Brain Computer Interfaces for the Control of Robotic Swarms

by

Georgios Konstantinos Karavas

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved June 2017 by the
Graduate Supervisory Committee:

Panagiotis Artemiadis, Chair
Spring Berman
Hyunglae Lee

ARIZONA STATE UNIVERSITY

August 2017

ABSTRACT

A *robotic swarm* can be defined as a large group of inexpensive, interchangeable robots with limited sensing and/or actuating capabilities that cooperate (explicitly or implicitly) based on local communications and sensing in order to complete a mission. Its inherent redundancy provides flexibility and robustness to failures and environmental disturbances which guarantee the proper completion of the required task. At the same time, human intuition and cognition can prove very useful in extreme situations where a fast and reliable solution is needed. This idea led to the creation of the field of *Human-Swarm Interfaces (HSI)* which attempts to incorporate the human element into the control of robotic swarms for increased robustness and reliability. The aim of the present work is to extend the current state-of-the-art in HSI by applying ideas and principles from the field of *Brain-Computer Interfaces (BCI)*, which has proven to be very useful for people with motor disabilities. At first, a preliminary investigation about the connection of brain activity and the observation of swarm collective behaviors is conducted. After showing that such a connection may exist, a hybrid BCI system is presented for the control of a swarm of quadrotors. The system is based on the combination of motor imagery and the input from a game controller, while its feasibility is proven through an extensive experimental process. Finally, speech imagery is proposed as an alternative mental task for BCI applications. This is done through a series of rigorous experiments and appropriate data analysis. This work suggests that the integration of BCI principles in HSI applications can be successful and it can potentially lead to systems that are more intuitive for the users than the current state-of-the-art. At the same time, it motivates further research in the area and sets the stepping stones for the potential development of the field of *Brain-Swarm Interfaces (BSI)*.

ACKNOWLEDGMENTS

This has been an incredible journey, which taught me a lot and helped me grow both as a person and as a professional. I had the luck to meet and work with some incredible people who guided me and supported me throughout this, while also trusting that I will succeed.

First and foremost, I would like to thank my advisor Professor Panagiotis Artemiadis who gave me a unique opportunity to work on this project and whose guidance provided me the key to unlock my potential. He pushed me to constantly improve myself and he helped me exceed my capabilities for which I will be always in debt to him.

I would also like to acknowledge my colleagues and friends in the HORC lab whose friendship and support created an amazing working environment that allowed me to thrive. I am especially thankful to Daniel Larsson, Zahi Kakish and Chuong Nguyen whose help during the course of this project was invaluable and without whom I would not be able to achieve my goals. I am also greatly thankful to Bryan Whitsell, Jeffrey Skidmore, Mark Ison, Justin Hunt, Varun Nalam, Keivan Mojtahedi and Mohammad Alzorgan, and others for their very important inputs and feedback, but also for their friendship and support.

Last but not least, I would like to thank my parents and my brother for their support, their understanding and because they let me follow my dreams even if that meant that I would be so far away from them.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

PREFACE

The following thesis is the result of three academic years of research in the Human Oriented Robotics and Control Lab at Arizona State University. The research performed during this time has currently resulted in 2 published papers in peer-reviewed conference proceedings. These contributions are outlined in Appendix A.

Some of the following chapters contain portions of these papers which are correspondingly cited within the text. These portions from previously published papers have been included, with permission, because they each contribute to the investigation of using Brain-Computer Interface ideas and principles in order to control a robotic swarm, which is the main topic of this thesis. Permission has been granted for using copyrighted material in this thesis (see Appendix B), and all co-authors have given permission to include material from co-authored papers (see Appendix C).

More specifically, Chapter 3 discusses the effects of swarm collective behaviors on human brain activity. This chapter uses material from the publication in the proceedings of the 2015 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (Karavas and Artemiadis, 2015).

Finally, Chapter 4 presents a hybrid BCI for the control of a small team of quadrotors. This chapter uses material from the oncoming publication on the proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (Karavas *et al.*, 2017).

Chapter 1

INTRODUCTION

A *robotic swarm* (Beni, 2005) can be defined as a large group of simple, inexpensive robots with limited sensing and/or actuating capabilities that rely on local communication and sensing in order to perform a specific mission. The members of a swarm can also be referred to as *agents*. The communication between agents can be *explicit*, e.g. using wireless technology, or *implicit*. In the latter case, the agents interact with their environment and detect possible changes in it realized by the other agents in the group in order to adjust their behavior accordingly. In addition, the control of the swarm can be performed in a *decentralized* manner, where each agent acts solely on its on-board sensing and logic, or in a *centralized* one, where there is a global controller that assigns roles and commands to each agent.

In general, a system like this has two main advantages. First, its redundancy allows for flexibility and robustness to failures and external disturbances during task execution. That in turn guarantees the completion of the mission. If one or more members of the team fail, the rest of them can still complete it. Second, the overall behavior of the swarm emerges in most cases from the behaviors of its individual agents. This means that complex tasks can be executed by using simple algorithms which significantly reduces development and maintenance costs. Due to these advantages, swarm robotics has been used as a solution to many different applications, such as cooperative transport (Mellinger *et al.*, 2010) and multi-robot manipulation (Wang and Schwager, 2015), coordinated construction (Stein *et al.*, 2011), area coverage and area exploration (Leonard *et al.*, 2010) as well as pollutant source localization (Soares *et al.*, 2016) and others.

On the other hand, *Human-Swarm Interfaces (HSI)* have gained increased popularity over the past decade. The aim of HSI is to incorporate the human element

in the control of large teams of robotic swarms. The advantage of such an endeavor is increased robustness to failures and extreme situations where human perception, cognition and intuition might provide correct solutions in a faster and more reliable manner. An example of an HSI system can be found in the work of McLurkin *et al.* (2006), whose platform was able to handle various types of human-swarm interactions which varied from charging and deployment to programming and control. Another example can be found in Naghsh *et al.* (2008), where the authors discuss the design and implementation of a system for interactions between a swarm of robots and firefighters.

The main objective of this work is to extend the current state-of-the-art on HSI by incorporating elements from the field of Brain-Computer Interfaces (BCI) as an alternative option for interactions between humans and swarms. The idea is that a BCI system for the control of a robotic swarm can provide a more intuitive solution for human operators. They will be able to translate their thoughts directly into commands for the robots without the need of complicated interaction patterns. At the same time, such a system can also provide a way of monitoring the users' engagement in the task as well as their fatigue, which can help prevent accidents and increase the efficiency of mission execution.

Chapter 2

BACKGROUND AND RELATED WORKS

The present work is a heavily multidisciplinary project as it borrows ideas and methodologies from a lot of fields, such as signal processing, neuroscience, machine learning, swarm robotics, Human-Swarm Interfaces (HSI), Brain-Computer Interfaces (BCI) and even Silent Speech Interfaces (SSI) and imagined speech. This chapter presents previous works that are related mainly with the fields of HSI, BCI and imagined speech.

2.1 Human-Swarm Interfaces (HSI)

As explained in Chapter 1, HSI try to incorporate the human factor into the control of robotic swarms in order to leverage the unique characteristics of human cognition and guarantee the success of the mission. This comes at a cost though because human operators have limitations, such as due to fatigue or loss of concentration, while, at the same time, they can make mistakes that can jeopardize the mission of the swarm. This delicate balance between the positives of human understanding and intuition and the drawbacks of human limitations has been extensively studied and many solutions have been proposed. Two of them can be found in the works of McLurkin *et al.* (2006) and Naghsh *et al.* (2008) (see also Chapter 1). In fact, McLurkin *et al.* (2006) designed and implemented an interface for the iRobot platform which included both software and hardware components. The main merit of this design was that robots were treated and maintained as a team rather than as individuals, which greatly reduced maintenance and programming times, as well as cognitive load for the operators of the swarm.

Other approaches use hand gestures (Nagi *et al.*, 2014; Podevijn *et al.*, 2013) to interact with the swarm. The idea is that hand gestures are a very important part of

human communication and thus they offer a very intuitive solution when people are trying to "communicate" with robots. Other interfaces might include face engagement (Couture-Beil *et al.*, 2010) and gaze detection (Monajjemi *et al.*, 2013), while others may rely on ElectroMyoGraphic (EMG) devices (Stoica *et al.*, 2013), haptic devices (Nunnally *et al.*, 2013) or joysticks (Pollini *et al.*, 2009). The aim of these interfaces is the control of multiple aspects of the robotic swarm at once, while still remaining easy-to-use. For example, Pollini *et al.* (2009) use the idea of geometric abstraction in order to describe the swarm in terms of its shape, its position and its orientation. The first element is represented by the length of the axes of an ellipsoid that encloses the swarm, while the last two simply corresponds to the $x-$, $y-$, $\psi-$ coordinates of its planar motion. The regulation of these quantities is done through a joystick, which leads to the control of totally 5 Degrees of Freedom (DoF) with one single device.

On the other end, there are works that mainly focus on the interaction strategy itself rather than on the complete interface. For example, Kolling *et al.* (2013) discuss two different types of human-swarm interaction, namely intermittent and environmental by using two control methods as examples. In their work, the two methods are referred to as *selection* and *beacon* control, respectively. The first type refers to an approach where the user selects a group of agents/robots and commands them to execute a specific task. In the second case, the user places a virtual beacon on an area of interest and the agents' behavior depends on their distance from it. In Bashyal and Venayagamoorthy (2008), the users interact with the swarm via an *avatar*, i.e. an agent within the swarm that they can control completely. The users can interact with the swarm on a local level and they can incorporate their own perception about the swarm's state without centralized control. The local characteristics of this method also result in its scalability which makes it really attractive for bigger groups of agents or even for cases where multiple users interact with the robotic team. In

Crandall *et al.* (2005), the authors examine the concept of *Neglect Tolerance* which refers to the time that a user can stop issuing commands ("neglect") to the swarm before its performance starts to degrade. On the other hand, Walker *et al.* (2012) discuss the strategy of allowing the swarm's state to reach stability before applying a new command, a methodology that is termed in their work as *Neglect Benevolence*.

In other cases, researchers chose to examine how the level of interaction between a human operator and a swarm affects the performance of the system (Coppin and Legras, 2012; Cummings, 2004). Coppin and Legras (2012), for example, defined autonomy profiles based on Sheridan's ten levels of autonomy and combined them with a four stage model of information processing. The aim of this work was to analyze each level of interaction/autonomy in order to engage the user only when necessary, thus optimizing the efficiency of the interaction process. Finally, there were works that examined human-swarm interaction and defined measures to assess its quality. As an example, Harriot *et al.* (2014) proposed metrics such as *trust in the system*, *situation awareness* and *vigilance* in order to evaluate the quality of the interaction from the perspective of the human operator.

2.2 Brain-Computer Interfaces (BCI)

Brain Computer Interfaces (BCI) have gained increased popularity during the past two decades because they can provide people with motor disabilities or locked-in patients an alternative route of communication with their surrounding that bypasses the musculoskeletal system and is using the Central Nervous System (CNS) instead. In other words, people that are paralyzed due to amyotrophic lateral sclerosis (ALS), spinal cord injuries (SCI) or maybe a stroke have a better chance of interacting with their environment in a meaningful and effective way just by using their brain signals.

In this work, principles from BCI can be used to extend the current state-of-art in HSI and provide an alternative type of interaction for human operators.

In general, there are many flavors of BCI systems and many different technologies that can be used in order to extract neurophysiological information from the brain. These technologies can involve functional Magnetic Resonance Imaging (fMRI), functional Near Infrared Spectroscopy (fNIRS), ElectroEncephaloGraphy (EEG), MagnetoEncephaloGraphy (MEG) or ElectroCorticoGraphy (ECoG) (Lebedev and Nicolelis, 2006; Naseer and Hong, 2015; Wolpaw *et al.*, 2002). From all those methods, EEG is the most commonly used since it is non-invasive, while at the same time it provides very good temporal resolution with respect to the recorded brain activity. In addition to the technology used, there is also a great plethora of brain functions that BCIs might exploit in order to detect human intention. These may refer to P300 Event Related Potentials (ERP) (Farwell and Donchin, 1988), Slow Cortical Potentials (SCP) (Birbaumer *et al.*, 2000), Steady-State Visual evoked Potentials (SSVEP) (Herrmann, 2001) or Event-Related Desynchronization/Synchronization (ERD/ERS) (Neuper and Pfurtscheller, 2001). The latter type of brain activity has been extensively used in the literature and it is related to mental tasks that involve Motor Imagery (MI), i.e. imagination of limb movement.

BCI platforms can be divided into two main categories based on the type of training that they employ. The first one involves *operational conditioning*, where the users learn to control specific aspects of their brain signals, such as the amplitude. In the system created by Birbaumer *et al.* (2000), i.e. the Thought Translation Device (TTD), the users were performing binary decision tasks by controlling the amplitude of slowly changing signals in their brain, which are also known as Slow Cortical Potentials (SCP). In another case, the users were required to move a cursor in a 2D (Wolpaw and McFarland, 2004) or a 3D environment (Wolpaw *et al.*, 2010) by

controlling the amplitude of specific frequency bands that were chosen independently for each user. The main merit of this approach is that the users can control the speed of the cursor as a continuous variable along multiple directions, which consequently leads to the control of many DOF. Unfortunately, this advantage comes at a cost, since the user requires extensive training for several days (even months at times) in order to achieve an acceptable level of control.

On the other hand, there are approaches which involve training predictive models that can recognize the intent of the user based on previously collected data. These type of systems usually output discrete decisions. Most of the time, these decisions represent the high-level goals of the robotic platform, while the low-level control is performed using the on-board sensors and actuators. For example, in Millan *et al.* (2004), the user was moving a mobile robot through a maze-like environment by executing specific mental tasks each of one corresponding to a distinct action (move forward, turn left, turn right, stop). The robot avoided collisions with the walls of the maze based solely on its on-board sensing and without explicit input from the user. In another work (Bell *et al.*, 2008), the users were instructing a humanoid robot to move objects of interest to predefined locations by focusing their attention on specific parts of a monitor that were flashing at different frequencies. The system then would recognize their choice based on SSVEP. Most of those approaches rely on *machine learning* techniques in order to differentiate between different states of the users' brain and correctly identify their intent. An extensive overview of such techniques for EEG-based BCI applications can be found in the work by Lotte *et al.* (2007). The advantage of these methodologies is that they adapt to the user's brain activity and signals by exploring and extracting robust features that are automatically detected for each subject without lengthy training procedures. This leads into systems that are adaptive to each user, more robust to noise and easier to use. The only drawback

is that there are restrictions to the amount of brain states that these approaches can discriminate between.

In order to counter the restrictions imposed by the previous approaches, many works have examined the use of *hybrid BCI* platforms, a review of which has been done by Pfurtscheller *et al.* (2010) and Amiri *et al.* (2013). In general, hybrid BCI systems can be divided into three main categories. The first one refers to approaches that combine simultaneously or in sequence two different types of EEG modalities. As an example, Li *et al.* (2010) combined the ERD/ERS phenomena that occur during limb movement imagination with P300 potentials in order to control a 2D cursor on a monitor. The second category of hybrid BCI concerns systems that combine EEG signals with other types of biosignals such as ElectroMyoGraphs (EMG) (Leeb *et al.*, 2011) or ElectroOculoGrams (EOG) (Ma *et al.*, 2015). The last category involves approaches that combine EEG signals with assistive technologies (AT) (Millan *et al.*, 2010), such as a wheelchair, or external devices, such as mice and keyboards. In all those cases, the goal is either to enhance the accuracy of the final brain state classification or provide a type of "brain switch" that can help users execute more complex tasks. (Karavas *et al.*, 2017).

2.3 Imagined Speech

As explained previously, BCI systems are a very attractive solution for cases where people are not capable of overt movements, such as in cases of ALS, stroke or spinal cord injuries. Most BCI solutions involve Motor Imagery, i.e. imagination of limb movement, that results in ERD/ERS patterns in the brain. These phenomena are well studied and understood (Neuper and Pfurtscheller, 2001; Pfurtscheller and Lopes da Silva, 1999) and have been proven to provide a robust method for controlling robotic applications, such as prostheses (Muller-Putz *et al.*, 2005). In fact, LaFleur *et al.*

(2013) were able to control the motion of a quadrotor using scalp EEG validating the results across multiple subjects (Royer *et al.*, 2010).

