2022

Investigation into a Brain-Computer Interface System for Music with Speech Imagery

Tates Puetate, Alberto

http://hdl.handle.net/10026.1/18984

http://dx.doi.org/10.24382/724
University of Plymouth

UNIVERSITY OF PLYMOUTH

INVESTIGATION INTO A BRAIN-COMPUTER INTERFACE SYSTEM FOR MUSIC WITH SPEECH IMAGERY

by

ALBERTO TATES PUETATE

A thesis submitted to the University of Plymouth

in partial fulfillment for the degree of

RESEARCH MASTERS

School of Society and Culture

March 2022

ACKNOWLEDGEMENTS

The completion of this research master's degree is heartwarming as I could bring to reality an idea that started in a fictional manner, I could get in touch with people whose work is inspiring especially when it comes to approaching arts and technology in unthinkable ways. Apart from gaining technical and research skills, I have a better career perspective, thanks to the positive influence from the researching environment I could be involved with, even if COVID pandemic made unexpected plan changes.

All achievements that I have reached would not be possible without the unconditional support and love from my mother and father who are always happy and supportive when I have to decide for a new path, even being on the other side of the Atlantic Ocean.

Prof. Eduardo Miranda is simply one of the most admirable people I've had the chance to meet, his work inspired me to push my career beyond my expectations and to think really outside the box when it comes to scientific contributions. Since the first email I addressed to Eduardo, he has been entirely supportive and open to help. His strong experience with EEG-based systems helped me design optimal experiments for quick and better results. I am totally honored and proud to have had him as director of studies.
Being part of the Interdisciplinary Center of Computer Music Research (ICCMR) has been a dream-like experience, getting to know the contributions to art and science from different perspectives and goals was inspiring, the members I could meet were highly passionate about their works so I felt constantly pushed to improve my work.

The University of Plymouth principles of interdisciplinary contributions, and impact allowed this sort of program to happen, I couldn't achieve my research goals without these research cross paths that the institution has formed.
I am grateful to many friends and family back in Ecuador, that make them feel always close by, especially my cousins Chiqui and Vane.

AUTHOR'S DECLARATION

At no time during the registration for the degree of Research Masters has the author been registered for any other University award without prior agreement of the Doctoral College Quality Sub-Committee.

Work submitted for this research degree at the University of Plymouth has not formed part of any other degree either at the University of Plymouth or at another establishment.

Word count of main body thesis : 18.375

Signed: ……………………..

Date: ….30-03-22…..

INVESTIGATION INTO A BRAIN-COMPUTER INTERFACE SYSTEM FOR MUSIC WITH SPEECH IMAGERY

by

ALBERTO TATES PUETATE

ABSTRACT

Brain-computer interfaces (BCIs) aim to find a communication path with computers that dispenses with motor interaction using merely brain activity. This project investigates how BCI can have music as its destination, the system would allow people with motor impairments to get close to creative activities and offer musicians a novel dimension of performance. BCI systems present multiple challenges when attempting approaches in real-world circumstances. These challenges are encompassed in the transfer information rate and the deployability of such systems. BCI research is mainly medically focused therefore the brain activity measurement devices are designed under lab conditions, making them not portable for musical scenarios. This research focuses on electroencephalogram (EEG) as being one of the widest used methods, among the different EEG paradigms and decoding processes, this research evaluates the most feasible pipeline to implement a prototype that could activate a musical instrument.

Speech Imagery (SI) is the selected EEG paradigm, as in comparison with other known methods, it is a non-stimulus-based evoked potential, therefore the SI decoder aims to extract spatial information about the signal differences between attempts from imagine to pronounce vowels or phonemes aloud. The project proposes the use of 8 dry electrodes set to increment system portability. The research starts to evaluate between a traditional decoding pipeline with Common Spatial Patterns and Support Vector Machine, with a relatively new decoder based on Riemannian metric classification of Covariance Matrices. Selecting the latter as the most optimal method for online implementation, this research then evaluates a set of different imagery tasks to analyze the decoding performance and select the most optimal configurations.

The final prototype is built having as speech imagery tasks the rhythmic imaginations of vowel /e/, vowel /o/ and single imagery of word /mid/, the system converted 1 second of EEG into a covariance matrix which distance to each class centroid, obtained in offline analysis, was then compared in a decision three that selects a MIDI outcome to an external analogue synthesizer. The system achieved optimal accuracies, when detecting resting state against any imaginary task, of up to 87%, the overall accuracy of the online system for 4 classes was 57% ± 3 significantly higher than the 42% chance level, the system had a transfer rate of 3.4 bit/min that allowed the participant to have a novel musical experience.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION **11**

1.1 Motivation 11

1.2 Research Questions 12

1.3 Research Methodology 14

1.3.1 Wireless and dry EEG Headsets 14

1.3.1 EEG decoding methods 14

1.3.2 Programing and analysis tools 15

1.4 Thesis Structure 15

Chapter 2 15

Chapter 3 16

Chapter 4 16

Chapter 5 17

CHAPTER 2

BRAIN-COMPUTER INTERFACE METHODOLOGIES FOR MUSIC APPLICATIONS **18**

2.1 Overview 18

2.2 Introduction 19

2.2.1 Electroencephalogram (EEG) 20

2.3 Criteria of Comparison 21

2.3.1 Dimension of Control 21

2.3.2 Usability of the System 21

2.4 EEG-based paradigms for control BCI 22

2.4.1 P300 22

2.4.2 Steady-State Visual Evoked Potential 23

2.4.3 Event-Related Synchronization/Desynchronization 24

2.5 Comparison 25

2.6 Survey Conclusion 28

CHAPTER 3

INVESTIGATION OF EEG DECODING METHODS FOR SPEECH IMAGERY **29**

3.1 Overview 29

3.2 Introduction 28
3.2.1 BCI Control System 30
3.2.2 Brain Activity Acquisition 30
3.2.3 EEG Signal Decoding to Commands 32
Signal Preprocessing 32
Feature Extraction 33
Classification 35
3.2.4 Device to Control 37
3.3 Proposed methodology 37
3.4 Methodology Application 38
3.4.1 EEG experiments 38
3.4.2 Signal Processing 40
3.4.3 Results 45
3.5 Conclusion 46

CHAPTER 4

BRAIN-COMPUTER INTERFACE MUSIC SYSTEM USING SPEECH IMAGERY, A PROOF-OF-CONCEPT **49**

4.1 Overview 49
4.2 Introduction 50
4.3 Proposed Decoder 51
4.3.1 Feature Extraction 51
4.3.2 Classifier 52
4.4 Methodology 53
4.4.1 EEG Experiments 60
4.2 Offline Results 61
Table 3. MDM for two locations and 5 classes Accuracies 62
4.3 Real-time Implementation 62
4.4 Real-time Results 64
4.5 Conclusion 65

CHAPTER 5

CONCLUSION AND FUTURE WORK **66**

5.1 Overview 66
5.2 Contributions on Research Questions 66
5.3 Additional Contributions 69

5.4 Future Work 70
5.4.1 Suggested Research Questions 70
References **72**

LIST OF FIGURES

Fig 1. BCI Control System diagram 31
Fig 2. Difference between a raw EEG signal (left) vs band-pass filtered signal (right) 34
Fig 4. LDA hyperplane (Palaniappan 2014) 36
Fig 5. BCMI proposed methods. 38
Fig.6 EEG locations 39
Fig 7. Record Simulink modules 39
Fig 8. EEG record procedure 40
Fig. 11. CSP components for alpha-beta bands 43
Fig. 12. CSP components for gamma band signal 44
Fig. 13. Mean Covariance Matrices in Alpha-Beta bands 45
Fig. 14. Mean Covariance Matrices in Alpha-Beta bands 45
Fig. 15.BCMI Simulink module diagram design 48
Fig. 16. Riemannian distance from C1 to C2 53
Fig. 17. Speech Imagery BCMI scheme. 54
Fig.18 EEG locations 54
Fig. 19. Simulink model for filtered recording 55
Fig 20. Second set of experiments procedure 57
Fig 21. Time-frequency decomposition for /ee/, /ooo/ and rest classes in alpha-beta band 58
Fig 22. Time-frequency decomposition for /ee/, /ooo/, /mid/,/low/,/high/ in high gamma band 59
Fig. 23. Simulink Diagram of BCMI 62
Fig. 24 Decision tree diagram 63
Fig. 25. Time frequency representation of class /ee/ and /ooo/ for alpha-beta band 65

LIST OF TABLES

Table 1. BCI methods comparison 27
Table 2. Methods accuracy 46
Table 3. MDM for two locations and 5 classes Accuracies 60
Table 4. Online System Confusion matrix 64

ABBREVIATIONS

BCI Brain Computer Interface
CSP Common Spatial Patterns
ECoG Electrocorticography
EEG Electroencephalogram
ERD Event Related desynchronisation
ERP Event Related Potential
ERD Event Related synchronization
FFT Fast Fourier Transform
fNIRS Functional Near-Infrared Spectroscopy
FIR Finite Impulse Response
LDA Linear Discriminant Analysis
ICCMR Interdisciplinary Center for Computer Music Research
ICA Independent Component Analysis
MDM Minimum Distance to Mean
MIDI Musical Instrument Digital Interface
PSD Power Spectral Density
MI Motor Imagery
SVM Support Vector Machine
SEP Sensory Evoked Potentials
SMR Sensorimotor Rhythms
SSVEP Steady State Visually Evoked Potential
VEP Visually Evoked Potential

CHAPTER 1

INTRODUCTION

1.1 Motivation

I had to choose a dissertation project topic from an extensive list of options to finish my postgraduate program in artificial intelligence at the University of Essex. I wanted to feel engaged with whatever the topic was. This was the moment when I thought about merging what passionates me, music, with the computer science field I was studying. One way of doing so was any sort of self composing music, such approaches were already existent and were not really of my interest. While attending one of my lectures about robotics one statement from the professor stuck into my mind. He mentioned that any world phenomena could be measured and could be learnt from computers. Having that in mind I thought, If I could measure my brain activity when I am thinking about music then the computer could learn to differentiate the information and use it to produce music. I was lucky that one of the dissertation options was an opportunity to propose a project about Brain-Computer interfaces (BCI) to one of the supervisors. I came with the mentioned idea to Dr Ian Daily who kindly made it look clearer, introduced me to the fascinating field of BCI explaining the relationship between its different components as a point of combination between neuroscience, biomedical engineering, electronic science and computer science. He also introduced me to Prof Eduardo Miranda's work mentioning he was a specialist of BCI approaches with music as a new component and the leading pioneer of the Brain-Computer Music Interface (BCMI) concept.

I remember looking into Eduardo's work and feeling impressed and inspired by hi's science-fiction-like approaches going from interesting AI-generated music pieces to biocomputing instruments. However, the work that touched the most was related to assemblies of people with motor impairments being able to interact with music through BCI systems.

My dissertation goals then became clearer, I had noticed that current BCMIs approaches needed from a stimulus to work that came from external devices, I had the user experience in mind and used my interpretation of thinking about music as the paradigm. For me, the closest cognitive activity to music is the imagination of producing a sound with the mouth.

Therefore the aim of my previous work solidified to test the feasibility of a BCI system controlled by the attempt to imagine pronouncing a sound. That work yielded solid results that prove that Electroencephalogram (EEG) information from imagined vocalization can be clearly discerned from the rest state and therefore used for a control system. Such results were promising but the dissertation coverage wasn't wide enough to allow me to build a proof-of-concept of my idea.

I had the desire to bring the idea to reality and have a prototype built. Therefore I contacted professor Miranda who liked the BCMI approach with Speech Imagery and kindly invited me to apply for the Research Master in Computer Music and to be part of the Interdisciplinary Center of Computer Music Research at the University of Plymouth.

I was sure this new step would allow me to go deeper into the understanding of the brain's functionality, biological signals processing, EEG decoding methods and our brain's relationship with music.

1.2 Research Questions

This project conducts research at the intersection of technology and music through a non-traditional approach to music through BCI. Conventionally, in order to play music, we require motor interactions from our body, however, the project's outcome is to build a way to manipulate music that dispenses of the mentioned motor interactions. BCI technology encompasses different disciplines, neuroscience and biomedical engineering, the former studies how the brain and nervous system works while the latter how the information from the brain can be measured and interpreted. These two shape the core components of this research.

BCI has developed exponentially since its beginnings in the '90s, currently, there are multiple state-of-the-art methods to extract brain information, as there are different types of neural activity each BCI can focus on. The neural activity represents any cognitive task or state the human's brain is capable of managing e.g. visual recognition, awareness, motion or audio perception. The measurement methods and the aimed neural activity define the BCI paradigms alongside its limitations.

Some indicators of ideal BCI control approaches are: communication rate, portability and usability of the system. Thus, the objective of this research is to find an accurate BCI paradigm that enables the system to control music and enhances the user experience of previous approaches, therefore the research is addressed to answer the following questions.

RQ1 *Can we design a control BCMI system for instrument performance with a desirable musical experience?*
EEG is the most common and widely researched BCI method, this is because of its portability, optimal temporal resolutions and relatively cheap costs (Palaniappan 2014) (J. R. Wolpaw et al. 2000).

EEG measures the activity at the scalp, these measures are electromagnetic fields caused by the electrical communication of neurons, the years of EEG research have helped to distinguish certain potentials that have proven useful to build control BCI systems, that is the case of Visual Evoked Potentials, positive deviations of the signal caused by visual stimulus, researchers have used this paradigm to control devices as computer cursors, spellers or robotic limbs. This paradigm has achieved desired results when aiming to control music as it can allow a user to select among different sounds to play as in Flex on (Miranda and Eaton 2014) or Activating Memory in (Miranda and Castet 2014) where the users selected the next pieces to display on a digital sheet music, however there is one documented approach to perform over a single musical instrument, joyBeat on (Miranda and Eaton 2014), where the user could select the activated steps of a drum sequencer.

Event-related (de)synchronization (ERD/ERS) of sensory-motor signals is another BCI paradigm presented during preparation stages and at the moment of motor execution and Motor Imagery (MI). MI-based BCI is an attractive paradigm as it does not require any stimuli (Pfurtscheller et al. 2006), there are a couple of documented approaches for music with MI, one as an hybrid method with a visual potential in (Eaton, Williams, and Miranda 2015) that allowed to select between instruments and the other in (Deuel et al. 2017) where a user could select between two options to play the next note in a programmed sequence in a music generation software by imagining right hand opening or grasping.

To answer RQ1 the mentioned paradigms should be studied within a perspective to implement a BCMI to perform a single instrument, to differ from most of the previous BCMI approaches that use pre recorded sounds or sequences, and try to achieve an experience of desirable degrees of freedom as playing an instrument via conventional way.

RQ2 *Is it possible to build a proof-of-concept control BCMI system that enables the user to play a musical instrument with their attempt to imagine pronouncing vocal sounds?*

Even though MI-based BCI dispenses with stimulus, it is limited in the degree of freedom, having as options the possible movements of the different limbs, and it can be unintuitive in some cases, e.g. if we use imagining left-hand movement for turning off a music player.

Speech Imagery (SI) has been proposed as an alternative BCI paradigm to complement MI. SI is the mental activity of imagining to speak aloud without moving articulators or producing sounds, this paradigm can be intuitive as the imagined words could be associated with individual commands and, in theory, it has a wide range of options to classify into commands as there is a wide range of sound combinations that are possible to be pronounced with the mouth. There are promising results when it comes to decoding SI from EEG, some works as DaSalla in (DaSalla et al. 2009) or Deng in (Deng et al. 2010) achieved accuracies up to 70% when classifying vowels and phonemes, such results mark a theoretical possibility for the implementation of a control system which motivates to research SI paradigm further and answers RQ2.

1.3 Research Methodology

This section presents an overview of the research methodologies considered by the research. Firstly the brain activity detection method, second the EEG decoding methods and finally the programming and analysis tools used to build the final delivery.

1.3.1 Wireless and dry EEG Headsets

EEG sets come in different configurations depending on the manufacturer, in general, they can be dry and wet. Wet EEGs use gel to improve the conductivity between the scalp and the sensor while Dry electrodes usually have non-flat shapes to ensure surface coverage and dispense with any conductive liquid. Dry electrodes usually have poorer signal quality as their contact with the surface may change with any minor movement. This thesis adopts dry EEG as a method for the brain's signal acquisition. Medical-grade EEG headsets are prefered among BCI research because of their reliable performance, however, this research relies on a wireless headset taking into consideration to answer the research questions. Previous ICCMR projects have attempted BCMI approaches with portable headsets but usually had low performance due to poor communication rates. The research uses an 8-electrodes EEG set manufactured by G.tec Guger Technologies, considering usability criteria and thinking about the possible use of BCMI to be deployed in musical scenarios.

1.3.1 EEG decoding methods

EEG encodes information about how different brain areas communicate over time, therefore this information is based on the power variation and location of those signals, but the signal is not phase-locked to a certain frequency there is spectral information about neural activation that has driven the researchers to use statistical signal processing methods to try analyzing and extracting useful features. EEG features, as mentioned, come from three important sources. Spatial, spectral and temporal information. For each source, there have been different methods that proved useful to enhance overall BCI performance.

This research is testing different sets of methods to achieve optimal performance, the first spectral method in consideration is filtering, as it is known that different mental states are prominent between different frequency bands, filtering is also useful when reducing noise usually related to the power line (Eduardo Reck Miranda 2014).
BCI systems require two important steps in the decoding process: feature extraction and classifications, both performed by statistical methods. The feature extraction methods that his research adopts are Common Spatial Patterns (CSP) which is a spatial filtering algorithm that transforms the EEG data and allows better discrimination between different trials. Covariance Matrices (CM) are considered as well as a transformation for time-series data, in this case, the EEG single-trial raw signal, into a matrix which elements are the relationship between the single sources or

electrodes. These methods produce vectors, a feature that needs to be learnt from a statistical model for further classification and testing.

Classification methods to try are machine learning methods that proved useful for EEG signals. Linear Discriminant Analysis (LDA) is a popular algorithm that separates the data with the use of hyperplanes. Support Vector Machines (SVM) is a popular method among MI-based BCI that use support vectors to find an optimal hyperplane or discrimination line to separate different sets of data based on where they relay in a multidimensional space. The last classifier method adopted by this research is a Minimum Distance to Mean (MDM) classifier. MDM has a straightforward discrimination logic based on the distance that a data point to classify has to the average point of each class available this method use distance functions to perform the classification, because of CM characteristics of lay in a multidimensional space Riemannian metrics are considered for the MDM classification. A deeper explanation of the methods is presented in Chapter 3.

The appropriate selection of a pipeline with the mentioned methods that achieved high performance when analyzed offline is then taken to test online and answer RQ2.

1.3.2 Programing and analysis tools

In order to identify adequate methods to answer RQ1 and proceed into the attempt to build an online approach and answer RQ2. The first tool is selected to supply the way the EEG data is going to be collected, this is restricted to the headset manufacturer. G.Tec provides Matlab's Simulink interface to connect the Bluetooth device and do the recordings, therefore Matlab's Simulink is the main tool to develop the signal acquisition experiments and design an online simulation of the BCMI project. Matlab's Pyschtoolbox, used in this research, is a programming interface to do visual programming to present stimulus and design EEG experiments. Matlab's Eeglab toolbox is used for offline EEG analysis; it offers a wide range of functions to read, preprocess and analyze EEG data. Python environment is also used to further decode EEG epoched data. MNE module for python is adopted in this work as it has useful functions as CSP and filtering. The project uses the Scikit-learn module which has ready to use machine learning algorithms and pyRiemannian, a module with Riemannian metric functions for EEG data.

1.4 Thesis Structure

Including the current chapter, this thesis consists of five chapters that are described below.

Chapter 2

This chapter is a survey of the different BCI paradigms that have been used to create BCMI approaches for control systems. The survey evaluates 2 criteria, the dimensions of control and the extensiveness of the configuration of the method to be used in a subject-independent approach. The methods surveyed are P300, SSVEP and ERD/ERDS. Led by RQ1 the survey highlights ERD/ERS, even though it dispenses with stimulus, it is present with limitations because of the dimensions of control it can have, the survey introduces Speech Imagery as a complement, stating the activity induces ERD/ERS plus Speech Evoked Potentials (SEV), potentials that have not been totally understood yet.

Chapter 3

This chapter studies EEG decoders for Speech Imagery, driven by RQ2 the chapter mentions some of the most common methods to decode Motor Imagery from EEG that could be used to decode SI. After carrying the first set of EEG experiments, to a single participant due to ongoing COVID pandemic restrictions, using an electrode location focused on the motor-cortex area and recording user attempt to imagine vocalizing vowel /e/, the vowel /o/ and staying in resting state. Two decoding pipelines are tested with the obtained data, and a comparison is made based on the accuracies, and the complexity of the algorithms for implementing an online approach. The first pipeline covers a popular decoding group of methods, this is CSP as feature extraction procedure and both LDA and SVM separately as classifiers. The second pipeline has covariance matrices as features sent to a Riemannian MDM classifier. Both pipelines and results are obtained based on four different frequency bands alpha (7-16Hz), beta (16-28Hz) and low gamma (28-60 Hz) and high gamma (60-100Hz). The overall results are similar between the two pipelines but the latter is selected for the remainder of this thesis because of its ease of implementation for online running.

