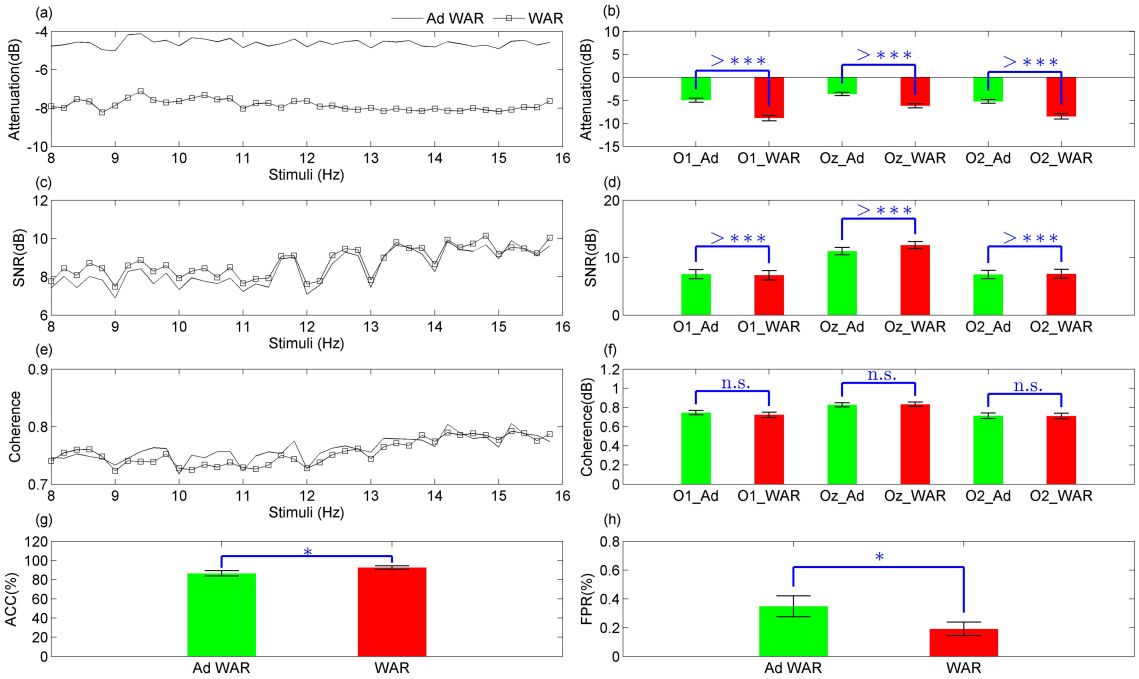

Figure 4.1: Comparison between the spatial filters WAR (red) and Ad WAR (green) through 40 stimuli using SSVEP. p values representation ($^{n.s.}$non-significant, $p > 0.05$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{>***}p < 0.0001$). a) and b) show the stimuli attenuation; c) and d) show the signal to noise ratio (SNR) between main components and neighbor frequencies, using the power spectrum of the pre-processing signals using FFT; e) and f) show the coherence analysis of the main components; g) accuracy of the BCI based on CCA to recognize 40-targets of SSVEP; h) false positive rate of the BCI to recognize SSVEP targets.

based on CCA, which was applied, in this case to recognize 40 targets using the locations O1, O2, Oz, PO3, PO4, PO5, PO6, POz, and Pz, also used in (CHEN et al., 2015). The WAR filter presented the highest values of accuracy ($ACC = 92.56 \pm 1.83\%$) and false positive rate ($FPR = 0.19 \pm 0.04\%$) with a significant difference ($p < 0.05$), while Ad WAR also showed good values of ACC ($86.44 \pm 2.81\%$) and FPR ($0.34 \pm 0.07\%$).

In general, the Ad WAR filter decreases significantly ($p < 0.0001$) the attenuation of SSVEP components, with similar values of coherence in comparison to the WAR filter.

4.2 BCIs for gait intention recognition

Six subjects (S1 to S6) were analyzed through electrodes located on the primary and supplementary motor areas, using three states: rest from stand position (36 segments, 2 s of duration), gait planning (24 segments, 1.5 s of duration), and gait stopping (24 segments, 1.5 s of duration).

Here, the statistical analysis was divided in two stages. Firstly, the proposed methods were analyzed in BCIs to recognize gait planning, using both classes: rest from stand position, and gait planning. Then, the best BCIs to recognize gait planning were used in the second stage of evaluation to recognize both gait planning and gait stopping. It is worth noting that all classifiers used in the designed BCIs, such as SVM and RDA, were adjusted here by the proposed method in Section 3.1.3.

4.2.1 Gait planning recognition

In order to select the best proposed method for feature selection (see Sections 3.1.2), BCIs based on the traditional spatial filter WAR and SVM linear were used in combination with the two proposed methods. Here, the supervised method that combines both RE and ACC was carried out with the classifier LDA. Then, covariance matrices (full, diagonal and spherical) for LDA were evaluated by the proposed method in Section 3.1.3, to obtain a setup of LDA that improves the feature selection.

The performance of both BCIs was compared with other BCI using the full feature set. Figure 4.2 shows a similar performance without significant differences for both BCIs using the proposed methods for feature selection. In addition, the proposed methods were capable of selecting those features that reproduce a similar performance in comparison to the full feature set, such as shown in Figure 4.2. A summary of ranges of selected features for each proposed method is presented in Table 4.1, showing that feature sets with lowest sizes were obtained from the method based on LDA. Notice that a large computational process is executed by methods based on classifiers during the feature selection. Although LDA has shown a low computational cost in comparison to other classifiers, its performance was very delayed. In contrast, the unsupervised method presented low computational cost, providing similar accuracy to other BCIs. Table 4.2 shows a summary of the BCI performance for six subjects, using the unsupervised method. This BCI presented $ACC > 67\%$, $TPR > 66\%$, $FPR < 33\%$ and $F1 > 56\%$ for all analyzed subjects.

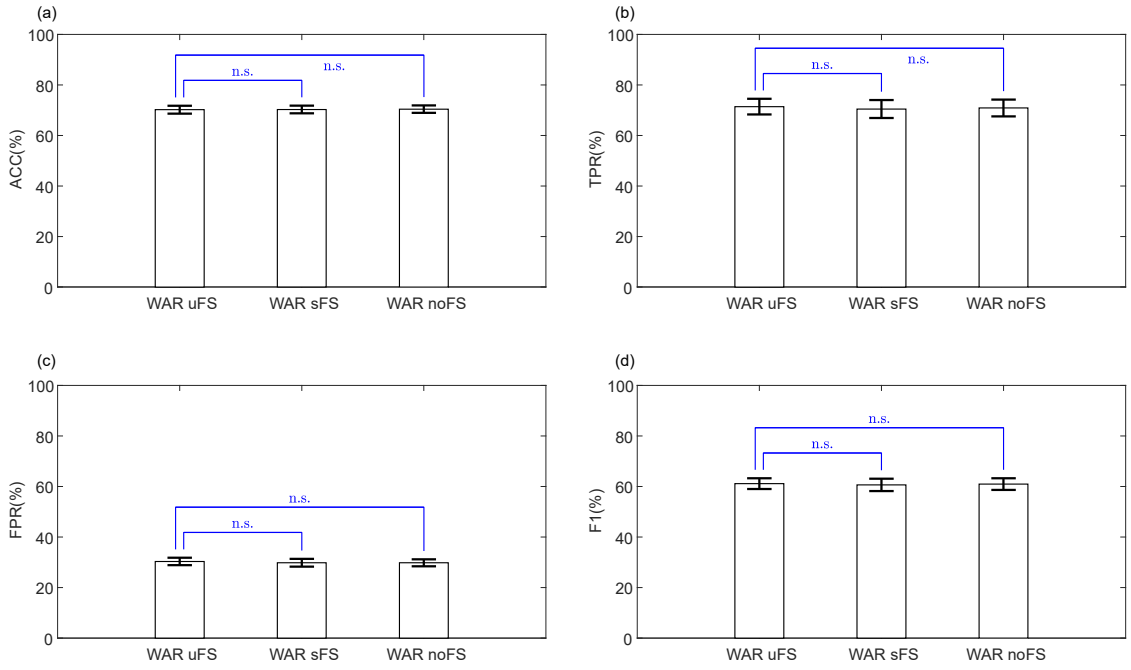

Figure 4.2: Representation of gait planning recognition through BCIs based on the traditional spatial filter WAR and SVM. uFS, sFS, and noFS means unsupervised feature selection, supervised feature selection, and non-feature selection, respectively. Here, both classes gait planning, and rest from stand position were considered. p values representation ($^{n.s}$non-significant, $p > 0.05$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{>***}p < 0.0001$). a) accuracy; b) true positive rate; c) false positive rate; d) F1 value.

Table 4.1: Output of the BCIs for the stage of feature selection, during the gait planning recognition.

Sub	WAR			WARb		WARc	
	Original Features (Size)	**Selected Features (Size)**	**SVM(C)**	**Selected Features (Size)**	**SVM(C)**	**Selected Features (Size)**	**SVM(C)**
S1	36	26-31	0.05	14-16	0.05	36	0.05
S2	36	21-24	0.01	16-23	0.5	36	0.01
S3	36	24-26	0.01	17-18	0.01	36	0.01
S4	36	20-27	0.01	17-25	0.01	36	0.01
S5	36	23-25	0.01	16-20	0.01	36	0.01
S6	36	20-24	0.01	15-20	0.01	36	0.01

Two classes were analyzed. Sub, subjects; SVM, Support Vector Machine linear; C, the adjusted box constraint for SVM linear; WAR, WARb and WARc representing BCIs based on the following methods for feature selection, respectively: unsupervised with Representation Entropy (RE) and Maximal Information Compression Index (MICI), RE and Linear Discriminant Analysis (LDA), and non-feature selection.

Table 4.2: Performance of the BCI based on WAR, unsupervised feature selection and SVM linear, to recognize gait planning.

Subj	ACC(%)	TPR(%)	FPR(%)	F1(%)
S1	71.89 (6.83)	73.99 (16.57)	28.44 (11.07)	63.62 (8.28)
S2	68.78 (2.41)	67.80 (2.83)	30.71 (2.37)	60.54 (4.04)
S3	67.23 (4.03)	66.75 (4.30)	32.60 (5.89)	56.82 (2.39)
S4	67.58 (12.94)	66.43 (20.77)	31.62 (9.98)	56.90 (16.03)
S5	69.77 (5.52)	73.80 (13.95)	32.33 (4.18)	60.82 (9.95)
S6	75.94 (4.35)	79.67 (18.19)	26.26 (2.68)	68.13 (10.42)

Two classes were analyzed. Some values are presented as Mean (SD), SD, standard deviation; ACC, accuracy; TPR, true positive rate; FPR, false positive rate.

Thus, the unsupervised method for feature selection was adopted here to analyze in different BCIs, the performance of both approaches of spatial filters, in order to preserve the useful information during the gait intention recognition.

A comparison between LAR, Ad LAR, WAR, Ad WAR and Ad WARb filters was carried out in the pre-processing stage of different BCIs for gait planning recognition. Notice that Ad LAR and Ad WAR nomenclatures refer to combinations of the proposed model in Equation (3.10), with the proposed method based on CCC and standard deviation to select neighbor electrodes, while Ad WARb is used when the neighbor selection is based on Zscore. Here, these BCI used unsupervised feature selection plus SVM classifier with linear kernel.

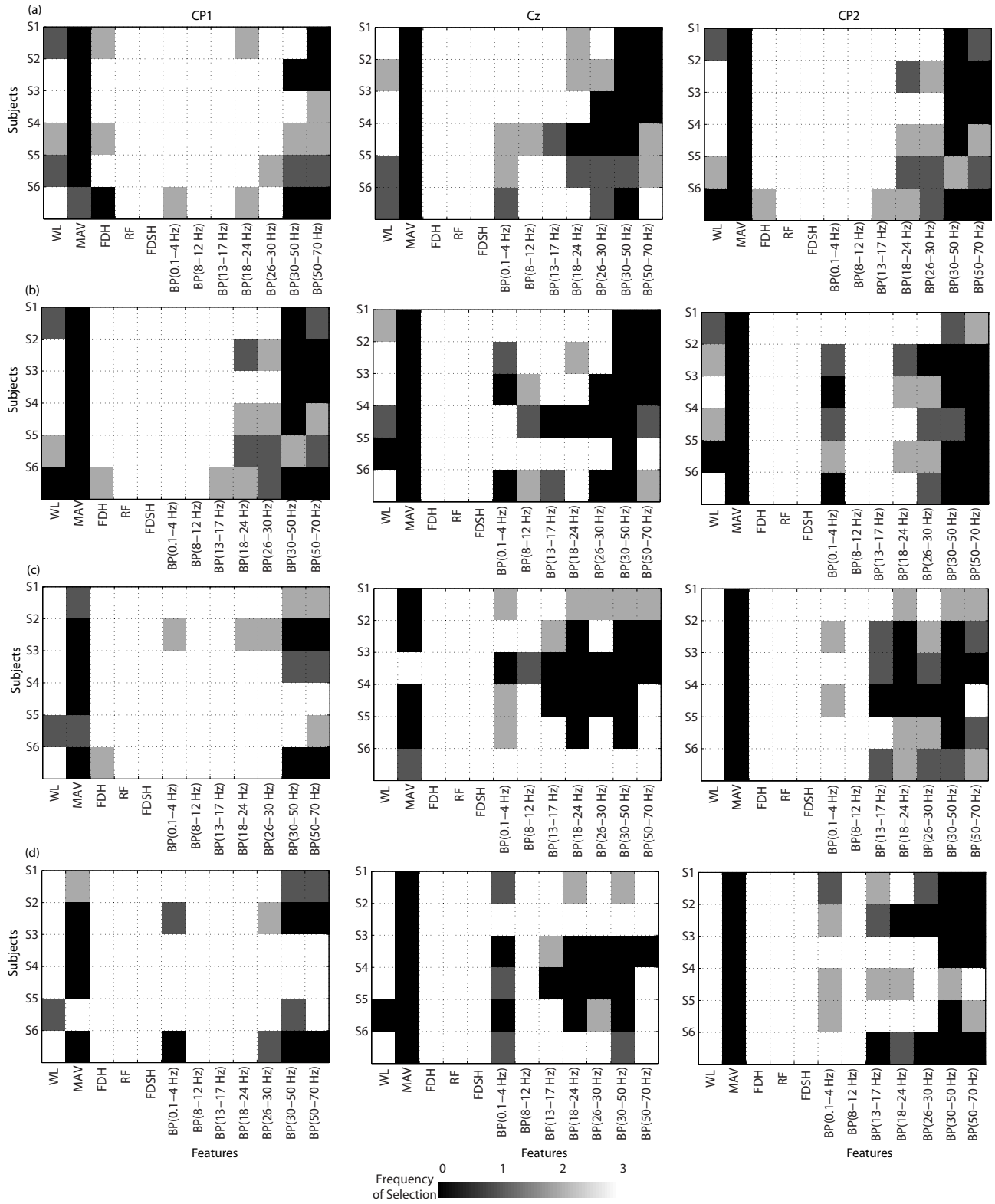

Figure 4.3: Representation of the BCI output during the feature selection, after applying Ad WAR, WAR, Ad LAR and LAR filters. Both classes gait planning, and rest from stand position were considered. The rows present, for the six subjects, the selected features on CP1, Cz and CP2. Rows a) Ad WAR; b) WAR; c) Ad LAR; d) LAR.