Although motor imagery has proven to be very useful for helping people with motor disabilities, it can be restrictive in cases where a human operator wants to use a joystick or any other hand-held device in combination with a BCI system. The reason is obvious since the operation of the joystick will activate the same regions as motor imagery, thus preventing a concrete distinction between these two actions. As a result, better solutions need to be found.

An easier approach could include the use of *visual imagery*, i.e. imagination of shapes or images (Contreras and Sundararajan, 2016; Esfahani and Sundararajan, 2012). Another solution could be the use of *speech imagery*, i.e. the imagination of pronouncing certain phonemes or words within one's mind without explicit vocalization of the sound. The latter option seems more attractive and more intuitive as speech is a process that is well developed in humans, while imagined speech, although less concrete, is still a task that many people perform during their life, especially when thinking.

BCI systems that rely on imagined speech for their operation belong to a larger set of systems called *Silent Speech Interfaces (SSI)*, a thorough review of which can be found in the work by Herff and Schultz (2016). Speech imagery-based systems can be divided into two main categories; those that involve the imagination of phonemes and sounds, such as vowels (e.g. /a/, /i/, /u/) and syllables (e.g. /ba/, /ku/) and those that involve the imagination of words, such as "up", "down" and "separate". In any of those cases, the user is imagining the pronunciation of the sounds or words without overtly saying them. In general, there is quite a plethora of works that relate to speech recognition and imagination. For example, Wester and Schultz (2006) examined the recognition of speech activity based on brain signals gathered by EEG sensors for five

different types of speech, namely overt, whispered, silent (i.e. moving only the lips), mumbled and imagined. The authors showed results about the feasibility of such a technology and they provided a good basis for further research on the subject.

In the work by D 'Zmura *et al.* (2009), the authors applied the Hilbert transform on appropriately filtered data in order to extract signal envelopes in the theta (3-8Hz), alpha (8-13Hz) and beta (13 - 18Hz) bands. These were later used to create matched filters that would classify the data into six different conditions, namely two syllables imagined repeatedly at three different rhythms each. Subsequently, Deng *et al.* (2010) classified between the syllables /ba/ and /ku/ imagined at different rhythms using the Hilbert-Huang Transform (HHT) (Huang and Shen, 2005; Huang *et al.*, 1998) and focusing on the frequency band of 3 -20Hz. In the work by Brigham and Kumar (2010), the authors also aimed to distinguish between two syllables, namely /ba/ and /ku/, but they only focused on one rhythm. Their method was applied on data provided by D 'Zmura *et al.* (2009) and their experiment involved 7 subjects. Their biggest challenge was the unsuitability of certain trials for training and testing due to noise. By applying the Hurst exponent as a method of automatic trial selection, they managed to significantly increase their classification scores, it affected though the strength of their findings. In the work by DaSalla *et al.* (2009a,b), the classification method was based on Common Spatial Patterns(CSP) and Support Vector Machines (SVM). The algorithm was able to successfully differentiate between three different brain states, namely imagination of the vowel /a/, the vowel /u/ and a resting state. Based on their CSP analysis, the most discriminant features originated from channels Fz, C3, Cz and C4, which suggests that their results were mostly attributed to motor cortex activation rather than pure speech imagery. It should be noted though that speech production also involves motor cortex activation which could be the reason of this outcome. On the other hand, Idrees and Farooq (2016) were able to clas-

sify between the same three tasks (/a/, /u/ and "rest") based on data provided by DaSalla *et al.* (2009a) by applying wavelet decomposition and extracting 11 features from the signals, namely mean, variance, skewness, kurtosis, geometric mean, harmonic mean, inter-quartile range, energy sum, entropy and, standard deviation and waveform length. The classification of the corresponding feature vectors was based on selecting specific features out of the extracted ones and applying Linear Discriminant Analysis (LDA) (Fisher, 1936; McLachlan, 2004). Kim *et al.* (2014) were able to classify between three different vowels focusing on the alpha band and utilizing spatial filters based on CSP as with previous works. Other features that have been proposed in the literature for speech imagery include Gabor Filters (Chi *et al.*, 2011) and Mel Frequency Cepstral Coefficients (MFCC) (Riaz *et al.*, 2015).

As mentioned before, another category involves classification of imagined words. An initial attempt was done by Suppes *et al.* (1997) in which the authors managed to classify 7 words during an auditory comprehension task using EEG signals. Some of the subjects were also performing speech imagery during the task. That work was later extended for sentences (Suppes *et al.*, 1998). Other works focused entirely on speech imagery. For example, Salama *et al.* (2014) were able to classify between imagination of the word "Yes" and the word "No" by employing different types of classifiers, such as Support Vector Machines (SVM), Discriminant Analysis (DA), Self-Organizing Map (SOM), Feed-Forward Back-Propagation (FFBP) or a combination of those. In the work performed by Wang *et al.* (2013), the authors were able to classify between two chinese characters (which corresponded to the words "left" and "one", respectively) and the rest state (thus, a total of 3 tasks) by applying CSP and SVM. The important distinction with respect to the other works was that they used two different montages, i.e. one that covered the whole brain and involved 30 channels, and another that included 15 channels which were covering Broca's and

Wernicke's area on the left hemisphere of the brain. Based on their results though, the choice of the montage did not create any significant effect on the classification accuracy. On the other hand, González-Castañeda *et al.* (2016) were able to distinguish between 5 different words, namely "up", "down", "left", "right" and "select", for 27 subjects. In their work, they applied two different methods to classify these words, namely *sonification* and *textification*. The first was based on the Discrete Wavelet Transform (DWT) and the Relative Wavelet Energy (RWE), while the second involved the combination of Bump modelling with the Bag-of-Words algorithm. They were able to achieve very high accuracies, reaching even 91.52% using the textification method. Finally, Mohanchandra and Saha (2016) were able to classify between the words "water", "help", "thanks", "food" and "stop" (also 5 words) by employing a one-against-all multiclass SVM combined with the Subset Selection Method (SSM). The latter was based on a set of Principal Representative Features (PRF) whose goal was to reduce the dimensionality of the data. Their accuracies ranged from 60% to 92% which lie way above the chance level for 5 words, i.e. 20%.

Chapter 3

NEURAL RESPONSE UNDER OBSERVATION OF SWARM COLLECTIVE BEHAVIORS

This chapter discusses whether changes in the collective behavior of a swarm can elicit brain responses to an individual that observes them. This was a preliminary investigation about the feasibility of a Brain-Swarm Interface (BSI) that would be based directly on the behaviors of the swarms rather than on other types of mental tasks, such as motor imagery, arithmetic operations and so on. In this analysis, the focus lay on the swarm's cohesion. This collective behavior is related to the direction of each agent in the swarm and how much it deviates from the average direction at which the swarm travels. In what follows, the experimental procedure and the signal preprocessing steps are described. The results of this study are also presented followed by final remarks and possible implications.

3.1 Experimental Procedure

In this experiment, two healthy subjects (S1-2, males, ages 26-35) performed a series of trials where they observed a simulated swarm of agents flying from one edge of a monitor to the other, as shown in 3.1. In each case, the swarm would begin at a specific cohesion level and it would either change to a different one or it would remain at the same level as the trial progressed. Some examples of trials are presented in Fig. 3.2 through 3.5. Each subject performed one session with multiple trials. Each trial lasted approximately 10-11 seconds. The sequence of the trials was randomized but it remained the same for both subjects. The experimental protocol was approved by the ASU IRB (Protocols: 1309009601, STUDY00001345) and both subjects signed an informed consent form prior to participating in the experiment.

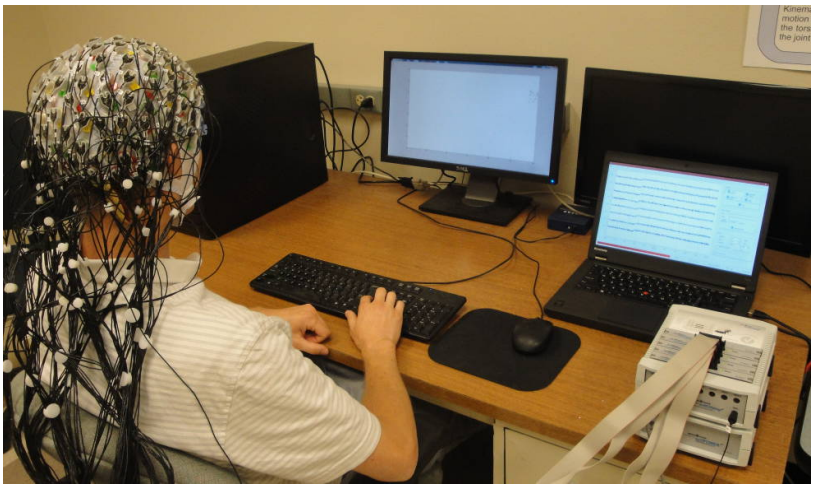

Figure 3.1: Subject in front of the monitor wearing the EEG cap and electrodes. (Karavas and Artemiadis (2015), ©2015 IEEE)

In the context of this investigation, the term *swarm cohesion* refers to how strictly parallel are the agents' paths with respect to each other or, in other words, how organized is the swarm while following a prescribed path. In more detail, a high level of cohesion denotes that the agents move almost on parallel paths with very small dispersion between each other, while a low level denotes that the paths can intersect with each other and produce a more random outspread of the swarm (Karavas and Artemiadis, 2015). The levels of cohesion that were used during the experiment were three, namely *high*, *medium* and *low*. The experimental procedure is described in more detail below.

In each trial, the swarm would begin its motion in a starting formation (not dependent on the cohesion level), where all the agents would be close to each other, and it would move on a straight line. At a random time between 20% and 40% of the screen width, the swarm's cohesion level would change gradually until 10% of the screen width later, where the cohesion would reach its final level and its change would stop. This gradual change lasted approximately 840 ms. After the cohesion level change, the swarm would continue to travel in a straight line as before. At

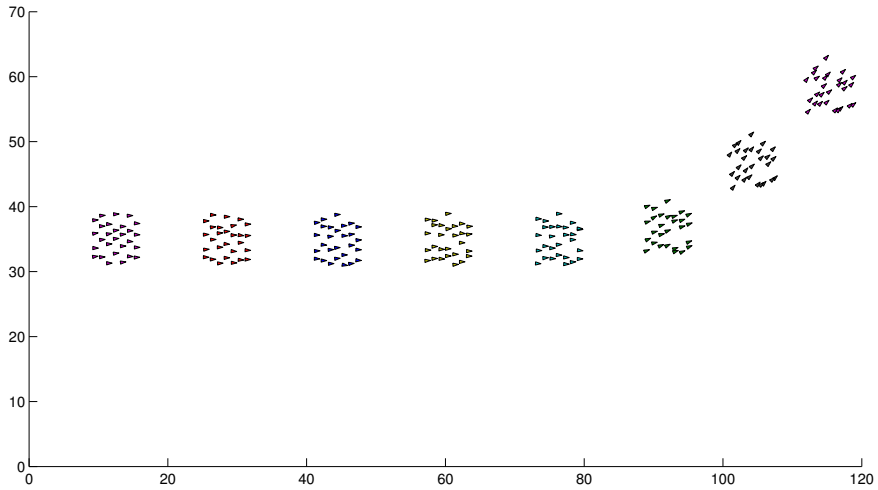

Figure 3.2: The swarm moves constantly at high cohesion level until it turns upwards. (Karavas and Artemiadis (2015), ©2015 IEEE)

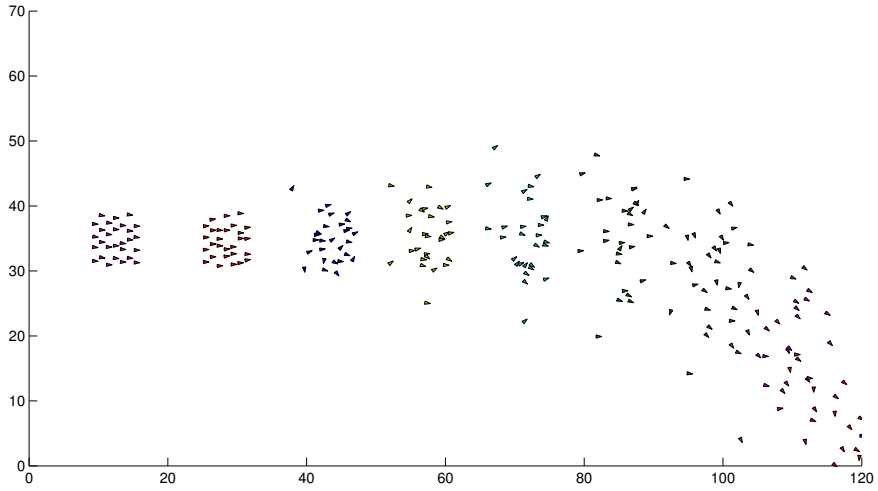

Figure 3.3: The swarm changes its cohesion level from high to low and moves downwards. (Karavas and Artemiadis (2015), ©2015 IEEE)

a random time between 60% and 80% of screen width, the swarm would gradually deviate from its path towards either the upper or the lower edge of the screen, as it is shown in Fig. 3.2. The subjects were instructed to press a key on the keyboard in front of them immediately when this direction change occurred. They were completely unaware of the cohesion change and its timing and no response was required by them on that. The key press action had a dual goal of a) keeping the subjects engaged in the experiment and b) providing a metric as to whether the swarm's cohesion level

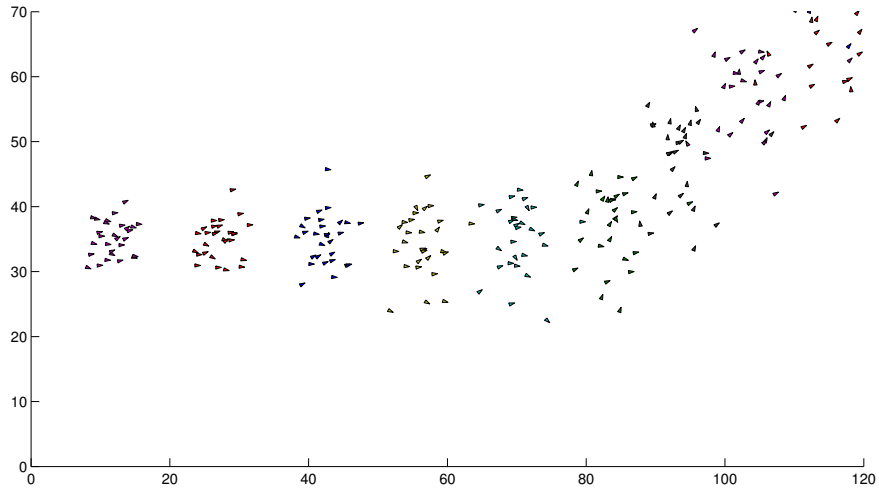

Figure 3.4: The swarm changes its cohesion level from medium to high and turns upwards. (Karavas and Artemiadis (2015), ©2015 IEEE)

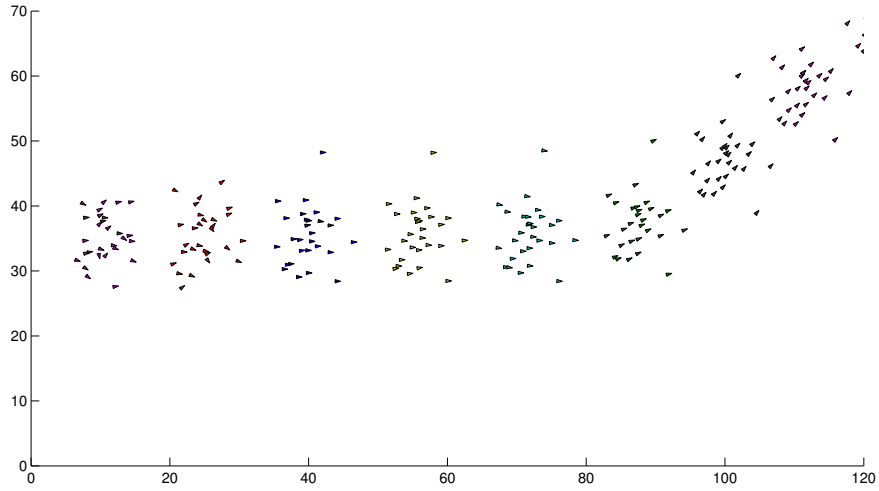

Figure 3.5: The swarm changes its cohesion level from high to low and turns upwards. (Karavas and Artemiadis (2015), ©2015 IEEE)

affected their response times with regard to the direction change. Since the main analysis of the data was going to be based on Event-Related Potentials (ERP), the times of the cohesion and the direction change, respectively, were chosen so as to avoid overlaps between the elicited ERPs from those two events/changes. Regarding the direction change on the swarm's path, this was chosen randomly across trials and it was completely independent from the cohesion level of the swarm or any cohesion change that could occur previously to that.