Chapter 4

This chapter is based on research between different sets of cognitive tasks to measure and between different electrode locations, similar to the previous chapter the research involves reaching the most optimal configuration of EEG that delivers the best accuracies. The research test between the previous EEG location and a new location that focuses on the left hemisphere as it is known as the responsible for language generation and comprehension, the latter location slightly higher accuracies and more discriminative information in the gamma band. The research in this chapter makes a variation on the speech imagery task, first, a rhythmic repetitive version of the vowel e imagined twice with high pitch /e e/, for vowel o the task is to imagine three times /o o o/. Second, the imagined speech of /high/, /mid/ and /low/ words.
The repetitive imagined speech led to higher accuracies while the analysis of the imagined words yielded the word “mid” to have more discriminative information when being pairwise classified. Therefore the final implementation of the proof-of-concept is based on an EEG experiment using the cognitive tasks of imaging

vocalizing word /mid/, twice high pitched sustained vowel e (/e e/) and three times vowel o (/o o o/). The prototype core implementation had a pairwise classification logic where the higher frequency bands decided whether the trial belonged to e class while the lower frequency bands whether it was o, mid or resting state. The overall classification of the online prototype wasn't optimal, however, it allowed to build a proof-of-concept.

Chapter 5

This chapter is the thesis conclusion and summarizes how the work approached the research questions, the chapter presents also a foundation for future work based on the limitations and possible variations to develop for future BCMI systems.

CHAPTER 2

BRAIN-COMPUTER INTERFACE METHODOLOGIES FOR MUSIC APPLICATIONS

2.1 Overview

Brain-computer interfaces (BCIs) aim to create a direct communication path between the user and a machine dispensing with muscular interactions. This type of system aims to give access and improve the lives of people with motor impairments. This chapter presents a survey of Brain-Computer Interface methods using Electroencephalography (EEG) to build a Music Control System, evaluating Visual Evoked Potential P300, Steady State Evoked Potential (SSVEP) and Event-Related Synchronization and Descincronisation (ERS/ERD). These methodologies have been proved useful to implement systems that have music as a component to control. The survey takes into consideration the limitations of BCI control systems such as usability, generalization capabilities and the ease of taking the system out of an investigation field. These investigations take as criteria the dimensions of control that the methods can encompass, the complexity of their preparation processes and their ease of adaptation for multiuser purposes. Each technique has different properties and advantages, the chapter concludes by reviewing the technique that helps to answer RQ2 and examines the possibilities to build a working BCMI prototype.

2.2 Introduction

Neuroscience and biomedical engineering are persistently looking into more optimal ways of communication between people and the environment while finding ways to enhance people's health. BCI research has contributed to other external disciplines, some EEG-based BCI has inquired into oneirology by studying dreams (Horikawa et al. 2013), affective science with emotional state recognition(Othman 2014), physiology with systems to assist motor rehabilitation (Al-Qazzaz et al. 2021), neurology to help to identify epilepsy (Jaseja and Jaseja 2012), to mention but some.

In June of 1999 was the first international meeting of Brain-Computer Interface (BCI) Technology, had 22 research groups from different countries shaping the definition of BCI as a *communication system that does not depend on the brain's normal output pathways of peripheral nerves and muscles,* therefore, BCI aims to give users a communication interface with computer systems straight from brain activity (J. R. Wolpaw et al. 2000).

Within different variations of BCI, one core element of their design is the measure of brain activity. Based on how this measurement is acquired, two BCI methods can be mentioned, invasive methods with electrocorticogram (ECoG) and non-invasive with electroencephalogram (EEG), magnetoencephalogram (MEG), positron emission tomography (PET) among others (Jonathan R. Wolpaw et al. 2006).

Based on the BCI's goal, two major kinds of BCI can be mentioned, active BCI where the user intentionally tries to control her/his brain activity to have the desired output on a system such as directing the computer mouse cursor (Molina-Cantero et al. 2021) or neuroprostheses (Muller-Putz et al. 2019), to mention but few. And passive BCI, on the other side, where the system reacts to the user's current brain state and activity, trying to adapt the system to another desired state, as in (Daly et al. 2014) where the system tried changing the current user mood depending on the current state being measured.

Over the years different approaches to control BCI have emerged, all of them aiming to give communication paths to people with motor impairments as locked syndrome affections. From early BCI destinations as a wheelchair (Diez 2018) or digital spellers (Li 2011) to modern control destinations such as virtual reality (Kim et al. 2021) or multimedia controllers (Tseng et al. 2015).

It was natural to expect music to be the destination of a control BCI. The idea to use the brain to control sound emerged in the 1960s where Alvin Lucier used electrodes on his own scalp and amplified its signals to loudspeakers that were directly placed to percussion instrument surfaces and membranes and made them vibrate (Eduardo Reck Miranda 2014). Such an approach inspired computer musicians with this new dimension of control, in 2005 Hitenberg and Bair presented a parametric sonification using EEG, they mapped different frequency bands to instruments using the musical instruments with digital interface (MIDI) communications. Current BCI findings allow new computer music approaches but challenges within BCI keep their use to real-life performing scenarios. One problem that BCI encounters is the quality of the signals measured from the brain, hits refer to a signal to noise ratio (SNR) and for EEG it requires highly confident devices plus processing algorithms that process the signals.

From the different BCI applications, Brain-Computer Music Interface (BCMI) looks for introducing music to the system whether as a stimulus or as a direct mapping of the signals for a passive practice or uses music as the final target to control on active systems, in the latter case, for example, controlling a musical device, or parameters of music samples while they are being played as well as deciding which arrange may be coming next.

This survey focuses on methods for active control of a music system that aims to find channels to the voluntary control of music parameters and give approaches to musical activities for people with motor impairments. The BCI methods surveyed in this paper use EEG as a brain activity acquisition tool. EEG is selected because of the background work and promising achievements on this technology and because of its preference from the actual research community and the advantages it has among other methods as being portable, and relatively economical to acquire.

2.2.1 Electroencephalogram (EEG)

Electroencephalogram (EEG) was first described in 1929 by Hans Berger and became one of the most used tools for the diagnosis of neurological disorders like epilepsy. Later in the 1970s, several scientists developed simple communication systems that were driven by this electrical activity, and now is the most used method for BCI systems because of its relative portability, cheaper costs and optimal temporal resolution (Eduardo Reck Miranda 2014) (J. R. Wolpaw et al. 2000).

EEG is a type of oscillating electrical potential recorded from the scalp surface, the electrical potentials are caused by neural activations on the brain, this activation creates electrical dipoles in different regions of the brain and, because of the brain's good conductivity, these signals can be captured from multiple locations in the scalp, this activations are of small amplitude in microVolts range and have an attenuation caused by the skull and scalp (Palaniappan 2014). EEG is usually made from different electrodes placed on the scalp. One standard configuration of electrodes is 10-20 electrode system called this way because there are 10 to 20 millimeters distances between the electrodes ("The Ten Twenty Electrode System: International Federation of Societies for Electroencephalography and Clinical Neurophysiology" 1961), with arrays of 32, 64 up to 256 electrodes. Each electrode uses a naming system based on their position in the cortical area they are covering (P for parietal, F for frontal, C for central and O for occipital) and followed by an even number for the right hemisphere and odd for the left.

Some changes on the EEG signal (Evoked potentials) that occur after certain events are remarkable as they are a general response of the brain to a stimulus or cognitive task activations. Evoked potentials highlight the anticipation and perception activity of the brain. These changes are helpful when trying to extract useful information from recordings. Some of the most used EEG signals potentials are Visual Evoked Potential (P300), steady-state visual evoked potential (SSVEP), event-related synchronization and desynchronization (ERD/ERS) (Palaniappan 2014). This chapter is directed towards choosing the appropriate EEG paradigm to design a BCMI and the review of possible technical limitations and their implications on helping to solve RQ2.

2.3 Criteria of Comparison

The section explains the framework used to compare EEG paradigms that have proved useful for BCMI applications and their examples while defining the link of the criteria with the research questions.

2.3.1 Dimension of Control

Any BCI control system aims to control a final device, depending on the device the system would consider what are the options or parameters to control. Basic BCI applications can present binary classifications problems as the left or right direction selection (Lee et al. 2019) or evidently the options to select from can be a wide range option depending on the applications going to from characters in spellers as (Yonghao Chen et al. 2021) with 160 targets or (Hubner, Schall, and Tangermann 2019) with 36 options in chess game application. However, the dimensions of control encompass the number of possible options, the need for a second dimension as parameters from a first choice is imminent. For example, a BCI controlled wheelchair as in (Olesen et al. 2021) requires an intensity measure in order to control the speed. Therefore such a problem can be taken as adding more options to the first dimension of control but an ideal system design would have a set of options to choose from and the capability to change the characteristics of those selections.

Considering the RQ2 of a system that controls an instrument, it is accurate to analyze that this involves multiple dimensions of control, having as main dimensions the musical note to play. Then many possibilities for control can be stated such as the duration of the note or its velocity going up to instrument parameters such as type or tuning.

Dimensions of control is a criterion of comparison for BCMI systems as it allows to define the limitations of each BCI paradigm in accordance to the number of options they select from, inaccurate and optimal performances.

2.3.2 Usability of the System

The usability of a BCI system starts with the consideration of how the brain information is taken, as aforementioned this survey is based on EEG methods. There is an existing invasive method of EEG, ECoG, where the electrodes are placed directly on the brain tissue, for reasons of safety non-invasive methods are preferred even if the skull causes a signal attenuation and quality reduction.

Usability also encompasses the capabilities of each BCI method to work properly in different environmental conditions, taking into consideration an application with music the response to the user to the possible sound feedback, ideally, should not interfere with the system performance. There could also be limitations for the visually evoked methods depending on the luminosity of the intended scenario to test the system. One important term of usability taken into account is the capability of the method to work with several users, EEG signals vary considerably between people, this is a major challenge in BCI research and one of the reasons that keep BCI under investigative conditions. Usability encompasses as well the user experience, if a BCI control system for music is intended, an ideal experience is to replicate the profound feeling of musical instrument performance, musical interfaces require a high degree of flexibility. This criterion traces the limitations of the EEG-based BCI paradigms as the visual evoked methods need the user to stare at flickering stimulus displayed on screens that may cause tiredness to some users, and in case of musical instrument performing may not present an appealing experience. And the imagery-based paradigms, on the other hand, have narrowed degrees of freedom as the overall system accuracy decreases with the number of imagery classes they can have as seen in (Krishna et al. 2018).

2.4 EEG-based paradigms for control BCI

2.4.1 P300

P300 potential is the brain's natural response to a stimulus, this signal is a positive deviation from baseline, time-locked around 300 and 600 milliseconds after stimulus onset, the signal latency and amplitude vary according to age (Polich, Ladish, and Burns 1990). However, it can also be provoked by motor, cognitive and sensor events. Different stimuli provoke this potential in different scalp topography (Kiyi, Oztura, and Yener 2021). Some studies presented that P300 approaches based on tactility were more optimal for identification than vision or sound (Mao et al. 2021; Z. Chen et al. 2020). P300 is present within the alpha band at 8Hz, usually, a target stimulus that the subject was expecting, but P300 can be elicited in relation to the *odd-ball* paradigm when an infrequent event happens after a user experiences a set of frequent events (Hucker 2007).

P300 event-related potential has maximal locations in central cortex areas as Fz, Cz and Pz however greater location densities provide higher accuracies (Palaniappan 2014) (E. R. Miranda et al. 2011). P300 has been used in BCI control systems because they represent the user's choice over stimuli. In this method, a set of options are presented on the screen and the participant counts how many times the desired option flashes, this selection requires repetitive trials and signal process averaging or components discrimination to identify the desired stimulus because a single P300 can be shadowed by normal ongoing brain activity (Grierson and Kiefer 2014). Visual P300 is optimal when presenting a good amount of options, as mentioned before, a P300 system may present as many suitable options as can be fit on screen.

P300 was adopted to build musical interfaces, Grierson and Kiefer (2014) researched the implementation of a P300 composer. Some good approaches for BCMI with P300 are mentioned in (Grierson and Kiefer 2014) where the set of options to choose are different musical parameters as in P300 Composer, where the options are an array of notes or P300 DJ choosing the next song being played. This technique offered a gaze-based control for musical instruments, however, the movement restrictions and the user's vision fixation may not be desirable. Considering a dimensionality of control, P300 allows unidimensional communication, therefore adding new dimensions for musical instruments could be implemented with system variations where more than one selection is needed.

Some BCI systems with visual P300 can be used without training, as the potentials are presented in animals and humans as a response to stimuli. Then P300 analysis, in essence, involves identifying whether there is a prominent fluctuation from baseline. However, as for all event-related potentials, EEG signals are subject-specific and tend to vary even between recording sessions, making it hard to design subject-independent decoders.

2.4.2 Steady-State Visual Evoked Potential

SSVEP, for short, is the repeated brain's response to stimuli, it belongs to the sensory-evoked potentials (SEP's). Unlike event-related potentials, SEP are phase-locked according to the stimulus sending frequency (Sun et al. 2021). For visual stimuli, when a flashing sign is presented at frequencies higher than 6 Hz, the interval between stimuli is substantially shorter than the duration of the response therefore the responses to individual stimuli overlap, presenting significant amplitudes at the stimulus frequency or its harmonics, or harmonics, patterns in the EEG signal (Norcia et al. 2015). SSVEP analysis takes advantage of these frequency-domain characteristics having maximal signal at the visual cortex on the occipital region. BCI applications with SSVEP have shown good performance with only 2 electrodes; however 4 to 12 electrodes achieved more reliable results (Kolodziej, Majkowski, and Rak 2016). SSVEP makes the BCI system dispose of the need for performing numerous trials since the responses are consistent and high in amplitude. SSVEP allows users to have active control by presenting different visual stimulus flashes at unique frequencies, as mentioned, the activity in a certain signal frequency can be detected with spectral analysis. The optimal information transfer rates (ITR) that SSVEP has achieved, have had it as a prefered methodology for control applications such as games (Martišius and Damaševičius 2016), robot navigation (Farmaki, Christodoulakis, and Sakkalis 2016), rehabilitation (Chu et al. 2014), to mention but few techniques.

Miranda et al.(2011) incorporated SSVEP so a user could choose between an array of musical notes (Eduardo Reck Miranda and Castet 2014). Another interesting approach for a control system is Flex on (Eaton, Williams, and Miranda 2015) where the user was presented with 4 options as a flickering stimulus on the screen, each of the options had assigned a different soundscape instrument and the duration of the selection was defined the length being played.

One advantage that SSVEP presents is that the user does not need any physical effort but to haze at the desired stimulus (Palaniappan 2014). Another is that SSVEP can bring a second dimension of control with the duration that is being held while gazing at the stimulus. A longer period of time elicits higher amplitude responses, which can be added as intensity information useful to bring more options to command the final destination.

2.4.3 Event-Related Synchronization/Desynchronization

ERD and ERS are potentials related to sensorimotor activities, these activations are signal responses to the action of sensorimotor tasks that generally have a non-conscious mechanism (Jeannerod 1995). These rhythms are prominent on alpha and beta bands (7-30 Hz), ERD is a negative deviation during the preparatory stage of body movement, reaching its maximal at the motor onset, while ERS is its counterpart, a positive deviation occurring at the same time. The location of these potentials depend on the location of the limb intended to move, this is, if a right-handed user moves the right hand an ERS will occur on the right hemisphere motor cortex area, while an ERD would occur on the opposite. Is known that movement imagery (MI) presents a planning stage causing ERD/ERS(Palaniappan 2014).

One advantage of having ERD/ERS is that it dispenses with any sort of stimulus, its dependence relies on the user's imagination. However it is prone to low performance related to user concentration, as environmental and user internal factors play an important role. This method is more sensitive to artifacts like ocular and muscle movements than its visual counterparts, as the movement imagination can lead to minor ocular or muscular movement which is also captured by EEG (Yao Chen, Zhang, and Wu 2019). There are consistent results when identifying the limb's movement side which has led to present interesting control approaches as a wheelchair with 5 options from MI tasks(Jiang et al. 2014) or to control prosthetic limbs as (Iwatsuki et al. 2020).

In music an approach is mentioned in (Eaton, Williams, and Miranda 2015) where a BCMI system adds another dimension of options with ERD/ERS to the one already present for SSVEP, so the selection on the latter would be the pattern for an instrument to play while the former would choose between 2 instruments.

MI-based BCI is attractive as it does not require stimulus from additional equipment. However it may be limited in the degree of freedom when using it to control a device and, while it is intuitive for movement control, it can be unintuitive in the case of other applications, e.g. if we use imagining left-hand movement for turning on a heater.

Speech Imagery (SI) has been proposed as an alternative BCI paradigm and a complement for MI. SI is the mental activity of imagining to vocalize words or sounds aloud without moving articulators (Fujimaki et al. 1994). This paradigm could be intuitive as the imagined words can be associated with individual commands and, in theory, it has a very wide range of options to classify into commands as there is a wide range of sound combinations that are possible to be pronounced with the mouth.

As with MI, the information of SI is encoded on the power and location variation of the EEG signals; these potentials have been described as Speech Related Potentials

(SRP) and their characteristics, unlike MI - ERD/ERS, are not well understood, however, studies as in (Kovacs, Winkler, and Vicsi 2017) show that SRP of phones differ from each other depending on the manner of articulation both in amplitude and scalp location and (Wang et al. 2013) showed ERD and ERS potentials in preparatory stage for speech. This could provide a novel approach to BCMI systems for music controlling when the speech imagined, properly classified, can be mapped to musical parameters.

Some BCI researches with speech imagery have reached promising classification accuracies when analyzed offline as in (Riaz et al. 2014) that compares different classifiers from regressors to neural networks. Some of the current best results usually come from complex machine learning classifiers that present challenges for online practices because of their computational cost that adds undesirable delay to the overall response of the BCI.

2.5 Comparison

This section presents an in-depth comparison of the 3 types of EEG-based BCI paradigms for music control. EEG is an optimal methodology of brain activity detection mainly because of its safety, portability and relatively easy configuration steps. EEG configuration varies between each methodology, methods such as P300 and SSVEP based on visual stimulus can work properly with a reduced number of electrodes achieving optimal results as work in (Sozer 2018) with single electrodes. However higher densities enhance the overall system accuracy. For MI and SI methods, the topography of electrode placement is usually set in the motor cortex area and around the left hemisphere, Broca and Wernicke's area as they studied to be responsible for language production and comprehension (Friedman et al. 1998; Binder 2015). For ERS/ERD higher electrode density induces better results as the potentials spread around all scalp because of the brain's high conductivity.

In terms of the environmental condition of system deployment, visually evoked methods require controlled light conditions, specifically dimly lit scenarios. which is a constraint for musical approaches if deploying in live musical scenarios. For MI and SI, the environmental conditions should ensure the user's concentrations (Zapała et al. 2018), which may also represent a constraint for BCMI deployments based on this method.

Another usability context to compare is the user experience. As mentioned P300 and SSVEP require a visual stimulus that comes from an external display, the extensive use of such a method can cause user fatigue and depending on the options to choose can be non-intuitive. On the other hand, the sensorimotor rhythms such as MI and SI can be more intuitive especially in the case of SI, as any word could be associated with a command, these later methods dispense with the need for any external devices as their basis is the user imagination.

When analyzing subject-independent generalization capabilities, P300 and SSVEP present advantages as these potentials are natural responses of the brain to a stimulus which means, in theory, the same system design is capable of being used by multiple users, however, the response time of the evoked potential is subject to change based on users age among with its frequency response (Oikonomou et al. 2017). In the system, EEG installation stages may not be complicated depending on the number of electrodes on the headset. The user performance may dispense with training as the only cognitive task to perform is merely the act of gazing. On the other hand for MI and SI as the system requires every user to be trained, and the decoding models would learn from recordings of motor imagery or speech imagery tasks respectively, that in general vary from frequency predominant values, then the approaches for these methods are strictly subject-independent, with the need of training and fitting stages before the actual performance (Joadder et al. 2019). However new statistical models alongside artificial intelligence algorithms have shown capabilities for generalizing inter-subject trials of MI and SI (Kaongoen, Choi, and Jo 2021). Some studies have presented the use of statistical models alongside Riemannian metric to generalize the correlation between channels for single-trials along different users using covariance matrices, this proposes to represent a generalized matrix for a certain class that can be used as initial information to further subjects, a concept known as transfer learning (Congedo, Barachant, and Bhatia 2017), the use of deep learning models such as recurrent neural networks have proven useful to build subject independent classifiers (Tibrewal, Leeuwis, and Alimardani, n.d.), however, the computational resources needed for those algorithms may be a constraint for BCMI applications as rapid time responses are expected when performing with music.