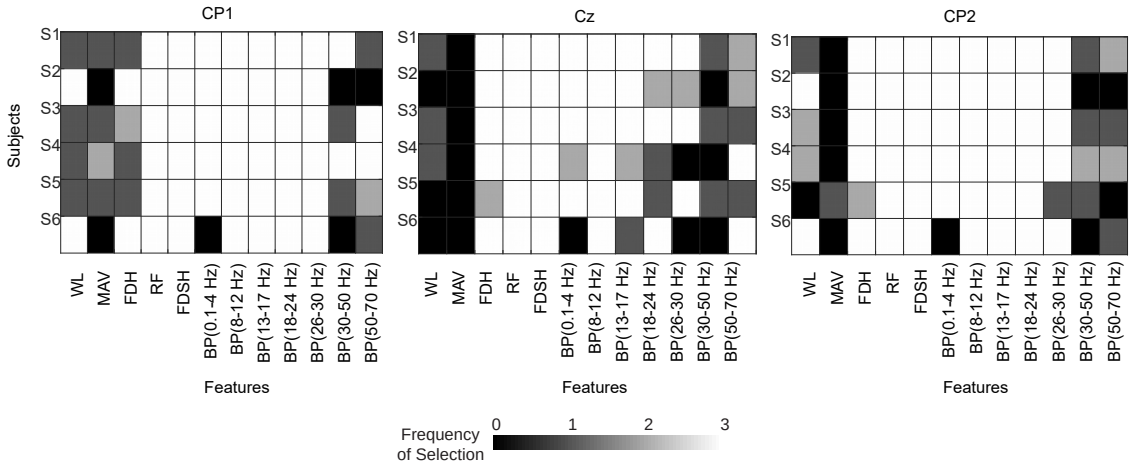

Figure 4.4: Representation of the BCI output during the feature selection, after applying the Ad WARb filter.Both classes gait planning, and rest from stand position were considered. The rows present, for the six subjects, the selected features on CP1, Cz and CP2.

Once obtained the reference-free EEG (RF), the feature vectors were computed and analyzed through an unsupervised method for feature selection. Figures 4.3a (first row), 4.3b (second row), 4.3c (third row), and 4.3d (fourth row) show the BCIs outputs after applying the feature selection on six subjects. Furthermore, Figure 4.4 shows the selected features after applying the Ad WARb filter. For all subjects, when applying these aforementioned filters, a good contribution was obtained on all channels using features in the time domain, such as FD, RF, and FDSH. However, the feature MAV was little considered. In addition to MAV, BP (0.1-4 Hz), computed on the locations Cz and CP2, was also little considered on almost all subjects, using WAR and LAR filters. It is worth noting that, when applying both Ad WAR and Ad WARb, BP (0.1-4 Hz) and other features were more selected on almost all subjects. Curiously, features from CP1 were highly selected, followed by features from CP2.

Figures 4.5 and 4.6 show a summary of the BCIs performance for the six subjects during the gait planning recognition, applying Ad LAR, Ad WAR, Ad WARb, LAR, WAR filters, and the other RK, CSP, and FBCSP methods. Figure 4.6 shows two parameters termed latency and continuous recognition, which were used to analyze the performance of the BCIs throughout all segments studied. Here, each gait intention was considered as planning command (or not failed intention), for similar patterns or states recognized throughout 88 ms (35 samples, at 400 Hz of sampling rate). This way, the latency was defined as the minimum delay that the motion intention was recognized as planning command. For each recognized segment, true epochs

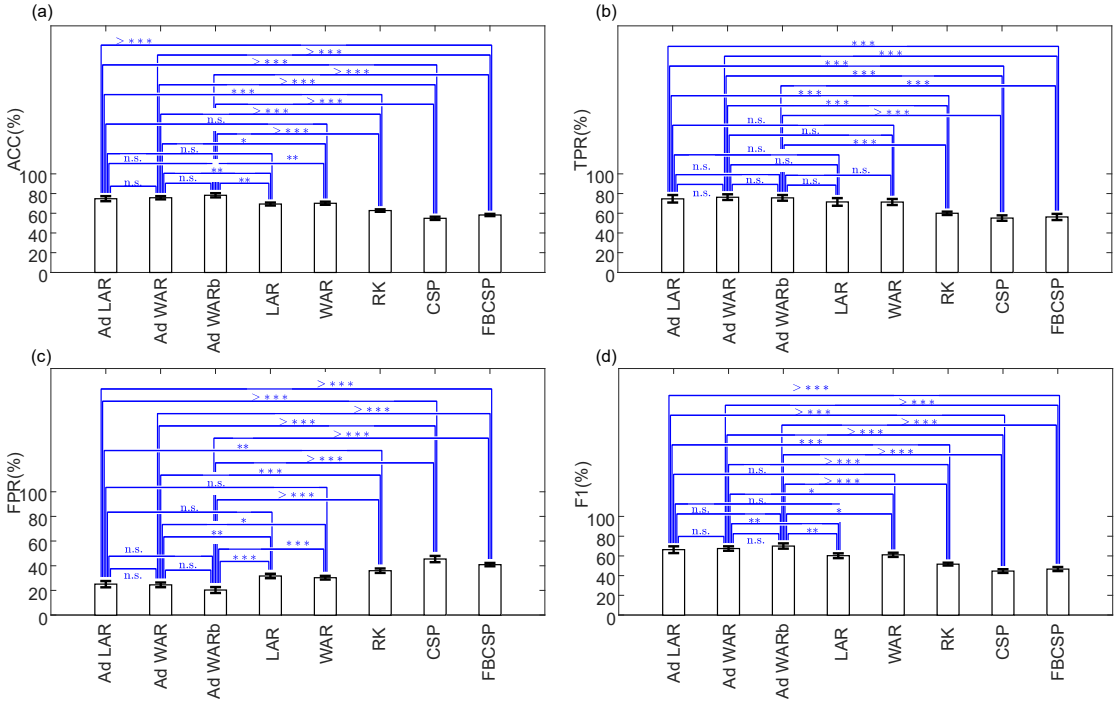

Figure 4.5: Performance of the BCI during the gait planning recognition, applying Ad LAR, Ad WAR, Ad WARb, LAR, WAR, RK, CSP and FBCSP filters. p values representation ($^{n.s}$non-significant, $p > 0.05$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{>***}p < 0.0001$). a) accuracy; b) true positive rate; c) false positive rate; d) F1 value.

that achieved pattern recognition above 88 ms were selected. Thus, for each gait planning, the minimum and maximum intervals of patterns recognized continuously were respectively computed on the epoch set, such as the median and maximum values. It is possible to see in Figures 4.5 and 4.6 that Ad LAR, Ad WAR, Ad WARb presented the best performance, showing mean values of $ACC \geq 74.79\%$, $TPR \geq 74.66\%$, $FPR \leq 25.06\%$, $F1 \geq 66.27\%$, latency of movement anticipation $\leq$ -1084.90 ms, and continuous recognition of planning patterns from 566.70 to 781.0 ms.

Tables 4.3, 4.4 and 4.5 show a summary of BCIs based on Ad WAR, Ad LAR, and Ad WARb, respectively, which presented the best performance for the six subjects during gait planning recognition. These filters Ad LAR, Ad WAR, and Ad WARb improved significantly the performance of the well-known methods RK, CSP and FBCSP (BLANKERTZ et al., 2008; ANG et al., 2012a; BARACHANT et al., 2013). Furthermore, Ad LAR improved slightly the performance of LAR. Additionally, both Ad WAR and Ad WARb presented higher values of ACC, FPR and F1 than LAR and WAR, with a significant difference. For instance, the BCIs based on the adap-

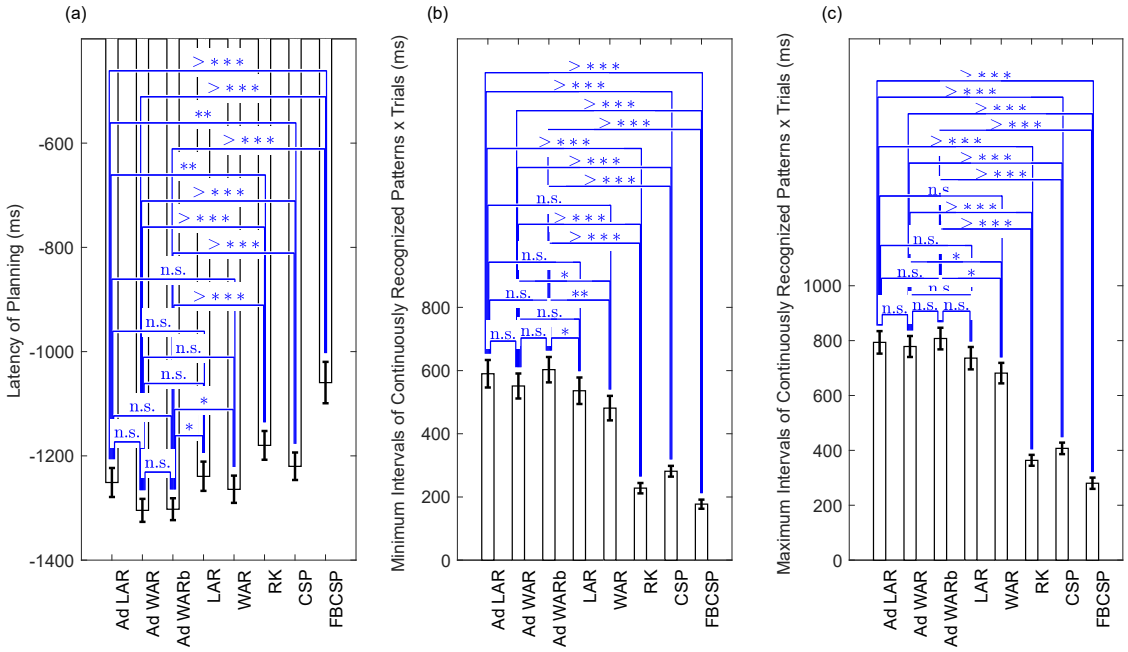

Figure 4.6: Performance in the time domain of the BCI during gait planning recognition, applying Ad LAR, Ad WAR, Ad WARb, LAR, WAR, RK, CSP and FBCSP filters. p values representation ($^{n.s}$non-significant, $p > 0.05$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{>***}p < 0.0001$). a) latency to recognize gait planning; b) minimum interval of continuously recognized patterns during gait planning; c) maximum interval of continuously recognized patterns during gait planning.

Table 4.3: Performance of the BCI based on Ad WAR to recognize gait planning.

Subj	**SVM (C)**	**ACC(%)**	**TPR(%)**	**FPR(%)**	**F1(%)**	**Latency (ms)**
S1	0.05	79.40 (3.51)	79.67 (12.57)	20.14 (8.92)	72.09 (3.28)	−1228.50 (320.50)
S2	0.01	82.50 (7.20)	83.02 (12.26)	17.93 (4.53)	76.72 (11.03)	−1368.60 (106.80)
S3	0.01	73.53 (7.31)	71.47 (8.56)	25.38 (8.93)	63.75 (7.00)	−1315.00 (211.30)
S4	0.01	74.54 (4.80)	74.36 (19.95)	25.14 (2.56)	64.51 (10.45)	−1318.20 (266.50)
S5	0.1	67.41 (7.89)	72.52 (17.59)	35.27 (6.61)	58.55 (12.90)	−1297.70 (241.30)
S6	0.01	76.81 (2.94)	76.86 (5.37)	23.04 (5.22)	68.88 (3.36)	−1304.10 (154.40)

Two classes were analyzed. Some values are presented as Mean (SD); SD, standard deviation; SVM, Support Vector Machine; C, box constraint; ACC, accuracy; TPR, true positive rate; FPR, false positive rate; Latency, delay in recognizing gait planning.

tive filters, and classifiers SVM or RDA, achieved accuracy higher than 73% for all subjects, with the best performance obtained on subjects S1 ($TPR = 83.58 \pm 10.81\%$, $FPR = 15.66 \pm 7.49\%$), S2 ($TPR = 82.83 \pm 6.56\%$, $FPR = 18.37 \pm 3.57\%$), S4 ($TPR = 75.24 \pm 9.92\%$, $FPR = 16.37 \pm 1.89\%$), S5 ($TPR = 83.56 \pm 13.65\%$, $FPR = 20.08 \pm 6.79\%$) and S6 ($TPR = 81.71 \pm 18.68\%$, $FPR = 12.19 \pm 5.25\%$). In contrast, LAR and WAR filters (MCFARLAND et al., 1997) showed accuracy higher than 67.23%, with the best performance obtained for subjects

Table 4.4: Performance of the BCI based on Ad LAR to recognize gait planning.