Table 3.1: Cohesion change cases. (adapted from Karavas and Artemiadis (2015), ©2015 IEEE)

Case #		1	2	3
Cohesion Levels	Start	high	medium	low
	End			

Case #		4	5	6
Cohesion Levels	Start	high	high	medium
	End	medium	low	high

Case #		7	8	9
Cohesion Levels	Start	medium	low	low
	End	low	high	medium

One of the main aspects of this investigation is the analysis of the ERP waveforms of the subjects at different levels of swarm cohesion and at cases where the level of cohesion changed during a single trial. Considering the three basic cohesion levels and all the possible types of transitions between them (i.e. high to medium, high to low, low to medium etc.) results in 9 cases which are listed for clarity in Table 3.1. For each of the 9 cases, the subjects performed 60 trials which resulted in a total of 540 trials.

3.2 Data Acquisition and Preprocessing

For the EEG data collection, a BrainProducts ActiCHamp amplifier system was used. The data were recorded at 1000Hz using 128 electrodes which were placed on the subjects according to the International 10/20 system (Klem *et al.*, 1999).

The preprocessing stage was performed using the EEGLAB (Delorme and Makeig, 2004) and ERPLAB (Lopez-Calderon and Luck, 2004) packages for Matlab. The data were bandpass filtered between 0.1 and 40 Hz using a low-pass and a high-pass Butterwoth filter of 6^{th} order in order to remove any high frequency noise (e.g. EMG noise) and low frequency trends. An additional notch filter at 60 Hz was then applied to guarantee the proper suppression of any remaining 60 Hz line noise. Next, the filtered data were re-referenced to a Common Average Reference (CAR). For each trial, an epoch starting at 0.5s before and ending at 1.5s after the onset of the cohesion change was extracted.

In order to remove artifacts due to eye blinks and eye movements, an Independent Component Analysis (ICA) was performed on the epoched data. For each independent component (IC), an equivalent dipole was fitted using the *dipfit* function from EEGLAB. Those ICs whose dipoles were located near the eyes or outside of the head region were treated as noise and were removed. The remaining components were remapped on the original electrode space which created the final clean data that were used for the subsequent analysis.

Finally, for each case of Table 3.1, the clean data from all corresponding trials were gathered and an average signal across them was computed. This signal represented the ERP for that particular case. The aim of the ERP computation was to examine whether there were any significant effects in the brain signals of the subjects during changes of the cohesion level in comparison to instances where no changes had occurred. This comparison is presented more analytically in the next section.

3.3 Results

The results of this analysis can be divided into two main parts. The first one involves the comparison of the users' reaction times in response to the swarm's di-

Table 3.2: p-values for reaction time pairs where each pair corresponds to two different cohesion levels. (adapted from Karavas and Artemiadis (2015), ©2015 IEEE)

	Subject 1	Subject 2
High vs. Medium	5.34×10^{-4}	8.19×10^{-12}
High vs. Low	1.86×10^{-7}	1.27×10^{-17}
Low vs. Medium	4.2×10^{-3}	2.54×10^{-2}

rection change across the three different levels of cohesion. This includes both cases where the level changed and those where it did not. The second part of the results refers to the comparison of ERP responses between cases where the cohesion level remained constant and those where it changed.

3.3.1 Reaction Times to Direction Change

The reaction times for both subjects at each level of swarm cohesion are shown in Figure 3.6. In this figure, it is shown that the cohesion level of the swarm affects greatly the reaction times of the subjects. In particular, the lower the cohesion level was, the longer it took the subjects to perceive any changes in the swarm's direction and appropriately react to it. In order to prove that this increase in response time was statistically significant, separate paired t-tests were performed for each possible pair of cohesion levels (i.e. high - medium, high - low, medium - low) and for each subject. The corresponding probabilities (p-values) are shown in Table 3.2. As it is shown, all p-values lie below the 5% confidence level, which validates the previous observation.

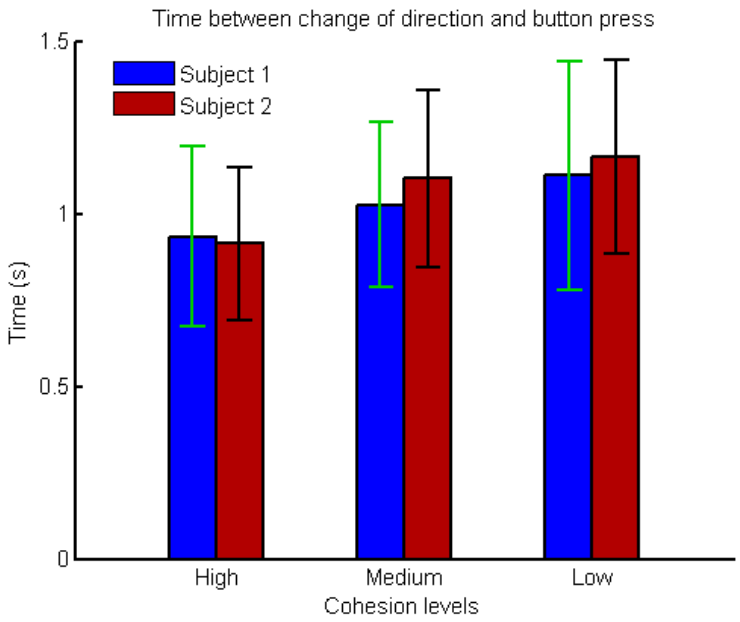

Figure 3.6: Supervisors' reaction times (mean and standard deviation) for different levels of swarm cohesion. (Karavas and Artemiadis (2015), ©2015 IEEE)

3.3.2 ERP Waveforms Comparison (Cohesion Change vs. Baseline)

For the case of the ERP comparison, the rationale was the following. As explained in Section 3.1 and shown in Table 3.1, in some of the cases studied, the cohesion level did not change within the trial. From those trials, an average ERP was extracted and was treated as a baseline for the particular cohesion level that it corresponded to. For the other cases, where the cohesion level changed, an ERP was also extracted and compared to the corresponding baseline ERP. For example, according to Table 3.1, case 1 corresponds to the baseline of the high cohesion level. The ERP that was computed for that case was then compared to cases 4 and 5 in which the cohesion level changes from high to medium and high to low, respectively. This means that for each cohesion level there are 2 comparisons (one for each of the remaining levels) which leads to a total of 6 comparisons. These are shown in Figures 3.7 to 3.12 for subject S1 and Figures 3.16 to 3.21 for subject S2.

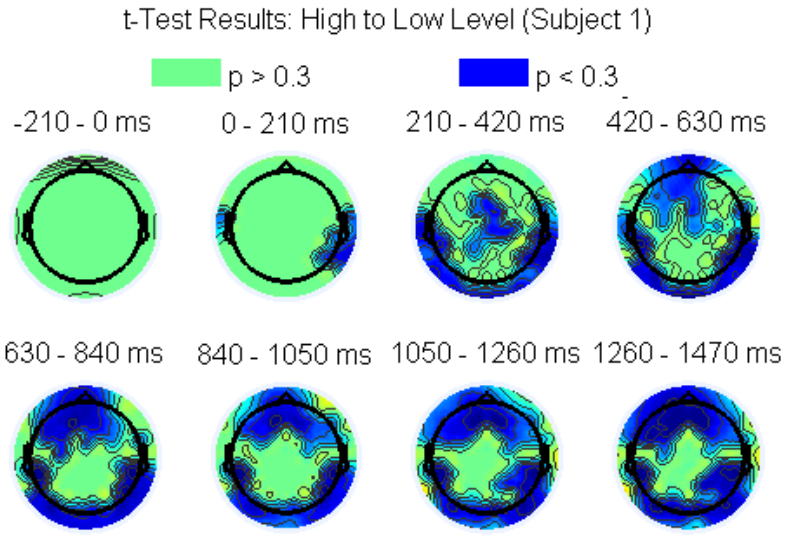

Figure 3.7: *t*-test values for a high-to-low cohesion change over a scalp map (subject S1). (adapted from Karavas and Artemiadis (2015), ©2015 IEEE)

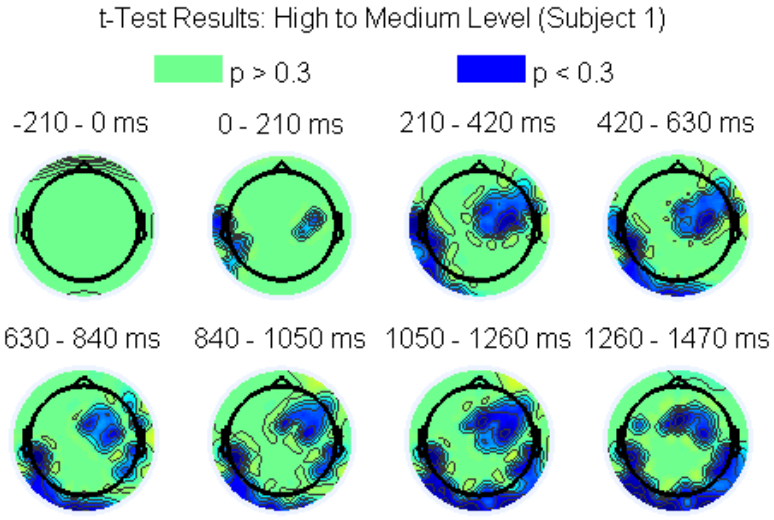

Figure 3.8: *t*-test values for a high-to-medium cohesion change over a scalp map (subject S1). (adapted from Karavas and Artemiadis (2015), ©2015 IEEE)

In more detail, the ERP comparison for each subject was performed as follows. For each cohesion level, for each channel and at each time instant, a paired t-test was performed between the baseline EEG signal values (one value per trial) for that level and the corresponding values for each change (e.g. high v.s. high to medium and high v.s. high to low). Then, an average of the p-values of the test across non-overlapping windows of 210ms was computed. The result was plotted as multiple scalp maps, as shown in Figure 3.7. For the sake of clarity, all channels that had p-values below 30% during a particular time window are shown in blue color. For the rest of the channels, the corresponding values on the map were set to 0 (zero) and are plotted with green color. The areas of map that did not contain any channels were colored based on an interpolation procedure among the p-values available. As already mentioned, the ERP analysis included 0.5s before stimulus (cohesion change) onset and 1.5s after. Correspondingly, the cohesion change started at $t = 0ms$ and, according to Section 3.1, it ended approximately 840ms later.

In Figures 3.7 and 3.8, the comparison is performed between the baseline of the high cohesion level and the cases where the level changes from high to low and from high to medium, respectively. The regions that showed the highest difference across the compared conditions were located at the visual association area and the primary visual cortex. This was expected since the type of the stimulus was visual. On the other hand, there was also significant activation over the prefrontal areas, premotor area and the primary motor cortex. One thing that is important to notice is that the difference between the compared conditions seems to continue to be significant even after the end of the stimulus ($t > 840ms$).

Similar observations about the areas of interest and the persistence of the statistical significance between conditions after the stimulus can be made by observing Figures 3.9 and 3.10, as well. In those, the scalp maps correspond to comparison of

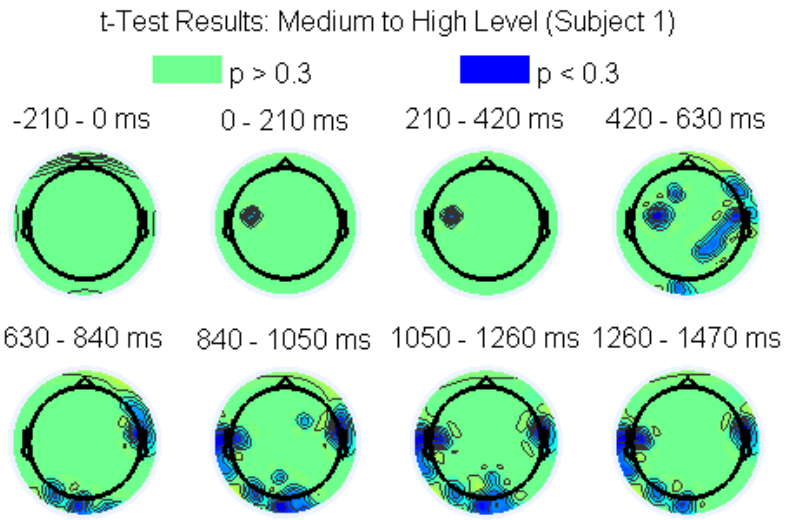

Figure 3.9: *t*-test values for a medium-to-high cohesion change over a scalp map (subject S1). (adapted from Karavas and Artemiadis (2015), ©2015 IEEE)

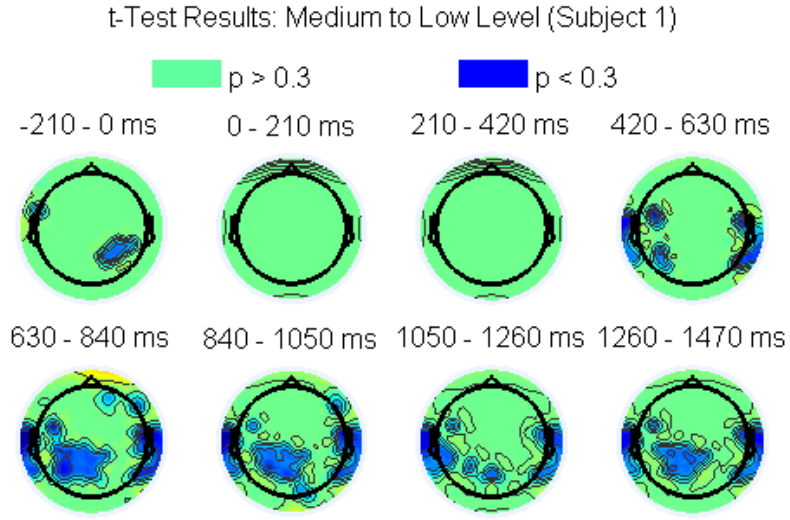

Figure 3.10: *t*-test values for a medium-to-low cohesion change over a scalp map (subject S1). (adapted from Karavas and Artemiadis (2015), ©2015 IEEE)

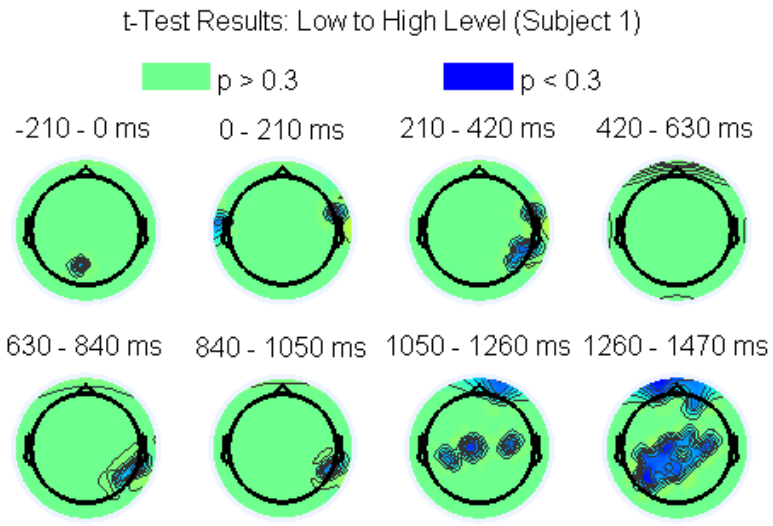

Figure 3.11: *t*-test values for a low-to-high cohesion change over a scalp map (subject S1). (adapted from Karavas and Artemiadis (2015), ©2015 IEEE)

the medium cohesion level with the cohesion changes from medium to high and from medium to low, respectively. Here, significance levels are lower, while there seems to be a delay between the stimulus onset and the time when the ERP for the medium-to-low change significantly deviates from the medium cohesion level baseline. Also, in that case only the visual association and spatial association areas are activated more, while there is no significant difference for the primary visual area. A delay of significance in the primary visual area can also be observed in 3.9.