To compare dimensionalities of control, the options each method can have as classes to discern are taken into consideration as a first dimension capability. P300 and SSVEP can theoretically have as many options as they can be displayed on a screen or flickering device in the case of SSVEP, however, the options need to fit adequately in order not to cause visual fatigue (Gembler, Stawicki, and Volosyak 2016). For MI, the options are limited to the limbs intended to move, e.g. right foot, left shoulder, etc.

For SI the system can have as many commands as the models are capable of learning, theoretically, a SI approach may have as many options as many sounds combinations can be pronounced with the mouth. For these two paradigms, the current studies reveal that the increment of options leads to an accuracy decrease.

Only SSVEP has been proved to add another dimension of control, as mentioned before, the duration of the staring has been used to control parameters for musical approaches as work in (Eaton, Williams, and Miranda 2015). Some studies on SI have had optimal accuracy when decoding rhythmic repetitions of imagined speech, one possibility of extra dimensionality for SI may be to classify the number of repetitions of a certain vowel or phoneme being imagined.

Table 1 below, summarizes the characteristics of the three methods with the proposed criteria

	Dimensions of control	Preparation steps	inter-user generalization
P300	1 dimension selected from as many options as can non-invasively fit the screen	Need of repetitive trials for ensuring decision. Extensive training is not necessary.	Can work for many users without prior data collection. Could need some adaptation based on user's brain response time
SSVEP	2 dimensions (choice and duration) selected from as many options as can non invasively	Extensive training is not necessary.	Can work for many users without prior data collection. Could need some adaptation based on user's brain response time
ERD/ERDS	2 dimensions (choice and duration) selected from the number of options that the system could accurately learn	Need of data collection stage for system fitting. User training is necessary.	Need of data collection and fitting stage for each user. Each user added to the system can decrease the overall accuracy

Table 1. BCI methods comparison

2.6 Survey Conclusion

This chapter surveyed the different EEG-based BCI methods for control that can be used in musical applications. Some BCMI approaches look promising and have achieved the goal of drawing near people with motor impairments close to musical control activities. However, every investigated method presents some limitations, because of the subjective nature of the brain and EEG ("Genesis of MEG and EEG" 2019), the systems struggle to build general prototypes that would work for many users, most BCI approaches need stages of signal filtering, source separation, feature extraction and classification to detect useful information from EEG, which represents extensive off-line analysis with complex mathematical operations. This processing pipeline can be a constraint when building online systems because of the computing time, especially music-related as ideally the music control is desired not to have any delay.

Although the mentioned constraints, SSVEP approach is a feasible method for a music system, its two dimensions of control gives some possibilities of design that may bring ideal user musical experiences, this could be the case of a system that shows some notes to play as options on one side and some instruments to the other, therefore a user may select the note to be played and the length of the perform would depend on their gazing time, while have also the option to select between different instruments, This method has the advantage of easy configuration stage that may be suitable for subject independent approaches, working with low electrode density makes it an ideal prospect for new BCMI concepts. However, it has optimal proven performance and can have 2 dimensions of control, it depends on the use of an external device to present the stimulus that may not give a pleasant musical experience.

ERD/ERS methods are interesting because of the fact that they are purely imagery oriented, with no need for stimulus which may bring different musical experiences. Within this method. Speech Imagery is attractive because sustained single vowel imagery may be linked with musical imagery or singing and a robustly trained BCMI could encompass a good quantity of options to control, with an intuitive way of performance. The latter method is investigated in the remainder of this thesis as a first approach to Brain-Computer Interfacing using Speech imagery, even though SSVEP is proven to achieve more solid results, SI has a promising expectation regarding a musical experience.

CHAPTER 3

INVESTIGATION OF EEG DECODING METHODS FOR SPEECH IMAGERY

3.1 Overview

This chapter explores methods to decode Speech Imagery that are feasible to build a proof-of-concept brain-computer music interface (BCMI). The research adopts dry and wireless electroencephalogram (EEG) headsets because of its popularity among other BCI methods. EEG experiments were designed with Matlab's Simulink software, headset manufacturer G.Tec introduced an interface for Simulink alongside the EEG amplifiers that easily enables the Bluetooth connection with the measurement device. Psychtoolbox for Matlab is used to control the display and allows the user to present graphic instructions to the user. The user was asked to imagine pronouncing sustained vowels /e/ and /o/ during. This chapter discusses the performance of methods that proved useful to decode Speech Imagery, on one hand, Common Spatial Patterns (CSP), an algorithm well known to extract spatial feature vectors, which are used as input for SVM and LDA classifiers. On the other hand, covariance matrices as features are classified by a Minimum Distance to Mean (MDM) classifier using the Riemannian metric. Three different frequency bands (7-28Hz, 28-60Hz, 60-100Hz) were evaluated to increase the classifier accuracies. Higher accuracies on pairwise classifications were 62%$\pm$ 2 for CSP+LDA, 63% $\pm$ 2 for CSP+SVM and 65% $\pm$ 5 for MDM. The chapter concludes that the use of MDM is prefered to further developments because of the simplicity of the discrimination algorithm.

3.2 Introduction

Technology develops from human needs shaping people's daily activities including how we interact with it. Art with no exception adopts technology while finding new ways of producing and showing pieces, as it can be seen today, computer art is a powerful means for contemporary expression (Malina 2002). For music, it is difficult now to dispense with digital tools. Computer music production technology has become the core technology of sound manipulation of new media (Zhang and Hou 2021) thus music as a human expression is attached to the development of technology and the exploration of human-technology interactions.

This chapter evaluates methods that can help shape a proposition of a non-traditional approach to music through Brain-Computer Interface (BCI), an interdisciplinary growing field of science.

Starting from an imaginative science fiction idea of a machine that reads the human's mind and produces the sounds that one is thinking about, to define methods proved to be functional for a proof of concept and conclude the feasibility of such approach in a non-theoretical delivery.

BCI, a definition shaped in 1991 in what was the first international meeting of this technology, setting its definition as a communication system that does not depend on the brain's normal output pathways of peripheral nerves and muscles, therefore, BCI aims to give users an interaction channel with computers straight from the brain activity (J. R. Wolpaw et al. 2000). BCI term emerged around the 1970s at the University of California but was Hans Berger, a German psychiatrist, in 1923, singularly responsible for the discovery and development of human brain wave recording (Vaque and La Vaque 1999).

BCI applications can have as goal controlling devices such as neuroprosthesis or the mouse cursor (Knězík and Drahanský 2007), some applications, to help people with neuromuscular impairments in the recovering process (Cincotti et al. 2012) and others studying complex brain functions as semantics in (Mitchell et al. 2008).

From the control BCI applications ideas, the one of having music as a destination emerged around 1960 where Alvin Lucier used electrodes on his scalp and amplified its signals to loudspeakers that directly placed to percussion instrument surfaces and membranes made them vibrate.

BCMI adds Music to BCI applications whether to direct musical mapping brain data, a process known as *sonification*, exploring neural activities involved in music listening or imagery or direct control musical parameters of a device or instrument (Eduardo Reck Miranda 2014).

3.2.1 BCI Control System

As mentioned above, a BCI control system aims to actively manipulate a final device from decoding brain activity, such a system has three main components as shown in figure 1. The brain activity acquisition component is essential in all BCI applications, involving what is the brain data to be extracted and the appropriate measurement methods. The decoding to commands component involves analyzing and processing the brain activity to extract useful information that can be used as parameters to send commands for the final component consisting of the device to control.

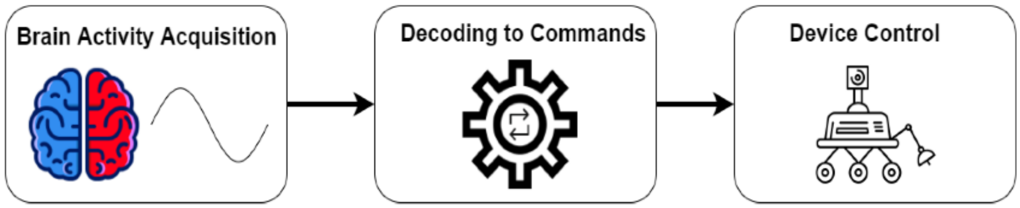


Fig 1. BCI Control System diagram

This chapter analyzes the feasibility to implement a BCMI application, therefore, a BCI system that aims to control a musical instrument as a final device, the next sections in this chapter would describe what are the possible components of such a system and an understanding of why certain methodologies are selected and evaluated within each component.

3.2.2 Brain Activity Acquisition

The brain's complexity both on the density of neurons and the working structure in spatial and time domains makes it difficult to study with precision. Current BCI techniques use three different means of measurement: neural electrical field activations, the blood oxygenation flow known as *hemodynamics*, or the brain's chemical interactions (Orrison et al. 2017). For hemodynamics, there are broad-study methods as Functional Magnetic Resonance Imaging (fMRI) and Near-infrared spectroscopy (NIRS) that imaging the blood oxygenation flow changes in response to neural activity. Positron Emission Tomography (PET) is another innovative imaging method based on short-term radioactive material used in brain chemical reactions; however, this research focuses on the brain's electrical signals as the evaluation marked in chapter 2 highlights the portability capacity present in the measure of electrical activity.

Neural activations on the brain create electrical dipoles with potentials or electrical fields that can be measured by electrodes in different regions of the brain because of its high conductivity. Electroencephalogram (EEG) is the recording of such potentials, of microVolts amplitude, with electrodes placed in the scalp surface, EEG is one of the most used BCI methods because of its portability, relatively cheap costs and optimal temporal resolution detecting changes on a millisecond-level (Palaniappan 2014)(Palaniappan 2014; J. R. Wolpaw et al. 2000). Magnetoencephalography (MEG) can also record electrical activity with the use of highly sensitive devices instead, Superconducting Quantum Interference Devices (SQUID) that can measure subtle changes in electrical fields and give a well defined electrical activity image as the outcome (Schnitzler and Gross 2010).

BCI methods to detect electrical activity can also be invasive as electrocorticogram (ECoG) where electrodes are placed directly on the brain surface through medical surgery, resulting in better spatial resolution.

From the aforementioned methods, EEG is a preferred method, first because of its wider use for BCI research with a solid community and currently achieved advances, second because of its possibility to be taken to a practical, non-theoretical approach. Because it has portability capacities as evaluated in chapter 2 and lastly because of an 8 electrode EEG equipment is available to conduct this research.

Brain activity acquisition also known as signal acquisition has a core procedure that is selecting what cognitive tasks, of which activity produced, is desired to be recorded. It is clear now that different regions in the brain are responsible for different cognitive tasks therefore a pre-selection of the brain region or regions to study is needed while selecting the appropriate acquisition method, for example, the acquisition experiments should focus on the occipital region of the brain if we are interested in a visual activity or they should focus on the temporal regions if the interest is the activation related to an audio task (Sitaram et al. 2008).

The cognitive task that addressed this study is Speech Imagery which brain activity recording is focused on the sensorimotor and language production regions of the brain, as evaluated in Chapter 2, among some of the studied electrical responses from the brain to different cognitive tasks, Speech Imagery can lead to the development of novel ways of BCI systems for control, the next section explains what is the brain activity expected from this task and what methods can be useful for developing this BCMI proposal.

3.2.3 EEG Signal Decoding to Commands

To understand the brain activations related to different cognitive tasks, the focus is on where this activity happens in the brain and how it evolves in time. Different signal acquisition methods have different features and techniques for decoding the data that reflects the spatial and temporal activations, the techniques are called signal processing, and their global goal is to clean the obtained raw data, unveil from the whole the most useful pieces of data, a process known as feature extraction, and use computational models to try to classify whether the data or signal being analyzed comes from the cognitive task of study.

For EEG data a general decoding pipeline has three main processes: signal pre-processing, feature extraction and classification.

Signal Preprocessing

As an electrical potential signal, the EEG data relies on power, frequency and time information, but what EEG's electrodes detect is not just cerebral activity, the sensors are also sensitive to electrical changes from another part of the body. Data that is not from brain sources is denominated *artifact* and can have as well an external-body origin such as line power noise or radioelectric signals (Eichele et al. 2010).

Filtering is an effective preprocessing procedure as it removes power from undesired frequency bands, it is known that different frequency bands represent different mental states, Delta (1-4 Hz) associated with sleep states, Theta (4-7 Hz) related to drowsiness, Alpha (7-12 Hz) restful states, Beta (16-24 Hz) active mind states, and Gamma (30-100 Hz) related to concentration states. Thus filtering helps to analyze certain frequency bands and catch for more evident information.

It is common to filter out noise related to the power line that goes on 50 or 60Hz depending on countries with what is called a notch filter that powers off all signals on those frequencies (Eduardo Reck Miranda 2014).

Other preprocessing techniques include statistical methods to try to separate different source components of data and remove the ones related to artifacts; this is the case of Independent Component Analysis (ICA)(Eduardo Reck Miranda 2014; Palaniappan 2014).

Figure 2 shows an example of the difference between raw EEG signals with jaw movement artifacts (left) and a clearer signal applying a band-pass fir filter from (7-100Hz) and notch filtering at 50Hz for line noise reduction (right).

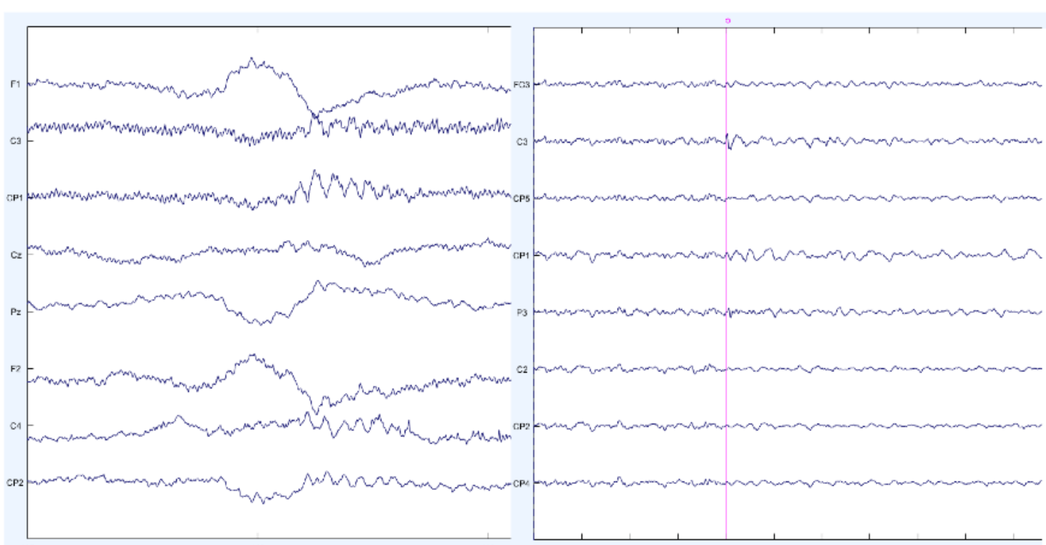


Fig 2. Difference between a raw EEG signal (left) vs band-pass filtered signal (right)

Feature Extraction

Feature Extraction is a group of methods that aim to represent raw EEG rhythms into a compacted number of relevant values, in general decreasing the data dimensionality and increasing important information (Fabien Lotte and Congedo 2016).
EEG has three important sources of information that commit features :

- **Spatial Information:** Features describing where in space the relevant signals are coming from. This is equivalent to selecting specific EEG channels where a certain signal is meant to be originated.
- **Spectral (frequential) Information:** Features that describe how power in a specific frequency band varies.
- **Temporal Information:** Features describing how relevant signals are with respect to time or selecting different time windows to analyze.(Eduardo Reck Miranda 2014)

There are EEG signals which feature patterns that are evoked after certain stimulus or cognitive tasks have been identified and proven useful for control BCI approaches, this evoked potentials highlight anticipation and perception activity in the brain this is the case of :

Visual Evoked Potentials (P300), a positive deviation from baseline and time-locked at signals around 300 and 600 milliseconds after a visual stimulus that the subject was expecting, usually at 8Hz on central regions of the brain.

Steady-state visual evoked potentials (SSVEP) is the repeated response of the brain to a visual stimuli, when a visual stimuli is presented to frequencies higher than 6Hz the response between stimuli overlaps, amplifying the signal at the stimulus frequency. These signals are maximal at the occipital region of the brain.

Event-related synchronization and desynchronization (ERD/ERS) are potentials related to motor tasks, prominent in alpha and beta bands, ERD is a negative deviation during preparatory stages of body movement and ERD is the counterpart that occurs at the same time, the location of this potentials depend on the limb intended to move, right side body limbs would elicit ERS on the right hemisphere and vice-versa. These signals are prominent in the motor cortex area of the brain in the central region (Palaniappan 2014).

The use of data science techniques have been proved useful for feature extraction on EEG, for spatial information, Common Spatial Patterns (CSP) is a spatial filtering algorithm that applies a linear transformation to the data resulting in a set of spatial filters where the signal variance is maximal for one class while minimal for the other.

For Motor Imagery (MI), CSP has proven solid results as it helps identify the spatial differences in the signal between two imagery tasks (Gao 2014).

Covariance Matrix, also known as Symmetric Positive Definite (SPD) matrix is an array representation of the relationship between the different variables of the data in this case the EEG channels, the elements of the matrix are a value of how much the signal from two channels vary together, this is known as covariance, then the matrix dimensions are defined by the number of channels where the main diagonal is the covariance between the one channel and itself, it is symmetric because as shown in figure 3 its transpose does not change the matrix and positive because the variance is computed from square distance to the variable mean (Congedo, Barachant, and Bhatia 2017). Figure 3 is the mean covariance matrix from multiple 2 second trials of an 8 channel EEG signal represented in a heatmap, lighter color indicates stronger correlation between the channels along the trial's time.

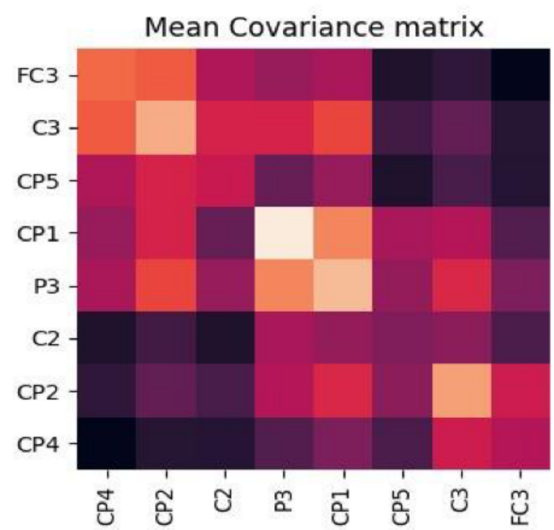


Fig 3. Covariance Matrix of 8 EEG channels represented in a heatmap

Classification

The classification step of the decoding process aims to label what is the cognitive task that is generating the neural activity being analyzed thus the label can be used to select further commands to control for the last component of a BCI system. This classification is done with the use of machine learning algorithms that can learn how to identify the class of a certain arrangement of features therefore is important on how discriminative the features are and what is their dimensionality to reach optimal classifier accuracies (Eduardo Reck Miranda 2014).

Linear Discriminant Analysis (LDA) is a popular algorithm that aims to use hyperplanes to separate the data, essentially tracing a line and label the class depending on which side of the line the datapoint to classify relays as shown in figure 4 (F. Lotte et al. 2018). Finding an optimal line secures the classifier accuracy.

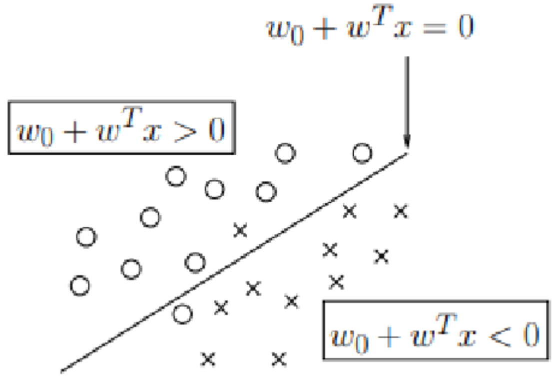


Fig 4. LDA hyperplane (Palaniappan 2014)

Support Vector Machines (SVM) is another popular algorithm that discriminates between classes with the use of a hyperplane, using the same logic of LDA but with the help of support vectors that help to mark the optimal hyperplane, support vectors are lines marking where is the nearest data point so the optimal hyperplane has the same distance to the support vectors (F. Lotte et al. 2018).

Minimum Distance to Mean (MDM) is a classifier that would select the data class depending on the distance it has to the average point of each from the available data; this method has proved useful when comparing covariance matrix distances. Covariance matrices are marked in a non-euclidean space because of their characteristic of being symmetric and positive lying in a cone-shape space where the 3D axis are the covariance between the two channels and each individual channel variance, therefore Riemannian geometry is needed to find the distance that represents a geodesic in this case. Riemannian geometry is used for 3D spaces assuming that a line is not the shortest distance between two points but a geodesic curve (Congedo, Barachant, and Bhatia 2017).