Subj	**SVM (C)**	**ACC(%)**	**TPR(%)**	**FPR(%)**	**F1(%)**	**Latency (ms)**
S1	0.1	75.02 (8.37)	75.92 (10.78)	24.88 (17.09)	67.46 (7.24)	−1257.50 (295.90)
S2	0.01	82.03 (4.35)	82.83 (6.56)	18.37 (3.57)	76.45 (6.35)	−1325.60 (179.80)
S3	0.1	66.41 (1.23)	66.23 (3.82)	33.53 (0.50)	55.90 (3.42)	−1084.90 (422.20)
S4	0.05	60.67 (12.55)	57.02 (28.16)	37.18 (6.44)	47.28 (19.26)	−1122.00 (401.9)
S5	0.01	81.17 (8.50)	83.56 (13.65)	20.08 (6.79)	73.87 (12.79)	−1355.40 (180.80)
S6	0.5	83.49 (2.73)	82.42 (13.03)	16.32 (6.95)	76.66 (5.27)	−1337.50 (162.20)

Two classes were analyzed. Some values are presented as Mean (SD); SD, standard deviation; SVM, Support Vector Machine; C, box constraint; ACC, accuracy; TPR, true positive rate; FPR, false positive rate; Latency, delay in recognizing gait planning.

Table 4.5: Performance of the BCI based on Ad WARb to recognize gait planning.

Subj	**SVM (C)**	**ACC(%)**	**TPR(%)**	**FPR(%)**	**F1(%)**	**Latency (ms)**
S1	0.05	83.85 (2.31)	83.58 (10.81)	15.66 (7.49)	77.58 (1.26)	-1265.70 (250.00)
S2	0.1	82.59 (5.02)	80.94 (3.47)	16.49 (6.48)	76.73 (6.59)	-1357.30 (153.00)
S3	0.01	73.81 (4.54)	68.14 (6.13)	23.47 (3.62)	62.69 (5.30)	-1222.10 (320.90)
S4	0.1	80.88 (4.57)	75.24 (9.92)	16.37 (1.89)	71.59 (7.23)	-1307.70 (288.80)
S5	0.05	62.59 (5.66)	64.10 (12.66)	37.76 (9.79)	52.33 (5.68)	-1299.60 (156.10)
S6	0.01	86.03 (3.49)	81.71 (18.68)	12.19 (5.25)	78.80 (8.94)	-1330.00 (223.00)

Two classes were analyzed. Some values are presented as Mean (SD); SD, standard deviation; SVM, Support Vector Machine; C, box constraint; ACC, accuracy; TPR, true positive rate; FPR, false positive rate; Latency, delay in recognizing gait planning.

S1 ($TPR = 73.99 \pm 16.57\%,\ FPR = 28.39 \pm 11.07\%$), S5 ($TPR = 77.11 \pm 16.56\%,\ FPR = 36.75 \pm 3.14\%$), and S6 ($TPR = 79.67 \pm 18.19\%,\ FPR = 26.25 \pm 2.68\%$).

A comparison was carried out with BCIs based on RK (BARACHANT et al., 2013), FBCSP (ANG et al., 2012a), CSP (BLANKERTZ et al., 2008; MÜLLER-GERKING; PFURTSCHELLER; FLYVBJERG, 1999), and, which have been successfully used to recognize imagery motor tasks during on-line processing. Here, the proposed BCIs using Ad LAR, Ad WAR and AdWARb significantly improved their performance in comparison to the BCIs based on RK, CSP, and FBCSP.

It is worth mentioning that the BCIs based on RE (for feature selection) and RK (for feature extraction) presented the best performance, in accordance with the hypothesis that unsupervised methods may be more appropriated to analyze patterns with high uncertainty such as gait planning.

Figure 4.7, and Table 4.6 show the output of the BCIs at the stages of feature extraction (for CSP, FBCSP, and RK) and feature selection (for RE). For the BCIs based on RE, feature

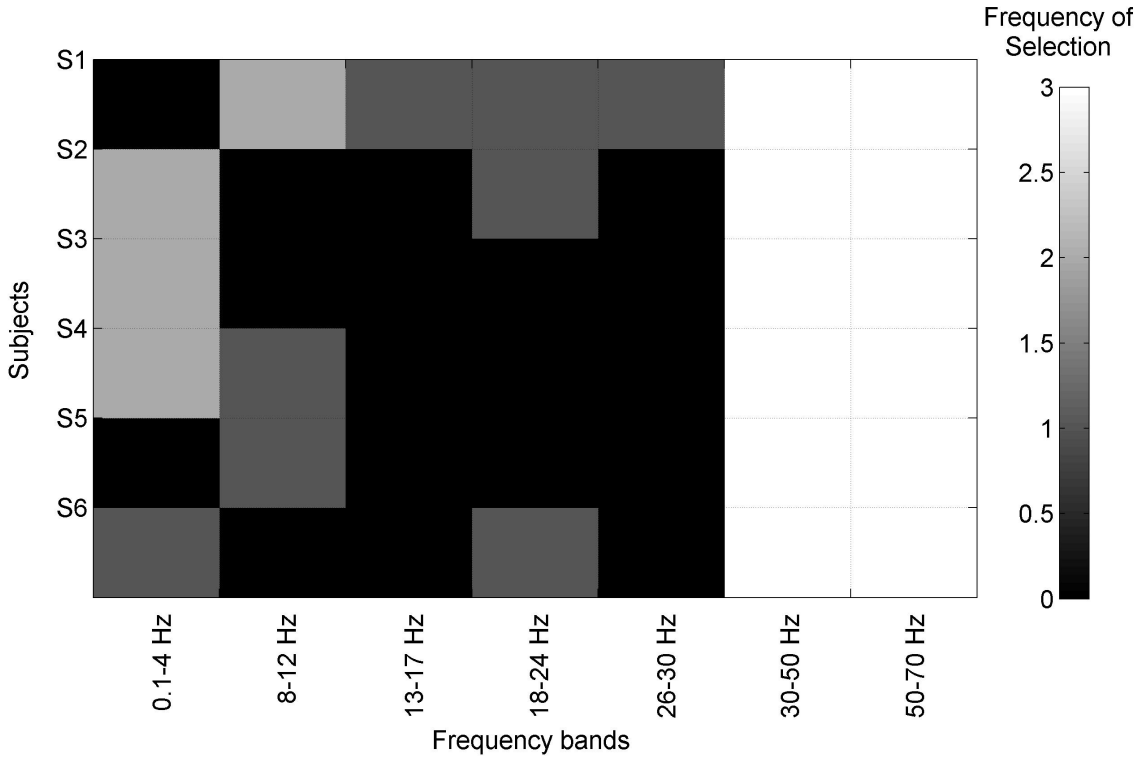

Figure 4.7: Representation of the BCIs outputs on the feature extraction stage applying Filter-Bank Common Spatial Filter. Two classes were analyzed. The rows represent, for the six subjects, selected features on FC1, FC2, C3, C4, Cz, CP1, and CP2.

Table 4.6: Output of the BCIs for the stage of feature extraction and selection, during gait planning recognition.

Sub	**FFS**	**AdLAR**	**AdWAR**	**AdWARb**		**CSP**		**FBCSP**		**RK**
	(Size)	**SF (Size)**	**SF (Size)**	**SF (Size)**	m	**SF (size)**	m	k	**SF (size)**	**SF (size)**
S1	36	23-34	24-29	25-34	4	8	3-4	12	24-32	36
S2	36	20-25	24-26	26	4	8	3-4	12	16-24	36
S3	36	21-22	27-28	24-34	3-4	6-8	4	12	16-32	36
S4	36	24-25	20-28	23-31	4	8	4	12	16-24	36
S5	36	26-30	21-32	21-31	3-4	6-8	4	12	16-24	36
S6	36	27-32	20-24	22-24	3-4	6-8	3-4	12	16-34	36

Two classes were analyzed. FFS, full feature set; SF, selected features; Ad LAR, Adaptive Local Average Reference; Ad WAR, Adaptive Weighted Average Reference; RE, Representation Entropy; CSP, Common Spatial Pattern; FBCSP, Filter-Bank Common Spatial Pattern; m, the first and last m rows of the projection matrix; k, the first best individual features; RK, Riemannian Kernel.

vectors were selected, with sizes from 20 to 34. Moreover, for these BCIs based on RE, good accuracy ($ACC \geq 75\%$) was obtained on the subjects S1, S2, and S6. Additionally, for CSP and FBCSP, a total of 8 channels was adopted to recognize gait planning. For CSP and FBCSP, m values (first and last rows of the projection matrix) were 3 and 4 for almost all subjects.

Figure 4.7 shows that gamma bands (30-50 Hz, and 50-70 Hz) were highly considered for all subjects, using FBCSP.

Here, the BCI based on RK showed on all subjects better values of ACC ($62.76 \pm 1.28\%$), TPR ($59.95 \pm 1.63\%$), FPR ($36.0 \pm 1.79\%$), F1 ($51.66 \pm 1.47\%$) in comparison to CSP and FBCSP, which agrees with (BARACHANT et al., 2013). The best performance of the BCI based on RK was obtained on subjects S1 ($TPR = 62.11 \pm 3.81\%,\ FPR = 31.26 \pm 9.95\%$), S2 ($TPR = 66.38 \pm 2.95\%,\ FPR = 37.21 \pm 7.09\%$) and S3 ($TPR = 61.36 \pm 5.1\%,\ FPR = 36.55 \pm 5.30\%$). It can be observed in Tables 4.3, 4.4, and 4.5 that the BCIs based on Ad WAR, Ad LAR and Ad WARb presented high performance on subjects S1 and S2. Thus, the virtual distance defined in Equation (3.10), combined with results from application of Equations (3.11) and (3.12), may be suitable to preserve the neural information related to gait intention. In contrast, both methods did not present an equivalent performance on subjects S3-S6, which may be related to the first stage of the proposed method. Alternatively, for Ad LAR, Ad WAR and Ad WARb, good performance was obtained for subjects S1, S2, S4, S5, and S6. For Ad LAR, Ad WAR, and Ad WARb filters, it is possible to observe that different adjusts of C were obtained among subjects, with small values (0.01, 0.05 and 0.1) fixed in the last k-fold for almost all subjects. Figure 4.6a, and Tables 4.3 and 4.5, also show that the Ad WAR and Ad WAR filters, significantly improved the latency (from -1222.10 to -1368.60 ms) to recognize gait planning. In addition, Figures 4.5 and 4.6 show that Ad WAR and Ad WARb slightly improved the performance of Ad LAR, however, a non-significant difference was obtained between them from the analyzed parameters. Figure 4.5, and Table 4.5 show that the BCI based on Ad WARb filter improved the recognition of gait planning on almost all subjects, achieving promising results ($ACC > 80\%$) for 4 subjects. It is worth mentioning that the Ad WARb filter is based on Zscore analysis to reject, with low computational cost, undesirable neighbor electrodes during EEG processing. Moreover, this method is more easy to implement than the other approach developed for Ad WAR.

In order to obtain a BCI with good accuracy and speed, classifiers such as SVM linear, LDA, and RDA were evaluated using the Ad WARb filter in the pre-processing step. Figure 4.8 shows that classifiers of low computational cost, such as RDA and LDA ,may achieve similar

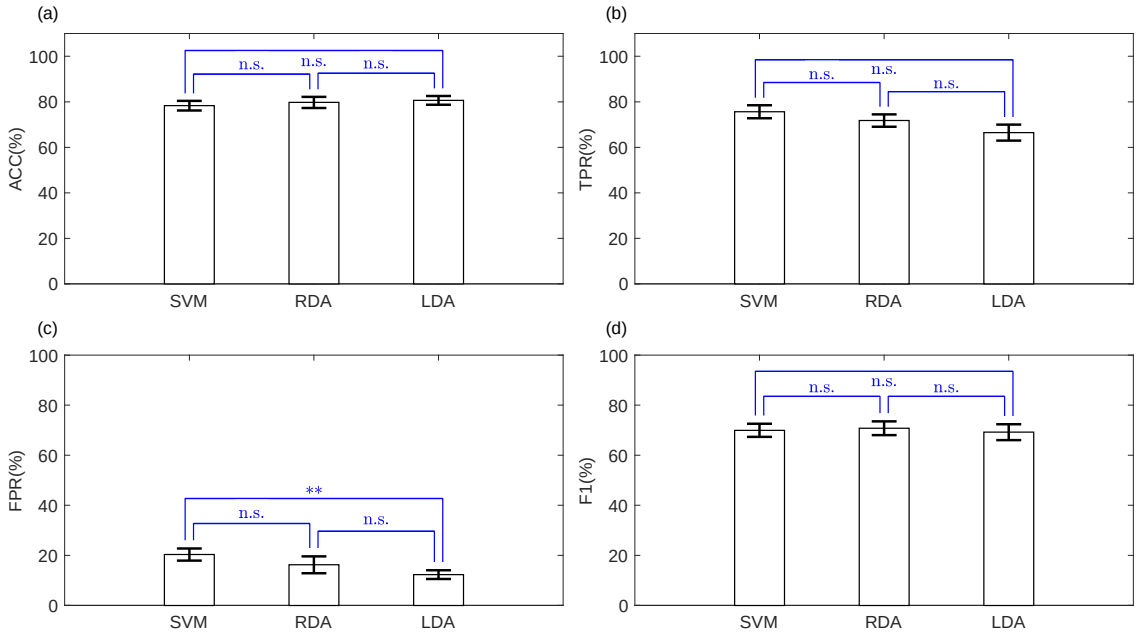

Figure 4.8: Performance of BCIs based on the Ad WARb filter for classifiers SVM, RDA and LDA during the recognition of both gait planning and rest state. p values representation ($^{n.s}$non-significant, $p > 0.05$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{>***}p < 0.0001$). a) accuracy; b) true positive rate; c) false positive rate; d) F1 value.

Table 4.7: Performance of the BCI based on Ad WARb and RDA to recognize gait planning.