The corresponding comparisons for the low cohesion level are shown in Figures 3.11 and 3.12. These are only presented for completeness since no statistical significance can be observed in those cases. Even though statistical differences can be observed from the baseline in the low-to-high case of cohesion change (Fig. 3.11), these appear 210ms after the end of the stimulus and might not be related to the stimulus directly. Finally, Figures 3.13 through 3.15 show the brain maps of the ERP components for each cohesion level and the corresponding cohesion changes for subject S1. In these

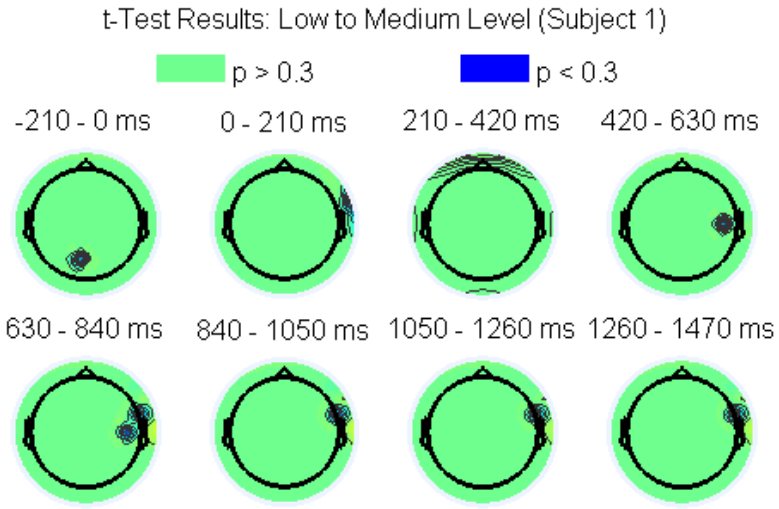

Figure 3.12: *t*-test values for a low-to-medium cohesion change over a scalp map (subject S1). (adapted from Karavas and Artemiadis (2015), ©2015 IEEE)

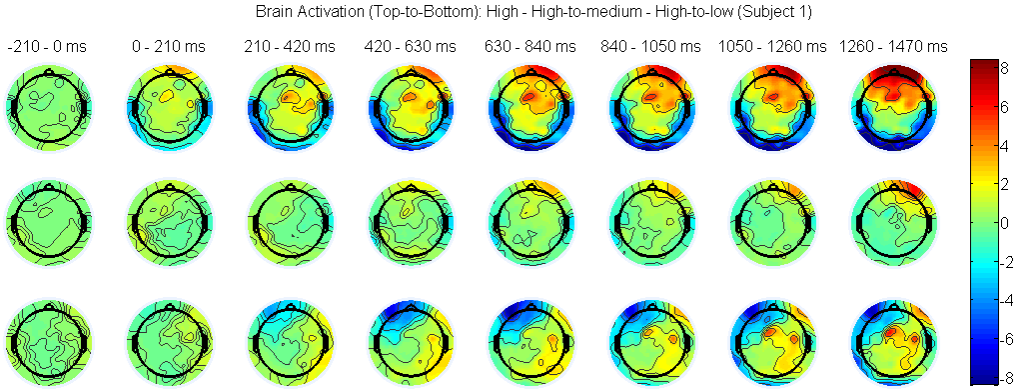

Figure 3.13: ERP activations for cases where the starting cohesion level is *high* (subject S1). The first row denotes no cohesion change, while the last two correspond to the high-to-medium and high-to-low changes, respectively. (Karavas and Artemiadis (2015), ©2015 IEEE)

figures, it can be seen that the activations are greater for bigger step in the change of the cohesion level.

For completeness, the corresponding comparison graphs for subject S2 are shown in Fig. 3.16 to 3.21. Unfortunately, these graphs cannot provide similar arguments for the relation between swarm cohesion change and brain activity, since the corresponding significance levels are much lower and the areas that appear to be more

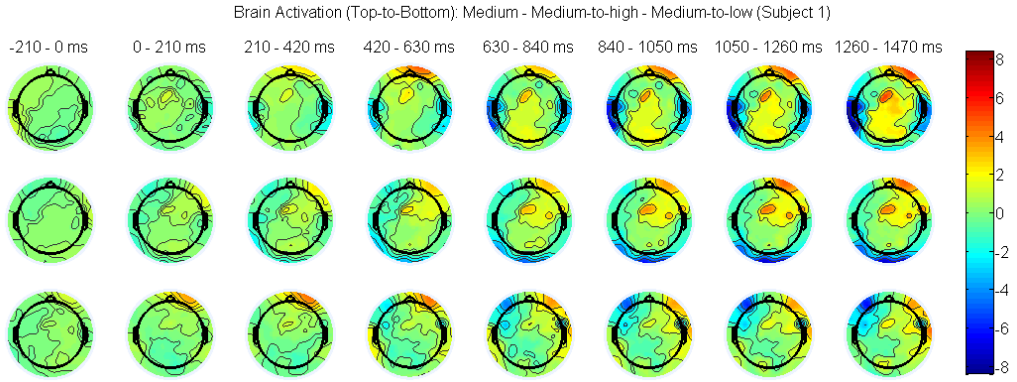

Figure 3.14: ERP activations for cases where the starting cohesion level is *medium* (subject S1). The first row denotes no cohesion change, while the last two correspond to the medium-to-high and medium-to-low changes, respectively. (Karavas and Artemiadis (2015), ©2015 IEEE)

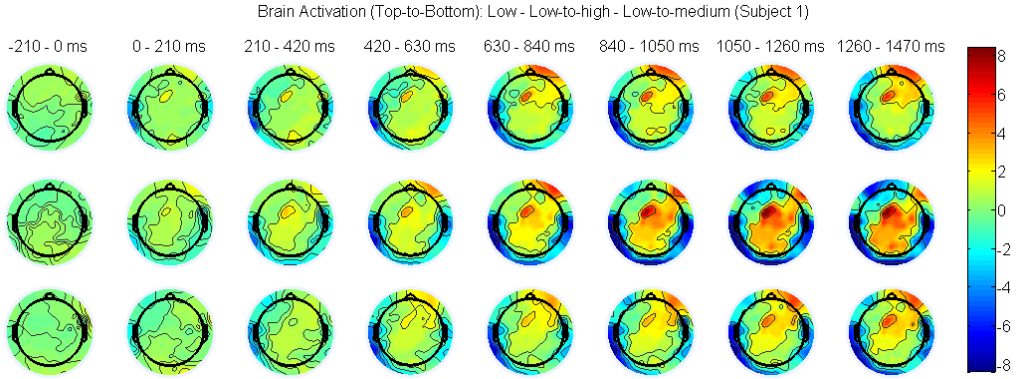

Figure 3.15: ERP activations for cases where the starting cohesion level is *low* (subject S1). The first row denotes no cohesion change, while the last two correspond to the low-to-high and low-to-medium changes, respectively. (Karavas and Artemiadis (2015), ©2015 IEEE)

influenced are not the same as the ones for subject S1. The only valid observation is that there are indeed some areas that are activated differently on this subject and, thus, any hypothesis relating brain activity and swarm cohesion changes can still be valid.

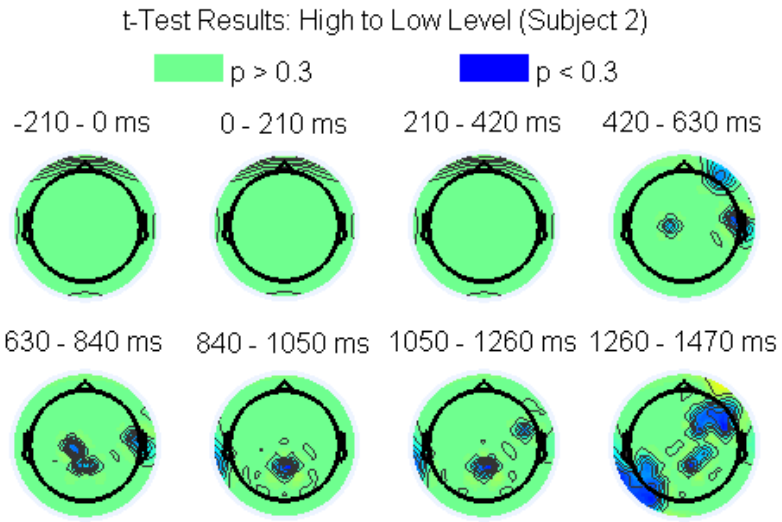

Figure 3.16: *t*-test values for a high-to-low cohesion change over a scalp map (subject S2). (adapted from Karavas and Artemiadis (2015), ©2015 IEEE)

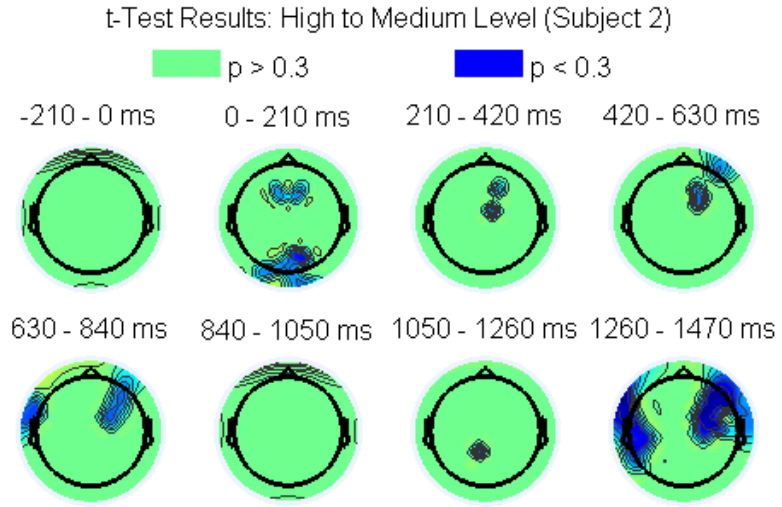

Figure 3.17: *t*-test values for a high-to-medium cohesion change over a scalp map (subject S2). (adapted from Karavas and Artemiadis (2015), ©2015 IEEE)

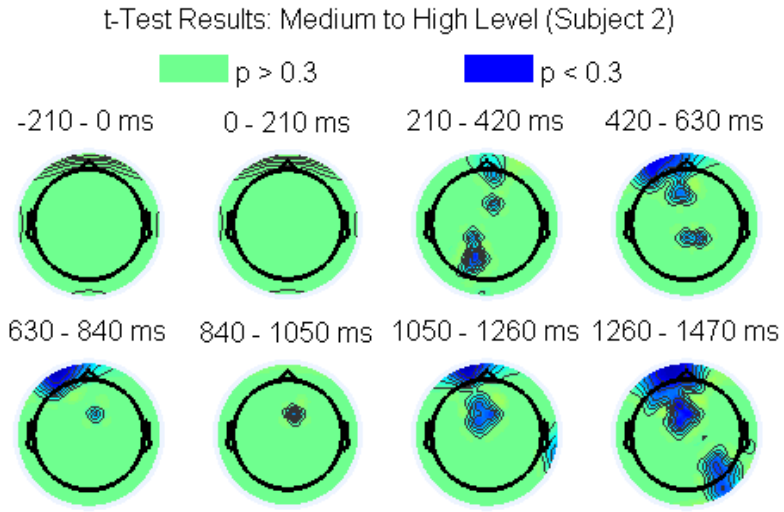

Figure 3.18: *t*-test values for a medium-to-high cohesion change over a scalp map (subject S2). (adapted from Karavas and Artemiadis (2015), ©2015 IEEE)

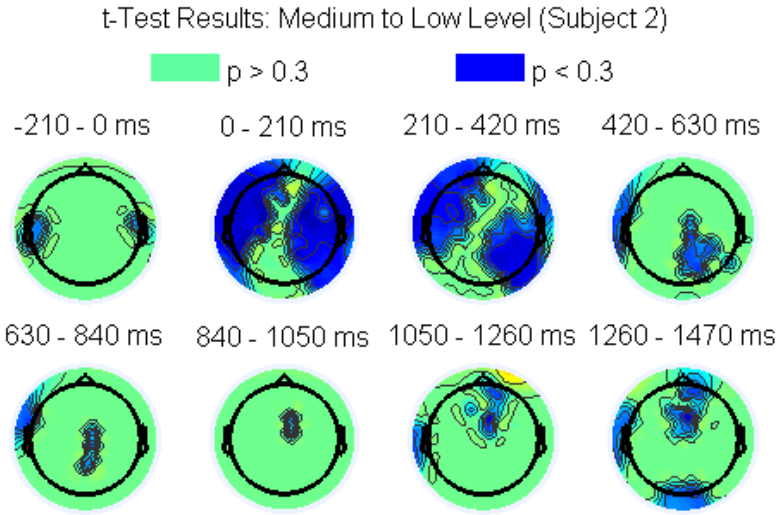

Figure 3.19: *t*-test values for a medium-to-low cohesion change over a scalp map (subject S2). (adapted from Karavas and Artemiadis (2015), ©2015 IEEE)

Figure 3.20: *t*-test values for a low-to-high cohesion change over a scalp map (subject S2). (adapted from Karavas and Artemiadis (2015), ©2015 IEEE)

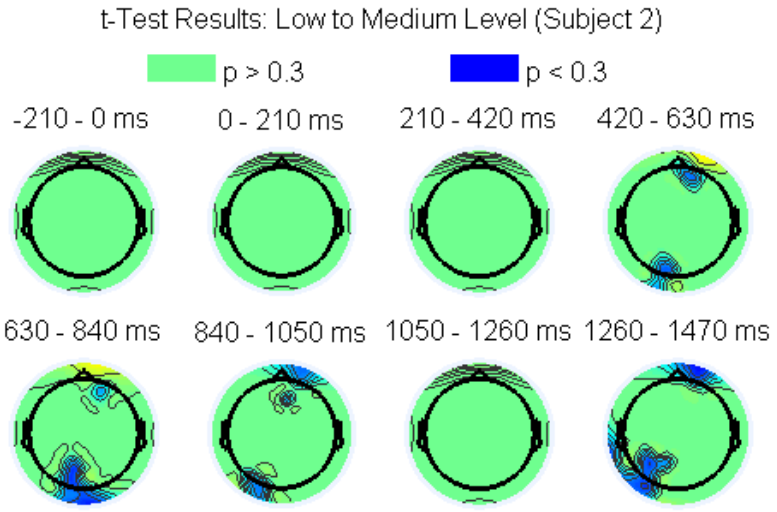

Figure 3.21: *t*-test values for a low-to-medium cohesion change over a scalp map (subject S2). (adapted from Karavas and Artemiadis (2015), ©2015 IEEE)

3.4 Conclusions

This chapter presented an experiment that aimed to investigate the connection between a swarm's collective behavior and the activity in a human brain. The collective behavior that was examined was the cohesion of the swarm, which was related to the level at which its agents are moving in parallel or not (see Section 3.1 for more details). The experimental procedure and the data analysis were also presented. The results suggested that a connection between swarm collective behavior and brain activity might be present, which is a big step forward in the realm of BCI applications for HSI as discussed in Chapter 2. In fact, the effect seems to be more distinct for larger degrees of cohesion change. Nevertheless, the number of subjects (i.e. two) and the confidence level used for the graphs in Section 3.3 (i.e. 30%) do not guarantee that the extracted conclusions are completely valid and, thus, more research is needed, both in terms of number of subjects and in terms of types of collective behavior. Despite all that, this investigation provides a good starting point for further evaluation of this problem.

Chapter 4

CONTROLLING A SWARM OF QUADROTORS USING A HYBRID BCI

This chapter presents a novel Brain Computer Interface (BCI) platform for the control of a swarm of quadrotors. The system belongs to the general category of hybrid BCI because it combines principles of motor imagery (MI) and ERD/ERS phenomena in the brain with the input from a game controller. The chapter analyzes every step of this process starting from the experimental procedure and the data preprocessing steps, continuing to the feature extraction procedure and the applied classification methods and finishing with describing how the output of the EEG system is translated into commands for the quadrotors. Some details about the robotic platform and its characteristics are also given. Finally, the chapter concludes by presenting the results of using the platform under two different scenarios and discussing advantages and drawbacks, as well as future work and possible improvements.