3.2.4 Device to Control

The last component of a BCI control system could be any sort of device for daily use in which parameters are instructed by coming commands from the previous component; the grade of control depends on how many classes the system is enabled to select. SSVEP approaches, for example, can have as many classes or options as can be shown on a display, as the speller in (Ansari and Singla 2016) with 25 letters as options.

For MI approaches the range of options is narrowed to how many imagery tasks can be discerned by the system's classifier as in (Choi and Cichocki 2008) where the system had 4 controlling options for a wheelchair.

It is easy to define that for BCMI the output device could be a musical instrument, in this case, the options to play are every musical note available on the instrument, or a recording device with a certain number of ready to play pieces.

3.3 Proposed methodology

This research proposes a BCMI control system using Speech Imagery (SI) to play a musical instrument in a basic manner. Therefore the user would play one note by imagining pronouncing a sustained phoneme, when the user imagines vowel " e " the system plays a high pitch note, when the user imagines the vowel " o " the system plays a low pitch note and the instrument does not actuate if the user is in resting state.

SI is a field of interest within BCI as its approach aim to help people with speech impairments, it is known that speech imagery induces ERD/ERS signals in the motor cortex area related to tongue movement (Wang et al. 2013), and neural activity in the left hemisphere's Broca's and Wernicke's areas associated with language production and comprehension, respectively (Friedman et al. 1998). Wernicke's area has been of interest for SI, as the use of semantics when asking participants to imagine about words have shown discriminative information that differs from phoneme imagination as seen in (Rybář, Poli, and Daly 2021; Kaongoen, Choi, and Jo 2021). Some BCI studies on SI have reached optimal classification accuracy between different SI tasks (Bakhshali et al. 2020; Kaongoen, Choi, and Jo 2021; Brigham and Vijaya Kumar 2010), these results are promising and push further this research.

SI is selected as the mental activity for this research firstly because as an imagery task, unlike the visual evoked approaches as evaluated in Chapter 2, it dispenses with any sort of stimulus giving a different user experience and second because the research assumes a relationship between vocalization imagery and musical activities. Thus the signal acquisition would be focused on a mixture of left and motor cortex regions to try covering activity from the areas previously mentioned.

The ongoing research uses merely quantitative data to test the feasibility of the proposed BCMI system based on an experiment/test-driven methodology with the procedures described in section 3.2.3. Therefore, selecting the best methods for decoding depends on the classifier accuracies in offline analysis and their complexity to be implemented in an online approach.

As shown in figure 5, the methods to test are bandpass filters (Alpha, Beta, Gamma), CSP and Covariance matrices for feature extraction and SVM, LDA and MDM as classifiers.
Once the optimal decoding pipeline is selected the further research challenge is to build the prototype with the chosen methods.

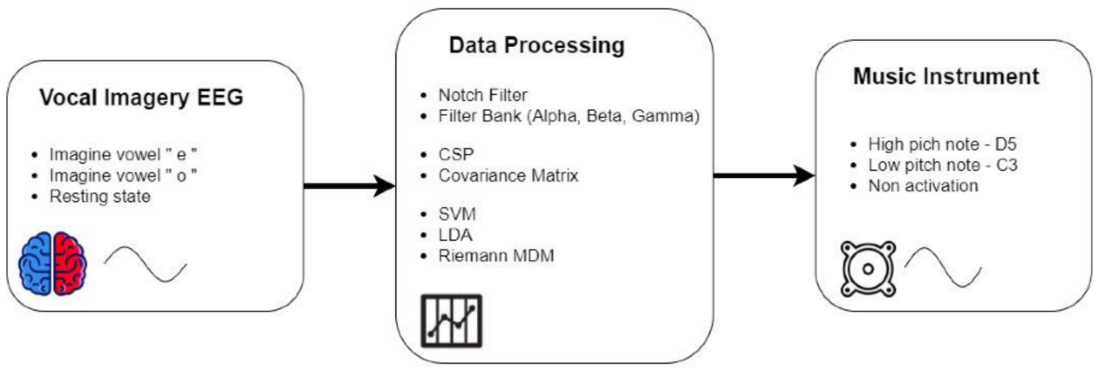


Fig 5. BCMI proposed methods.

3.4 Methodology Application

This section describes first how the EEG experiments were conducted, the application of the evaluated methodologies and finally the results obtained within the different combination sets.

3.4.1 EEG experiments

Signal acquisition experiments were held on a single subject. The experiment results may be biased as the participant has previous knowledge and experience in speech imagery and EEG experiments, therefore his imagery attempts have a persistent desire to produce high quality signal recordings.

Current accessibility restrictions due to the COVID pandemic make the participation of more subjects unfeasible. The measurement device, a dry wireless EEG headset g.MOBILlab+ by G.tec with 8 channels that records at 16 bit resolution and 256 Hz sample rate. Locations were placed as shown in figure 6. Electrode names by standardization take a letter based on the brain regions, an odd number for the left and even for the right hemisphere.

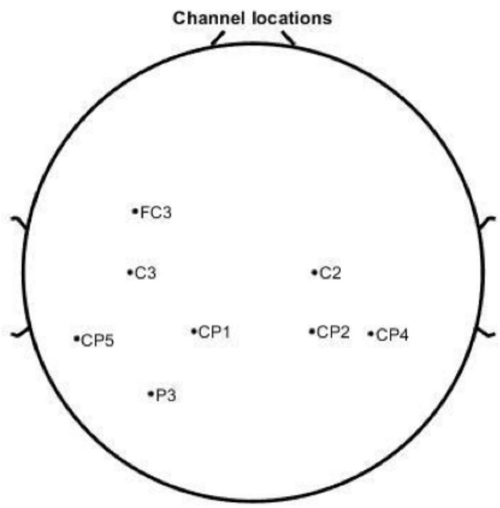


Fig.6 EEG locations

As mentioned earlier G.Tec provides an interface to connect and access g.MOBILlab+ through Matlab Simulink modules. The current experiment was designed to run the Simulink simulation for each time the participant was asked to perform the imagery task, therefore the Simulink blocks shown in figure 7, connected to the amplifier and saved the 3 seconds EEG record into a .mat file with a differentiable name based on the class of the speech imagery task on duty.

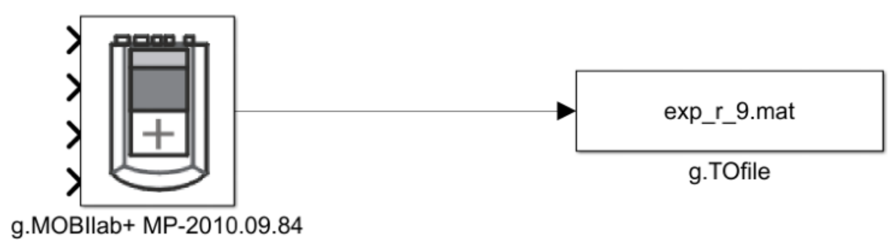


Fig 7. Record Simulink modules

After a practice period, where the participant was asked to perform the speech tasks aloud following the instructions from the screen. The recording experiments were divided into two blocs, each bloc recorded 20 trials for each of the three classes. Within each trial, the participant was presented with an instruction followed by a fixation cross on the screen that marked the beginning of the trial, after 3 seconds a pause indication was shown marking the start of the resting state for another 3 seconds before the new trial started. For resting state, an instruction captioned “rest”, was presented so the subject knew he had to keep the resting state of an empty mind. The experimental procedure is shown in figure 8. The experiment resulted in 40 trials of 3 seconds for each class.

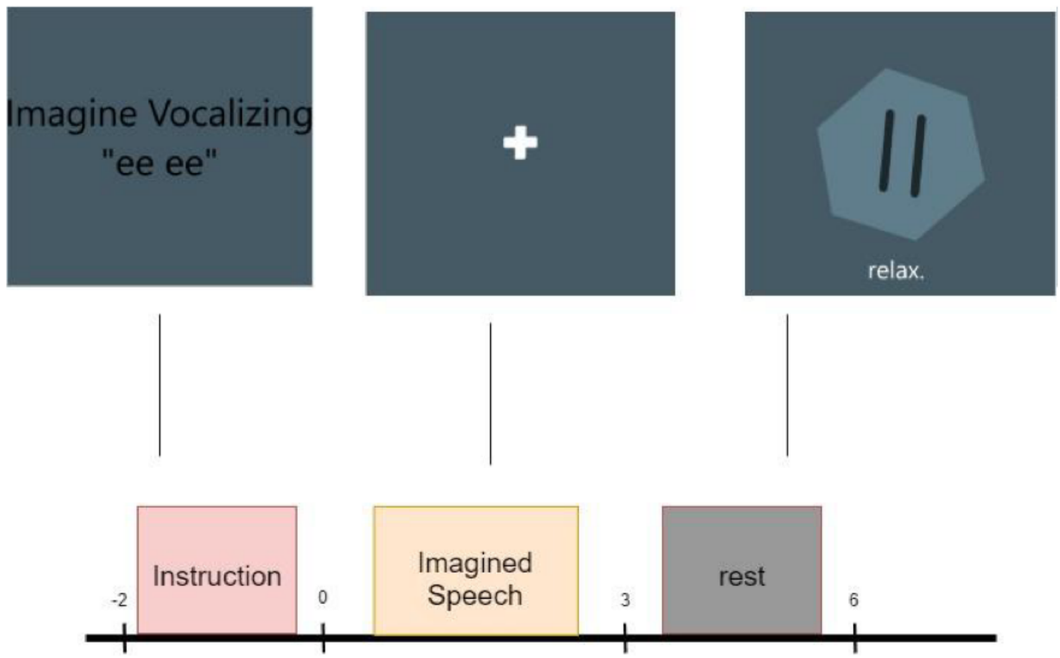


Fig 8. EEG record procedure

3.4.2 Signal Processing

The first pre-processing step is to visually analyze through all the signal recordings, to inspect for clearly noticeable muscular or ocular artifacts, as well as noisy channels; 11 bad trials are rejected after visual inspection, leaving 29 total trials. The signal was trimmed one second before the trigger and one seconds after it, that is 2 seconds trials.

The Power Spectral Density (PSD) for the average of the 70% trials is shown in figure 9. The power noise is clearly present at 50Hz, two butterworth bandpass filters are applied, one to notch 50Hz line noise (order 2, from 49-51Hz), and the second one, of order 20, to bandpass with lower cut at 7Hz and high cut at 100Hz to keep the relevant frequencies to evaluate. Butterworth is the filter design of choice as g.MOBILlab+ provides filter modules of this kind, easy to implement for online simulations. The filtered PSD is shown in figure 10.

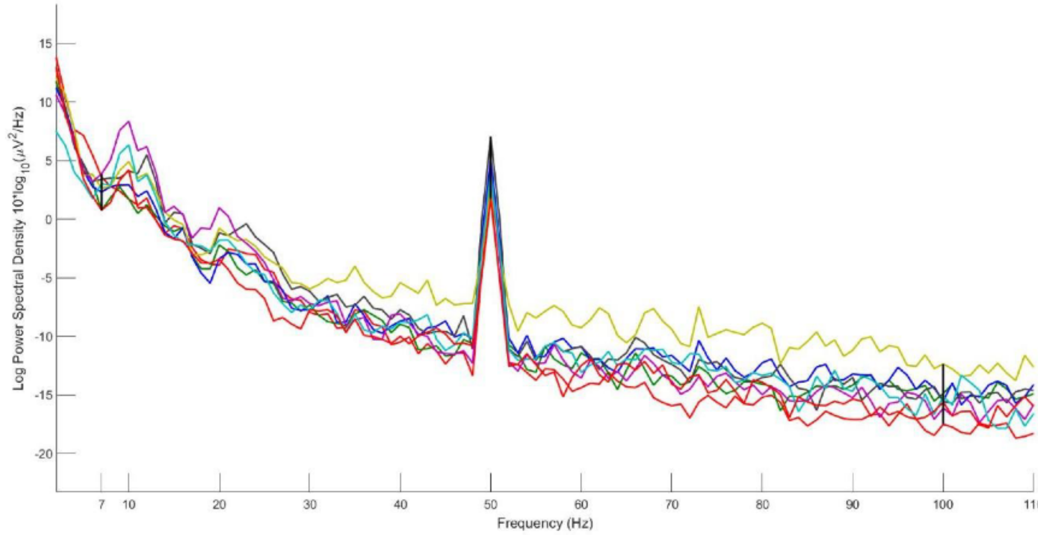


Fig. 9 Power Spectral Density of average raw EEG

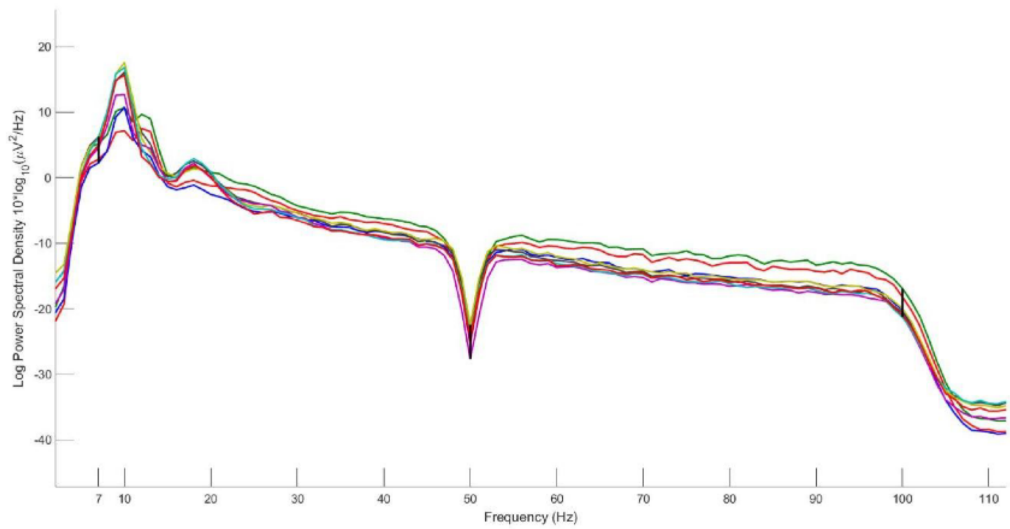


Fig. 10 Power Spectral Density of filtered averaged EEG

The filtered data is divided into 2 frequency bands, one covering alpha and beta bands (7-26 Hz) and other for gamma bands (27-100 Hz), each band would be the input signal for the following classification pipelines.

1. The CSP algorithm, from the MNE library (Gramfort et al. 2013), is set to extract 4 components for pairwise calculation, figure 11 shows components computed for alpha and beta bands, these components represent spatial filters that maximize the signal variance for one class and minimize it for the other when applied to the signal.

As it can be seen in figure 11, the filters against rest state show similarities around the same areas, in CSP0 and CSP1, the filter would maximize the signal variance of the rest class for channels close to the frontal area enhancing discriminative information. Figure 12, shows the CSP filters for gamma bands, where filters against “e” class show similar topography in CSP1 and CSP3, the filter would minimize signal variance around P3.

The CSP components are used to transform the signal that is then used as input data for LDA and SVM; these classifier algorithms are used from scikit-learn Python functions keeping their default configuration.

2. Covariance matrices are computed for each trial, for each class the average is calculated using pyRiemann module of Python; its averaging function gives a centroid matrix for each of the three classes. The covariance centroids for alpha-beta bands are presented in figure 13 where most of the channel’s covariance don't seem to differ between classes, however, covariance between FC3, C3 and CP5 is stronger for resting states. For the gamma band shown in figure 14, variance for channel FC3 in class “e” is higher than in the other classes, overall there is discriminative information between the different mean covariance matrices. The obtained centroids are used to classify single trials using the MDM algorithm defined on the Riemannian distance between the upcoming covariance matrix and different class centroids, therefore the shortest distance would label the upcoming signal.