Sub	λ	γ	**ACC**(%)	**TPR**(%)	**FPR**(%)	**F1**(%)
S1	0.6	0.2	86.45 (2.45)	79.08 (16.26)	9.34 (5.78)	79.26 (4.99)
S2	0.5	0.7	82.18 (4.71)	76.61 (1.89)	14.74 (6.55)	75.36 (6.07)
S3	0.9	0.8	76.39 (6.82)	58.46 (3.95)	15.10 (7.87)	61.87 (7.29)
S4	1.0	0.1	83.49 (4.03)	65.65 (10.67)	7.88 (4.15)	71.65 (8.07)
S5	0.8	0	60.91 (7.70)	68.93 (7.54)	42.59 (14.44)	53.40 (3.66)
S6	0.6	0.6	88.95 (1.44)	81.90 (8.84)	7.78 (6.14)	83.20 (0.56)

Here, both classes gait planning, and rest from stand position were considered. Some values are presented as Mean (SD); SD, standard deviation; λ and γ, adjusted parameters for RDA; ACC, accuracy; TPR, true positive rate; FPR, false positive rate.

performance to SVM without significant difference ($p < 0.05$), being RDA the classifier more suitable to recognize both rest state and gait planning. Table 4.7 shows a summary of the BCI based on RDA to recognize gait planning. Although SVM and RDA do not present significant difference, SVM is more appropriate than RDA to recognize motor planning, such as shown in Tables 4.5 and 4.7. Nevertheless, RDA may be used to reduce false positives.

The next section aims to evaluate the aforementioned classifiers in BCI based on Ad WARb to

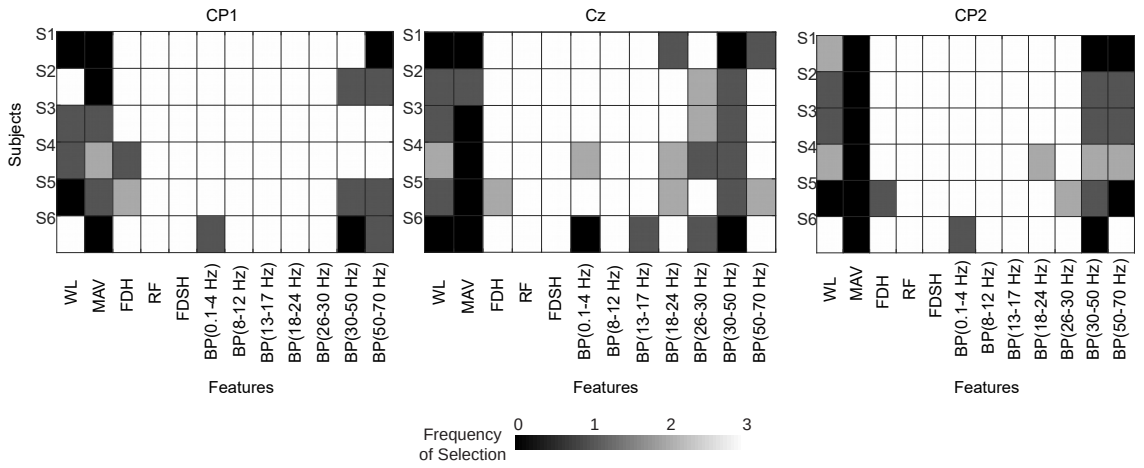

Figure 4.9: BCI output obtained from a cluster of three classes (rest, gait planning, and gait stopping) during the feature selection, after applying the Ad WARb filter. The rows present, for the six subjects, the selected features on CP1, Cz and CP2.

recognize gait planning/stopping.

4.2.2 Gait planning/stopping recognition

Six subjects were analyzed to evaluate the performance of different BCIs based on Ad WARb, and classifiers such as SVM, LDA, and RDA, in order to recognize the following three classes: rest state, gait planning, and gait stopping.

Figure 4.9 shows the output of the feature selection for BCI based on Ad WARb to recognize the following three classes: rest state, gait planning, and gait stopping. It is possible to observe that features in time domain, such as FDH, RF, and FDSH presented good contribution for all subjects. Additionally, features in frequency domain correspondent to slow cortical potentials, mu and beta bands were highly selected for all subjects.

Figures 4.10 and 4.11 show that similar performance without significant difference may be obtained for all subjects, using three classifiers. Notice that the recognition of classes, such as rest and gait stopping was better than gait planning, as both classes rest state and gait planning are very close.

Regarding SVM and RDA, whose performance is shown in Tables 4.8 and 4.9, respectively, RDA presented promising results, improving the performance for almost all subjects. Specially, good accuracy (ACC $>$76%, TPR$>$68% and FPR$<$16%) was obtained on Subjects S1, S2 and S6,

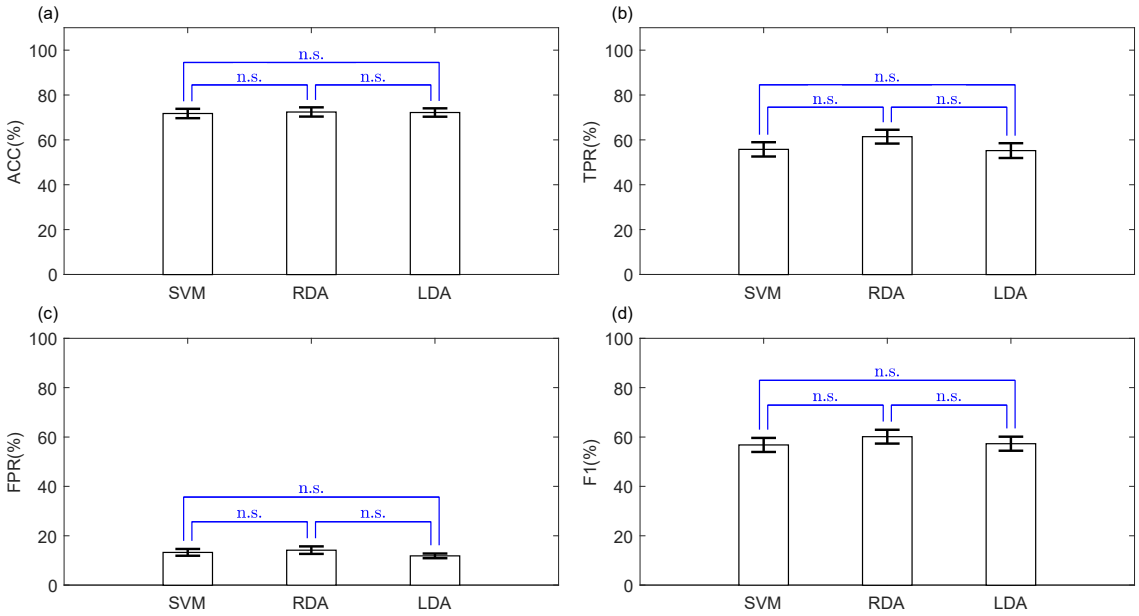

Figure 4.10: Performance of BCIs based on the Ad WARb filter for classifiers SVM, RDA and LDA during the gait planning recognition. Here, three states such as rest state, gait planning, and gait stopping were considered. p values representation ($^{n.s}$non-significant, $p > 0.05$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{>***}p < 0.0001$). a) accuracy; b) true positive rate; c) false positive rate; d) F1 value.

Table 4.8: Performance of the BCI based on the Ad WARb and SVM, for the following three states: rest, gait planning, and gait stopping.

Sub			**Planning**			**Stopping**	
	SF	**C**	**ACC (%)**	**TPR (%)**	**FPR (%)**	**TPR (%)**	**FPR (%)**
S1	25-26	0.05	77.92 (5.71)	61.37 (18.41)	10.72 (5.03)	75.85 (10.10)	8.34 (3.82)
S2	25-34	0.01	76.93 (1.67)	60.94 (7.57)	11.10 (2.97)	76.94 (6.29)	6.54 (2.58)
S3	26-34	0.01	67.26 (1.54)	45.46 (6.31)	14.56 (1.67)	64.89 (9.74)	10.48 (1.21)
S4	22-34	0.01	71.93 (4.18)	57.04 (1.58)	12.91 (2.64)	66.08 (9.52)	13.36 (5.45)
S5	21-30	0.1	56.16 (4.53)	42.63 (9.87)	21.84 (8.31)	66.63 (16.03)	17.45 (7.13)
S6	23-28	0.01	80.40 (2.00)	67.28 (18.10)	8.49 (3.29)	77.34 (0.91)	7.41 (2.33)

Some values are presented as Mean (SD); SD, standard deviation; SF, selected features; C, adjusted box constraint for Support Vector Machine;ACC, accuracy; TPR, true positive rate; FPR, false positive rate.

such as shown in Table 4.9. Notice that RDA was automatically adjusted, searching values very close to LDA setup (λ=1, and γ=0).

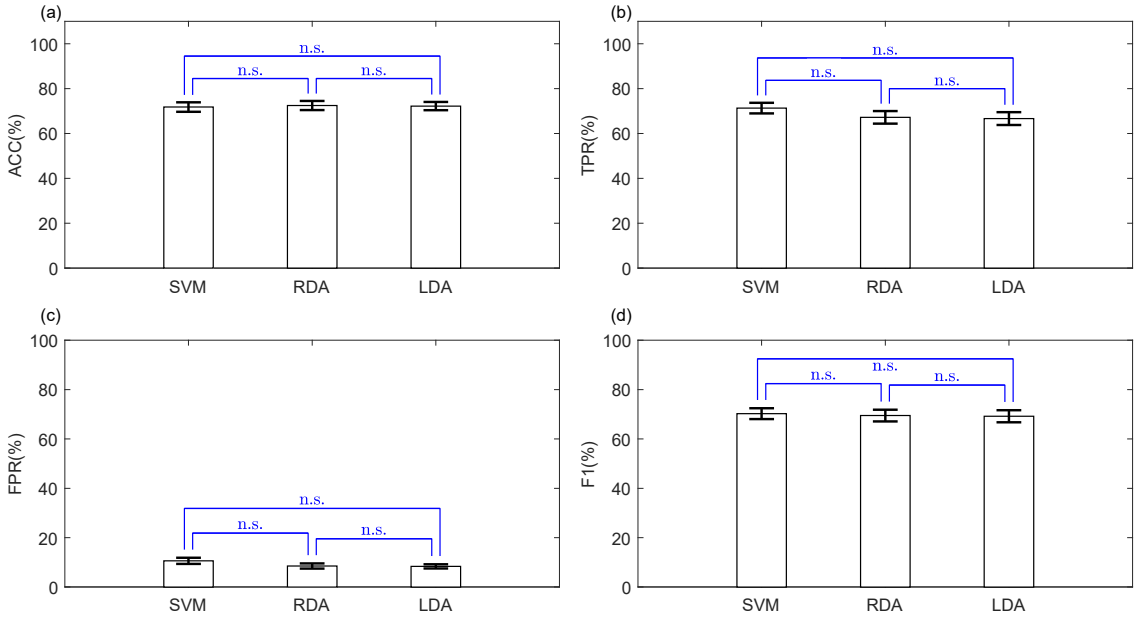

Figure 4.11: Performance of BCIs based on the Ad WARb filter for classifiers SVM, RDA and LDA during the gait stopping recognition. Here, three states such as rest state, gait planning, and gait stopping were considered. p values representation ($^{n.s}$non-significant, $p > 0.05$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{>***}p < 0.0001$). a) accuracy; b) true positive rate; c) false positive rate; d) F1 value.

Table 4.9: Performance of the BCI based on the Ad WARb and RDA, for the following three states: rest, gait planning, and gait stopping.

Sub				**Planning**		**Stopping**		
	SF	λ	γ	**ACC (%)**	**TPR (%)**	**FPR (%)**	**TPR (%)**	**FPR (%)**
S1	25-26	0.9	0.3	79.07 (5.65)	71.82 (11.29)	12.92 (4.98)	68.82 (7.75)	5.37 (3.08)
S2	25-34	0.8	0	76.79 (2.02)	70.01 (1.23)	15.15 (1.51)	73.90 (6.48)	3.78 (2.81)
S3	26-34	0.9	0.5	66.37 (1.33)	42.55 (9.19)	13.53 (2.88)	57.56 (13.68)	8.20 (2.65)
S4	22-34	1	0	73.61 (2.57)	59.00 (4.68)	11.89 (0.88)	58.34 (12.65)	9.14 (4.00)
S5	21-30	0.9	0	58.11 (5.66)	54.37 (11.16)	24.29 (8.45)	66.65 (14.53)	15.23 (3.18)
S6	23-28	0.9	0.4	80.68 (1.00)	70.96 (8.50)	7.14 (2.69)	77.88 (4.34)	9.36 (0.79)

Some values are presented as Mean (SD); SD, standard deviation; FS, Selected features; λ and γ, adjusted parameters for RDA; ACC, accuracy; TPR, true positive rate; FPR, false positive rate.

4.3 Discusion

The best performance of the BCIs was obtained with unsupervised methods, such as RK (BARACHANT et al., 2013) for feature extraction, and the proposed method based on RE for feature selection. Thus, unsupervised methods for feature extraction and feature selection may be more appropriate to analyze patterns of high uncertainty, such as motor planning. Nevertheless, amplitude values throughout the raw EEG during the rest state and gait planning

may be too close. For this reason, the performance of the BCI based on RK may be affected. The continuous recognition throughout the planning interval was improved on almost all subjects using the adaptive filters proposed here, as shown in Figures 4.6b-4.6c. It is worth noting that both filters improved the selection of the feature BP (0.1 to 4 Hz) on all channels, which is in accordance with (JIANG et al., 2015; SBURLEA; MONTESANO; MINGUEZ, 2015; VELU; SA, 2013), as low frequency is relevant to improve the gait planning recognition. Nevertheless, these components may be highly affected by EOG ($\leq$ 0.5 Hz) and movement artifacts ($\leq$ 20 Hz) (JIANG et al., 2015; VELU; SA, 2013; PUTHUSSERYPADY; RATNARAJAH, 2005). In contrast, features such as MAV, BP (30-50 Hz), and BP (50-70 Hz) were rarely selected. However, other EEG studies have reported that the gamma band ($>$ 30 Hz) plays an important role during walking (SEEBER et al., 2015). Additionally, in our work, a high contribution from cortical parietal areas (CP1 and CP2) was obtained, which agrees with (SBURLEA; MONTESANO; MINGUEZ, 2015; LEW et al., 2012).