4.1 Experimental Procedure

This section presents the experimental procedure for the evaluation of the hybrid BCI system and its feasibility for the control of robotic swarms. For the testing of the system, two experimental protocols were employed. The first one involved a two-dimensional disk in a virtual environment that the subjects had to manipulate in order to complete a specific task. The second protocol involved a team of quadrotors whose positions and relative distance to each other were controlled based on commands coming from the proposed BCI system.

In the first experiment, 5 healthy subjects (S1-S5, males, all right-handed, ages between 22 and 27) were controlling the position and the size of a two-dimensional circular disk in a virtual environment that was presented to them on a monitor. Fig. 4.1 shows the experimental setup. The experimental protocol was approved by the

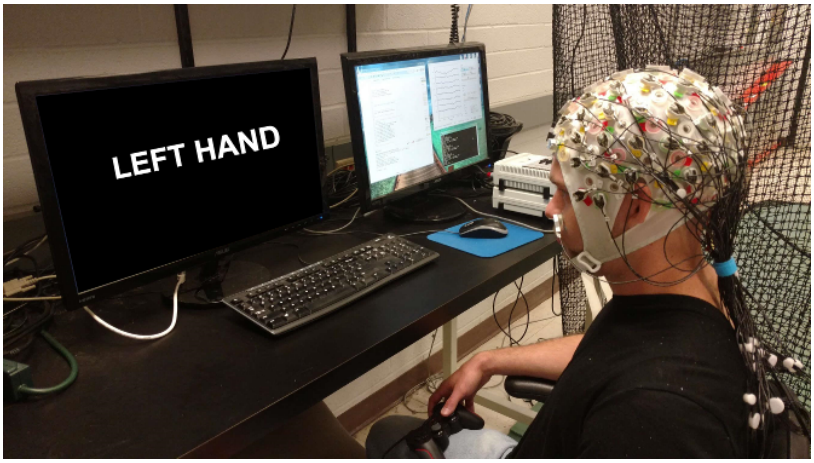

Figure 4.1: The subject is wearing the cap with EEG electrodes and looks at a monitor a few inches away. The monitor shows the task that the subject must execute. In this case the subject must imagine moving his left hand. (Karavas *et al.* (2017), ©2017 IEEE)

ASU IRB (Protocols: 1309009601, STUDY00001345) and was comprised of three phases, namely a data collection phase, a model training phase and a control phase, which are described in more detail below.

During the data collection phase, the subjects were asked to relax, stay still and refrain from any overt movement. The subjects were staring at a monitor in front of them which provided specific instructions depending on the mental task that they needed to perform. They were also given a game controller that they were instructed to hold with their right hand. The instructions were provided to the users through text messages on the screen and they comprised the phrases "Right Hand", "Left Hand" or "Rest". In the first case, the subjects were instructed to make random movements with the right joystick of the controller using their right hand. In the second case, the subjects needed to imagine closing their left hand to a fist. In order to ensure the proper activation of the corresponding brain areas during that task, the subjects were further instructed to imagine the feeling of their fingers touching their palm, i.e. the kinesthetic feeling of clenching their fist. In the third case, the subjects needed to relax their muscles completely and focus on their breathing or the

tip of their nose in order to get their brain into an idling / resting condition. In the literature, this can be referred to as *volitional rest* (LaFleur *et al.*, 2013; Royer *et al.*, 2010).

Each trial had three stages as shown in Fig. 4.2. In the first stage (preparation), the desired task was shown as a text on the monitor screen in red color to denote the start of the trial. After a period of 3s, the main trial would begin (second stage). The text would remain visible on the screen but it would change to a white color to denote that the main trial has started. The subjects had to start performing the desired task as soon as the message changed its color. The main trial would last for 4s and the subjects were instructed to repeat the task shown on the monitor repeatedly during that period of time. At the end of the trial, the message would disappear, a pause of 3s would follow and the next trial would start. Each of the subjects completed 20 trials per task, i.e. 60 trials in total in a single session. The trials were randomized across the different mental tasks. In order to ensure the good quality of the collected data, the subjects were further instructed to avoid any eye blinks during the main trial. The preparation stage and the pause at the end though provided ample time to the subjects to blink, swallow and perform minor movements that would ease the discomfort of the experiment.

During the model training phase, an algorithm that detected unique frequency bands related to the motor command tasks was applied to the previously collected data. More specifically, the algorithm was detecting those frequency bands for which the ERD/ERS phenomena corresponding to limb movement imagination or actual limb movement were more distinct for the subject of interest. In the next step, the same data were used to train the classifiers that would make the distinction between the three different tasks examined here (see Section 4.3 for details).

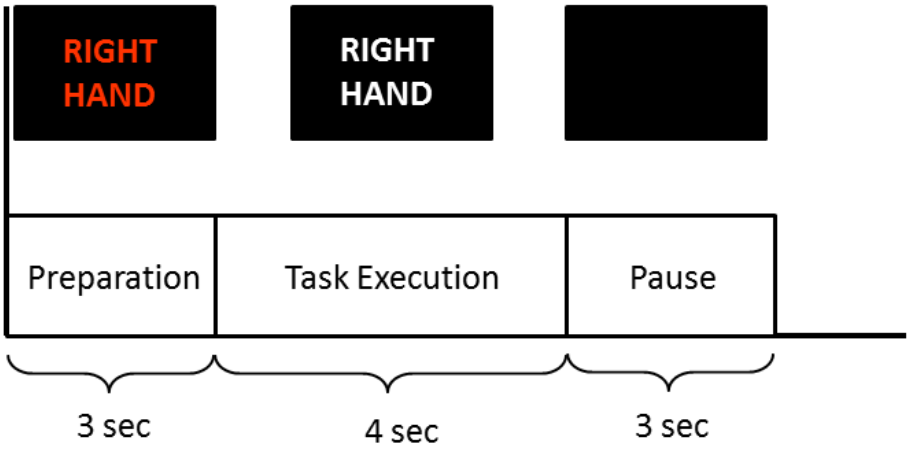

Figure 4.2: Each trial is comprised of a preparation period, the main trial where the subject performs the task at hand (here moving the joystick with the right hand) and a pause period. (Karavas *et al.* (2017), ©2017 IEEE)

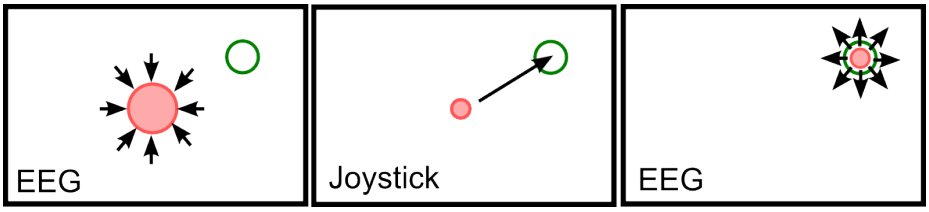

Figure 4.3: During the control phase, the user has to shrink the solid circle, move it inside the hollow target circle and then expand it again to match the size of the target circle. (Karavas *et al.* (2017), ©2017 IEEE)

During the control phase of the experiment, the subjects needed to achieve a simple goal that required the combination of the three tasks previously trained by the system. They had to move a circular disk inside a hollow circle and change the size of the disk so that it matches the size of the circle, as is shown in Fig. 4.3. In order to achieve that goal, the subjects had to either reduce the size of the disk by imagining clenching their left hand to a fist ("Left Hand" task), or increase the size of the disk by performing volitional rest ("Rest" task) or change the position of the disk by moving the joystick on the controller ("Right Hand" task). The size of the target circle was always smaller than the initial size of the disk in order to force the users to reduce its size first before attempting to fit it into the circle. At the same time, to make things easier, once the disk was inside the circle, it could not get out and the

subjects just needed to change its size without worrying about its position any more, since it was filling the space of the circle appropriately as it was expanding. The trial was deemed successful as soon as the disk fitted perfectly inside the circle. If the user was not able to achieve the goal in 60 seconds, the trial was deemed unsuccessful and a new trial began. Between trials, there was a pause of 5s.

Each subject completed two sessions of 40 trials each (80 trials in total), with a 10-20 minutes break in between. In a session, the 40 trials were randomized with respect to both the position and the size of the target circle, as well as the initial size of the circular disk. On the contrary, in all cases the position of the disk was always initialized at the center of the screen. The sequence of trials of the first session was kept the same for the second session as well, in order to compare the performance of the subjects between sessions as they were getting more acquainted to the system and the task. At the same time, this repetition provided to the user more time for training and learning to control their brain activity.

In the second experiment, one of the subjects had to move a team of quadrotors through a rectangular hoop. The phases of the experiment were kept exactly the same, i.e. there was still a data collection phase, a model training phase and a control phase. In fact, the goal of the control phase was very similar to the one in the previous experiment. The quadrotors started initially on a line formation and at a certain distance to each other. The subject had to reduce their distance by performing a "Left Hand" task, move the quadrotors through the hoop by using the joystick ("Right Hand" task) and return the quadrotors to their initial spacing by performing a "Rest" task. Thus, the position of the swarm was controlled using the joystick, while the distance between the robots was controlled using brain signals. The subject had to complete the task once, which served as a proof of concept regarding the feasibility of BCI for swarm-related tasks.

4.2 Data Acquisition and Preprocessing

The EEG signals were acquired at a sampling rate of 500Hz using a BrainProducts ActiCHamp amplifier from 64 electrodes placed according to the 10/20 International system (Klem *et al.*, 1999). Before performing any further analysis, the following preprocessing steps took place. The data were passed through a 5th order Butterworth bandpass filter with low and high cutoff frequencies of 1 and 40Hz, respectively, in order to remove low frequency trends, possible EMG-related artifacts and line noise. In order to take into account any volume conduction effects, which are frequent in scalp EEG (Holsheimer and Feenstra, 1977), a large Laplacian filter was applied to each channel of interest. In particular, for each channel the mean of its next nearest neighbors was subtracted from the original signal. The large Laplacian filter was chosen because, according to McFarland *et al.* (1997), it performs significantly better than other spatial filter types, such as small Laplacian and Common Average Reference (CAR). At the same time, its computational cost is low, which is crucial for real-time application such as this one.

In this investigation, the focus of the analysis was placed on the sensorimotor cortex and specifically on the channels C3, Cz, C4, FC3, CP3, C1, FCz, CPz, C2, FC4, and CP4, as shown in Fig. 4.4. This happened in order to further reduce the effects of EMG artifacts due to eye or jaw movements, which influence mainly the frontal and temporal electrodes. At the same time, the choice of those channels made more sense considering that the underlying mechanisms used for the function of the system were based on the activation of the motor area.

As explained, in both experiments there was a data collection phase which required off-line processing and a control phase where the processing of the signals was done on-line. In both cases, the preprocessing steps were mostly the same. The only

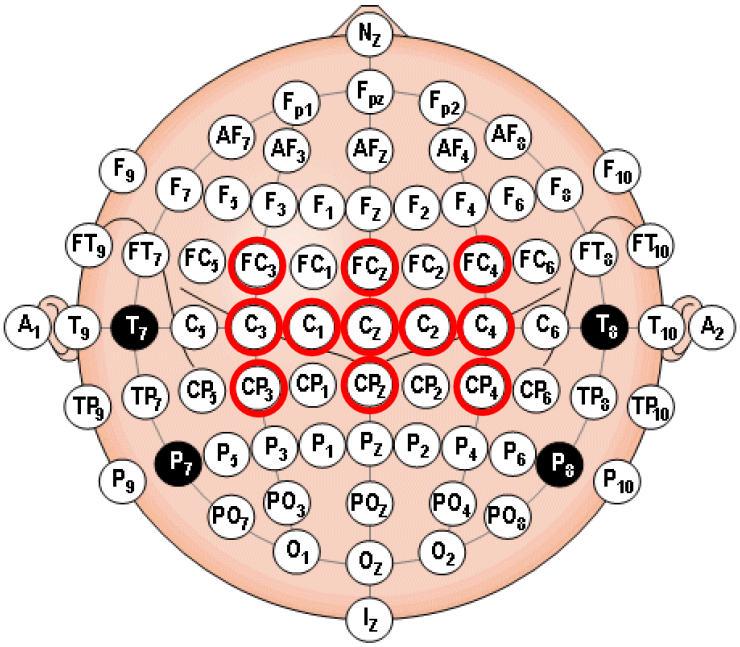

Figure 4.4: The channels chosen for the analysis of the EEG signals are shown here with red circles. These are channels C3, Cz, C4, FC3, CP3, C1, FCz, CPz, C2, FC4, and CP4 according to the 10/20 International System (adapted from Klem *et al.* (1999)).

difference was that in the control phase (on-line) an additional ElectroOculoGram (EOG) artifact removal algorithm (He *et al.*, 2004a, 2007) was applied before the large Laplacian referencing in order to completely eliminate any artifacts from eye blinks and eye movements.

4.3 Feature Extraction

This section discusses the feature extraction procedure that preceded the classification step. First, a Fast Fourier Transform (FFT) was applied to the data in order to extract frequency domain information. The FFT was applied on a window of 256 datapoints that was sliding every 25 datapoints (50ms). Before the application of the FFT, a Hanning window was applied to the data in order to reduce any aliasing effects due to the initial windowing procedure.

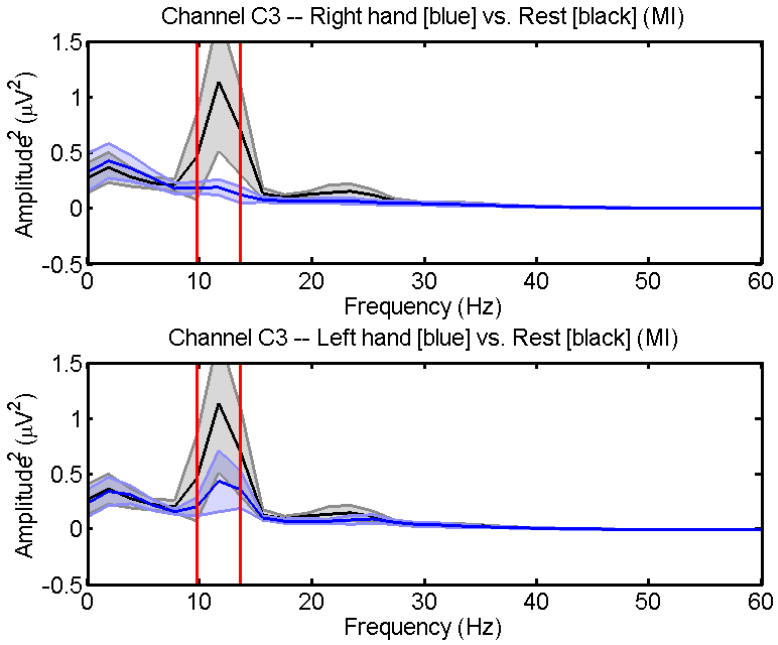

Figure 4.5: Frequency band detection for Subject S4 and channel C3. The red vertical lines denote the band of interest where the frequency components for right (top) or left (bottom) hand are different from the ones in "Rest" condition. (Karavas *et al.* (2017), ©2017 IEEE)

At the next step, an algorithm selected automatically a frequency band of interest for each channel (see example in Fig. 4.5). These frequencies were different for each subject. The aim of the algorithm was to find for each channel the frequency band that was activated the most. The term "activation" in this context is explained as follows. During the experiments, the subjects performed either actual limb movement (by using the joystick) or limb movement imagination ("Left Hand" task). In both cases, ERD/ERS phenomena were occurring inside the brain of the user. These phenomena were mostly concentrated in the alpha (α) and beta (β) bands, namely between the frequencies of 7 and 30 Hz. On the other hand, the "Rest" task was creating an idling mode inside the brain, which was expressed as higher amplitude of the alpha band. Thus, the algorithm was looking for reduction of the FFT spectrum amplitude with respect to the "Rest" state in the alpha (α) band (corresponding to an ERD) and/or an increase of the spectrum in the beta (β) band (corresponding to

an ERS). This happened in three steps. First, the average of all FFT coefficients was computed across a whole trial ($T = 4s \times 500Hz = 2000datapoints$). Next, the grand average of the FFT spectrum was computed across all trials for each separate task (20 trials per task). Finally, the FFT spectrum for each of the "Right Hand" and "Left Hand" tasks was compared to the one for the "Rest" task and, for each channel, the FFT coefficient with the largest difference was chosen to represent the activated band. The final FFT coefficients that corresponded to the frequency band of interest included the chosen coefficient and the coefficients immediately before and after that. This is also shown in the example of Fig. 4.5. These coefficients were later summed together for each channel of interest to form the final FFT vector $F \in \Re^{m \times 1}$, where $m = 11$ is the number of channels used in the analysis.