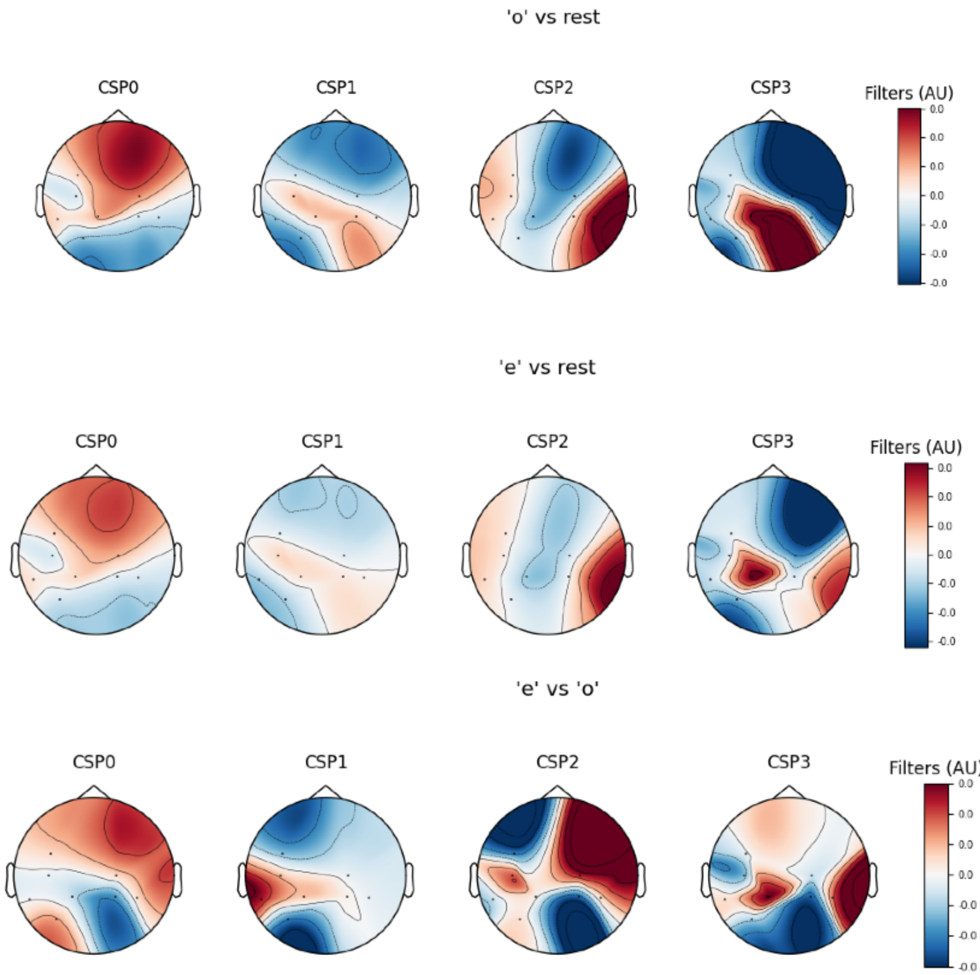


Fig. 11. CSP components for alpha-beta bands

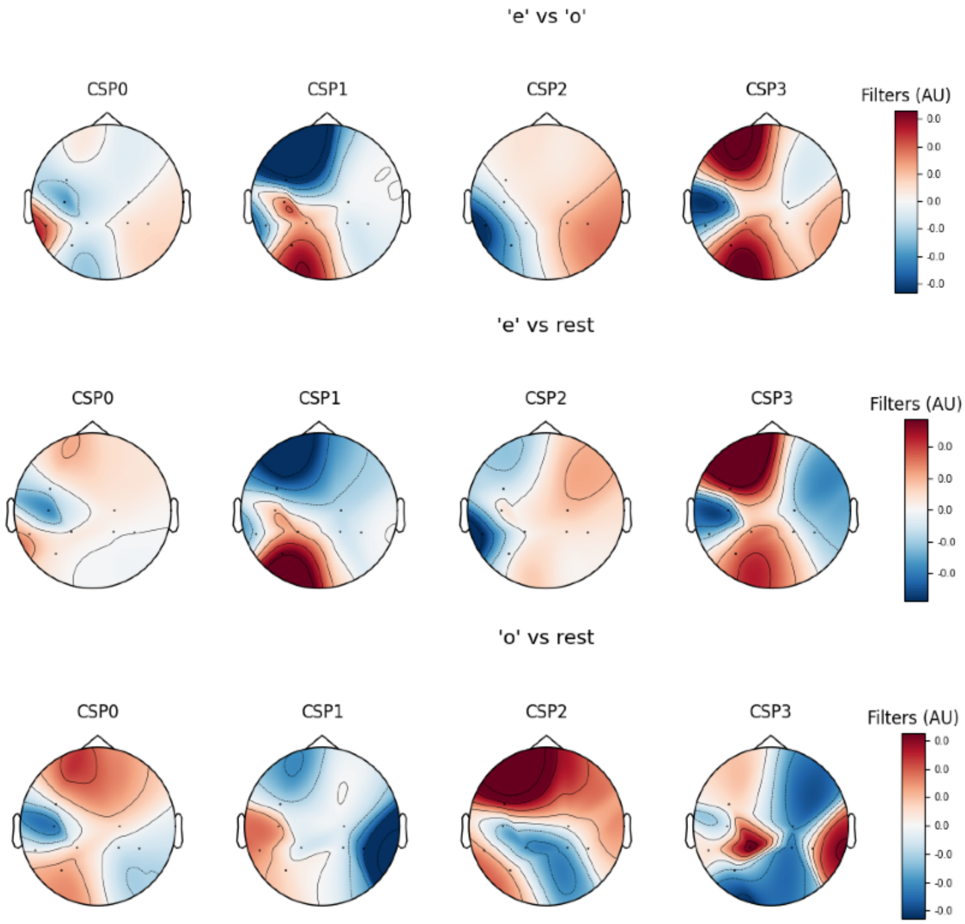


Fig. 12. CSP components for gamma band signal

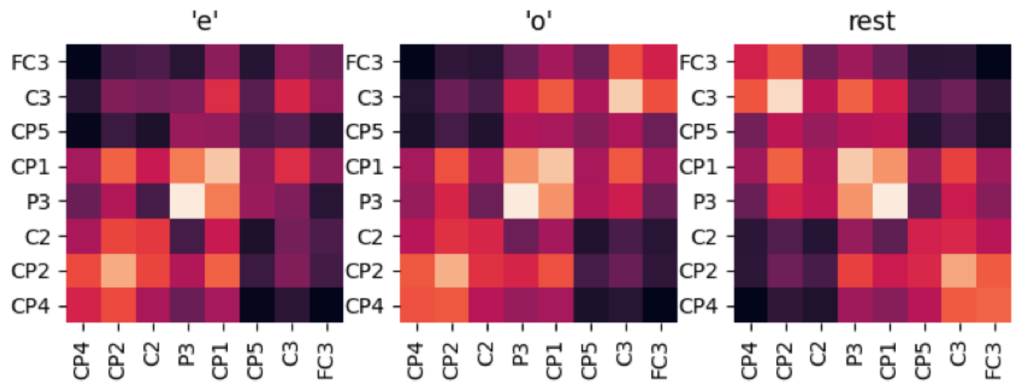


Fig. 13. Mean Covariance Matrices in Alpha-Beta bands

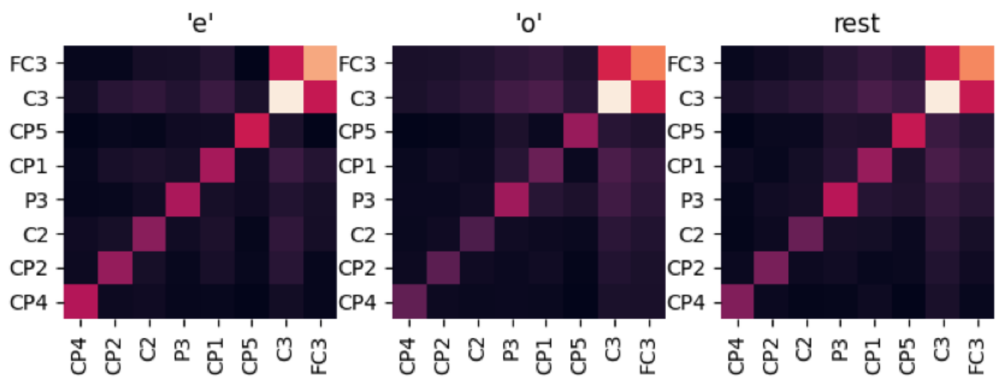


Fig. 14. Mean Covariance Matrices in Alpha-Beta bands

3.4.3 Results

Using a 10 fold cross-validation method that randomly divides 90% as training data and the other 10% for testing, each of the decoding pipelines are trained and tested multiple times.

The following results led to the addition of one more frequency band to consider, narrowing the Gamma band led to different results especially for CSP classifiers. Therefore, the Gamma band is divided and the two new bands are used as input to run the two pipelines to add its results. High Gamma and Low Gamma produced different accuracies thus having different discriminative information. The obtained average accuracies and respective variance for the 3 pairs of classes and the 3 frequency bands are shown in table 2. As 11 trials were rejected there is not an even number of classes, leaving 8 trials for class “o”, 8 for class “e”, and 13 for rest-state which defines the chance label of the classifications results.

The results are expected to highlight one of the three pipeline performances, depending on each of the divided frequency bands, and the pair of classes they were used with. The obtained accuracies that are significantly higher than chance level are considered for comparison and selection of the method to further research.

	CSP + LDA	CSP + SVM	MDM
“o” vs “e” 7-28Hz	0.54 ± 0.01	0.53 ± 0.01	0.58 ± 0.02
28-60Hz	0.56 ± 0.01	0.56 ± 0.02	0.59 ± 0.02
60-100Hz chance level = 0.50	0.59 ± 0.02	0.60 ± 0.01	$\mathbf{0.60 \pm 0.03}$
“o” vs rest 7-28Hz	0.93 ± 0.02	$\mathbf{0.97 \pm 0.01}$	0.92 ± 0.03
28-60Hz	0.68 ± 0.01	0.70 ± 0.01	0.70 ± 0.02
60-100Hz chance level = 0.60	0.67 ± 0.01	0.65 ± 0.02	0.64 ± 0.03
“e” vs rest 7-28Hz	$\mathbf{0.91 \pm 0.01}$	$\mathbf{0.91 \pm 0.01}$	0.90 ± 0.02
28-60Hz	0.64 ± 0.01	0.63 ± 0.02	0.69 ± 0.02
60-100Hz chance level = 0.61	0.66 ± 0.02	0.64 ± 0.01	0.68 ± 0.03

Table 2. Methods accuracy

3.5 Conclusion

As shown in the results table, the accuracy from the three classification procedures is very similar. It is highlighted that the alpha-beta powerband is more discriminative with an imaginary task against the resting state, while between imaginary tasks, the high gamma band (60 - 100Hz) has more discriminative information. One important variation that led to better results when classifying against the resting state, was to ask the participant to have his eyes closed during the resting state, for this, the participant knew that after rest instruction he had to close his eyes for 2 to 3 seconds so the time window of interest, 1 second after the trigger, was recorded while eyes closed, the reduction of visual activity in the alpha band is significant and allows high discrimination with imaginary events.

When selecting the optimal classification method there is a clear advantage on the Riemannian MDM above the two others that depend on CSP which is an algorithm designed to extract features between two classes only, however, a variation possibility can be considered feasible to work, this is because of the structure of the proposed BCMI system, as it has three classes it would need two classification process, first against the resting state and then between the imagery classes in a decision three manner.

The classifier method's accuracy does not significantly vary, thus the criteria to decide one for prototyping an online approach depends on its implementation complexity. As explained, CSP calculates filters, the ones obtained in the previous offline analysis, these filters would be applied to the online signal and the resulting signal would be classified by SVM because of its slightly better accuracy than LDA. For an SVM online prototype, the discriminant values, that represent the vector functions, obtained in the offline analysis would be used to implement the decision module for online discrimination.

For an MDM prototype, the Riemannian Means obtained in the offline analysis with the python module would be the centroids used to classify the upcoming covariance matrix of the online signal, based only on which distance to the centroid is the shortest.

There is no significant difference in the computational costs of CSP in comparison with covariance matrices, but the MDM discriminant process which consists only in computing and comparing the Riemannian distances to the means, which is a logarithmic difference of the covariance matrix's eigenvalues, is simpler than the discriminative function used by SVM. Riemannian MDM works with a smaller feature space dimension as the upcoming input is just the 8x8 covariance matrix, while the result from SVM is a filtered time-series signal. MDM method is therefore easier to implement needing few lines of code and 4 covariance matrices centroids as parameters (alpha “ o ”, alpha rest, high gamma “o”, high gamma “e”).

The BCMI control system Simulink design is as shown in figure 15 where the signal obtained by MobiLab+ amplifier has a notch filter for noise reduction, this is divided for filtering in alpha and high gamma bandpass modules. As shown in offline results the proposed prototype would discriminate against resting-state first using alpha division of the signal to discriminate against resting state and second to compare between the two imaginary classes centroids. MDM processing module involves covariance matrices and distances calculation, such values are passed to the MIDI module that would decide a control instruction to be sent to an external instrument through Musical Instrument with Digital Interface (MIDI). The instruction is computed in decision three based on the input distances where, the instrument receives MIDI message to play note D5 if the classification label is vowel /e/, note C3 if the classification label is vowel /o/ while no MIDI message is sent if classification label is rest.

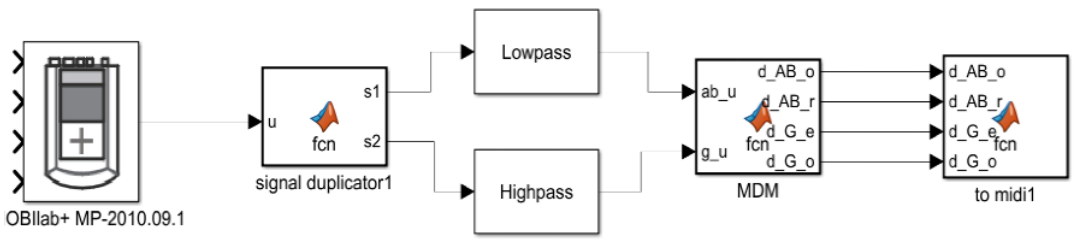


Fig. 15.BCMI Simulink module diagram design

CHAPTER 4

BRAIN-COMPUTER INTERFACE MUSIC SYSTEM USING SPEECH IMAGERY, A PROOF-OF-CONCEPT

4.1 Overview

This chapter delivers a proof-of-concept Brain Computer Music Interface system that enables the user to play an analogue synthesizer with their imagination to produce a sound with the mouth. The proof-of-concept takes as the foundation the results from Chapter 3 that selected as an optimal EEG decoding method for online Speech Imagery, the conversion of the EEG trials into Covariance Matrices that are labeled by a Minimum Distance to Mean classifier using Riemannian geometry in its distance function. This chapter evaluates first between two different electrode locations, one focused on the left hemisphere because of its relation with language comprehension and production, while the second location, takes Chapters 3 approach to cover the motor cortex and central-left region. Speech imagery tasks are also evaluated between one set of EEG experiments where the user is asked to imagine the vowels in a rhythmic manner, having /e/ repeated two times and /o/ three times during one second of recording, and another set of EEG recordings, using words, then /high/, /mid/ and /low/ words represented the classes, the later experiment didn't show any significant accuracy improve, however /mid/ word was adapted to be a new class for the prototype. At the end of this chapter the online prototype is built proving right the proposed architecture; the proof-of-concept achieved an accuracy of 31% within 4-class classification, above a 25% chance level.

4.2 Introduction

The emergent brain-computer interface (BCI) technology enables people with severe neuromuscular disabilities to directly use their brains to communicate or command external environments and devices. Event-Related Synchronization and Desynchronization (ERD/ERS) are electrical activations in response to sensorimotor tasks, tasks of limbs movement that are present in general as unconscious mechanisms (Pfurtscheller and da Silva 1999). These rhythms have been broadly studied as responses to movement preparation and Motor Imagery.

The variability of ERS/ERD intensity or power in particular frequency bands can be used to distinguish different motor imagery signals. These rhythms are presented in alpha and beta frequency bands and can be induced by motor imaginations of left and right hand, foot as well as tongue (Pfurtscheller and da Silva 1999; Iijima et al. 2009). Chapter 2 evaluates the advantages in these paradigms over BCI designs that require selective attention to cued sensory stimuli, even though other methods present higher information transmission rates for multiclass, in binary decision terms every noninvasive EEG technique is on the order of one decision every 2-10s (Marque 2012). However, MI has various challenges in its way to building an efficient BCI system, one important challenge is the existence of fewer recognizable motor types or stages (Yuan and He 2014).

Speech Imagery (SI) is proposed as an alternative BCI methodology as well as a complement for MI. SI is the mental activity of imagining to speak aloud without moving articulators or producing sounds (Fujimaki et al. 1994). As the imagined words can be associated with commands this paradigm could be intuitive and theoretically have as many possible choosing options as sound combinations can be done with the mouth.

As with MI, the information of SI is encoded on the power and location variation of the EEG signals.

Decodification of SI has achieved some promising results, starting from early studies that found the event-related potentials from single vowel imagination differ significantly from resting-state EEG (Fujimaki et al. 1994). DaSalla in (DaSalla et al. 2009) performed a classification from EEG data recorded when subjects imagined vowels /a/ and /u/ or stayed in resting state, using common spatial patterns (CSP) and reached classification accuracies of up to 70% with support vector machines (SVM).

A later study (Deng et al. 2010) used Hilbert spectrum analysis of the EEG to classify between syllables /ba/ and /ku/ imagined in different rhythms where the classification of the rhythm variation yielded a 58% classification accuracy, significantly higher than the 42% chance level. Later research by (Wang et al. 2013) used speech imagery of two monosyllabic Chinese characters that represent two different meanings and had different shapes and pronunciations. Using CSP and SVM, classification accuracies of up to 71% and 90% accuracy was achieved when classified against a rest state condition.

CSP is widely used in motor imagery to extract EEG features, it has shown good performance for two classification tasks, however, it depends on high dimensionality of data for better performance. As mentioned in Section 3, there is a relatively new method that has performed well to decode spatial information from EEG, Riemannian geometry has attracted attention (Guan, Zhao, and Yang 2019). The Riemannian framework is mathematically sharp as it involves abstract definitions of multidimensional shapes, however is simple to apply in practice, both algorithmically and computationally, which allows designs for online decoding machines that suit real-world operating conditions.

This Chapter evaluates different speech imagery tasks, that are decoded using covariance matrix Minimum Distance to Mean classifier with Riemannian metrics gives the highest accuracies, to a later online implementation of this decoding pipeline. The first set of imagery tasks are based on a rhythmic imagination of vowels /e/ and /o/ meanwhile the second set of experiments is based on three words that may be related to the analogue synthesizer notes (low, mid, high). The final delivery of this chapter is a working prototype with a combination of the previously mentioned tasks, having rhythmic /e/, /o/ and the word mid as options.

4.3 Proposed Decoder

This section describes the pipeline implemented to classify speech imagery tasks both in offline and online deliveries, with Covariance Matrices as feature extractors and Minimum Distance to Mean with Riemannian metrics.

4.3.1 Feature Extraction

The use of statistical techniques has been proved useful for the extraction of spatial and temporal information. Covariance Matrices, a Symmetric Positive Definite (SPD) matrix is an array representation of the relationship between the different variables of the data, in this case, the EEG channels. The elements of the matrix are a value of how much the signal from two channels varies together, this is known as covariance, where the main diagonal is the covariance of each channel with itself, known as a variance. These covariance representations encode spatial information of the signal. Let us take a 2 channel EEG signal, where $x_1(t)$ and $x_2(t)$ are the time-series recorded for 1 trial. Their covariance matrix C is

$$C = \begin{pmatrix} var\,(x_1) & cov(x_1\,,\,x_2) \\ cov(x_2\,,\,x_1) & var\,(x_2) \end{pmatrix}$$

The matrix is symmetric as $cov(x_1, x_2) = cov(x_2, x_1)$ and is positive because the variance is computed from square distance to the variable mean and because of Cauchy-Schwarz inequality $|cov(x_1, x_2)|^2 \leq var(x_1)\, var(x_2)$ (Congedo, Barachant, and Bhatia 2017).

4.3.2 Classifier

The classification step of the decoding process aims to label what is the cognitive task that is generating the neural activity being analyzed thus the label can be used to select further commands to control for the last component of a BCI system. This classification is done with the use of machine learning algorithms that can learn how to identify the class of a certain arrangement of features therefore is important on how discriminative the features are and what is their dimensionality to reach optimal classifier accuracies (Eduardo Reck Miranda 2014).

Minimum Distance to Mean (MDM) is a classifier that would select the data class label depending on the distance it has to the average point of each class from the available data; this method has proved useful when classifying covariance matrices from EEG trials with a distance metric based on Riemannian geometry.
Let us take as an example two covariance matrices C_O, C_E one representing the mean of trials for class O and the other one, the mean for class E. Then for a new trial covariance matrix C_K the class would be O if $\delta(C_K, C_O) < \delta(C_K, C_E)$ where $\delta(.,.)$ is the appropriate distance function (Yger, Berar, and Lotte 2017).
Covariance matrices are marked as a point that relay in multidimensional cone-shape space because of their symmetric positive characteristic. This hypercone shape, has its dimensions defined by $N(N + 1)/2$where N is the number of channels, therefore Riemannian geometry is needed to find the distance, as it represents a geodesic in this case, a hypercone shape. Therefore the distance between two matrices is defined by

$$\delta(C_1, C_2) = ||Log(C_1^{-1} C_2)|| = \sqrt{\sum_{n=1}^{N} log^2 \lambda_n}$$

where λ_n are the N eigenvalues of a matrix $C_1^{-1} C_2$ and a mean matrix G of a group $\{C_1,..., C_K\}$ would be the point where dispersion is minimal, a central distribution of distances, therefore

$$\arg\min \frac{1}{k} \sum_{k=1}^{k} \delta^2 (C_k , G)$$

Riemannian geometry is used for multidimensional spaces from the principle that a line is not the shortest distance between two points on a sphere, but a geodesic curve (Congedo, Barachant, and Bhatia 2017).
Riemannian distance is calculated by projecting flat spaces from the curved space as shown in figure 16.

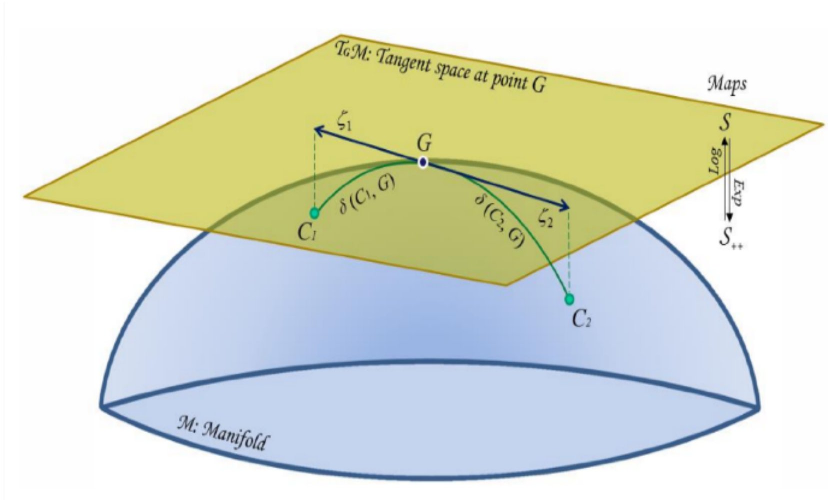


Fig. 16. Riemannian distance from C1 to C2

The robustness of Riemannian MDM gives it an advantage over other commonly used algorithms such as Common Spatial Patterns (CSP) because of its capabilities when handling noisy data that is common in EEG-based BCI, Riemannian metric proves advantages when finding data means, deviating less in the presence of outliers in comparison with arithmetic metric (Congedo, Barachant, and Bhatia 2017). It is clear to see better performance of Riemannian pipelines in comparison with CSP and linear classifiers as shown in (Congedo and Sherlin 2011) (Barachant et al. 2012).

4.4 Methodology

This chapter is divided into the offline and online sections below, as the offline section first would determine what are the theoretically optimal selections of

- EEG locations: left hemisphere (Broca's and Wernicke's areas) and left and central motor-cortex area
- Speech Imagery task variations: repetitive imagination of vowel /e/ and /o/, imagined speech of words "low", "mid" and "high"

Figure 17 shows the elements for the online prototype after selecting the parameters from the offline approach.

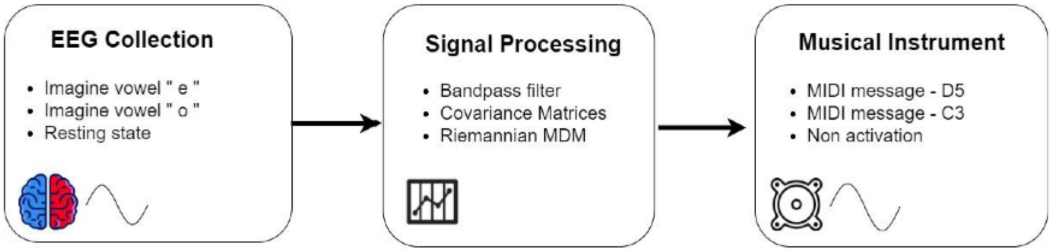


Fig. 17. Speech Imagery BCMI scheme.