In the literature, the recognition of self-paced lower-limb movements has been little studied. Some studies have been focused on recognizing gait starting, using amplitude features of low frequency (JIANG et al., 2015; LEW et al., 2012) or spectral features (HORTAL et al., 2016; VELU; SA, 2013). Jiang et al (JIANG et al., 2015) presented a BCI of single channel (Cz) to recognize gait starting using MRCP templates of 1 s in length, from which half a second before the peak negativity of the MRCP is used, reporting TPR and FPR values of 76.9% and 2.93 ± 1.09 per minute, respectively. Hortal et al. (HORTAL et al., 2016) proposed a BCI to recognize gait starting ($TPR \geq 54.8\%$) and stopping ($TPR \geq 56.1\%$), using spectral features from 9 pre-processed channels using a Laplacian filter (PFURTSCHELLER; NEUPER; BERGER, 1994), achieving TPR values of 54.8% and 56.1% to recognize both states, respectively, using SVM with Gaussian kernel. However, these previous works did not focus on decoding pre-movement states (from -1.5 to 0 s before the footswitch or leg angle release), and only one (HORTAL et al., 2016) focused oo recognizing gait stopping. Thus, a direct comparison with these aforementioned works is not easy. Another study proposed the combination of features based on amplitude at low frequency (MRCP from 0.1 to 1 Hz) and spectral information (from 8 to 13 Hz) to recognize gait planning, using a total of 10 channels (SBURLEA;

MONTESANO; MINGUEZ, 2015), and fastICA as pre-processing stage on the BCI, achieving accuracy of 70%. Similarly, they suggest that these combined features can increase the BCI performance.

Although the proposed filters achieved promising results, (ACC $\geq$ 74.79%, TPR $\geq$ 74.66%, FPR $\leq$ 25.06%, F1 $\geq$ 66.27%) to recognize gait planning, these methods only presented similar behavior (or similar performance) for the subjects S1 and S2. In fact, different performances of Ad LAR, Ad WAR, and AdWARb on the same subject may be produced, as this method provides the highest weighted indices to those neighbor electrodes far in information from the target electrode, in order to preserve the neural information. This strategy is sensible to artifacts, thus, it depends highly on the first stage introduced in the method to select the appropriate electrodes, in order to reduce undesirable artifacts. For now, the standard deviation and Zscore analysis were adopted as reference for this purpose. However, future works will be carried out to improve its selectivity.

Chapter 5

Conclusion

Lower-limb robotic exoskeletons may be used to assist patients with difficulty in voluntarily initialize gait. At UFES/Brazil, the exoskeleton ALLOR (Advanced Lower-limb Orthesis for Rehabiliation) was developed, which may be commanded by a BCI, in order to reproduce more natural movements after a voluntary intention, thus, providing neuroplastic changes in the cortical pathways on subjects. For this purpose, oscillations, such as MRCP and ERD/ERS potentials, can be used to recognize pre-movement (motor planning) and post-movement actions on the ongoing EEG activity, during real and imaginary motor tasks. However, patterns related to pre-movement and post-movement actions, which can be used to close the loop control between user and exoskeleton, may be affected by several factors, such as intra-subject reaction time variability, EEG variability, event duration, among others. In addition, the latency and time delay of these events (MRCP and ERD/ERS) can be affected for each subject, depending on his/her reaction time, and intrinsic neural network. Additionally, motor planning begin before the movement onset, thus, the training stage of BCIs may be affected by the intuitive labeling process, as it is not possible to use reference signals, such as goniometer or footswitch, to select those time periods really related to motor planning. Therefore, the gait planning recognition is a challenge due to the high uncertainty of selected patterns.

This Doctoral Thesis was focused on obtaining a BCI based on unsupervised methods of low computational cost, to convey control commands for the exoskeleton ALLOR, in order to assist

patients with difficulty to walk voluntarily.

New unsupervised methods to improve the performance of different BCIs during gait planning/ stopping recognition were addressed. For instance, a new adaptive spatial filter for on-line processing based on the Concordance Correlation Coefficient (CCC) was proposed to preserve the useful information on EEG signals, while rejecting neighbor electrodes around the electrode of interest. Here, two methods for electrode selection were proposed. First, both standard deviation and CCC between target electrodes and their correspondent neighbor electrodes are analyzed on sliding windows to select those neighbors that are highly correlated. Second, Zscore analysis is performed to reject those neighbor electrodes whose amplitude values presented significant difference in relation to other neighbors.

Furthermore, another method that uses the representation entropy and the maximal information compression index was proposed for feature selection, which may be robust to select patterns, as only it depends on cluster distribution. In addition, a statistical analysis was introduced here to adjust, in the training stage of BCIs, regularized classifiers, such as support vector machine (SVM) and regularized discriminant analysis (RDA).

BCIs based on the proposed methods shown promising results in the following two scenario: recognition of two classes, such as rest from stand position and gait planning, and the recognition of the aforementioned classes including gait stopping.

The proposed filters based on similarity analysis improved the feature selection on MRCPs, mu and beta bands, which play an important role during the motor intention recognition. However, although BCIs using our adaptive spatial filters achieved promising results, these methods only presented similar behavior (or similar performance) for some subjects. In fact, different performances of Ad LAR, Ad WAR, and AdWARb filter may be produced on the same subject, as this method provides the highest weighted indices to those neighbor electrodes far in information from the target electrode, in order to preserve the neural information. This strategy is sensible to artifacts, thus, it depends highly on the first stage (Neighbor selection), which is introduced in the method to select the appropriate electrodes, in order to reduce undesirable artifacts. For now, the standard deviation and Zscore analysis were adopted as reference for this purpose. However, future works will be carried out to improve the selectivity of this stage.

Our unsupervised approach for feature selection may be used in multi-classes classification issues. For instance, it has shown good performance in BCIs to recognize gait planning/stopping, although other applications may be explored, such as feature analysis on EEG clusters obtained from motor imagery. In addition, the proposed method to adjust classifiers may be used in other classification problems for different regularized classifiers, such as SVM with kernel Radial Basic Function (or Gaussian), and K-Nearest Neighbors. For example, our lab has experience in other works of pattern recognition, such as myoelectric pattern recognition for upper-limb forearm amputees. On the other hand, experiences reached with powerful tools, such as CSP, FBCSP and RK methods may be used in our lab to recognize on-line motor imagery, applying it to command the lower-limb exoskeleton ALLOR. Additionally, the proposed method may be explored in other applications that use array electrodes, for example, measuring surface myoelectric signals (sEMG) to preserve the useful information.
We presented here the following papers published during the research conducted in this PhD thesis:

1. Denis Delisle-Rodriguez, Ana Cecilia Villa-Parra , Teodiano Bastos-Filho, Alberto López-Delis, Anselmo Frizera-Neto, Sridhar Krishnan, Eduardo Rocon, Adaptive spatial filter based on similarity indices to preserve the neural information on EEG signals during on-line processing, Sensors, 2017, accepted.

2. Denis Delisle Rodriguez, Ana Cecilia Villa Parra, Anselmo Frizera Neto, Sridhar Krishnan and Teodiano Bastos, Optimización de una Interfaz Cerebro-Computador para Reconocer Patrones de Planificación de la Marcha, IBERDISCAP2017.

3. Denis Delisle-Rodriguez, Ana Cecilia Villa-Parra, Anselmo Frizera-Neto, Sri Krishnan, Teodiano Bastos, Automatic Selection of classifiers for gait planning recognition through a BCI, Simpósio Brasileiro de Automação Inteligente (SBAI 2017).

4. Denis Delisle Rodriguez, Ana Cecilia Villa Parra, Alberto López-Delis, Teodiano Bastos, Anselmo Frizera Neto and Eduardo Rocon, Non-supervised Feature Selection: Evaluation in a BCI for Single-Trial Recognition of Gait Preparation/Stop, Converging Clinical and

Engineering Research on Neurorehabilitation II, Springer International Publishing, 2017, 1115-1120.

5. Denis Delisle Rodriguez, Ana Cecilia Villa Parra, Alberto López-Delis, Anselmo Frizera Neto, Eduardo Rocon and Teodiano Bastos, Non-supervised cluster partition and outliers remotion: evaluation in a BCI for gait planning recognition, XXV Brazilian Congress on Biomedical Engineering (CBEB 2016).

5.1 Acknowledgements

we would like to thank CNPq (304192/2016-3), CAPES (88887.095626/2015-01), and FAPES (72982608) from Brazil, and Emerging Leaders in the America's Program (ELAP/from Canada).

5.2 Publications

1. Denis Delisle-Rodriguez, Ana Cecilia Villa-Parra , Teodiano Bastos-Filho, Alberto López-Delis, Anselmo Frizera-Neto, Sridhar Krishnan, Eduardo Rocon, Adaptive spatial filter based on similarity indices to preserve the neural information on EEG signals during on-line processing, Sensors, 2017, Status: accepted.

2. Denis Delisle Rodriguez, Ana Cecilia Villa Parra, Anselmo Frizera Neto, Sridhar Krishnan and Teodiano Bastos, Optimización de una Interfaz Cerebro-Computador para Reconocer Patrones de Planificación de la Marcha, IBERDISCAP2017, Status: accepted.

3. Ana Cecilia Villa-Parra *, Denis Delisle Rodriguez, Jessica Paola Souza Lima, Anselmo Frizera-Neto, Teodiano Freire Bastos,Knee Impedance Modulation to Control an Active Orthosis Using Insole Sensors, Sensors, 2017, Status: pending editor decision.

4. Ana Cecilia Villa Parra, Denis Delisle-Rodriguez, Thomaz Botelho, John Jairo Villarejo Mayor, Alberto López Delis, Ricardo Carelli, Anselmo Frizera Neto, Teodiano Freire Bastos, Control of a Robotic Knee Exoskeleton for Assistance and Rehabilitation Based

on Motion Intention from sEMG, Research on Biomedical Engineering, 2017. Status: in revision

5. Denis Delisle-Rodríguez, Teodiano Freire Bastos, Álvaro Costa García, Enrique Hortal Quesada, Juan C. Moreno, Juan Carlos Alcázar, Guillermo Herrera, Sergio Casco, Antonio del-Ama, Interfaz hombre-máquina, Editores: José María Azorín Poveda, José Luis Pons Rovira, Anselmo Frizera Neto, Thomaz Rodrigues Botelho, Ángel Manuel Gil Agudo, Javier Orlando Roa Romero, Título: Exoesqueletos Robóticos para Rehabilitación y Asistencia de Pacientes con Daño Neurológico. Experiencias y Posibilidades en Iberoamérica, ISBN: 978-84-15413-29-5, 2017, Editorial: Programa Iberoamericano de Ciencia y Tecnología para el Desarrollo (www.cyted.org), Lugar de publicación: España

6. Denis Delisle-Rodriguez, Ana Cecilia Villa-Parra, Anselmo Frizera-Neto, Sri Krishnan, Teodiano Bastos, Automatic Selection of classifiers for gait planning recognition through a BCI, Simpósio Brasileiro de Automação Inteligente (SBAI 2017).

7. Denis Delisle Rodriguez, Ana Cecilia Villa Parra, Alberto López-Delis, Teodiano Bastos, Anselmo Frizera Neto and Eduardo Rocon, Non-supervised Feature Selection: Evaluation in a BCI for Single-Trial Recognition of Gait Preparation/Stop, Converging Clinical and Engineering Research on Neurorehabilitation II, Springer International Publishing, 2017, 1115-1120.

8. Denis Delisle Rodriguez, Ana Cecilia Villa Parra, Alberto López-Delis, Anselmo Frizera Neto, Eduardo Rocon and Teodiano Bastos, Non-supervised cluster partition and outliers remotion: evaluation in a BCI for gait planning recognition, XXV Brazilian Congress on Biomedical Engineering (CBEB 2016), accepted.

9. A.C. Villa-Parra, D. Delisle-Rodríguez, A. López-Delis, T. Freire Bastos, A. Frizera-Neto, Exploración de patrones de EMGs/EEG orientada al control de exoesqueletos de extremidad inferior, Cognitve Area Network, vol 3, n1, 2016, ISSN 2341-4243.

10. Parra, A. C. V. ; Rodriguez, D. D. ; Delis, A. L. ; Bastos Filho, T. F. ; Sagaró, Roberto Zamosa; Frizera Neto, Anselmo . Towards a robotic knee exoskeleton control based on

human motion intention through EEG and sEMG signals. Procedia Manufacturing, v. 1, p. 1-9, 2015.

11. Parra, A. C. V. ; Vazquez, L. B. ; Rodriguez, D. D. ; Sagaró, Roberto Zamosa ; Bastos Filho, T. F. ; Frizera Neto, Anselmo . Design of active orthoses for a robotic gait rehabilitation system. Frontiers of Mechanical Engineering, v. 1, p. 1-13, 2015.

12. Denis Delisle Rodriguez, Ana Cecilia Villa Parra, Alberto López-Delis, Anselmo Frizera Neto, Eduardo Rocon and Teodiano Bastos, Towards a Wearable Robot for Lower Limb Rehabilitation through Human Motion Intention, ICNR 2016, accepted.

13. Lopez-Delis, A.; Delisle-Rodriguez, D.; Villa-Parra, A. C. ; Bastos-Filho, T. . Knee motion pattern classification from trunk muscle based on sEMG signals. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2015, Milan. 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2015. p. 2604-2607.

14. Delisle-Rodriguez, D.; Villa-Parra, A. C.; Bastos, T. ; Frizera-Neto, A. ; López-Delis, A. . An Exploration of the Erector Spinae Muscle for Knee Exoskeleton Control. IFMBE Proceedings. 1ed.: Springer International Publishing, 2015, v. 51, p. 994-997.

15. Delisle, D. ; Parra, A. C. V. ; Teodiano Freire Bastos-Filho ; Frizera Neto, Anselmo ; Delis, A. L. . Protocol to Acquire sEMG and EEG Signals for Control of an Active Knee Orthosis Based on Detection of Movement Intention. In: 1st International Workshop on Assistive Technology, 2015, Vitória. Proceedings of the 1st International Workshop on Assistive Technology, 2015. p. 74-77.