The final features were extracted by applying Principal Component Analysis (PCA) on the FFT features that were chosen based on the previous approach. In particular, all FFT features for all tasks ("Right Hand", "Left Hand", "Rest") across all trials of the data collection phase were gathered together in one structure and PCA was applied to that. The Principal Components (PCs) that would describe 90% or more of the data variance were selected as the final features. This resulted in 4 to 5 components depending on the subject. Finally, these features were collected in data sequences that would be later used for the training of the machine learning models.

4.4 Classification Procedure

The classification of the EEG signals was based on the method of Hidden Markov Models (HMM). An HMM provides a methodology to describe stochastic processes. In particular, an HMM uses probability distributions over observation sequences which are emitted from an underlying (*hidden*) process in order to predict its future outputs (Karavas *et al.*, 2017). As a method, it was originally developed by Rabiner (1989)

for applications that involved speech recognition. Nonetheless, it has also been used for the classification of EEG signals (Argunsah and Cetin, 2010; Lee and Choi, 2003; Obermaier *et al.*, 2001) in order to deal with their non-stationarity and the involved periodicities.

An HMM comprises a state transition matrix A, an observations probability model B and an initial state probability π (Rabiner, 1989). The sequences of observations that were used in order to classify between the different brain states of the user contained continuous variables (see Section 4.3 for more information about the features). This led to the natural choice of modeling the probability distribution of the observations B as multiple Gaussian mixtures (McLachlan, 2000). In addition, for each of the tasks involved ("Right Hand", "Left Hand" and "Rest"), a separate HMM was trained, thus using the method in its *generative* form. It was assumed that the underlying processes for the tasks were fairly similar to each other and so the number of members in the Gaussian mixtures and the number of hidden states were chosen to be the same for all of the models.

The training of the classifier was performed as follows. The feature vector sequences of each task were fed into the corresponding HMM. The sequences had a length of 20 datapoints (which corresponded to data collected during 1 second) and were created by sliding a 20-point window on the series of feature vectors created in the feature extraction step point by point. The Baum-Welch algorithm (Huang *et al.*, 1990; Rabiner and Juang, 1993) was used in order to estimate the parameters of each HMM. In addition, because of variability among subjects, the number of members in the Gaussian mixtures and the number of hidden states of the HMM were different for each subject and each was chosen as follows.

At the beginning, the entire dataset was divided into a training, a validation and a testing set. For each different pair of number of states and number of members in

the Gaussian mixtures, a separate series of models was trained using the training set. For the training, only 5 iterations of the Baum-Welch algorithm were run in order to keep the processing time low. In order to choose the best pair, two criteria were used. The first one involved the overall accuracy of the classifier on the validation data. The second criterion involved the relative differences between the accuracies of the three trained HMMs, as those were evaluated on the validation set. This criterion ensured that the resulting classifier would provide good performance across all types of tasks, i.e. it would not provide high accuracy for one task but significantly lower for another. A pair would receive a high score if it had high accuracy or low relative differences, respectively. Each pair was ranked based on these criteria separately and the final ranking was based on the summation of the corresponding scores. At the end, the pair with the highest score was chosen as the final candidate and the HMMs were retrained using 10 iterations. The number of iterations in both cases of training (initial and final) was chosen empirically based on the convergence rate of the algorithm in previous experiments. A final testing of the trained models was run on the testing set.

Finally, during the control phase of the experiment, the classification between the tasks was performed in the following manner. A sequence of feature vectors was fed into each HMM, a log-likelihood value was computed for each of them using the Forward algorithm (Rabiner, 1989) and the data were classified according to the maximum of these likelihood values.

4.5 EEG System Output Generation

The final output of the EEG system comes out of the combination of three elements. The first one is related to the amplitude of the frequency spectrum of the EEG signals at a specified band, the second is related to the classification decision of the

HMM classifier with respect to the brain state of the user, while the third element is the input from the game controller. The result is a command vector whose elements correspond to the control of specific DOF of the robotic system. In order to get that command vector, the following procedure was followed.

At each iteration k, the raw activation v_k was calculated as:

$$v_k = (\bar{F}_{CP3,Rest} - F_{CP3,k}) + (\bar{F}_{CP4,Rest} - F_{CP4,k}) \tag{4.1}$$

The variables $F_{CP3,k}$, $F_{CP4,k}$ represent the summed spectral power of the frequency band of interest, i.e. the summed FFT coefficients before the application of PCA (chosen according to the methodology described in Section 4.3) at channels CP3 and CP4, while $\bar{F}_{CP3,Rest}$, $\bar{F}_{CP4,Rest}$ represent the mean of the summed spectral power over the same channels during the "Rest" task as recorded during the data collection phase. The value v_k was then passed through an exponential filter for smoothing purposes in order to make easier its control by the user. The exponential filter can be written as:

$$\tilde{v}_k = (1 - \alpha)v_k + \alpha\tilde{v}_{k-1} \tag{4.2}$$

where α is a smoothing constant ($\alpha = 0.9418$). The value of α was chosen during preliminary tests based on reports from the subjects in order to achieve a smooth change on the size of the solid disk in the virtual interface. Finally, a thresholding procedure was applied to $\tilde{v}_k$ in order to filter out any misclassifications and ensure the smooth operation of the system during the control phase. For this procedure, two different thresholds were used at the same time, a *high* t_H and a *low* t_L, and they were computed separately for each subject according to the following equations:

$$t_H = p * [E(f_{CP3,Rest}) + E(f_{CP4,Rest}) - min(f_{CP3,LH}) - min(f_{CP4,LH})] \tag{4.3}$$

$$t_L = \sqrt{Var(f_{CP3,Rest}) + Var(f_{CP4,Rest})} \tag{4.4}$$

where p is a weighting factor ($p = 0.85$). The value of p was the same for all subjects and it was chosen during preliminary tests in order to ensure that the thresholding procedure would successfully negate any misclassification commands. The statistics $E()$, $Var()$ and $min()$ refer to the expected value (mean), the variance and the minimum values of the data, respectively. The subscripts LH and $Rest$ refer to the "Left Hand" and "Rest" tasks, respectively. The thresholds were computed based on the data from the initial data collection phase.

At each iteration k, the final activation variable Δq_k was then computed based on the relation of $\tilde{v}_k$ to t_L and t_H and the classification decision D_k provided by the HMM classifier. The decision D_k could take three possible values, namely $D_k = 0$ for the "Right Hand" task, $D_k = 1$ for the "Left Hand" task and $D_k = 2$ for the "Rest" task. According to that:

$$\Delta q_k = \begin{cases} 0 & \textbf{if } t_L < \tilde{v}_k < t_H \textbf{ OR } D_k = 0 \\ -(\tilde{v}_k - t_H) & \textbf{if } \tilde{v}_k > t_H \textbf{ AND } D_k = 1 \\ -(\tilde{v}_k - t_L) & \textbf{if } \tilde{v}_k < t_L \textbf{ AND } D_k = 2 \end{cases} \tag{4.5}$$

At this point, it should be mentioned that the "Rest" state, as defined here, does not correspond to a brain state where the subject does not do anything but rather to the state of volitional rest (LaFleur *et al.*, 2013; Royer *et al.*, 2010), which corresponds to the intentional focus on some other part of the body (e.g. on the nose) or on the breathing rather than on the hands as in the other tasks. Thus, although it seems intuitive to assign $\Delta q_k = 0$ to the "Rest" state ($D_k = 2$), the latter is treated as a separate mental task (not as "switching-off" of the brain activity) and the current choice can still make sense for the user.

The final command vector $u_k = [\Delta q_k, \Delta x_k, \Delta y_k]^T$ consisted of the activation variable Δq_k and the joystick input $c_k = [\Delta x_k, \Delta y_k]$, which were all varied in a continuous

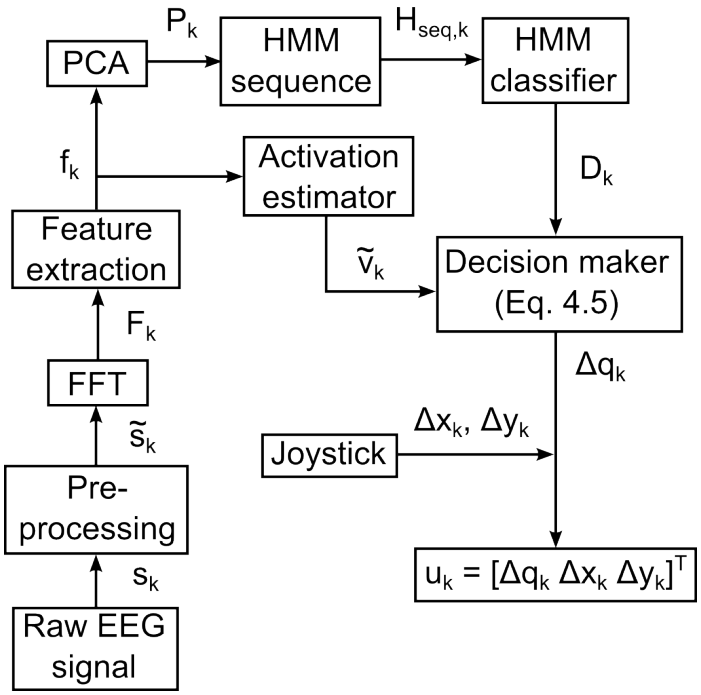

Figure 4.6: The diagram describes the entire procedure for computing the system output that is used in the control phase. Here, s_k is the raw EEG signal, $\tilde{s}_k$ is the clean/preprocessed signal, F_k represents the FFT coefficients, while f_k refers to the spectral features extracted from F_k. Also, P_k refers to the PCA features and $H_{seq,k}$ represents the HMM sequence. Finally, $\tilde{v}_k$ is the activation parameter, D_k is the classification decision, Δq_k refers to the EEG command, Δx_k, Δy_k refer to the joystick input, while the subscript k denotes the current iteration. (Karavas *et al.* (2017), ©2017 IEEE)

manner. For each of the experiments described in Section 4.1, each element of u_k was responsible for changing one DOF of the system. More specifically, in the first experiment, Δq_k controlled the size of the solid disk, while Δx_k and Δy_k controlled its position on the screen. In the second experiment, Δq_k controlled the distance between the agents and Δx_k and Δy_k controlled the swarm position along an axis normal to the line formation on the horizontal plane and along its height, respectively. As previously explained (Section 4.1), the users would imagine clenching their left hand to a fist to reduce the size of the solid disk or the distance between the quadrotors and would perform volitional rest to apply the opposite actions. When the users were using the joystick ($D_k = 0$), the size of the solid disk or the distance between the

agents were not changing ($\Delta q_k = 0$). A diagram of the entire procedure is also given for clarity in Fig. 4.6.

4.6 EEG Output to Quadrotor Control

This section describes the procedure under which the final output of the hybrid BCI system was used in order to control the team of quadrotors in the second experiment (Section 4.1). As mentioned, the quadrotors were initially placed in a line formation and at specific distance to each other. They were custom built based on an FPV 250mm frame and featured the APM 2.6+ autopilot for lower level autonomous control. The onboard autopilot received command signals from a ground station that was translating the output of the BCI system into desired roll, pitch and yaw angles. The methodology with which this was happening is described below.

The first step of this procedure involved the translation of the BCI system output into desired quadrotor positions. This was based on a planning scheme that took into account the absolute position of the agents with respect to a global frame, the desired change in their distance Δq_k and the inputs Δx_k, Δy_k from the joystick and provided the new reference positions for the quadrotors. A 4-camera optical motion capture system (Bonita, Vicon Inc) was used for the tracking of the vehicles. The second step involved taking these desired final positions and calculating the desired roll, pitch and yaw angles and the appropriate thrust input that would move the vehicles to their destination, i.e. the high-level controller of the quadrotors. Using the desired roll, pitch and yaw angles and the desired thrust, the on-board low-level controller of the quadrotors would drive them to the desired position. The diagram presented in Fig. 4.7 shows the interaction between the user, the BCI system and the quadrotors. In Fig. 4.8, the corresponding diagram for the experiment where the user interacts with the virtual interface (first experiment) is shown for completeness.

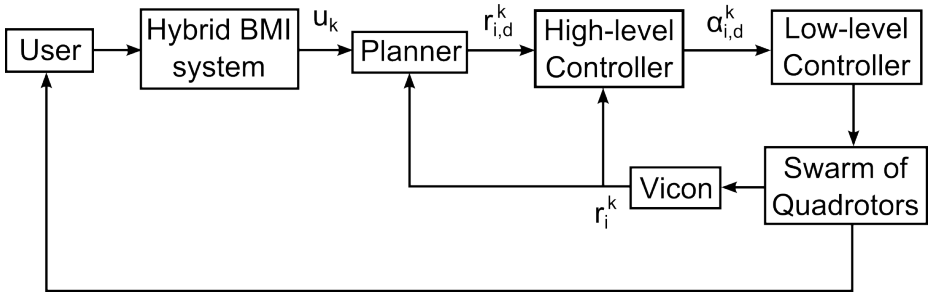

Figure 4.7: System block diagram in the case of the swarm control. The command vector u_k is fed into the planning algorithm. At each iteration k, the planner provides for each quadrotor i the desired position $r^k_{i,d}$ which is translated through the high-level controller into the desired attitude vector $\alpha^k_{i,d} = [\phi^k_{i,d}, \theta^k_{i,d}, \psi^k_{i,d}, T^k_{i,d}]^\top$, where ϕ, θ, ψ and T represent the roll, pitch, yaw angles and thrust, respectively. The Vicon system provides for each agent the current position r^k_i. (Karavas *et al.* (2017), ©2017 IEEE)

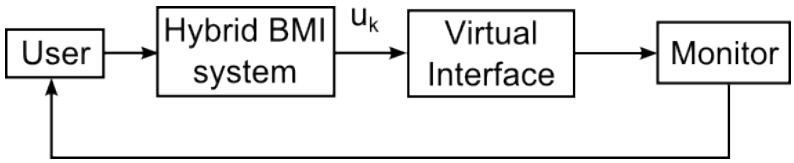

Figure 4.8: System control diagram during the experiment where the user interacts with the virtual interface.

The planning scheme that transformed the BCI system output vector into desired positions calculated a desired position vector $\mathbf{r}_d$ for each quadrotor in Cartesian space based on the following equation:

$$\mathbf{r}_{d,i} = \mathbf{r}_i + \mathbf{r}_{joy} + \mathbf{r}_{eeg} \tag{4.6}$$

where $i = 1 \dots l$ represents the index of each quadrotor, l is the number of quadrotors (in this case $l = 3$), $\mathbf{r}_{d,i}$ is the desired position of quadrotor i, $\mathbf{r}_i$ is its current position, $\mathbf{r}_{joy}$ represents the change in the desired position due to the joystick input and $\mathbf{r}_{eeg}$ the corresponding change due to the EEG input. The above equation can be written in more detail as:

$$\mathbf{r}_{d,i} = \mathbf{r}_i + \begin{bmatrix} \Delta x_k \\ 0 \\ \Delta y_k \end{bmatrix} + (\mathbf{r}_c - \mathbf{r}_i) \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} w \; sign(d_{m-1} - d_m) \tag{4.7}$$

where Δx_k, Δy_k are the controls given by the joystick at iteration k, $\mathbf{r}_c = 1/l \sum_i \mathbf{r}_i$ represents the coordinates of the swarm center in 3D space, w is a weighting factor, d_m represents the distance level of the swarm and $sign()$ is defined as follows:

$$sign(a) = \begin{cases} 1 & \text{if } a > 0 \\ -1 & \text{if } a < 0 \\ 0 & \text{if } a = 0 \end{cases} \tag{4.8}$$

The change of distance level d_{m-1} to d_m was executed in steps. The activation variable Δq_k was added at each iteration k to an accumulator $S_{q,k}$ with starting value $S_{q,0}$. Every time the absolute difference $|S_{q,k} - S_{q,0}|$ became equal or greater than a multiple of 0.5, the distance level changed from its previous value d_{m-1} to its current one $d_m = round(S_{q,k})$, where the function $round()$ rounds its input to the nearest integer. This was done in order to avoid accidental change of the distance level due to misclassification from the EEG system. At the same time, it allowed the users some error which in turn relieved some of the performance pressure and helped them use the system more efficiently.