4.4.1 EEG Experiments

Signal acquisition experiments were held to the same participant as in Chapter 3, participation of more participants was unfeasible due to the COVID pandemic. Using the 8 channel EEG amplifier, g.MOBIlab+, with the two different locations to experiment with as shown in figure 18. Location in right is set to the left hemisphere while the location in left is set to the central-parietal left region (FC3, C3, CP1, CP5, P3, C2, CP2, CP4)

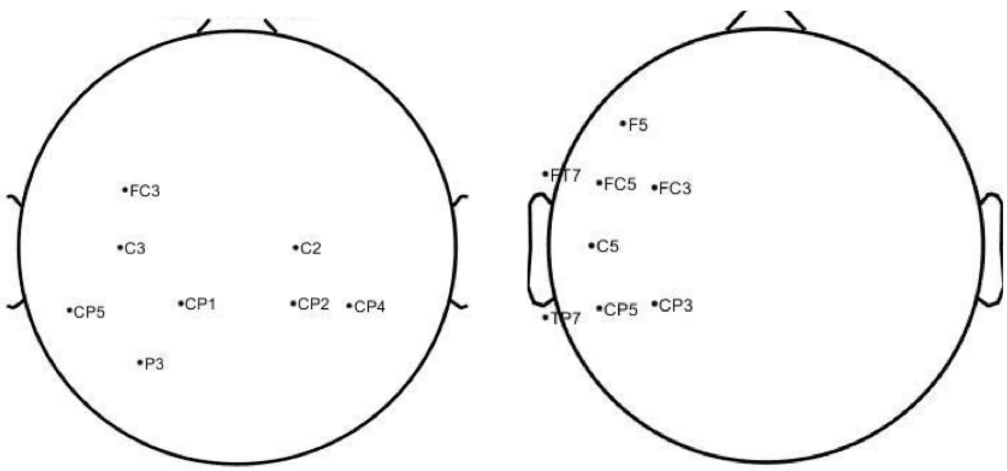


Fig.18 EEG locations

As mentioned in Chapter 3, G.Tec provides an interface to connect and access g.MOBILlab+ through Matlab Simulink modules, the current experiment was designed to run the Simulink simulation for each time the participant was asked to perform the imagery task. Simulink blocks that recorded each trial on this experiment are shown in figure 19. The interface module at the left connects to the amplifier, then the signal is notch filtered at 50Hz for the next module to cut power line noise as identified in Chapter 3, with the frequency bands of interest, to later bandpass filter the signal from 7 to 100 Hz with the butterworth filter modules provided by the amplifier manufacturer. Applying this filtering while recording the signal shortens the preprocessing process and gives a picture of how the online processing may perform

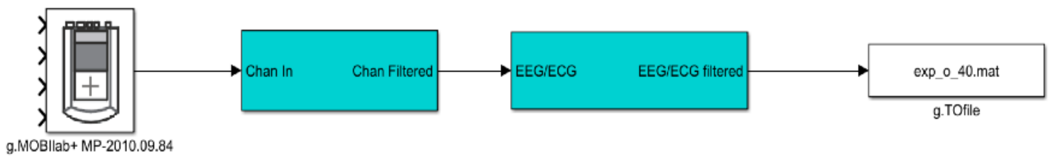


Fig. 19. Simulink model for filtered recording

The recording experiments were divided into two blocs, each bloc recorded 30 trials for each of the four classes

- Imagine Vocalize rhythmically two times vowel “ e ” - “e e”
- Imagine Vocalize rhythmically three times vowel “ o ” - “o o o”
- Imagine Vocalize word “low”
- Imagine Vocalize word “mid”
- Imagine Vocalize word “high”

After a practice period, where the user was asked to perform the speech tasks aloud and then to practice just imagining the words or phonemes shown on the screen. The recording experiments were divided into two blocs, each bloc recorded 20 trials for each of the three classes. Within each trial, the participant was presented with an instruction followed by a fixation cross on the screen that marked the beginning of the trial, after 2 seconds a pause indication was shown marking the start of the resting state for another 2 seconds before the new trial started. The experimental procedure is shown in figure 20. The experiment resulted in 40 trials for each EEG placement location of 2 seconds for each class.

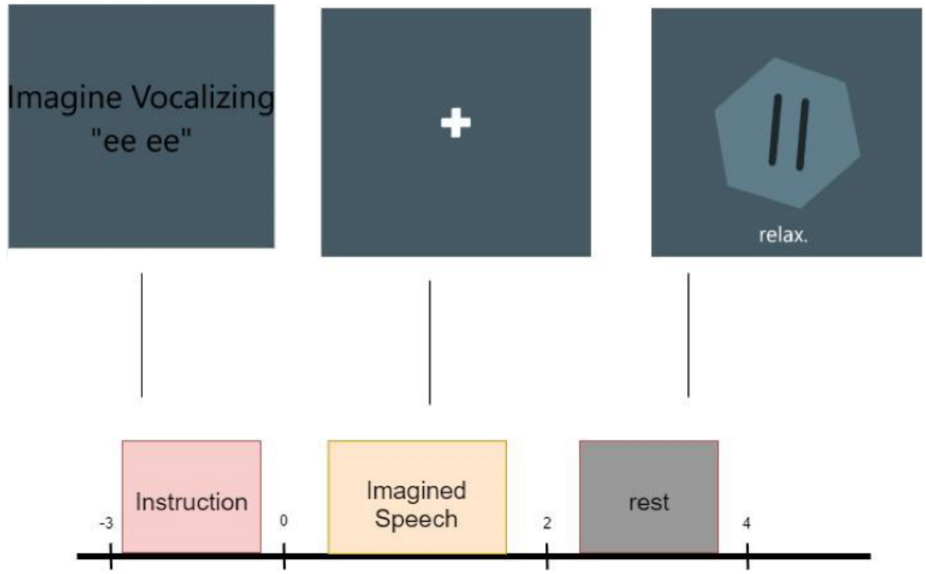


Fig 20. Second set of experiments procedure

Signal Processing

The first pre-processing step is to analyze through the signal recordings, after eye inspection and consideration, 50 bad trials are rejected considering evident movement artifacts and channel noise; there are a total of 283 trials for the left location and 253 for the central-left location. For all trials the signal is trimmed one second before trial and one second after to have even 2 second trials.

In order to visualize the power variation across the different bands and try localizing any particular pattern, the signal is decomposed using Morlet time-frequency representation defining two logarithmic spaces of interest one for alpha (7-28Hz) and other high gamma (60-85Hz), MNE (Gramfort et al. 2013) library has a useful function to join timefrecuency-plots for all channels used to plot the figures below. Figure 21 shows the alpha and beta time-frequency activity for left hemisphere locations for rhythmic imagery tasks and rest state.

The negative/positive power clusters shown in figure 21 may mark the different repetitive patterns of the imagery tasks but there is no clear spatial difference, in contrast, figure 22, the plot join of gamma bands for the imager tasks, shows clear spatial difference in the activations on channels CP5 or FC5.

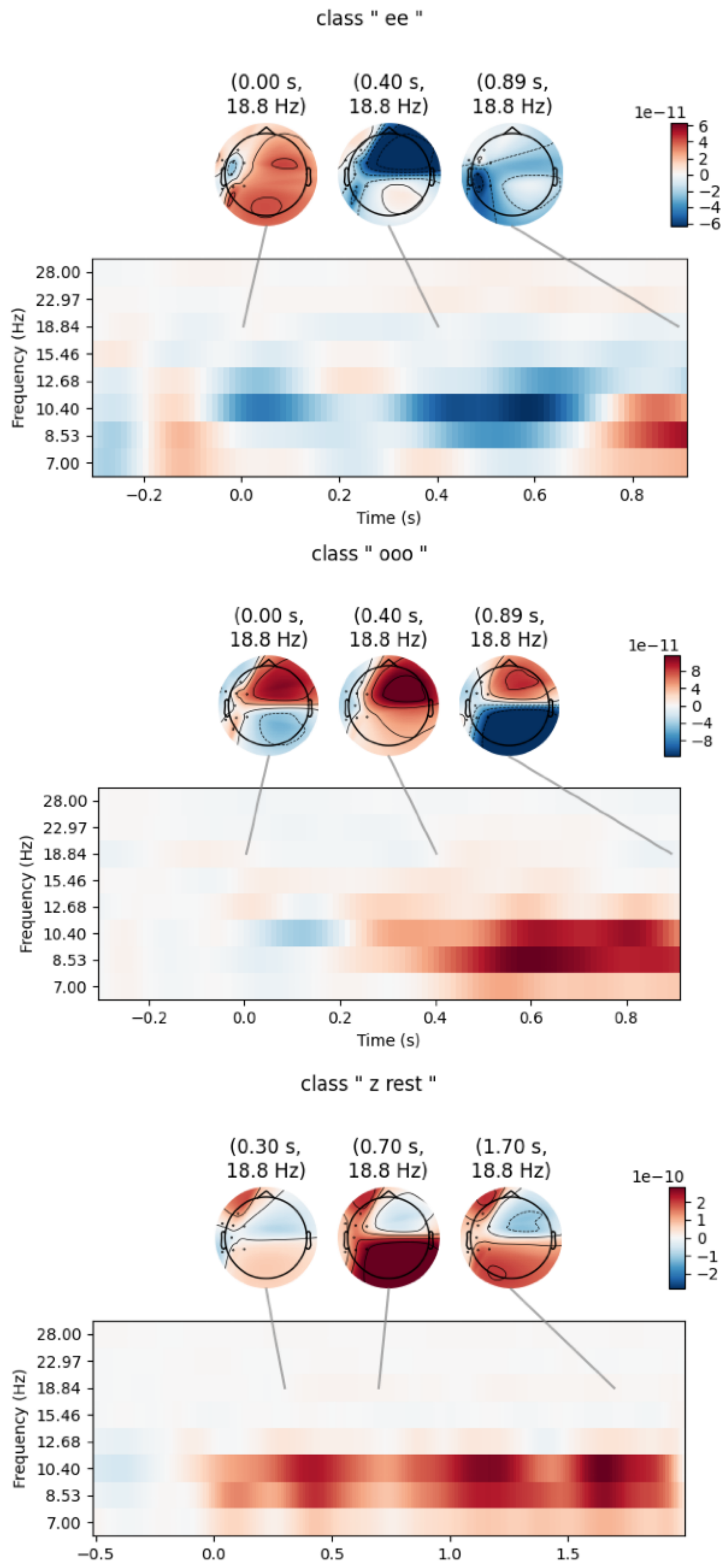
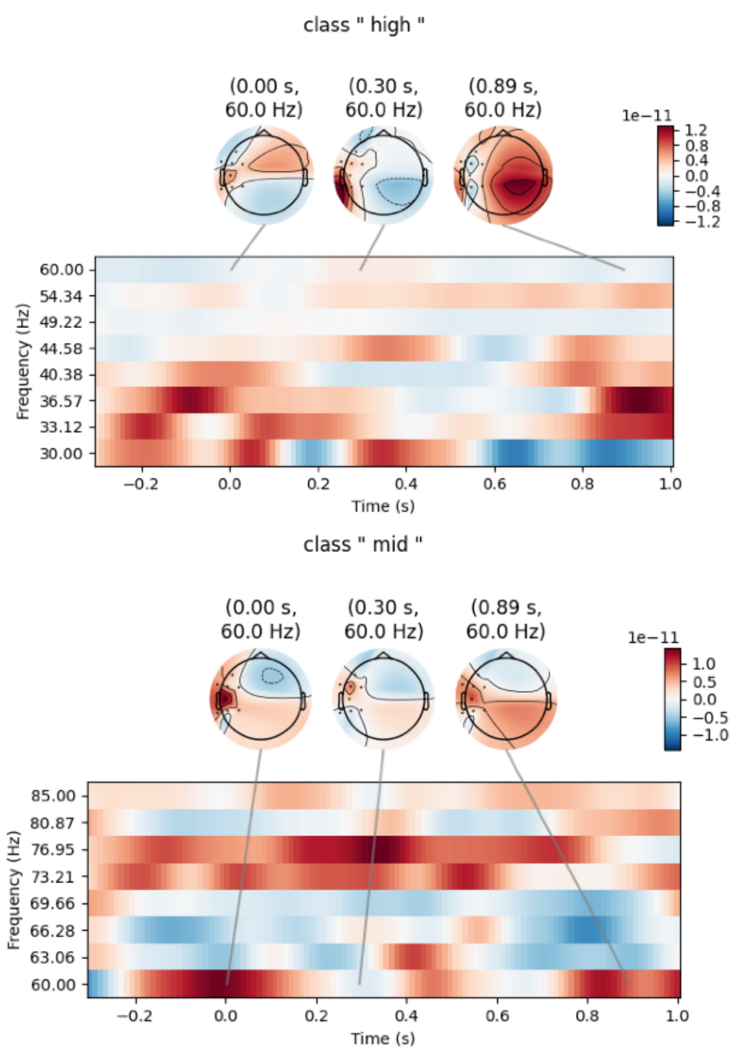


Fig 21. Time-frequency decomposition for /ee/, /ooo/ and rest classes in alpha-beta band

It can be seen in Figure 21, how signal power evolves in time along the trial. It is clear to difference the rest-state activity from the imagery tasks as there is a consistent positive power around electrode CP3 and CP5 close to the occipital region, while there is power variation around the same area for the imagery tasks, this presents clear evidence from the accuracies obtained in Chapter 3 when classifying against resting state was remarkably higher.

In Figure 22 there is some spatially discriminative information that can be seen happening on gamma band, at around 0.3 seconds after trigger on set, the signal changes broadly depending on the imagery task, at 60Hz, signal power focuses on FC5 for /mid/ while going back to CP5 for /ee/ or /low/. The spatial difference at this frequency represents evidence about the better classification performance show in Chapter 3 between imagery tasks



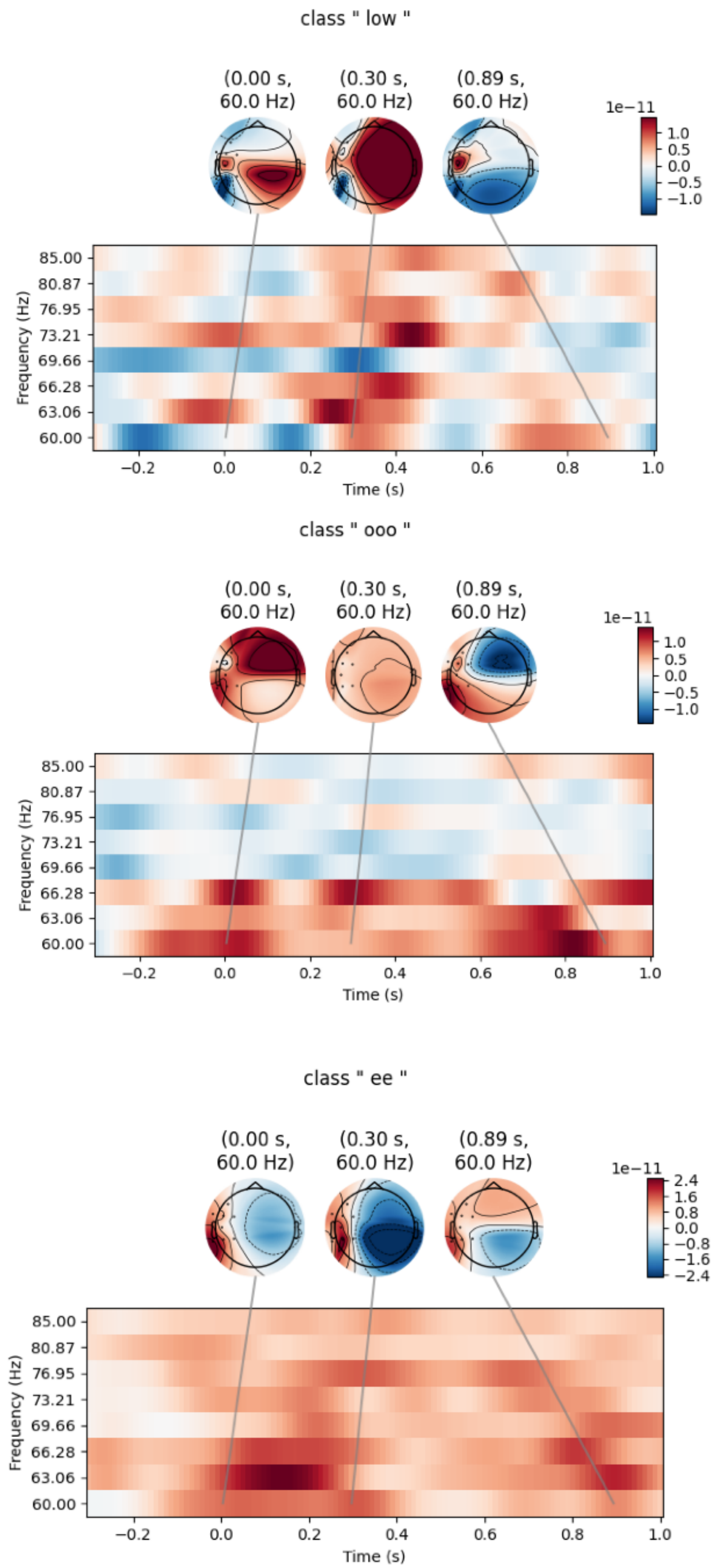


Fig 22. Time-frequency decomposition for /ee/, /ooo/, /mid/,/low/,/high/ in high gamma band

Following the time-frequency analysis that backs up Chapter 3 results, the signal is divided into 2 frequency bands for Riemannian mean computation and MDM classification, alpha-beta (7-28 Hz) band and high gamma (60-85Hz).

4.2 Offline Results

Using a 10-fold cross-validation method, the MDM classifier is tested between each possible pair of classes, the obtained accuracies and respective variances are shown in table 3. As in Chapter 3, the results are expected to highlight firstly what set of electrode locations give best accuracies, second what imagery task based on its pairwise classification could be taken as an option to build the online implementation, this based on accuracies that are significantly higher than chance level.

		Left hemisphere	Central left region
“ooo” vs “ee” chance level = 0.52	7-28Hz	0.37 ± 0.03	0.40 ± 0.06
	60-85Hz	0.63 ± 0.04	0.59 ± 0.02
“ooo” vs rest chance level = 0.73	7-28Hz	0.90 ± 0.02	0.62 ± 0.01
	60-85Hz	0.64 ± 0.03	0.57 ± 0.04
“ee” vs rest chance level = 0.71	7-28Hz	0.96 ± 0.01	0.71 ± 0.01
	60-85Hz	0.48 ± 0.03	0.63 ± 0.02
“low” vs “mid” chance level = 0.51	7-28Hz	0.62 ± 0.03	0.57 ± 0.03
	60-85Hz	0.53 ± 0.01	0.52 ± 0.02
“high” vs “mid” chance level = 0.50	7-28Hz	0.40 ± 0.01	0.38 ± 0.02
	60-85Hz	0.55 ± 0.03	0.52 ± 0.02

“low” vs “mid” 7-28Hz	0.42 ± 0.01	0.43 ± 0.01
60-85Hz chance level = 0.58	0.63 ± 0.04	0.56 ± 0.01
“mid” vs “ee” 7-28Hz	0.58 ± 0.01	0.56 ± 0.01
60-85Hz chance level = 0.51	0.74 ± 0.04	0.71 ± 0.01
“mid” vs “ooo” 7-28Hz	0.51 ± 0.01	0.51 ± 0.01
60-85Hz chance level = 0.50	0.62 ± 0.04	0.58 ± 0.01

Table 3. MDM for two locations and 5 classes Accuracies

The highest accuracies for discerning between two imaginary tasks are reached with the left hemisphere location and with the repetitive imaginary tasks, it is remarkable as well that the information from beta and alpha band is highly discriminative against the resting state with any imagery task confirming Chapter 3 results.

Discrimination accuracy between words is not higher from the rhythmic vowels, however, /mid/ word reaches the highest accuracies, specially when classified against /ee/, then a hybrid design taking /mid/ word is worth trying for the online prototype.

Therefore the online implementation of the system would use the left hemisphere location and the bandpass division to categorize firstly if the signal is coming from a resting state, in the case covariance distance is closer to an imaginary class, the system would classify it again whether the signal represents /ee/, /ooo/ or /mid/ imaginary tasks.

4.3 Real-time Implementation

This implementation designed on Matlab Simulink sends a MIDI signal out for an external analogue synthesizer. Because of the results in offline analysis, the simulation prototype would split the signal into two channels for high and low pass filters as shown in the Simulink module diagram shown in figure 23.

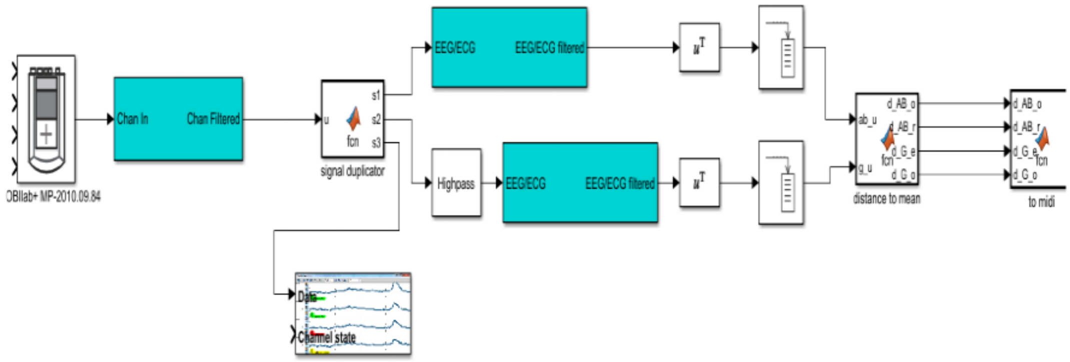


Fig. 23. Simulink Diagram of BCMI

The divisions of the signal are buffered into 2-sec blocks which are converted into alpha-beta and gamma covariance matrices using the Matlab function for covariance matrices within the “distance to mean" module. Thus the matrices are passed to the MIDI module as parameters.