16. Parra, A. C. V. ; Delisle-Rodriguez, D. ; Lotero, F. A ; Valadão, Carlos ; Bastos, T. ; Frizera Neto, Anselmo . Robotic Systems for Gait Rehabilitation. In: Adriano de Oliveira Andrade, Alcimar Barbosa Soares, Alexandre Cardoso, Edgard Afonso Lamounier. (Org.). Tecnologias, Técnicas e Tendências em Engenharia Biomédica. 1ed.: Canal6 Editora, 2014, v. , p. 264-279.

17. Delisle, D.; Parra, A. C. V. ; Garcia, J. F. C. ; Moraes, C. ; Bastos, T. F. ; Frizera Neto, Anselmo ; Delis, A. L. Development of an EEG and sEMG Wireless System for a Robotic Walker. In: Congresso Brasileiro de Engenharia Biomédica (CBEB), 2014, Uberlândia. Anais do Congresso Brasileiro de Engenharia Biomédica (CBEB), 2014. p. 4 pg.

18. Delisle-Rodriguez, Denis; Castillo-Garcia, Javier F.; Bastos-Filho, Teodiano; Frizera-Neto, Anselmo ; Lopez-Delis, Alberto. Using linear discriminant function to detect eyes closing activities through alpha wave. In: 5th IEEE Biosignals and Biorobotics conference (BRC 2014), 2014, Salvador. 5th ISSNIP-IEEE Biosignals and Biorobotics Conference (2014): Biosignals and Robotics for Better and Safer Living (BRC), 2014. p. 1-4.

19. Vazquez, L. B.; Parra, A. C. V.; Delisle, D. ; Sagaró, Roberto Zamosa; Bastos Filho, T. F. ; Frizera Neto, Anselmo . Proposal of an Assisted-Motion System for Gait Rehabilitation. In: International Workshop on Wearable Robotics, 2014, Baiona. Proc. of the International Workshop on Wearable Robotics, 2014. v. 1. p. 1-6.

20. Garcia, J. F. C. ; Delisle, D. ; Pomer-Escher, A. ; Frizera Neto, Anselmo ; Lima, E. R. ; Bastos, T. F. . Neurofeedback Tool to Improve the Onset Delay and Sequential Movements During Motor Tasks. In: Congresso Brasileiro de Engenharia Biomédica (CBEB), 2014, Uberlândia. Anais do Congresso Brasileiro de Engenharia Biomédica (CBEB), 2014. p. 4 pg.

21. Delisle, D.; Garcia, J. F. C.; Bastos Filho, T. F.; Frizera Neto, Anselmo ; Delis, A. L. . Detection of Eyes Closing Activities through Alpha Wave by VariabilityAnalysis. Graz. Proc. of the 6th International Brain-Computer Interface Conference 2014. p. 1-6.

22. Castillo-Garcia, Javier; Parra, A. C. V. ; Bastos, T. ; Longo, Berthil ; Delisle-Rodriguez, D. . EEG-BCI Tecnologies for Neurorehabilitation. In: Adriano de Oliveira Andrade, Alcimar Barbosa Soares, Alexandre Cardoso, Edgard Afonso Lamounier. (Org.). Tecnologias, Técnicas e Tendências em Engenharia Biomédica. 1ed.: Canal6 Editora, 2014, v. , p. 343-364.

23. Castillo-Garcia, Javier ; Cotrina, Anibal ; Benevides, Alessandro ; Delisle-Rodriguez,

Denis ; Longo, Berthil ; Caicedo, Eduardo ; Ferreira, Andre ; Bastos, Teodiano . Adaptive BCI based on software agents. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2014, Chicago. 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2014. p. 5458-5461.

24. Garcia, J. F. C. ; Delisle, D. ; Pomer-Escher, A. ; Caicedo, E. ; Frizera Neto, Anselmo ; Lima, E. R. ; Bastos, T. F. . Propuesta de una ICC como Herramienta de Neuro-Realimentación. In: VI Jornadas AITADIS de Rehabilitación y Tecnologías de Apoyo a la Discapacidad, 2014, Asunción. Memorias de las VI Jornadas AITADIS de Rehabilitación y Tecnologías de Apoyo a la Discapacidad, 2014. p. 4 pg.

Appendices

Appendix A

A.1 Common Spatial Pattern

For Common Spatial Pattern, the CSP function, available at the toolbox biosig (SCHLÖGL; BRUNNER, 2008), was adopted for the spatial feature extraction on the training $\mathbf{A}_{ch,t,i}$ and validation $\mathbf{B}_{ch,t,i}$ set obtained for each fold, following the algorithm. Here, ch is the number of channels, t is the number of EEG samples per channel, i is the number of trials, and T denotes the transpose operator on the CSP projection matrix $\mathbf{P}$ of dimension $ch \times ch$. The index Idx described in Equation (3.13) was used on the training stage of the CSP method, in order to obtain the projection matrix that improves the performance of the BCI.

Algorithm:

1. Let $\mathbf{A}_{ch,t,i}$ and $\mathbf{B}_{ch,t,i}$ be the training and validation set, respectively.
2. Define a cross-validation 10-fold on the full training set $\mathbf{A}_{ch,t,i}$, to obtain combinations of new training $\mathbf{X}$ and testing $\mathbf{Y}$ set
3. For $kfold = 1$ to 10
4. $\mathbf{X}_{ch,t,i}$ and $\mathbf{Y}_{ch,t,i}$ are the bandpass filtered EEG.
5. Get patterns labeled as class 1 and 2 from $\mathbf{X}$
6. $\mathbf{P}^T$ = CSP(class1,class2); Computing the CSP projection matrix

7. For $m = 2$ to 4
8. $\overline{\mathbf{P}}^T = \mathbf{P}^T$; holding the first and last m rows
9. $\mathbf{cf_X} = log\left[diag\left(\overline{\mathbf{P}}^T\mathbf{X}\mathbf{X}^T\overline{\mathbf{P}}\right) / \text{tr}\left(\overline{\mathbf{P}}^T\mathbf{X}\mathbf{X}^T\overline{\mathbf{P}}\right)\right]$; normalized common feature
10. $\mathbf{cf_Y} = log\left[diag\left(\overline{\mathbf{P}}^T\mathbf{Y}\mathbf{Y}^T\overline{\mathbf{P}}\right) / \text{tr}\left(\overline{\mathbf{P}}^T\mathbf{Y}\mathbf{Y}^T\overline{\mathbf{P}}\right)\right]$; normalized common feature
11. Apply the normal normalization of both $\mathbf{cf}_X, \mathbf{cf}_Y$, using the mean and standard deviation values of $\mathbf{cf}_X$ to normalize $\mathbf{cf}_Y$
12. Apply Idx= LDA($\mathbf{cf}_X, \mathbf{label}_X, \mathbf{cf}_Y, \mathbf{label}_Y$);
13. Hold the $\mathbf{P}^T$ that increases Idx (see Equation (**??**)), improving the BCI performance
14. Repeat until $m = 4$
15. Repeat until $kfold = 10$
16. Filter $\mathbf{A}$ and $\mathbf{B}$
17. $\mathbf{cf_A} = log\left[diag\left(\overline{\mathbf{P}}^T\mathbf{A}\mathbf{A}^T\overline{\mathbf{P}}\right) / \text{tr}\left(\overline{\mathbf{P}}^T\mathbf{A}\mathbf{A}^T\overline{\mathbf{P}}\right)\right]$;
18. $\mathbf{cf_B} = log\left[diag\left(\overline{\mathbf{P}}^T\mathbf{B}\mathbf{B}^T\overline{\mathbf{P}}\right) / \text{tr}\left(\overline{\mathbf{P}}^T\mathbf{B}\mathbf{B}^T\overline{\mathbf{P}}\right)\right]$;

A.2 Filter-Bank Common Spatial Pattern

For Filter-Bank Common Spatial Pattern (FBCSP), the CSP function, provided in the toolbox biosig (SCHLÖGL; BRUNNER, 2008), was also adopted for the spatial feature extraction on the training $\mathbf{A}_{ch,T,i}$ and validation $\mathbf{B}_{ch,t,i}$ set obtained for each fold, following the algorithm. Here, ch is the number of channels, t is the number of EEG samples per channel, i is the number of trials, T denotes the transpose operator on the CSP projection matrix $\mathbf{P}_b$ of dimension $ch \times ch$, and b is the number of band filters. The index Idx described in Equation (3.13) was used on the training stage of the FBCSP method, in order to obtain the projection matrix that improves the performance of the BCI.

Algorithm:

1. Let $\mathbf{A}_{ch,T,i}$ and $\mathbf{B}_{ch,T,i}$ be the training and validation set, respectively.
2. Define a cross-validation 10-fold on the full training set $\mathbf{A}_{ch,T,i}$, to obtain combinations of new training $\mathbf{X}_{ch,T,i}$ and testing $\mathbf{Y}_{ch,T,i}$ set
3. For $kfold = 1$ to 10
4. For b = 1 to 7
5. $\mathbf{X}_{b,ch,T,i}$ and $\mathbf{Y}_{b,ch,T,i}$ are the bandpass filtered EEG.
6. Select classes 1 and 2 from $\mathbf{X}$
7. $\mathbf{P}_b^T$ = CSP(class1,class2); Computing the CSP projection matrix
8. For $m = 2$ to 4
9. $\overline{\mathbf{P}}_{b,m}^T = \mathbf{P}_{b,m}^T$; holding the first and last m rows
10. $\mathbf{cf_X} = log\left[diag\left(\overline{\mathbf{P}}_{b,m}^T \mathbf{X}_b \mathbf{X}_b^T \overline{\mathbf{P}}_{b,m}\right) / \text{tr}\left(\overline{\mathbf{P}}_{b,m}^T \mathbf{X}_b \mathbf{X}_b^T \overline{\mathbf{P}}_{b,m}\right)\right]$; normalized common feature
11. $\mathbf{cf_Y} = log\left[diag\left(\overline{\mathbf{P}}_{b,m}^T \mathbf{Y}_b \mathbf{Y}_b^T \overline{\mathbf{P}}_{b,m}\right) / \text{tr}\left(\overline{\mathbf{P}}_{b,m}^T \mathbf{Y}_b \mathbf{Y}_b^T \overline{\mathbf{P}}_{b,m}\right)\right]$; normalized common feature
12. $\mathbf{cf}_{X,m} = [\mathbf{cf}_{X,m} \;\; \mathbf{cf}_X]$ holding the features
13. $\mathbf{cf}_{Y,m} = [\mathbf{cf}_{Y,m} \;\; \mathbf{cf}_Y]$ holding the features
14. Repeat until $m = 4$
15. Repeat until $b = 7$
16. Evaluate of the feature set for each m
17. For $m = 2$ to 4
18. Apply the normal normalization of both $\mathbf{cf}_{X,m}, \mathbf{cf}_{Y,m}$, using the mean and standard deviation values of $\mathbf{cf}_{X,m}$ for $\mathbf{cf}_{Y,m}$
19. Rank based on the best individual features on $\mathbf{cf}_{X,m}$, using mutual information (ANG et al., 2012a; JAIN; DUIN; MAO, 2000)

20. Look the best first k features
21. For $k = 1$ to $k = 2 \times log2(7 \times 2 \times m)$
22. Apply Idx= LDA($\mathbf{cf}_{X,m}, \mathbf{label}_X, \mathbf{cf}_{Y,m}, \mathbf{label}_Y$);
23. Hold the $\overline{\mathbf{P}}_{b,m}^T$, and the best individual features that increases Idx (see Equation (**??**)), improving the BCI performance
24. Repeat until $k = 2 \times log2(7 \times 2 \times m)$
25. Repeat until $m = 4$
26. Repeat until $kfold = 10$
27. For b = 1 to 7
28. Filtering **A** and **B**
29. $\mathbf{cf_1} = log\left[diag\left(\overline{\mathbf{P}}_{b,m}^T \mathbf{A}_b \mathbf{A}_b^T \overline{\mathbf{P}}_{b,m}\right) / \text{tr}\left(\overline{\mathbf{P}}_{b,m}^T \mathbf{A}_b \mathbf{A}_b^T \overline{\mathbf{P}}_{b,m}\right)\right]$; normalized common feature
30. $\mathbf{cf_2} = log\left[diag\left(\overline{\mathbf{P}}_{b,m}^T \mathbf{B}_b \mathbf{B}_b^T \overline{\mathbf{P}}_{b,m}\right) / \text{tr}\left(\overline{\mathbf{P}}_{b,m}^T \mathbf{B}_b \mathbf{B}_b^T \overline{\mathbf{P}}_{b,m}\right)\right]$; normalized common feature
31. $\mathbf{cf_A} = [\mathbf{cf_A cf_1}]$ holding $\mathbf{cf_1}$
32. $\mathbf{cf_B} = [\mathbf{cf_B cf_2}]$ holding $\mathbf{cf_2}$
33. Repeat until $b = 7$
34. Selecting on $\mathbf{cf_A}$ and $\mathbf{cf_B}$ the best individual features obtained from the cross-validation

A.2.1 Mutual information-based best individual feature algorithm

The Mutual Information-based Best Individual Feature (MIBIF) algorithm (JAIN; DUIN; MAO, 2000) is based on the filter approach. The mutual information of each feature is computed and sorted in descending order. The first k features are then selected. The MIBIF algorithm is described as follows:

Algorithm:

1. To initialize the feature set $\mathcal{F} = [\mathbf{f}_1^\dagger, \mathbf{f}_2^\dagger, \ldots, \mathbf{f}_{9*2m}^\dagger] = \overline{\mathbf{V}}$ from Equation (2.20), and the true label set $\mathcal{C} = \overline{\mathbf{y}}$ from Equation (2.21) whereby $\mathbf{f}_j^\dagger \in \mathbb{R}^{n\times 1}$ is the jth column vector of $\overline{\mathbf{V}}$, and the true label of each trial $\overline{\mathbf{V}}_\mathbf{i} \in \{1, 2\}$.