Next, the desired position $\mathbf{r}_{d,i}$ was translated into desired roll, pitch and yaw angles (ϕ, θ and ψ respectively) and desired thrust (T) in order to move the vehicles appropriately. This was accomplished by a Proportional - Integral - Derivative (PID) controller that drove the position and velocity of a quadrotor to their desired values. The control laws were based on the translational dynamics of the quadrotors shown below:

$$m \begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix}^i = \begin{bmatrix} 0 \\ 0 \\ -mg \end{bmatrix}^i + \mathbf{R}_b^i \begin{bmatrix} 0 \\ 0 \\ \sum_{n=1}^{4} T_n \end{bmatrix}^b \tag{4.9}$$

where m represents the mass of the quadrotor, T_n denotes the thrust from motor n, g is the gravitational acceleration, $[\ddot{x}, \ddot{y}, \ddot{z}]^T$ are the inertial x, y and z accelerations and $\mathbf{R}_B^I$ is the rotation matrix between the inertial frame and the body frame which is placed on the vehicle. The superscripts I and B denote quantities that are related to the inertial and the body frame, respectively. Assuming that the angles ϕ, θ and ψ remain small, as has been done in other works (Mellinger *et al.*, 2012), a PD controller was formulated and it regulated the desired quadrotor attitude based on the position and the velocity. The controller can be written as follows:

$$\ddot{\mathbf{r}}_d = \mathbf{K}_P \mathbf{e}_p + \mathbf{K}_D \mathbf{e}_v \tag{4.10}$$

where $\ddot{\mathbf{r}}_d$ is the desired acceleration vector, $\mathbf{K}_P$, $\mathbf{K}_D$ are diagonal gain matrices for the proportional and derivative terms, respectively, and $\mathbf{e}_p$, $\mathbf{e}_v$ represent the errors in position and velocity. In this work, the desired velocity was set to zero (hover) and the desired yaw angle to its initial value. Then, equation 4.9 was solved for the desired attitude of the vehicle, taking into account that the controller in 4.10 computed the desired translational acceleration. This led to the desired roll, pitch, yaw and thrust values as follows:

$$\phi_d = \frac{1}{g}[\ddot{r}_{x,d} cos\psi + \ddot{r}_{y,d} sin\psi] \tag{4.11}$$

$$\theta_d = \frac{1}{g}[\ddot{r}_{x,d} sin\psi - \ddot{r}_{y,d} cos\psi] \tag{4.12}$$

$$\psi_d = \psi_0 \tag{4.13}$$

$$T_d = m(g + \ddot{r}_{z,d}) \tag{4.14}$$

where ϕ_d, θ_d, ψ_d, T_d are the desired roll, pitch, yaw and thrust, respectively, ψ is the current yaw angle, ψ_0 is the initial yaw angle and $\ddot{r}_{x,d}$, $\ddot{r}_{y,d}$, $\ddot{r}_{z,d}$ represent the desired

linear accelerations along the $x-$, $y-$ and $z-$ axes, respectively. The quantities calculated in 4.11 - 4.14 were finally sent wirelessly via Bluetooth to the quadrotors, which regulated their attitude based on their on-board controllers. More details about the position and attitude control of the quadrotors can be found in Larsson (2016).

4.7 Results

This section presents the results of the experiments described in Section 4.1. As explained, in order to prove the feasibility of the proposed approach and evaluate its performance two experiments took place. In the first one, the subjects had to manipulate the size and position of a two-dimensional disk in a virtual environment in order to complete a task. Its purpose was to verify the applicability of the algorithm in a simulated control environment with multiple subjects. The second experiment involved the control of the position and the relative distance of a team of quadrotors using the proposed BCI system. Its purpose was to assess the system's behavior during the control of an actual robotic swarm. The results of these experiments are presented below.

4.7.1 Virtual Environment Experiment

In order to measure the performance of the BCI system for the first experiment, three different metrics were employed. These were the completion rate, the completion time and the accuracy of the system output. The corresponding results for each of those metrics are presented in the form of boxplots. Each box corresponds to data from all subjects within a group of 10 trials. The $\times$ represents the mean value of the shown metric for each box, while the black asterisks indicate statistically significant differences at the 5% confidence level between each 10-trial group and the one at the beginning of the experiment. The green asterisks show statistically significant

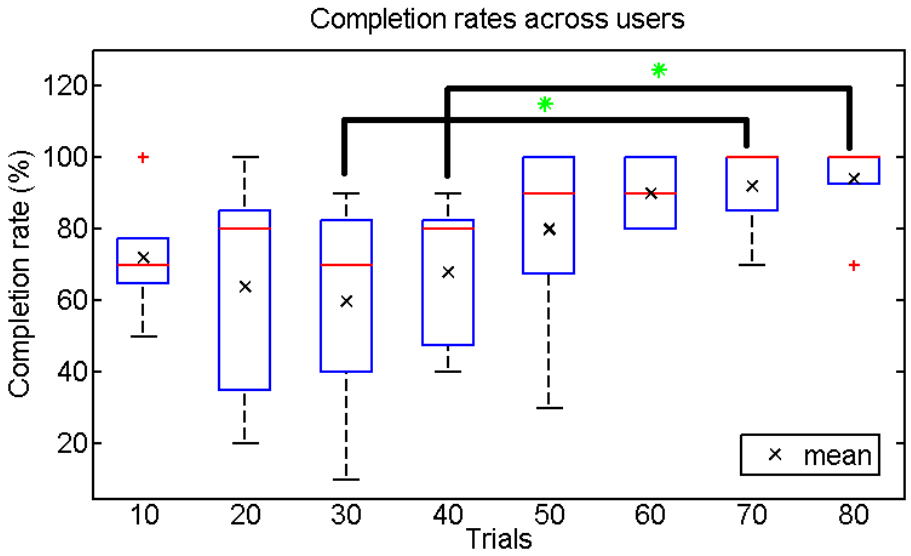

Figure 4.9: Completion rates across all subjects. The × shows the mean value for each group of measurements. The black asterisk above each box shows statistical significance of each group of trials with respect to the first 10 trials. The green asterisk refers to the statistical significance between trials from session 1 (trials 1-40) and the corresponding trials of session 2 (trials 41-80). For example, we compare trials 1-10 to trails 41-50, trials 11-20 to trials 51-60 and so on. (Karavas *et al.* (2017), ©2017 IEEE)

differences at the 5% confidence level between groups of trials from session 1 (trials 1-40) and the corresponding trials of session 2 (trials 41-80). For example, there is a comparison between trials 1-10 and trials 41-50, correspondingly between trials 11-20 and trials 51-60 and so on.

Referring back to the metrics, the completion rate corresponded to the amount of trials that the subject completed successfully, and it is presented here as a percentage of completed trials out of every 10 trials of the experiment. As it is shown in Fig. 4.9, the completion rates for this experiment were increasing, reaching even 100%, as the subjects became more accustomed to the system. In order to prove that this increase was statistically significant, a left-tailed paired t-test was performed. The test concluded that such an increase was indeed statistically significant which means that the subjects were getting better as they were using the BCI system. The same

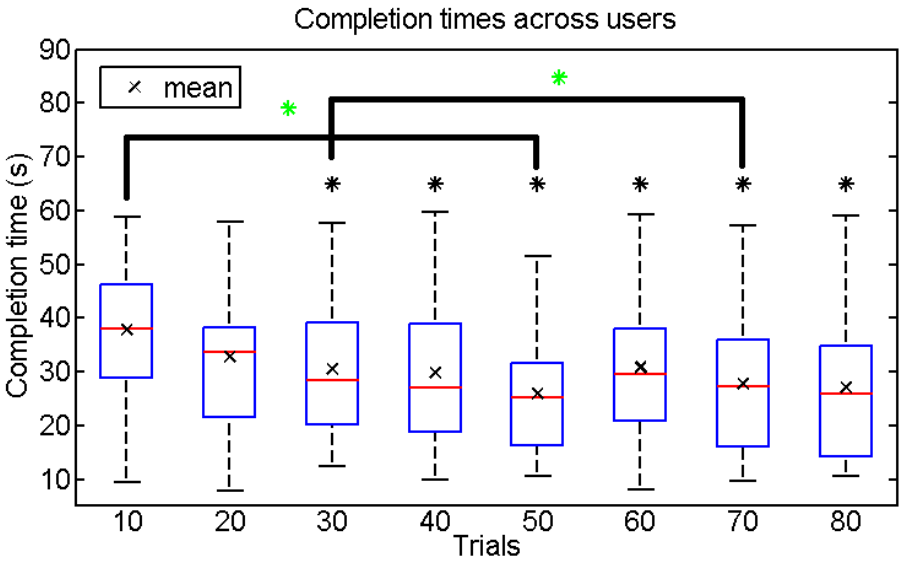

Figure 4.10: Completion times across all subjects. The × shows the mean value for each group of measurements. The black asterisk above each box shows statistical significance of each group with respect to the first 10 trials. The green asterisk refers to the statistical significance between trials from session 1 (trials 1-40) and the corresponding trials of session 2 (trials 41-80). For example, we compare trials 1-10 to trails 41-50, trials 11-20 to trials 51-60 and so on. (Karavas *et al.* (2017), ©2017 IEEE)

type of test was performed in order to compare groups of trials from the first session of the experiment to their matching ones from the second. Indeed, the test showed statistically significant differences on the 5% confidence level between the groups of trials 21-30 and 61-70 and the groups of trials 31-40 and 71-80 (green asterisks). This also proved an improvement in the completion rates (from session 1 to session 2 this time) which further validated that there was an increase in the performance of the system.

The graph in Fig. 4.10 shows completion times, which refer to the time it took for the users to complete the task in seconds. Here, only successful trials were taken into consideration. As seen from the figure, the completion times were getting smaller as the experiment progressed. A right-tailed paired t-test was performed on these data, in order to prove a decrease in the completion time metric. The test did indeed show

that the observed decrease was statistically significant on the 5% confidence level when comparing the first group of trials with each of the next ones (black asterisks). Similar conclusions can be drawn when comparing the groups of trials 1-10 with 41-50 and 21-30 with 61-70, as shown in Fig. 4.10 with green asterisks.

The metric of accuracy is defined as the ratio of correct state transitions n_c of the disk over all of its total state transitions n_t. A state transition is defined as either a size increase, a size decrease or movement. During each trial, at each iteration step, the corresponding accuracy measuring algorithm would check whether a state transition occurred and, if it did, it would also check its type and the location of the disk relative to the target circle. Every time a state transition occurred, the value of n_t was increased by 1. Each time this transition was correct, the value of n_c was also increased by 1. A transition was deemed correct if it helped the completion of the task. For example, if the disk was outside the target, the user had to decrease its size in order to make it fit. Similarly, if the disk was inside the target, the user had to increase its size. For the case of disk movement, the transition was deemed correct if the size of the disk did not change while it was moving. That corresponded to the system correctly identifying that the user was using the joystick. If the user did not perform the intended transition according to the disk's relative location to the target, the value of n_c was not increased.

In this investigation, the focus lay on the accuracy rate of the system rather than the misclassification rate for two reasons. The first one was to incorporate in one metric the performances of both the algorithm and that of the subjects. The second one related to the fact that the subjects were free to complete the task in a self-paced way, meaning that the series of state transitions for the disk was not given a priori to them and they were free to choose which ones to apply and in what order. Thus, it would not be appropriate to consider misclassification rates because there were no

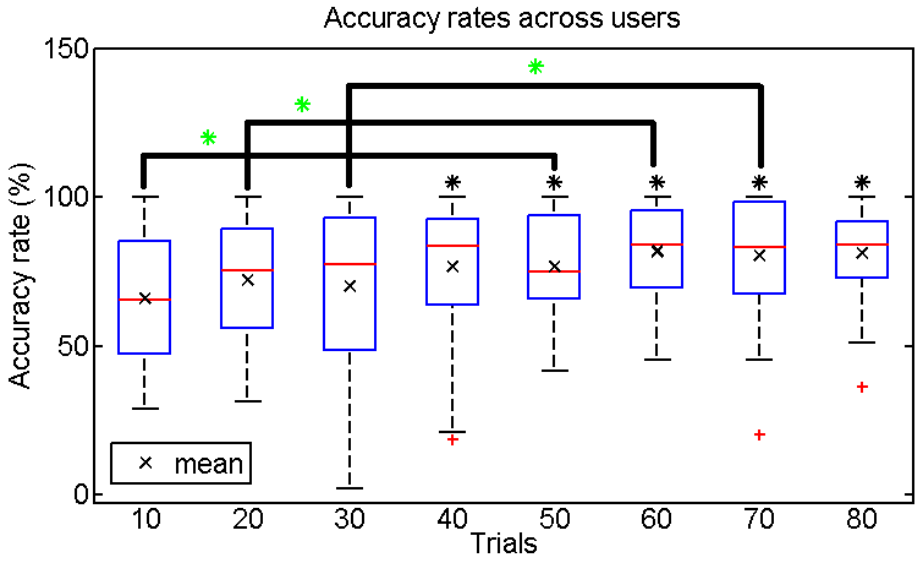

Figure 4.11: Accuracy rates across all subjects. The × shows the mean value for each group of trials. The black asterisk above each box shows statistical significance of each group with respect to the first 10 trials. The green asterisk refers to the statistical significance between trials from session 1 (trials 1-40) and the corresponding trials of session 2 (trials 41-80). For example, we compare trials 1-10 to trails 41-50, trials 11-20 to trials 51-60 and so on. (Karavas *et al.* (2017), ©2017 IEEE)

desired states that the disk had to go through at each specific time, which means that no comparison could be done.

The overall accuracy of the system against the trials/time is presented in Fig. 4.11. Also, the corresponding accuracy rates for each brain state separately are presented in Fig. 4.12 - Fig. 4.14. The overall accuracy over all trials was in general above 60% on average. At the same time, there was an increase in the accuracy rates as the subjects interacted more with the system. The difference between trials 1-10 and trials 31-80 is significant on the 5% confidence level. For these data, a left-tailed paired t-test was performed in order to prove an increase in the accuracy. There is also a statistically significant increase on the 5% confidence level when comparing the first 30 trials of session 1 and the first 30 of session 2 (trials 41-70). These observations indicate that the accuracy levels were indeed increasing with time as the subjects were getting more familiar with the system and the task.

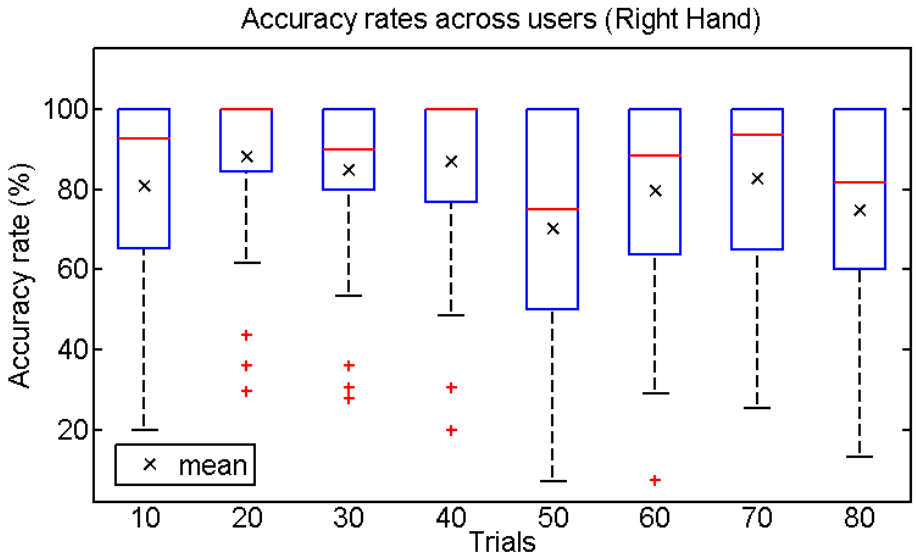

Figure 4.12: Accuracy rates across all subjects for the "Right Hand" task. The × shows the mean value for each group of measurements. (Karavas *et al.* (2017), ©2017 IEEE)

Finally, for the accuracy rates of the individual brain states, a left-tailed paired t-test on the 5% confidence level was performed (the same as for the overall accuracy rates). Based on that, the following conclusions can be drawn. First, there was no statistical difference across trials when considering the "Right Hand" task (Fig. 4.12). This was expected since the users were moving the joystick naturally without modulating explicitly any aspects of their brain signals. Thus, no training was involved and the activations should have remained the same. On the other hand, there was explicit manipulation of brain activity for both the "Left Hand" and the "Rest" task, which explains the differences observed in those cases, validated as significant by using the same test as for the case of the "Right Hand" task.