MIDI module has the selection step. From the alpha-beta covariance matrix, the distance to rest alpha-beta and /ooo/ alpha-beta is first compared, this is because /ooo/ and rest centroids were closer to each other than /ee/ or /mid/ this led to false positives between /ooo/ and rest states, this first comparison ensured that most of the times that signal did not come from rest state. Following a straightforward discriminative logic, then if the distance is closer to rest then nothing would happen. If the distance is closer to /ooo/ then the distance from /ee/ gamma, /ooo/ gamma and /mid/ gamma to the upcoming gamma covariance matrix is compared depending on which is closer, the module would send a MIDI message using the audio toolbox, the midi note to be sent is D5 in case of /ee/, G4 in case of /mid/ and C3 in case of /ooo/. The decision login is shown in figure 24. The performance of the real-time implementation is shown in the confusion matrix on Table 3.

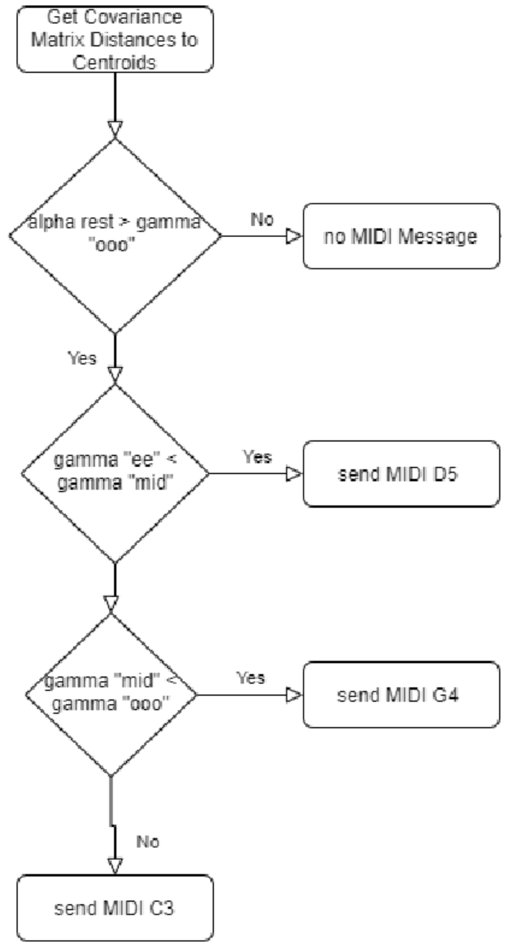


Fig. 24 Decision tree diagram

4.4 Real-time Results

Six sessions were taken in place varying the day and time of the system test, for each a new set of training data had to be recorded. Each session tried the system during a 20 min period, where 100 trials gave the following average confusion matrix.

	class “ee”	class “ooo”	class mid	class rest
class “ee”	11 ± 3	2 ± 1	4 ± 5	2 ± 1
class “ooo”	2 ± 4	8 ± 4	4 ± 3	5 ± 4
class mid	5 ± 2	3 ± 2	9 ± 2	2 ± 2
class rest	3 ± 1	9 ± 4	4 ± 2	27 ± 5

Table 4. Real-time System Confusion matrix

The system performance variation between days was clear, however accuracies did not scale significantly, with overall accuracy of the system being 0.57 ± 3. The performance was lower than the theoretically expected from offline analysis in Chapter 3, this was caused because of the chance decreasing from the decision tree, classification against resting state had accuracies similar to offline but the accuracy decreased on each compression step producing more false negative and low accuracy for the imagery tasks.

4.5 Conclusion

This proof-of-concept is an approach that proves Riemannian classification methods are promising for BCI control applications, the easy discretion process of MDM classifier makes the online system easy to build needing just the centroids as parameters to compare against. Because of the Riemannian metric being a distance value there is a variation that was applied to enhance the system accuracy up to 69% for classification of the pairs /ooo/ and /ee/, whenever an /ooo/ class was wrongly selected instead of /ee/, the difference between these two distances was smaller than 0.03 meanwhile when a true positive /ooo/ was selected in average this difference was greater than 0.08, that marked a threshold that could help to discriminate further. Is worth mentioning that the resting state was an eyes-closed resting-state because the prominence of alpha and beta bands, in this case, made the signal more discriminative in lower frequencies as shown in the results than a rest state with eyes open in previous studies. There are some other Riemannian classification methods that can reach higher accuracies than MDM as Correntropy Spectral Density or Tangent Space (Bakhshali et al. 2020) classifiers that could be tried for music approaches.

The results of this prototype could vary between days and the performance was better whenever the MCM centroids were taken the same day than the system testing, averaging data from multiple days and times resulted in a decrease in the overall accuracy. Same day data for train and test led to small variation on overal real-time system as seen in Table 3.

There are still clear challenges for Speech Imagery multiclass classification because of the nature of EEG signals, but the results obtained in the offline analysis showed spatially differentiable activity on the studied Broca's and Wricken's areas for gamma-band, it is interesting as well noticing the accuracy increment with the repetitive imagery task in concern with the single task, as covariance matrixes encode the variation of power and space of the time series, repetitive imagery activity causes better discrimination.

The time-frequency analysis shows that classes including the "e" vowel show positive deviations from the baseline in alpha and beta bands while it is negative for classes using the " o " vowel around the same electrode location, these results can be seen in figure 21, as well as in figure 25 where records were averaged from different recording session , these deviations may mark certain rebounds for these different tasks as it can be seen they are caused in different channel locations.

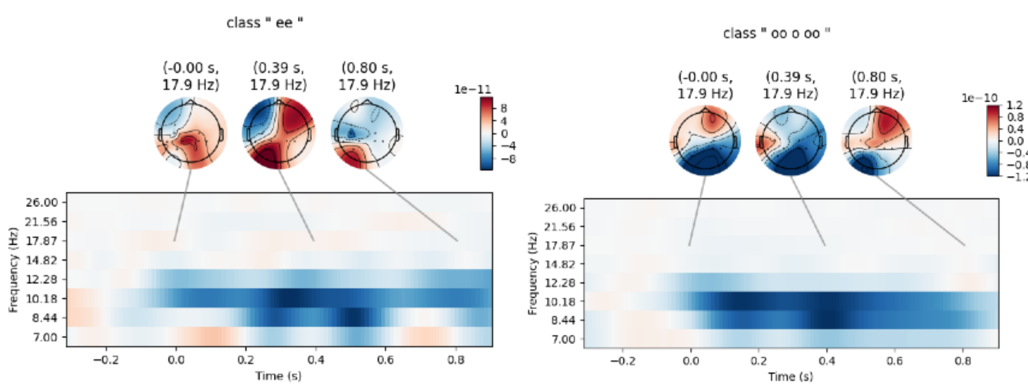


Fig. 25. Time frequency representation of class /ee/ and /ooo/ for alpha-beta band

This prototype had an overall good user experience allowing the participant to play three notes from a synthesizer without the need of stimulus and using a PC display to mark only a fixation cross as a settling point. There are different possibilities for extension of this work if a higher number of classes can be discerned and as mentioned in Chapter 3, this music approach can be generalized to control any other sort of device.

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 Overview

This thesis encompasses the delivery of a proof-of-concept BCI system for music control with Speech Imagery. This delivery is beneficial for people who suffer from motor disabilities, as a new way to connect with creative activities, it presents a novel dimension of control for musicians adventurous into computer music. This chapter is a summary of various conclusions that have emerged in this research's journey. It has presented different state-of-art decoding methods for Speech Imagery and selected the most optimal for a later presentation of a proof-of-concept system that enables the user to play an instrument with the imagination about vowel pronunciations. There were additional deductions based on the final delivery performance which are discussed in this chapter. The late sections offer different pathways to analyze for further work and research in this interdisciplinary field.

5.2 Contributions on Research Questions

This section presents how the research introduced an answer to the research questions that addressed this research.

RQ1 *Can we design a control BCMI system for instrument performance with a desirable musical experience?*

The second chapter in this thesis conducted a survey about methods that proved useful for control BCMI analyzing two criteria, the dimension of control that each method can have and their usability. Based on usability it was clearly stated that EEG is the optimal way of brain activity measurement, therefore three EEG techniques were evaluated, P300, SSVEP and ERD/ERS where SSVEP was selected as the most suitable method for BCMI performance based on the dimensionality of control criterion. However, SSVEP evaluation indicates that the overall user experience for musical applications is not the best because the method needs to display visual stimulus and needs the users to constantly gaze at them, and there has not be any approach that controls a signal instrument and triggered notes, therefore speech imagery related ERD/ERS is proposed as a new exploratory technic whew the user commands the system by thinking about speak. However the built real-time prototype did not answer the question as fluid performance of the instrument couldn't be obtained, because of its classifying accuracy.

Contribution 1.1 *Using low density dry and wireless electrode headsets*

Although Speech Imagery has activities that can be measured by EEG in different regions of the brain, this experiment conducts experiments with an 8 dry electrode array headset manufactured by G.Tec. The research deduced that a placement focused on the left hemisphere with electrodes at F5, FT7, FC5, FC3, C5, TP7, CP5, CP3 achieved workable results with a short set-up time, around 5 minutes, due to not extra conductive-fluid was needed. The headset connects through Bluetooth and can be used rapidly with the manufacturer interface for Simulink Matlab. As shown in chapters 3 and 4 the headset presents optimal reliability while being portable and easy to use for SI-based BCMI applications.

Contribution 1.2 *ERD/ERS allows a new BCMI experience.*

The implemented BCM has a novel methodology used for the first time in a Musical application; all of the current BCMI applications have either sets of composed music and pre-recorded sounds that were activated by user selection, generally based on a visual stimulus. Other BCMI applications have an instrument whose musical pieces were already set and the brain activity changed its parameters. Whilst the implemented prototype actuated an analogue synthesizer, note by note, so the pitch of the note was selected based on the speech imagery task the user was attempting, this interaction happened regardless of the sight direction of the user or its standing position. Even if the user had just 3 options to control, and the overall accuracy of the system was not optimal, the musical experience was described as desirable, as the instrument would only actuate based on the user's concentration to imagine. One of the options was the instruction to the device not to actuate that had as a mental activity an eye-closed resting state. The musical properties controlled were fundamentally basic reassembling the performance of a classical instrument, using just single tone notes and the instrument didn't produce any sound as soon as the user closed the eyes.

RQ2 *Is it possible to build a proof-of-concept control BCMI system that enables the user to play a musical instrument with their attempt to imagine pronouncing vocal sounds?*

This project links the cognitive activity of imagining vocalizing sounds with music imagery, as some people experience musical imagery by the act of silent humming. Therefore Speech Imagery as a mental activity to measure from EEG led to the investigation about the feasibility to implement a system to control music. There have been different studies that showed the use of statistical tools with computer algorithms can help to extract useful information and decode SI-related activity, this is the case of the evaluated methods, CSP that is useful to extract variations related to spatial differences of the signals, Covariance Matrices that represent a correlation of the signal between the different channels. And the evaluated classification methods that label if the EEG trails belong to a speech imagery attempt, SVM and LDA are machine learning classifiers widely used in BCI and Riemannian MDM, a novel approach to EEG classification. The evaluation concludes that the accuracies of the tested methods were mostly the same, but a set of covariances matrices classified by Riemannian MDM was relatively simpler and direct to implement, this led to building a first prototype that reached the question's answer and enabled the user to play an analogue synthesizer by their imagination of producing a sound with the mouth.

Contribution 2.1 *State-of-the-art ERD/ERS decoding methodologies are useful for Speech Imagery classification*

Chapter 3 of this thesis did an evaluation of three popular methods to decode Speech Imagery: a combination of CSP feature extraction method with SVM and LDA classifiers and Covariance Matrices classified by a Riemannian MDM. In order to test these methods, a set of EEG experiments took place for three speech imagery classes (sustained vowel o, sustained vowel e and resting-state condition) where CSP reached accuracies of up to 65%, LDA accuracies of up to 64% and Riemannian MDM accuracies of up to 68% for pairwise classification that were significantly high than the 53% chance level. The evaluation confirmed that these methods are useful to decode Speech Imagery information and therefore can be used for the implementation of an online Speech Imagery BCI.

Contribution 2.2 *Changes between Speech Imagery tasks frequency, duration and meaning led to improved results*

Chapter 4 of this thesis evaluated the use of different speech imagery tasks, the second set of experiments conducted were based on rhythmic imagination of vowels which EEG feature extraction and classification were tested with the Riemannian MDM method as selected in Chapter 2 led to this BCMI investigation. Repetitive imagination of letter /e/ two times and /o/ three times led to an increase in accuracy of up to 70%. Evaluation on different power bands led to significant differences in the classification, alpha and beta bands were more discriminative for comparing against resting state, while gamma bands led to better classifications between the imagery. The third set of EEG experiments was conducted using words instead, having as classes the words /high/, /mid/ and /low/, accuracies were significant but there was evident discriminative information for word /mid/ as it led to the highest accuracies when classified pairwise. Based on those experiments a last of EEG experiments were conducted in order to try a mixed set of imagery tasks having rhythmic vowels /e e/ and /o o o/, word /mid/ and resting state condition, where the classification method reached an accuracy up to 72% for pairwise condition and a 31% in 4-class classification above 25% chance level.

Contribution 2.3 *MDM Riemannian classifier is can be implemented to build a BCMI system that controls a musical instrument with Speech Imagery*

Chapter 4 of this thesis proved that the simple discriminations functions of Riemannian MDM classification of covariance matrices made implementation of the BCMI system feasible, allowing it to process and classify 3 seconds of EEG data in less than 500 milliseconds. Based on chapter 4 offline accuracy results, the final workflow of the system was a pairwise classification three comparing between the two different frequency bands evaluated. The overall accuracy of the final subject-independent prototype was 51% for the three classes and a resting state condition class. The music notes that the system actuated depending on the imagery label was D5 for /e e/, G4 for /mid/ and C3 for/o o o/. Musical information was sent using a MIDI function of Matlab toolbox for audio processing through a USB interface connecting the analogue synthesizer MIDI plug.

5.3 Additional Contributions

In addition to the research questions, the following contributions were deducted about the use of Speech Imagery for BCI systems.

Contribution 3.1 *Rhythmic differences of speech imagery tasks can be discriminative in lower EEG frequencies*

Chapter 4 presented time-frequency analysis between different frequency bands to compare activation rates between the imagery tasks, in the case of repetitive /e e/ the PSD plots showed marked power clusters around 7 to 20Hz in comparison with shorter length imagery classes as /mid/ word, single repetition of vowels or sustain imagination of vowels. Consequently, when the classification was set to the mentioned frequency bands, the accuracy was incremented, however, it did not change significantly when comparing between single time imagery tasks.

Contribution 3.2 *Eyes closed resting state led to better classification accuracies*

One of the variations between the different EEG experiments conducted in chapter 4 was the modification of regular resting state with eye closed ones. An eye closed resting state was evaluated to significantly increase the classification accuracies going from 68% with eyes open resting state up to 90% accuracy on its counterpart. Spectral density topographical maps showed that this modification led to a significant decrease of the activations related to the occipital region, better known to be responsible for vision-related activities.

5.4 Future Work

This project combines elements of computer music, neuroscience and biomedical engineering. This mixture of disciplines offers high potential for future research. Below there are two pathways that can be adopted for future work in the field.

Pathway 1 - *Use of different EEG configurations*

This project used an 8 electrode headset focusing on the sensorimotor and left hemisphere regions of the Brain. The speech imagery signals are expected to be maximal in the mentioned regions, however, the use of higher electrode densities can detect relations among other brain regions and therefore lead to better classifications, wet electrodes would also enhance signal quality, some 64 electrodes set placed among all areas can be considered for further investigation of speech-related potentials (Sree et al. 2021). The cost-benefit should be analyzed for a higher electrode density as the configuration time increases in relation to the number of electrodes.

Pathway 2 - *Use of hybrid BCI measurement methods and deep learning classifiers*

This project used EEG as a decoding method because of its good characteristics as being a non-invasive method, however combining multiple non-invasive measurement tools is a sensible consideration.

Even though EEG decoding is achieving promising results the method is prone to electrical noise and motion artifacts, therefore, fNIRS has been suggested and, subsequently, demonstrated to improve the overall performance of an EEG-based BCI system. fNIRS offers a reasonable spatial resolution and shares the safety and portability characteristics of EEG (Gratton et al. 2005). The optical intensities measured by fNIRS detectors indicate oxygenated hemoglobin concentration changes (HbO), which can be used as features to identify different neural activations (Rezazadeh Sereshkeh et al. 2019).

Hybrid EEG-fNIRS classification techniques can either use separate classifiers for each technique that would be joined by a probabilistic model to make predictions as proposed in (Sereshkeh et al. 2019) or they can have the features of both techniques as inputs to deep learning methods as shown in (Chiarelli et al. 2018).

Deep learning techniques are known to learn from complex forms of data, some of the latest studies have explored subject-independent SI classification. For example, work by Saha (Saha, Abdul-Mageed, and Fels 2019) proposed a hierarchical deep learning methodology combining spatial and temporal convolutional neural networks achieving an accuracy of 83% for six phoneme classes from 14 participants. The use of these classifiers is notably important to be researched for BCMI applications, however, the complexity of such algorithms has a constraint when implementing real-time applications.

5.4.1 Suggested Research Questions

The mentioned pathways raise the following research questions.

- Can a set of 64 wet electrodes EEG headset be feasible for implementing Speech Imagery BCMI systems with confident accuracy?

- Can a hybrid EEG-fNIRS BCI system increase the accuracy of the current state-of-art EEG decoding methods for Speech Imagery?

- Can the use of Deep Learning algorithms be implemented for a BCMI system with optimal time responses?

- Can we use Riemannian based classifiers to build subject-independent classifiers of a high number of Speech Imagery classes?

References

Al-Qazzaz, Noor Kamal, Zaid Abdi Alkareem Alyasseri, Karrar Hameed Abdulkareem, Nabeel Salih Ali, Mohammed Nasser Al-Mhiqani, and Christoph Guger. 2021. "EEG Feature Fusion for Motor Imagery: A New Robust Framework towards Stroke Patients Rehabilitation." *Computers in Biology and Medicine* 137 (August): 104799.

Ansari, Irshad Ahmad, and Rajesh Singla. 2016. "BCI: An Optimised Speller Using SSVEP." *International Journal of Biomedical Engineering and Technology*. https://doi.org/10.1504/ijbet.2016.078988.

Bakhshali, Mohamad Amin, Morteza Khademi, Abbas Ebrahimi-Moghadam, and Sahar Moghimi. 2020. "EEG Signal Classification of Imagined Speech Based on Riemannian Distance of Correntropy Spectral Density." *Biomedical Signal Processing and Control*. https://doi.org/10.1016/j.bspc.2020.101899.

Barachant, A., S. Bonnet, M. Congedo, and C. Jutten. 2012. "Multiclass Brain–Computer Interface Classification by Riemannian Geometry." *IEEE Transactions on Biomedical Engineering*. https://doi.org/10.1109/tbme.2011.2172210.

Binder, Jeffrey R. 2015. "The Wernicke Area." *Neurology*. https://doi.org/10.1212/wnl.0000000000002219.

Brigham, Katharine, and B. V. Vijaya Kumar. 2010. "Imagined Speech Classification with EEG Signals for Silent Communication: A Preliminary Investigation into Synthetic Telepathy." *2010 4th International Conference on Bioinformatics and Biomedical Engineering*. https://doi.org/10.1109/icbbe.2010.5515807.

Chen, Yao, Chao Zhang, and Xiaopei Wu. 2019. "To Assess the Influence of Artifacts on Motor Imagery Based BCI." *2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP)*. https://doi.org/10.1109/siprocess.2019.8868896.

Chen, Yonghao, Chen Yang, Xiaochen Ye, Xiaogang Chen, Yijun Wang, and Xiaorong Gao. 2021. "Implementing a Calibration-Free SSVEP-Based BCI System with 160 Targets." *Journal of Neural Engineering* 18 (4). https://doi.org/10.1088/1741-2552/ac0bfa.

Chen, Zongmei, Jing Jin, Ian Daly, Cili Zuo, Xingyu Wang, and Andrzej Cichocki. 2020. "Effects of Visual Attention on Tactile P300 BCI." *Computational Intelligence and Neuroscience* 2020 (February): 6549189.

Chiarelli, Antonio Maria, Pierpaolo Croce, Arcangelo Merla, and Filippo Zappasodi. 2018. "Deep Learning for Hybrid EEG-fNIRS Brain–computer Interface: Application to Motor Imagery Classification." *Journal of Neural Engineering*. https://doi.org/10.1088/1741-2552/aaaf82.