   Initialize set of selected features $\mathcal{S} = \varnothing$

2. To compute the mutual information $I(\mathbf{f}_j; \omega)$ of each feature $\mathbf{f}_j \in \mathcal{F}$ with each class label $\omega = \{1, 2\} \in \mathcal{C}$. Then, $I(\mathbf{f}_j; \omega)\ \forall\ j = 1, 2, ...(9{\cdot}2m)$ can be computed through the following expression:

$$I(\mathbf{f}_j; \omega) = H(\omega) - H(\omega|\mathbf{f}_j)\ , \tag{A.1}$$

   where $H(\omega) = -\sum_{\omega=1}^{2} P(\omega)\log_2 P(\omega)$.

   The conditional entropy can be computed by:

$$\begin{aligned} H(\omega|\mathbf{f}_j) &= -\sum_{\omega=1}^{2} p(\omega|\mathbf{f}_j)\log_2 p(\omega|\mathbf{f}_j) \\ &= -\sum_{\omega=1}^{2}\sum_{i=1}^{n} p(\omega|f_{j,i})\log_2 p(\omega|f_{j,i})\ , \end{aligned} \tag{A.2}$$

   where $f_{j,i}$ is the feature value of the ith trial from $\mathbf{f}_j$. The Bayes rule given in Equations (A.3) and (A.4) can be used to compute the probability $p(\omega|f_{j,i})$.

$$p(\omega|f_{j,i}) = \frac{p(f_{j,i}|\omega)P(\omega)}{p(f_{j,i})}\ , \tag{A.3}$$

$$p(f_{j,i}) = \sum_{\omega=1}^{2} p(f_{j,i}|\omega)P(\omega)\ , \tag{A.4}$$

   were $p(\omega|f_{j,i})$ is the conditional probability of the class ω given $f_{j,i}$, $p(f_{j,i}|\omega)$ is the conditional probability of $f_{j,i}$ given the class ω, $P(\omega)$ is the prior probability of class ω.

   The conditional probability $p(f_{j,i}|\omega)$ can be estimated using the Parzen Window (PARZEN, 1962) given by:

$$\hat{p}(f_{j,i}|\omega) = \frac{1}{n_\omega}\sum_{k\in I_\omega}\phi(f_{j,i} - f_{j,k}, h)\ , \tag{A.5}$$

   where n_ω is the number of data samples belonging to the class ω, I_ω is the set of indices of the training data trials belonging to the class ω, $f_{j,k}$ is the feature value of the kth

trial from $\mathbf{f}_j$, and ϕ is a smoothing kernel function with a smoothing parameter h given by Equations (A.6) and (A.8), being n the number of trials, σ the standard deviation of the distribution of y (see Equation (2.21)).

$$\phi(y,h) = \frac{1}{\sqrt[2]{2\pi}} \exp\left[-\frac{y^2}{2h^2}\right] \tag{A.6}$$

$$h^{opt} = \left(\frac{4}{3n}\right)^{1/5} \sigma \tag{A.7}$$

3. Sort all the features in descending order of mutual information computed in step 2 and select the first k features. Mathematically, this step is performed as follows till $|\mathcal{S}| = k$:

$$\mathcal{F} = \mathcal{F} \backslash \mathbf{f}_j, \mathcal{F}, \mathcal{S} = \mathcal{S} \cup \mathbf{f}_j | I(\mathbf{f}_j;\omega) = \max_{j=1,\ldots,(9*2m),\mathbf{f}_j \in \mathcal{F}} I(\mathbf{f}_j;\omega) \tag{A.8}$$

where k is the number of best features ($k = 2\log_2(9 * 2m)$), $\backslash$ denotes set theoretic difference; $\cup$ denotes set union, and $|$ denotes given the condition. Notice that since CSP features are paired, the corresponding pair of features is also included if it is not selected. Once the feature selection on $\overline{\mathbf{V}}$ is finished, the training data cane be expressed as $\overline{\mathbf{X}} \in \mathbb{R}^{n \times d}$, where d is the number of selected features.

A.3 Covariance matrices using Riemannian-based Kernel

For the Riemannian method, the functions such as *covariances*, *meancovariances*, and $Tangentspace$ (available at the website <https://github.com/alexandrebarachant>) were adopted here for the spatial feature extraction on the training $\mathbf{A}_{ch,t,i}$ and validation $\mathbf{B}_{ch,t,i}$ set obtained for each fold, following the described algorithm. Here, ch is the number of channels, t is the number of EEG samples per channel, i is the number of trials, and T denotes the transpose operator in the algorithm.

Algorithm:

1. Let $\mathbf{A}_{ch,t,i}$ and $\mathbf{B}_{ch,t,i}$ be the training and validation set, respectively.

2. $\mathbf{X}_{ch,t,i}$ and $\mathbf{Y}_{ch,t,i}$ are the bandpass filtered EEG from the training and validation set, respectively.

3. $\mathbf{Ctrain} = covariances(\mathbf{X})$;

4. $\mathbf{C} = meancovariances(\mathbf{Ctrain}, 'riemann')$; Computing the Riemannian mean

5. $\mathbf{Training\ set} = Tangentspace(\mathbf{Ctrain}, \mathbf{C})^T$; Spatial feature extraction

6. $\mathbf{Cvalidation} = covariances(\mathbf{Y})$;

7. $\mathbf{Validation\ set} = Tangentspace(\mathbf{Cvalidation}, \mathbf{C})^T$; Spatial feature extraction

A.4 Representation Entropy

Let the eigenvalues of the $d \times d$ covariance matrix of a feature set of size d be $\lambda_j,\ j = 1, \ldots, d$. Let $\widetilde{\lambda}_j$ has similar properties as probability, where $0 \leq \widetilde{\lambda}_j \leq 1$, and $\sum_{j=1}^{d} \widetilde{\lambda}_j = 1$. Hence, the entropy H_R can be computed by Equation (A.10).

$$\widetilde{\lambda}_j = \frac{\lambda_j}{\sum_{j=1}^{d} \lambda_j}\ , \tag{A.9}$$

$$H_R = \sum_{j=1}^{d} \widetilde{\lambda}_j \log \widetilde{\lambda}_j\ , \tag{A.10}$$

H_R, termed representation entropy (RE), takes the minimum value (zero) when all eigenvalues except one are zero. In other words, when all information are present along a single coordinate direction. it is a measure of the amount of information compression possible by feature selection, or dimensionality reduction. Thus, it expected that RE of individual clusters are as low as possible, while a final reduced feature set has low redundancies, and consequently, a high value of RE.

A.5 Maximal Information Compression Index

Let Σ be the covariance matrix of random variables x and y. In addition, maximal information compression index (MICI) is defined as $\lambda_2(x, y) =$ smallest eigenvalue of Σ (MITRA; MURTHY; PAL, 2002), given by:

$$2\lambda_2(x, y) = var(x) + var(y) - \sqrt{(var(x) + var(y))^2 - 4var(x)var(y)(1 - \rho(x, y)^2)} \ , \quad \text{(A.11)}$$

where ρ is the correlation coefficient, and $var(\cdot)$ is the variance operator.
The value λ_2 is zero when the features are linearly dependent, and it increases when the dependency between features decrease (MITRA; MURTHY; PAL, 2002). Then, λ_2 is nothing but the eigenvalue for the direction normal to the principle component direction of feature pair (x,y). High values ofMICI is achieved if a multivariate data is projected along its principal component direction. Therefore, it is a measure of the minimum amount of lost information or the maximum amount of information compression, possible.

A.6 SSVEP database

The database of EEG signals from 35 subjects (CHEN et al., 2015) was used to evaluate the preservation of SSVEP components after applying the adaptive filter.
Subjects of the database (from SV1 to SV35) were asked to focus on 40 characters flickering at different frequencies (8-15.8 Hz with an interval of 0.2 Hz) during six sessions. Each session was conducted for a total of 40 trials of 6 s in length that represent the full character set in a random selection. The trials were formed by two stages: (1) a visual cue (a red square) was presented on the selected character for a duration of 0.5 s, and (2) the corresponding stimulus was emitted to the subject for a period of 5 s. The screen was blank for 0.5 s before the next trial began, which provided for all subjects a short break between consecutive trials. Furthermore, all subjects were asked to shift their gaze to the target as soon as possible within the cue duration. The subjects were asked to avoid eye blinks throughout the stimulation period. Moreover, several minutes of rest were added between two consecutive sessions, in order to avoid visual fatigue.
The equipment Synamps2 (Neuroscan, Inc.) was used to acquire 64 channels of EEG signals, according to the international 10-20 system, with frequency range from 0.15 Hz to 200 Hz, notch filter at 50 Hz, and sampling rate of 1000 Hz. The ground electrode was placed midway between Fz and FPz, and the reference electrode was located on the vertex. The continuous EEG was segmented in epochs of 6 s (0.5 s pre-stimulus, 5.5 s post-stimulus onset), which were subsequently downsampled at 250 Hz. Such as aforementioned, this dataset was used here to

fit the proposed model through a group of subjects, evaluating on all subjects the ability of the adaptive filter to preserve the neural information at stimuli frequencies.

Bibliography

ALSHBATAT, A. I. N. et al. Eeg-based brain-computer interface for automating home appliances. 2014.

ANG, K. K. et al. Filter bank common spatial pattern algorithm on bci competition iv datasets 2a and 2b. *Frontiers in neuroscience*, Frontiers Media SA, v. 6, 2012.

ANG, K. K. et al. Mutual information-based selection of optimal spatial–temporal patterns for single-trial eeg-based bcis. *Pattern Recognition*, Elsevier, v. 45, n. 6, p. 2137–2144, 2012.

BAI, O. et al. Prediction of human voluntary movement before it occurs. *Clinical Neurophysiology*, Elsevier, v. 122, n. 2, p. 364–372, 2011.

BARACHANT, A. et al. Multiclass brain–computer interface classification by riemannian geometry. *IEEE Transactions on Biomedical Engineering*, IEEE, v. 59, n. 4, p. 920–928, 2012.

BARACHANT, A. et al. Classification of covariance matrices using a riemannian-based kernel for bci applications. *Neurocomputing*, Elsevier, v. 112, p. 172–178, 2013.

BLANKERTZ, B. et al. Optimizing spatial filters for robust eeg single-trial analysis. *IEEE Signal processing magazine*, IEEE, v. 25, n. 1, p. 41–56, 2008.

BULEA, T. C. et al. Sitting and standing intention can be decoded from scalp eeg recorded prior to movement execution. *Frontiers in neuroscience*, Frontiers Media SA, v. 8, 2014.

CASTRO, M. C. F.; ARJUNAN, S. P.; KUMAR, D. K. Selection of suitable hand gestures for reliable myoelectric human computer interface. *Biomedical engineering online*, BioMed Central, v. 14, n. 1, p. 30, 2015.

CHEN, X. et al. Filter bank canonical correlation analysis for implementing a high-speed ssvep-based brain–computer interface. *Journal of neural engineering*, IOP Publishing, v. 12, n. 4, p. 046008, 2015.

COSTA, Á. et al. Decoding the attentional demands of gait through eeg gamma band features. *PloS one*, Public Library of Science, v. 11, n. 4, p. e0154136, 2016.

COTRINA, A. et al. A ssvep-bci setup based on depth-of-field. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, IEEE, 2017.

DELISLE-RODRIGUEZ, D. et al. Non-supervised feature selection: Evaluation in a bci for single-trial recognition of gait preparation/stop. In: *Converging Clinical and Engineering Research on Neurorehabilitation II*. [S.l.]: Springer, 2017. p. 1115–1120.

DIEZ, P. F. et al. Commanding a robotic wheelchair with a high-frequency steady-state visual evoked potential based brain–computer interface. *Medical engineering & physics*, Elsevier, v. 35, n. 8, p. 1155–1164, 2013.

DO, A. H. et al. Brain-computer interface controlled robotic gait orthosis. *Journal of neuroengineering and rehabilitation*, BioMed Central, v. 10, n. 1, p. 111, 2013.

ERGIN, M. A.; PATOGLU, V. A self-adjusting knee exoskeleton for robot-assisted treatment of knee injuries. In: IEEE. *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*. [S.l.], 2011. p. 4917–4922.

FARRIS, R. J.; QUINTERO, H. A.; GOLDFARB, M. Preliminary Evaluation of a Powered Lower Limb Orthosis to Aid Walking in Paraplegic Individuals. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, v. 19, n. 6, p. 652–659, dez. 2011. ISSN 1534-4320.

FRIEDMAN, J. H. Regularized discriminant analysis. *Journal of the American statistical association*, Taylor & Francis, v. 84, n. 405, p. 165–175, 1989.

GALLEGO, J. Á. et al. A multimodal human–robot interface to drive a neuroprosthesis for tremor management. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, IEEE, v. 42, n. 6, p. 1159–1168, 2012.

GAMS, A. et al. Effects of robotic knee exoskeleton on human energy expenditure. *IEEE Transactions on Biomedical Engineering*, IEEE, v. 60, n. 6, p. 1636–1644, 2013.

GARRETT, D. et al. Comparison of linear, nonlinear, and feature selection methods for eeg signal classification. *IEEE Transactions on neural systems and rehabilitation engineering*, IEEE, v. 11, n. 2, p. 141–144, 2003.

GRAIMANN, B. et al. Visualization of significant erd/ers patterns in multichannel eeg and ecog data. *Clinical Neurophysiology*, Elsevier, v. 113, n. 1, p. 43–47, 2002.

GUERRERO-MOSQUERA, C.; NAVIA-VÁZQUEZ, A. Automatic removal of ocular artefacts using adaptive filtering and independent component analysis for electroencephalogram data. *IET signal processing*, IET, v. 6, n. 2, p. 99–106, 2012.

HASHIMOTO, Y.; USHIBA, J. Eeg-based classification of imaginary left and right foot movements using beta rebound. *Clinical neurophysiology*, Elsevier, v. 124, n. 11, p. 2153–2160, 2013.

HIGUCHI, T. Approach to an irregular time series on the basis of the fractal theory. *Physica D: Nonlinear Phenomena*, Elsevier, v. 31, n. 2, p. 277–283, 1988.

HORTAL, E. et al. Eeg-based detection of starting and stopping during gait cycle. *International journal of neural systems*, World Scientific, v. 26, n. 07, p. 1650029, 2016.

HUANG, C.-E.; CHEN, J.-S. On the implementation and control of a pneumatic power active lower-limb orthosis. *Mechatronics*, Elsevier, v. 23, n. 5, p. 505–517, 2013.

HUSSAIN, S.; XIE, S. Q.; LIU, G. Robot assisted treadmill training: Mechanisms and training strategies. *Medical Engineering & Physics*, v. 33, n. 5, p. 527–533, jun. 2011. ISSN 13504533. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S1350453310002973>.

HUSSAIN, S.; XIE, S. Q.; LIU, G. Robot assisted treadmill training: Mechanisms and training strategies. *Medical engineering & physics*, Elsevier, v. 33, n. 5, p. 527–533, 2011.

IKEDA, A. et al. Movement-related potentials recorded from supplementary motor area and primary motor area: role of supplementary motor area in voluntary movements. *Brain*, Oxford University Press, v. 115, n. 4, p. 1017–1043, 1992.

JAIN, A. K.; DUIN, R. P. W.; MAO, J. Statistical pattern recognition: A review. *IEEE Transactions on pattern analysis and machine intelligence*, Ieee, v. 22, n. 1, p. 4–37, 2000.

JAPKOWICZ, N.; SHAH, M. *Evaluating learning algorithms: a classification perspective*. [S.l.]: Cambridge University Press, 2011.

JIANG, N. et al. A brain–computer interface for single-trial detection of gait initiation from movement related cortical potentials. *Clinical Neurophysiology*, Elsevier, v. 126, n. 1, p. 154–159, 2015.

JUNG, T.-P. et al. Extended ica removes artifacts from electroencephalographic recordings. In: *Advances in neural information processing systems*. [S.l.: s.n.], 1998. p. 894–900.

JUNG, T.-P. et al. Removing electroencephalographic artifacts: comparison between ica and pca. In: IEEE. *Neural Networks for Signal Processing VIII, 1998. Proceedings of the 1998 IEEE Signal Processing Society Workshop*. [S.l.], 1998. p. 63–72.

KHATUN, S.; MAHAJAN, R.; MORSHED, B. I. Comparative study of wavelet-based unsupervised ocular artifact removal techniques for single-channel eeg data. *IEEE journal of translational engineering in health and medicine*, IEEE, v. 4, p. 1–8, 2016.

KILICARSLAN, A.; GROSSMAN, R. G.; CONTRERAS-VIDAL, J. L. A robust adaptive denoising framework for real-time artifact removal in scalp eeg measurements. *Journal of neural engineering*, IOP Publishing, v. 13, n. 2, p. 026013, 2016.

LAPLANTE, P. A. *Electrical engineering dictionary*. [S.l.]: CRC Press LLC, 2000.

LAWRENCE, I.; LIN, K. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, JSTOR, p. 255–268, 1989.

LEMM, S. et al. Spatio-spectral filters for improving the classification of single trial eeg. *IEEE transactions on biomedical engineering*, IEEE, v. 52, n. 9, p. 1541–1548, 2005.

LEW, E. et al. Detection of self-paced reaching movement intention from eeg signals. *Frontiers in neuroengineering*, Frontiers Media SA, v. 5, 2012.

LIN, C. et al. Libd3c: ensemble classifiers with a clustering and dynamic selection strategy. *Neurocomputing*, Elsevier, v. 123, p. 424–435, 2014.

LIN, Z. et al. Frequency recognition based on canonical correlation analysis for ssvep-based bcis. *IEEE Transactions on Biomedical Engineering*, IEEE, v. 54, n. 6, p. 1172–1176, 2007.

LIU, J. et al. Linear correlation between fractal dimension of eeg signal and handgrip force. *biological Cybernetics*, Springer, v. 93, n. 2, p. 131–140, 2005.

LOTTE, F. et al. A review of classification algorithms for eeg-based brain–computer interfaces. *Journal of neural engineering*, IOP Publishing, v. 4, n. 2, p. R1, 2007.

MCFARLAND, D. J. et al. Bci meeting 2005-workshop on bci signal processing: feature extraction and translation. *IEEE transactions on neural systems and rehabilitation engineering*, IEEE, v. 14, n. 2, p. 135–138, 2006.

MCFARLAND, D. J. et al. Spatial filter selection for eeg-based communication. *Electroencephalography and clinical Neurophysiology*, Elsevier, v. 103, n. 3, p. 386–394, 1997.

MCFARLAND, D. J.; WOLPAW, J. R. Sensorimotor rhythm-based brain-computer interface (bci): feature selection by regression improves performance. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, IEEE, v. 13, n. 3, p. 372–379, 2005.

MITRA, P.; MURTHY, C.; PAL, S. K. Unsupervised feature selection using feature similarity. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 24, n. 3, p. 301–312, 2002.

MULLEN, T. et al. Real-time modeling and 3d visualization of source dynamics and connectivity using wearable eeg. In: IEEE. *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*. [S.l.], 2013. p. 2184–2187.

MÜLLER-GERKING, J.; PFURTSCHELLER, G.; FLYVBJERG, H. Designing optimal spatial filters for single-trial eeg classification in a movement task. *Clinical neurophysiology*, Elsevier, v. 110, n. 5, p. 787–798, 1999.

MÜLLER, S. M. T.; BASTOS, T. F.; FILHO, M. S. Proposal of a ssvep-bci to command a robotic wheelchair. *Journal of Control, Automation and Electrical Systems*, Springer, v. 24, n. 1-2, p. 97–105, 2013.

NESHIGE, R.; LÜDERS, H.; SHIBASAKI, H. Recording of movement-related potentials from scalp and cortex in man. *Brain: a journal of neurology*, v. 111, p. 719–736, 1988.

NOLAN, H.; WHELAN, R.; REILLY, R. Faster: fully automated statistical thresholding for eeg artifact rejection. *Journal of neuroscience methods*, Elsevier, v. 192, n. 1, p. 152–162, 2010.

ONEN, U. et al. Design and actuator selection of a lower extremity exoskeleton. *IEEE/ASME Transactions on Mechatronics*, IEEE, v. 19, n. 2, p. 623–632, 2014.

PARZEN, E. On estimation of a probability density function and mode. *The annals of mathematical statistics*, JSTOR, v. 33, n. 3, p. 1065–1076, 1962.

PFURTSCHELLER, G.; FLOTZINGER, D.; KALCHER, J. Brain-computer interface—a new communication device for handicapped persons. *Journal of Microcomputer Applications*, Elsevier, v. 16, n. 3, p. 293–299, 1993.

PFURTSCHELLER, G. et al. Spatiotemporal patterns of beta desynchronization and gamma synchronization in corticographic data during self-paced movement. *Clinical neurophysiology*, Elsevier, v. 114, n. 7, p. 1226–1236, 2003.

PFURTSCHELLER, G.; NEUPER, C.; BERGER, J. Source localization using eventrelated desynchronization (erd) within the alpha band. *Brain Topography*, Springer, v. 6, n. 4, p. 269–275, 1994.

PFURTSCHELLER, G.; SILVA, F. L. D. Event-related eeg/meg synchronization and desynchronization: basic principles. *Clinical neurophysiology*, Elsevier, v. 110, n. 11, p. 1842–1857, 1999.

PHINYOMARK, A.; PHUKPATTARANONT, P.; LIMSAKUL, C. Feature reduction and selection for emg signal classification. *Expert Systems with Applications*, Elsevier, v. 39, n. 8, p. 7420–7431, 2012.

PHOTHISONOTHAI, M.; WATANABE, K. Optimal fractal feature and neural network: Eeg based bci applications. In: *Brain-Computer Interface Systems-Recent Progress and Future Prospects*. [S.l.]: InTech, 2013.

PONS, J. L. *Wearable robots: biomechatronic exoskeletons*. [S.l.]: John Wiley & Sons, 2008.

PUTHUSSERYPADY, S.; RATNARAJAH, T. H/sup/spl infin//adaptive filters for eye blink artifact minimization from electroencephalogram. *IEEE Signal Processing Letters*, IEEE, v. 12, n. 12, p. 816–819, 2005.

QIN, L.; DING, L.; HE, B. Motor imagery classification by means of source analysis for brain–computer interface applications. *Journal of neural engineering*, IOP Publishing, v. 1, n. 3, p. 135, 2004.

QUIROGA, R. Q. et al. Performance of different synchronization measures in real data: a case study on electroencephalographic signals. *Physical Review E*, APS, v. 65, n. 4, p. 041903, 2002.

ROCON, E. et al. Multimodal bci-mediated fes suppression of pathological tremor. In: IEEE. *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*. [S.l.], 2010. p. 3337–3340.

SALAZAR-VARAS, R. et al. Analyzing eeg signals to detect unexpected obstacles during walking. *Journal of neuroengineering and rehabilitation*, BioMed Central, v. 12, n. 1, p. 101, 2015.

SBURLEA, A. I. et al. Detecting intention to walk in stroke patients from pre-movement eeg correlates. *Journal of neuroengineering and rehabilitation*, BioMed Central, v. 12, n. 1, p. 113, 2015.

SBURLEA, A. I.; MONTESANO, L.; MINGUEZ, J. Continuous detection of the self-initiated walking pre-movement state from eeg correlates without session-to-session recalibration. *Journal of neural engineering*, IOP Publishing, v. 12, n. 3, p. 036007, 2015.

SBURLEA, A. I.; MONTESANO, L.; MINGUEZ, J. Advantages of eeg phase patterns for the detection of gait intention in healthy and stroke subjects. *Journal of Neural Engineering*, IOP Publishing, v. 14, n. 3, p. 036004, 2017.

SCHLÖGL, A.; BRUNNER, C. Biosig: a free and open source software library for bci research. *Computer*, IEEE, v. 41, n. 10, 2008.

SEEBER, M. et al. High and low gamma eeg oscillations in central sensorimotor areas are conversely modulated during the human gait cycle. *Neuroimage*, Elsevier, v. 112, p. 318–326, 2015.

SHAKEEL, A. et al. A review of techniques for detection of movement intention using movement-related cortical potentials. *Computational and mathematical methods in medicine*, Hindawi Publishing Corporation, v. 2015, 2015.

SHIBASAKI, H.; HALLETT, M. What is the bereitschaftspotential? *Clinical neurophysiology*, Elsevier, v. 117, n. 11, p. 2341–2356, 2006.

SWEENEY, K. T. et al. A methodology for validating artifact removal techniques for physiological signals. *IEEE transactions on information technology in biomedicine*, IEEE, v. 16, n. 5, p. 918–926, 2012.

SWEENEY, K. T.; MCLOONE, S. F.; WARD, T. E. The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique. *IEEE transactions on biomedical engineering*, IEEE, v. 60, n. 1, p. 97–105, 2013.

TOMA, K. et al. Generators of movement-related cortical potentials: fmri-constrained eeg dipole source analysis. *Neuroimage*, Elsevier, v. 17, n. 1, p. 161–173, 2002.

VASEGHI, S. V. *Advanced digital signal processing and noise reduction*. [S.l.]: John Wiley & Sons, 2008.

VELU, P. D.; SA, V. R. de. Single-trial classification of gait and point movement preparation from human eeg. *Frontiers in neuroscience*, Frontiers Media SA, v. 7, 2013.

VILLA-PARRA, A. et al. Towards a robotic knee exoskeleton control based on human motion intention through eeg and semgsignals. *Procedia Manufacturing*, Elsevier, v. 3, p. 1379–1386, 2015.

VITECKOVA, S.; KUTILEK, P.; JIRINA, M. Wearable lower limb robotics: A review. *Biocybernetics and Biomedical Engineering*, v. 33, n. 2, p. 96–105, jan. 2013. ISSN 02085216. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0208521613000065>.

WHO. *The impact of chronic disease in Brazil*. 2005. Disponível em: <http://www.who.int/chp/chronic_disease_report/media/brazil.pdf?ua=1>.

WHO. *Spinal cord injury*. 2013. Disponível em: <http://www.who.int/mediacentre/factsheets/fs384/en/>.

WHO. *Spinal cord injury: as many as 500 000 people suffer each year*. 2013. Disponível em: <http://www.who.int/mediacentre/news/releases/2013/spinal-cord-injury-20131202/en/>.

WHO. *Stroke, Cerebrovascular accident*. 2014. Disponível em: <http://www.who.int/topics/cerebrovascular_accident/en/>.

WHO. *Ageing and health*. 2015. Disponível em: <http://www.who.int/mediacentre/factsheets/fs404/en/>.

WHO. *Ageing and life-course*. 2015. Disponível em: <http://www.who.int/ageing/publications/world-report-2015/en/>.

WHO. *World report on ageing and health*. 2015. Disponível em: <http://apps.who.int/iris/bitstream/10665/186463/1/9789240694811_eng.pdf>.

WHO. *10 facts on ageing and health*. 2016. Disponível em: <http://www.who.int/features/factfiles/ageing/ageing_facts/en/>.

WHO. *Obesity and overweight*. 2017. Disponível em: <http://www.who.int/mediacentre/factsheets/fs311/en/>.

WHO. *Road traffic injuries*. 2017. Disponível em: <http://www.who.int/mediacentre/factsheets/fs358/en/>.

WOLPAW, J. R. et al. Brain-computer interface technology: a review of the first international meeting. *IEEE transactions on rehabilitation engineering*, IEEE, v. 8, n. 2, p. 164–173, 2000.

WOLPAW, J. R.; MCFARLAND, D. J. Multichannel eeg-based brain-computer communication. *Electroencephalography and clinical Neurophysiology*, Elsevier, v. 90, n. 6, p. 444–449, 1994.

WOLPAW, J. R. et al. An eeg-based brain-computer interface for cursor control. *Electroencephalography and clinical neurophysiology*, Elsevier, v. 78, n. 3, p. 252–259, 1991.

XU, R. et al. Enhanced low-latency detection of motor intention from eeg for closed-loop brain-computer interface applications. *IEEE Transactions on Biomedical Engineering*, IEEE, v. 61, n. 2, p. 288–296, 2014.