Choi, Kyuwan, and Andrzej Cichocki. 2008. "Control of a Wheelchair by Motor Imagery in Real Time." *Lecture Notes in Computer Science*. https://doi.org/10.1007/978-3-540-88906-9_42.

Chu, Yaqi, Xingang Zhao, Jianda Han, Yiwen Zhao, and Jun Yao. 2014. "SSVEP Based Brain-Computer Interface Controlled Functional Electrical Stimulation System for Upper Extremity Rehabilitation." *2014 IEEE International Conference on Robotics and Biomimetics (ROBIO 2014)*. https://doi.org/10.1109/robio.2014.7090671.

Cincotti, F., F. Pichiorri, P. Arico, F. Aloise, F. Leotta, F. de Vico Fallani, J. del R Millan, M. Molinari, and D. Mattia. 2012. "EEG-Based Brain-Computer Interface to Support Post-Stroke Motor Rehabilitation of the Upper Limb." *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. https://doi.org/10.1109/embc.2012.6346871.

Congedo, Marco, Alexandre Barachant, and Rajendra Bhatia. 2017. "Riemannian Geometry for EEG-Based Brain-Computer Interfaces; a Primer and a Review." *Brain-Computer Interfaces*. https://doi.org/10.1080/2326263x.2017.1297192.

Congedo, Marco, and Leslie Sherlin. 2011. "EEG Source Analysis." *Neurofeedback and Neuromodulation Techniques and Applications*. https://doi.org/10.1016/b978-0-12-382235-2.00002-0.

Daly, Ian, Etienne B. Roesch, James Weaver, and Slawomir J. Nasuto. 2014. "Machine Learning to Identify Neural Correlates of Music and Emotions." *Guide to Brain-Computer Music Interfacing*. https://doi.org/10.1007/978-1-4471-6584-2_5.

DaSalla, Charles S., Hiroyuki Kambara, Makoto Sato, and Yasuharu Koike. 2009. "Single-Trial Classification of Vowel Speech Imagery Using Common Spatial Patterns." *Neural Networks: The Official Journal of the International Neural Network Society* 22 (9): 1334–39.

Deng, Siyi, Ramesh Srinivasan, Tom Lappas, and Michael D'Zmura. 2010. "EEG Classification of Imagined Syllable Rhythm Using Hilbert Spectrum Methods." *Journal of Neural Engineering* 7 (4): 046006.

Deuel, Thomas A., Juan Pampin, Jacob Sundstrom, and Felix Darvas. 2017. "The Encephalophone: A Novel Musical Biofeedback Device Using Conscious Control of Electroencephalogram (EEG)." *Frontiers in Human Neuroscience* 11 (April): 213.

Diez, Pablo. 2018. *Smart Wheelchairs and Brain-Computer Interfaces: Mobile Assistive Technologies*. Academic Press.

Eaton, Joel, Duncan Williams, and Eduardo Miranda. 2015. "The Space Between Us: Evaluating a Multi-User Affective Brain-Computer Music Interface." *Brain-Computer Interfaces*. https://doi.org/10.1080/2326263x.2015.1101922.

Eichele, Tom, Matthias Moosmann, Lei Wu, Ingmar Gutberlet, and Stefan Debener. 2010. "Removal of MRI Artifacts from EEG Recordings." *Simultaneous EEG and fMRI*. https://doi.org/10.1093/acprof:oso/9780195372731.003.0006.

Farmaki, Christina, Georgios Christodoulakis, and Vangelis Sakkalis. 2016. "Applicability of SSVEP-Based Brain-Computer Interfaces for Robot Navigation in Real Environments." *Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference* 2016 (August): 2768–71.

Friedman, L., J. T. Kenny, A. L. Wise, D. Wu, T. A. Stuve, D. A. Miller, J. A. Jesberger, and J. S. Lewin. 1998. "Brain Activation during Silent Word Generation Evaluated with Functional MRI." *Brain and Language* 64 (2): 231–56.

Fujimaki, N., F. Takeuchi, T. Kobayashi, S. Kuriki, and S. Hasuo. 1994. "Event-Related Potentials in Silent Speech." *Brain Topography* 6 (4): 259–67.

Gao, Yu Yu. 2014. "Improved Common Spatial Patterns on EEG Feature Extraction." *Advanced Materials Research*.

https://doi.org/10.4028/www.scientific.net/amr.926-930.1814.
Gembler, Felix, Piotr Stawicki, and Ivan Volosyak. 2016. "Exploring the Possibilities and Limitations of Multitarget SSVEP-Based BCI Applications." *Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference* 2016 (August): 1488–91.
"Genesis of MEG and EEG." 2019. *Brain Signals*. https://doi.org/10.7551/mitpress/10745.003.0003.
Gramfort, Alexandre, Martin Luessi, Eric Larson, Denis A. Engemann, Daniel Strohmeier, Christian Brodbeck, Roman Goj, et al. 2013. "MEG and EEG Data Analysis with MNE-Python." *Frontiers in Neuroscience* 7 (December): 267.
Gratton, Enrico, Vladislav Toronov, Ursula Wolf, Martin Wolf, and Andrew Webb. 2005. "Measurement of Brain Activity by near-Infrared Light." *Journal of Biomedical Optics* 10 (1): 11008.
Grierson, Mick, and Chris Kiefer. 2014. "Contemporary Approaches to Music BCI Using P300 Event Related Potentials." *Guide to Brain-Computer Music Interfacing*. https://doi.org/10.1007/978-1-4471-6584-2_3.
Guan, Shan, Kai Zhao, and Shuning Yang. 2019. "Motor Imagery EEG Classification Based on Decision Tree Framework and Riemannian Geometry." *Computational Intelligence and Neuroscience*. https://doi.org/10.1155/2019/5627156.
Horikawa, T., M. Tamaki, Y. Miyawaki, and Y. Kamitani. 2013. "Neural Decoding of Visual Imagery during Sleep." *Science* 340 (6132): 639–42.
Hubner, David, Albrecht Schall, and Michael Tangermann. 2019. "Two Player Online Brain-Controlled Chess." *Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference* 2019 (July): 3018–21.
Hucker, Kyra. 2007. *Investigating Object Primes and Deception Using a P300 ERP Oddball Paradigm*.
Iijima, Mutsumi, Mikio Osawa, Ryuzo Mase, Ryosuke Ushijima, and Shinichiro Uchiyama. 2009. "103. Event-Related Synchronization and Event-Related Desynchronization of Gamma Band Activity in Visual Go/nogo Paradigm." *Clinical Neurophysiology*. https://doi.org/10.1016/j.clinph.2009.02.109.
Iwatsuki, Katsuyuki, Minoru Hoshiyama, Shintaro Oyama, Hidemasa Yoneda, Shingo Shimoda, and Hitoshi Hirata. 2020. "Electroencephalographic Functional Connectivity With the Tacit Learning System Prosthetic Hand: A Case Series Using Motor Imagery." *Frontiers in Synaptic Neuroscience* 12 (February): 7.
Jaseja, Harinder, and Bhanu Jaseja. 2012. "EEG Spike versus EEG Sharp Wave: Differential Clinical Significance in Epilepsy." *Epilepsy & Behavior*. https://doi.org/10.1016/j.yebeh.2012.05.023.
Jeannerod, M. 1995. "Mental Imagery in the Motor Context." *Neuropsychologia*. https://doi.org/10.1016/0028-3932(95)00073-c.
Jiang, Lijun, Eugene Tham, Mervyn Yeo, Zaixing Wang, and Bo Jiang. 2014. "Motor Imagery Controlled Wheelchair System." *2014 9th IEEE Conference on Industrial Electronics and Applications*. https://doi.org/10.1109/iciea.2014.6931221.

Joadder, Md A. Mannan, Joshua J. Myszewski, Mohammad H. Rahman, and Inga Wang. 2019. "A Performance Based Feature Selection Technique for Subject Independent MI Based BCI." *Health Information Science and Systems* 7 (1): 15.

Kaongoen, Netiwit, Jaehoon Choi, and Sungho Jo. 2021. "Speech-Imagery-Based Brain–computer Interface System Using Ear-EEG." *Journal of Neural Engineering*. https://doi.org/10.1088/1741-2552/abd10e.

Kim, Soram, Seungyun Lee, Hyunsuk Kang, Sion Kim, and Minkyu Ahn. 2021. "P300 Brain-Computer Interface-Based Drone Control in Virtual and Augmented Reality." *Sensors* 21 (17). https://doi.org/10.3390/s21175765.

Kiyi, Ilayda, Ibrahim Oztura, and Gorsev Yener. 2021. "The Evaluation of P300 Component Differences Across Sensory Modalities in Healthy Individuals Aged 50 and Over: Auditory and Visual." *International Journal of Psychophysiology*. https://doi.org/10.1016/j.ijpsycho.2021.07.394.

Knězík, Jan, and Martin Drahanský. 2007. "Simple EEG Driven Mouse Cursor Movement." *Advances in Soft Computing*. https://doi.org/10.1007/978-3-540-75175-5_66.

Kolodziej, Marcin, Andrzej Majkowski, and Remigiusz J. Rak. 2016. "Automatic Detection of SSVEP Using Independent Component Analysis." *2016 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*. https://doi.org/10.1109/spa.2016.7763612.

Kovacs, Annamaria, Istvan Winkler, and Klara Vicsi. 2017. "EEG Correlates of Speech: Examination of Event Related Potentials Elicited by Phoneme Classes." *2017 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. https://doi.org/10.1109/coginfocom.2017.8268226.

Krishna, D. Hari, D. Hari Krishna, I. A. Pasha, and T. Satya Savithri. 2018. "Multiclass Classification of Motor Imagery EEG Signals Using Ensemble Classifiers & Cross-Correlation." *International Journal of Engineering & Technology*. https://doi.org/10.14419/ijet.v7i2.6.10144.

Lee, Jaehyung, Kabmun Cha, Hyungmin Kim, Junhyuk Choi, Choonghyun Kim, and Songjoo Lee. 2019. "Hybrid MI-SSSEP Paradigm for Classifying Left and Right Movement toward BCI for Exoskeleton Control." *2019 7th International Winter Conference on Brain-Computer Interface (BCI)*. https://doi.org/10.1109/iww-bci.2019.8737319.

Li, Kun. 2011. *Advanced Signal Processing Techniques for Single Trial Electroencephalography Signal Classification for Brain Computer Interface Applications*.

Lotte, Fabien, and Marco Congedo. 2016. "EEG Feature Extraction." *Brain-Computer Interfaces 1*. https://doi.org/10.1002/9781119144977.ch7.

Lotte, F., L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger. 2018. "A Review of Classification Algorithms for EEG-Based Brain-Computer Interfaces: A 10 Year Update." *Journal of Neural Engineering* 15 (3): 031005.

Malina, Roger F. 2002. "The Stone Age of the Digital Arts." *Leonardo*. https://doi.org/10.1162/002409402320774286.

Mao, Ying, Jing Jin, Shurui Li, Yangyang Miao, and Andrzej Cichocki. 2021. "Effects of Skin Friction on Tactile P300 Brain-Computer Interface Performance." *Computational*

Intelligence and Neuroscience 2021 (February): 6694310.
Marque, P. 2012. "Non-Invasive Brain Computer Interfaces." *Annals of Physical and Rehabilitation Medicine*. https://doi.org/10.1016/j.rehab.2012.07.873.
Martišius, Ignas, and Robertas Damaševičius. 2016. "A Prototype SSVEP Based Real Time BCI Gaming System." *Computational Intelligence and Neuroscience* 2016 (March): 3861425.
Miranda, Eduardo Reck. 2014. "Brain–Computer Music Interfacing: Interdisciplinary Research at the Crossroads of Music, Science and Biomedical Engineering." *Guide to Brain-Computer Music Interfacing*. https://doi.org/10.1007/978-1-4471-6584-2_1.
Miranda, Eduardo Reck, and Julien Castet. 2014. *Guide to Brain-Computer Music Interfacing*. Springer.
Miranda, E. R., W. L. Magee, J. J. Wilson, J. Eaton, and R. Palaniappan. 2011. "Brain-Computer Music Interfacing (BCMI): From Basic Research to the Real World of Special Needs." *Music and Medicine*. https://doi.org/10.1177/1943862111399290.
Mitchell, Tom M., Svetlana V. Shinkareva, Andrew Carlson, Kai-Min Chang, Vicente L. Malave, Robert A. Mason, and Marcel Adam Just. 2008. "Predicting Human Brain Activity Associated with the Meanings of Nouns." *Science* 320 (5880): 1191–95.
Molina-Cantero, Alberto J., Juan A. Castro-García, Fernando Gómez-Bravo, Rafael López-Ahumada, Raúl Jiménez-Naharro, and Santiago Berrazueta-Alvarado. 2021. "Controlling a Mouse Pointer with a Single-Channel EEG Sensor." *Sensors* 21 (16). https://doi.org/10.3390/s21165481.
Muller-Putz, Gernot R., Rudiger Rupp, Patrick Ofner, Joana Pereira, Andreas Pinegger, Andreas Schwarz, Marcel Zube, Ute Eck, Bjorn Hessing, and Matthias Schneiders. 2019. "Applying Intuitive EEG-Controlled Grasp Neuroprostheses in Individuals with Spinal Cord Injury: Preliminary Results from the MoreGrasp Clinical Feasibility Study." *Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference* 2019 (July): 5949–55.
Norcia, Anthony M., L. Gregory Appelbaum, Justin M. Ales, Benoit R. Cottereau, and Bruno Rossion. 2015. "The Steady-State Visual Evoked Potential in Vision Research: A Review." *Journal of Vision* 15 (6): 4.
Oikonomou, Vangelis P., Anastasios Maronidis, George Liaros, Spiros Nikolopoulos, and Ioannis Kompatsiaris. 2017. "Sparse Bayesian Learning for Subject Independent Classification with Application to SSVEP-BCI." *2017 8th International IEEE/EMBS Conference on Neural Engineering (NER)*. https://doi.org/10.1109/ner.2017.8008423.
Olesen, Simon Dahl Thorsager, Rig Das, Mathias Dizon Olsson, Muhammad Ahmed Khan, and Sadasivan Puthusserypady. 2021. "Hybrid EEG-EOG-Based BCI System for Vehicle Control." *2021 9th International Winter Conference on Brain-Computer Interface (BCI)*. https://doi.org/10.1109/bci51272.2021.9385300.
Orrison, William W., Jeffrey Lewine, John Sanders, and Michael F. Hartshorne. 2017. *Functional Brain Imaging*. Elsevier Health Sciences.
Othman, Marini. 2014. *EEG Affective State Profiling for Understanding and Analyzing Children's Brain Development*.
Palaniappan, Ramaswamy. 2014. "Electroencephalogram-Based Brain–Computer Interface:

An Introduction.” *Guide to Brain-Computer Music Interfacing*. https://doi.org/10.1007/978-1-4471-6584-2_2.

Pfurtscheller, G., C. Brunner, A. Schlögl, and F. H. Lopes da Silva. 2006. “Mu Rhythm (de)synchronization and EEG Single-Trial Classification of Different Motor Imagery Tasks.” *NeuroImage* 31 (1): 153–59.

Pfurtscheller, G., and F. H. Lopes da Silva. 1999. “Event-Related EEG/MEG Synchronization and Desynchronization: Basic Principles.” *Clinical Neurophysiology*. https://doi.org/10.1016/s1388-2457(99)00141-8.

Polich, John, Christine Ladish, and Tim Burns. 1990. “Normal Variation of P300 in Children: Age, Memory Span, and Head Size.” *International Journal of Psychophysiology*. https://doi.org/10.1016/0167-8760(90)90056-j.

Rezazadeh Sereshkeh, Alborz, Rozhin Yousefi, Andrew T. Wong, and Tom Chau. 2019. “Online Classification of Imagined Speech Using Functional near-Infrared Spectroscopy Signals.” *Journal of Neural Engineering* 16 (1): 016005.

Riaz, Anaum, Sana Akhtar, Shanza Iftikhar, Amir Ali Khan, and Ahmad Salman. 2014. “Inter Comparison of Classification Techniques for Vowel Speech Imagery Using EEG Sensors.” *The 2014 2nd International Conference on Systems and Informatics (ICSAI 2014)*. https://doi.org/10.1109/icsai.2014.7009378.

Rybář, Milan, Riccardo Poli, and Ian Daly. 2021. “Decoding of Semantic Categories of Imagined Concepts of Animals and Tools in fNIRS.” *Journal of Neural Engineering* 18 (4). https://doi.org/10.1088/1741-2552/abf2e5.

Saha, Pramit, Muhammad Abdul-Mageed, and Sidney Fels. 2019. “SPEAK YOUR MIND! Towards Imagined Speech Recognition with Hierarchical Deep Learning.” *Interspeech 2019*. https://doi.org/10.21437/interspeech.2019-3041.

Schnitzler, Alfons, and Joachim Gross. 2010. “Magnetoencephalography (MEG).” *Encyclopedia of Movement Disorders*. https://doi.org/10.1016/b978-0-12-374105-9.00341-5.

Sereshkeh, Alborz Rezazadeh, Rozhin Yousefi, Andrew T. Wong, Frank Rudzicz, and Tom Chau. 2019. “Development of a Ternary Hybrid fNIRS-EEG Brain–computer Interface Based on Imagined Speech.” *Brain-Computer Interfaces*. https://doi.org/10.1080/2326263x.2019.1698928.

Sitaram, R., N. Weiskopf, A. Caria, R. Veit, M. Erb, and N. Birbaumer. 2008. “fMRI Brain-Computer Interfaces.” *IEEE Signal Processing Magazine*. https://doi.org/10.1109/msp.2008.4408446.

Sozer, Abdullah Talha. 2018. “Enhanced Single Channel SSVEP Detection Method on Benchmark Dataset.” *2018 15th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*. https://doi.org/10.1109/iceee.2018.8533933.

Sree, R. Anandha, R. Anandha Sree, B. Divya, and A. Kavitha. 2021. “Task Specific Brain Synchronization in EEG Based Speech and Speech Imagery Procedures.” *Journal of Physics: Conference Series*. https://doi.org/10.1088/1742-6596/1937/1/012044.

Sun, Qiang, Minyou Chen, Li Zhang, Changsheng Li, and Wenfa Kang. 2021. “Similarity-Constrained Task-Related Component Analysis for Enhancing SSVEP Detection.” *Journal of Neural Engineering* 18 (4).

https://doi.org/10.1088/1741-2552/abfdfa.
"The Ten Twenty Electrode System: International Federation of Societies for Electroencephalography and Clinical Neurophysiology." 1961. *American Journal of EEG Technology*. https://doi.org/10.1080/00029238.1961.11080571.
Tibrewal, Navneet, Nikki Leeuwis, and Maryam Alimardani. n.d. "The Promise of Deep Learning for BCIs: Classification of Motor Imagery EEG Using Convolutional Neural Network." https://doi.org/10.1101/2021.06.18.448960.
Tseng, Kevin C., Bor-Shing Lin, Alice May-Kuen Wong, and Bor-Shyh Lin. 2015. "Design of a Mobile Brain Computer Interface-Based Smart Multimedia Controller." *Sensors* 15 (3): 5518–30.
Vaque, T. J. La, and T. J. La Vaque. 1999. "The History of EEG Hans Berger." *Journal of Neurotherapy*. https://doi.org/10.1300/j184v03n02_01.
Wang, Li, Xiong Zhang, Xuefei Zhong, and Yu Zhang. 2013. "Analysis and Classification of Speech Imagery EEG for BCI." *Biomedical Signal Processing and Control*. https://doi.org/10.1016/j.bspc.2013.07.011.
Wolpaw, Jonathan R., Gerald E. Loeb, Brendan Z. Allison, Emanuel Donchin, Omar Feix do Nascimento, William J. Heetderks, Femke Nijboer, William G. Shain, and James N. Turner. 2006. "BCI Meeting 2005--Workshop on Signals and Recording Methods." *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society* 14 (2): 138–41.
Wolpaw, J. R., N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan. 2000. "Brain-Computer Interface Technology: A Review of the First International Meeting." *IEEE Transactions on Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society* 8 (2): 164–73.
Yger, Florian, Maxime Berar, and Fabien Lotte. 2017. "Riemannian Approaches in Brain-Computer Interfaces: A Review." *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society* 25 (10): 1753–62.
Yuan, Han, and Bin He. 2014. "Brain-Computer Interfaces Using Sensorimotor Rhythms: Current State and Future Perspectives." *IEEE Transactions on Bio-Medical Engineering* 61 (5): 1425–35.
Zapała, Dariusz, Piotr Francuz, Ewelina Zapała, Natalia Kopiś, Piotr Wierzgała, Paweł Augustynowicz, Andrzej Majkowski, and Marcin Kołodziej. 2018. "The Impact of Different Visual Feedbacks in User Training on Motor Imagery Control in BCI." *Applied Psychophysiology and Biofeedback* 43 (1): 23–35.
Zhang, Meili, and Kewen Hou. 2021. "Research on the Application of Computer Music Making Technology in New Media Environment." *Journal of Physics: Conference Series*. https://doi.org/10.1088/1742-6596/1871/1/012142.