The previously presented results provide a strong evidence that the system can be used successfully by multiple users, while achieving high performance and accuracy rates with minimum training (the entire control phase lasted on average less than an hour). It is important to note that the accuracy rates for most users at the very beginning of the experiment (first 10 trials) were still high enough to permit the

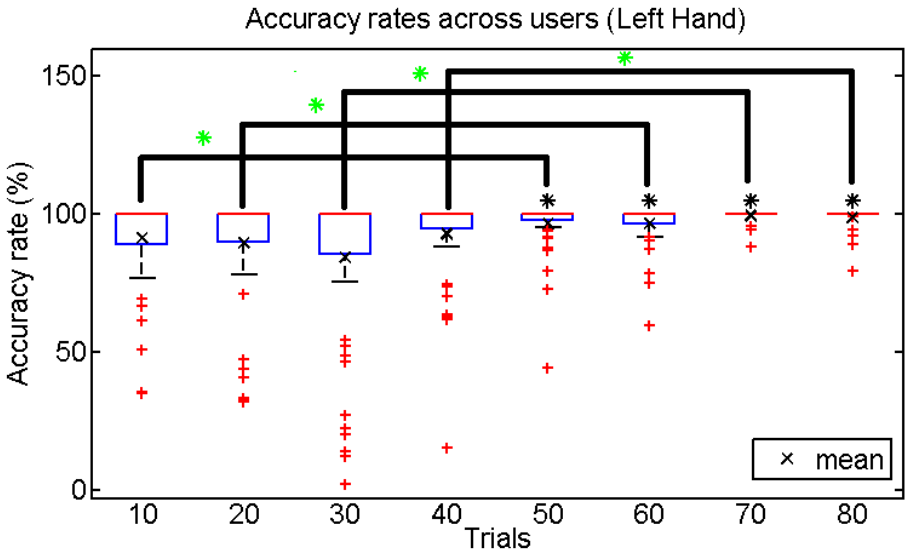

Figure 4.13: Accuracy rates across all subjects for the "Left Hand" task. The × shows the mean value for each group of measurements. The black asterisk above each box shows statistical significance of each group with respect to the first 10 trials. The green asterisk refers to the statistical significance between trials from session 1 (trials 1-40) and the corresponding trials of session 2 (trials 41-80). For example, we compare trials 1-10 to trails 41-50, trials 11-20 to trials 51-60 and so on. (Karavas *et al.* (2017), ©2017 IEEE)

Table 4.1: Model order parameters for all subjects.

	S1	S2	S3	S4	S5
# of PCs	4	4	4	4	5
# of states	5	4	7	5	4
# of GMs	8	6	4	5	5

completion of the task several times. It is worth mentioning again that the users were free to complete the tasks in any way they wanted, which means that they were able to combine brain signal control and joystick control seamlessly, at their own pace and without any reported difficulty. Finally, the system allowed the users to improve fast as they were interacting with the virtual environment.

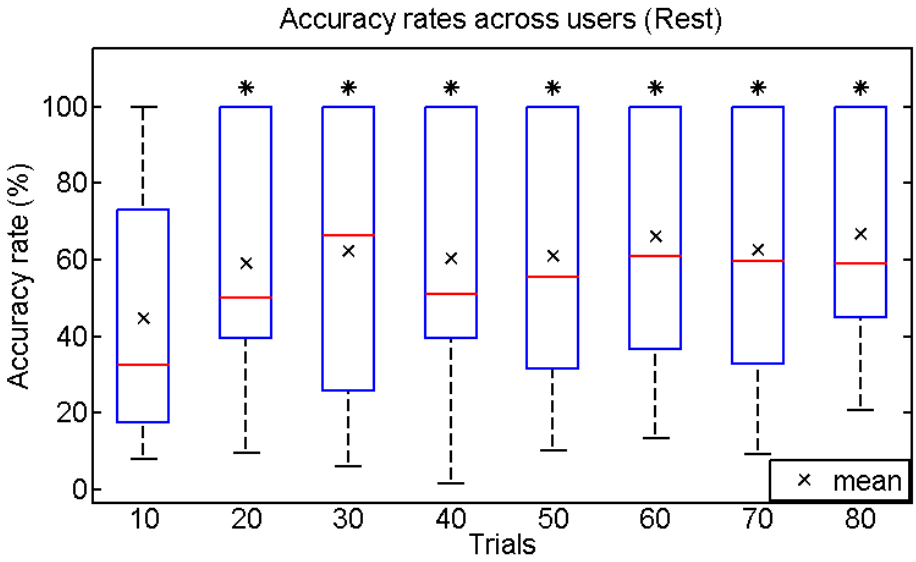

Figure 4.14: Accuracy rates across all subjects for the "Rest" task. The × shows the mean value for each group of measurements. The black asterisk above each box shows statistical significance of each group with respect to the first 10 trials. (Karavas *et al.* (2017), ©2017 IEEE)

4.7.2 Common Principal Components

Table 4.1 presents the number of parameters that the training algorithm extracted for each subject. As it can be seen, the model order for each subject was different, which was to be expected due to intra-subject variability. In order to demonstrate the generalization characteristics of the method, the Principal Components (PCs) that explained the variance of the spectral features across the different brain states of the subjects were examined. The goal was to find PCs that were similar across subjects. In order to do so, all the PCs that were chosen for the subjects where grouped together (see *Methods*) using the method of *k-means clustering* (Hartigan, 1975). Since the number of clusters k is not, in general, a-priori defined for a dataset (Halkidi *et al.*, 2002a,b), its value in this analysis was chosen based on the following procedure. A *k-means* clustering method was first applied for different values of k. In each case, the Sum of Squared Errors (SSE) between all data points and the centers

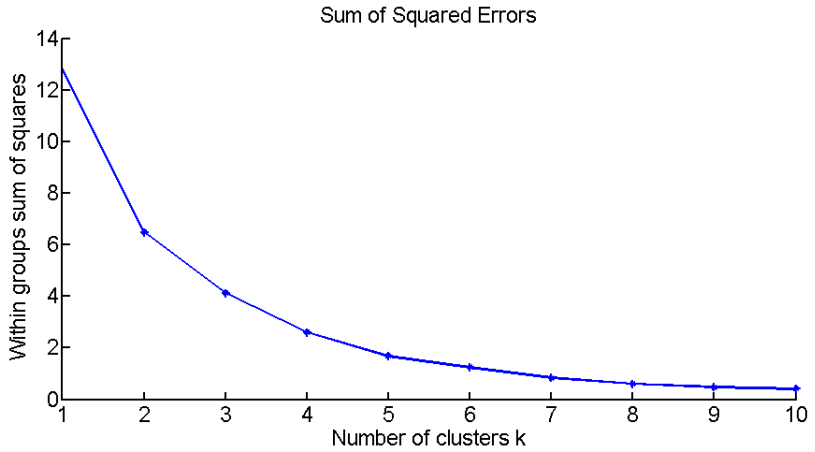

Figure 4.15: Sum of Squared Errors (SSE) versus number of clusters k.

of the clusters that they belonged to was calculated as follows:

$$SSE = \sum_{i=1}^{k} \sum_{x \in c_i} dist(x, c_i)^2 \tag{4.15}$$

where x is the cluster member, c_i represents the cluster centroid and $dist()$ represents the Euclidean distance between x and c_i (Fig. 4.15). Then, for each k, the relative difference between SSE_k and SSE_{k+1} with respect to the SSE for $k = 1$ was computed as:

$$\Delta SSE_k = \frac{SSE_{k+1} - SSE_k}{SSE_1} \tag{4.16}$$

where SSE_k, SSE_{k+1} represent the value of SSE for k and $k+1$, respectively, while SSE_1 is the SSE for $k = 1$. The value of ΔSSE_k represents the relative decrease in error with respect to the error for $k = 1$ as the number of clusters is increased by one. The goal was to choose the minimum number of clusters, while also having minimum clustering error. The minimum value of k for which $\Delta SSE_k < 10\%$ was finally chosen. According to Fig. 4.16, this value corresponded to $k = 5$.

As an additional measure to choose the best clusters and to avoid any outliers, the method of silhouettes (Rousseeuw, 1987) was also applied. The silhouette is a measure that is used in clustering analysis and it represents how similar a member is

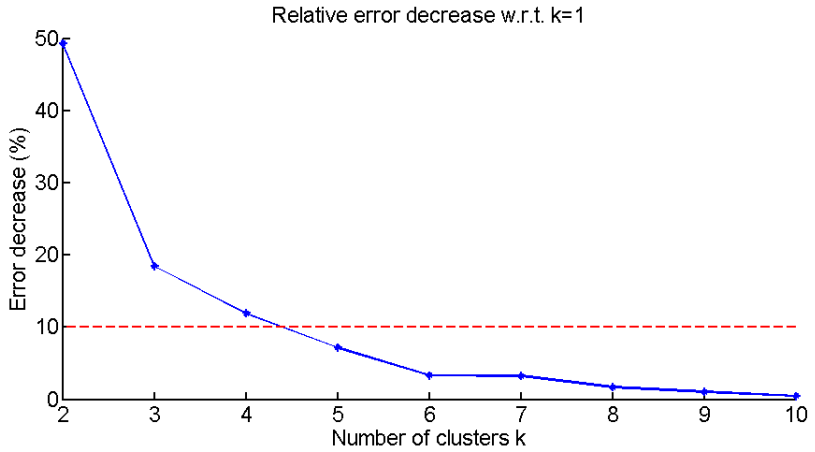

Figure 4.16: Relative Sum of Squared Errors (SSE) decrease against number of clusters k. Cut-off value of 10% is selected.

to its own cluster relative to the other clusters. It takes values in the range of $[-1, 1]$ with 1 representing perfect fit and -1 the opposite. From the clusters produced by the k-means clustering method, only those clusters that had an average silhouette value $s \geq 0.3$ were chosen as candidates. As a final step, in each cluster, if there were more principal components for the same subject, the method would select the one that was closer to the cluster center.

The results of this analysis are presented in Fig. 4.17 as scalp maps of common PCs and they are organized according to their type and the subject they belong to. The brain maps shown in this figure are created by interpolating the values of the PC weights that correspond to the 11 channels which the analysis was focused on (see *Methods*) considering all other channels as having a weight of zero. Three different types of PCs are shown with two (Type 1 and 2) being similar across four out of five subjects and one (Type 3) being similar across all of the subjects. That indicates that there was a strong correlation between the underlying brain states used for the control phase and the particular channels that had the biggest weights in those components. It also proves that the BCI methodology followed in this work can be greatly generalized without changes across multiple users.

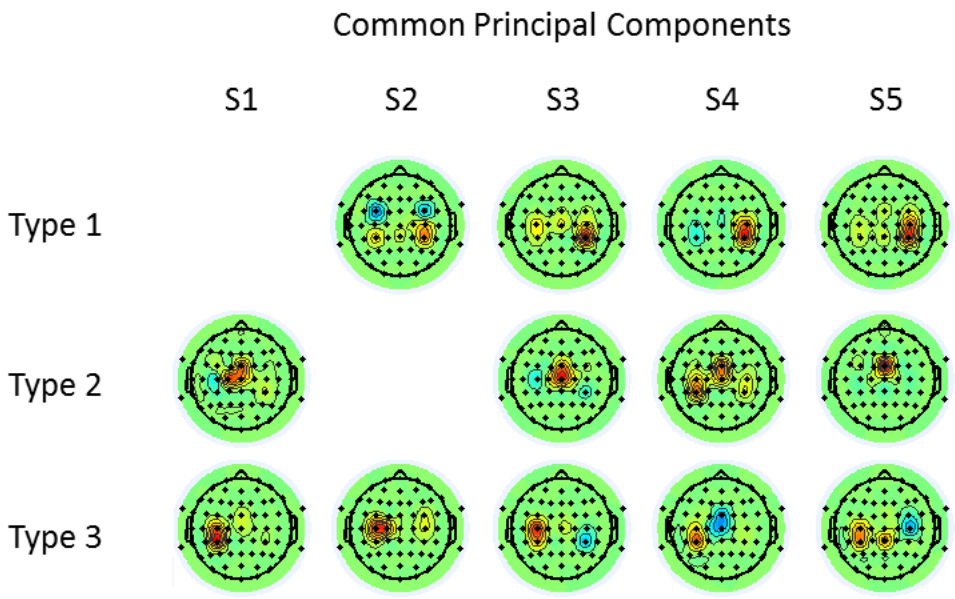

Figure 4.17: Common principal components across subjects. Each row refers to a different type, while each column refers to each of the subjects that participated in the experiment.

4.7.3 Control of a Swarm of Quadrotors

In the second experiment, the subject was controlling a swarm of 3 quadrotors using the proposed hybrid BCI system. As mentioned, the quadrotors were initially placed on a linear formation and at specific distance to each other. The task was to pass the quadrotors through a rectangular hoop. The subject had to bring the vehicles closer to the hoop, change the distance between them so that they can fit through the hoop using his brain signals, move them through the hoop using the joystick and finally return them to their original distances using brain activations again.

Fig. 4.18 shows snapshots of the experiment. A video of the experiment is included in HORC ASU (2016). The 3D position of the quadrotors during the experiment is shown in Fig. 4.19. In addition, Fig. 4.20 shows how the elements Δq (second subplot from the top) and Δy (bottom subplot) of the command vector u (see *Methods*)

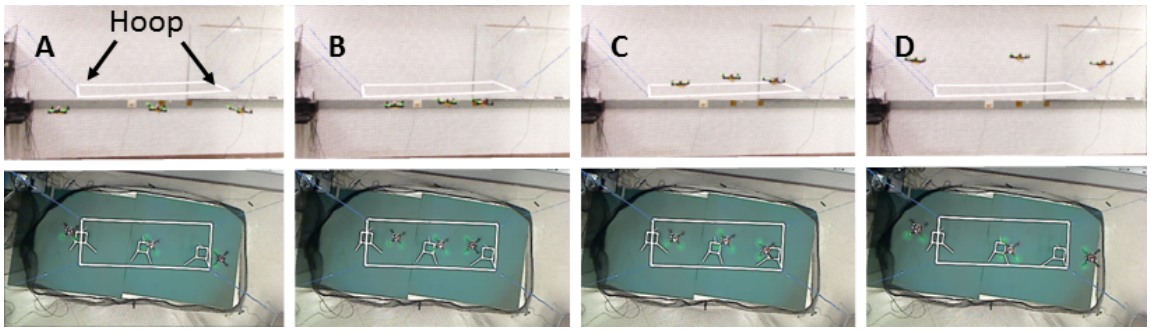

Figure 4.18: Snapshots of the three quadrotors passing through the rectangular hoop during the second experiment. The top row shows a side view of the motion of the swarm, while the bottom row shows the top view of the quadrotors. A: Initial formation, B: Change of formation, C: Passing the quadrotors through the hoop, D: Returning to initial formation. Video at HORC ASU (2016). (Karavas *et al.* (2017), ©2017 IEEE)

affected the position of the quadrotors in the line formation ($y-$ axis, top subplot) and their height ($z-$ axis, third subplot from top), respectively, in a cumulative way. The position and height of the quadrotors are expressed in meters with respect to the global frame provided by the localization system (Vicon system, see also Section 4.6).

As it can be seen in the figures, the subject was able to change the control input seamlessly from joystick input to EEG input and back with minimum error and without the vehicles changing their relative distance while passing through the hoop. This was a real-time demonstration of controlling a swarm of quadrotors using the proposed hybrid BCI which employed both EEG activations and joystick inputs.

4.8 Conclusions

This chapter presented a hybrid BCI system that combined brain activity with input from a game controller in order to control a small team of quadrotors. The system was extensively tested both in a virtual environment and on an actual robotic platform. The presented BCI system was simple to use and the users were quickly able to control it with high accuracy. Moreover, its application for the control of a swarm

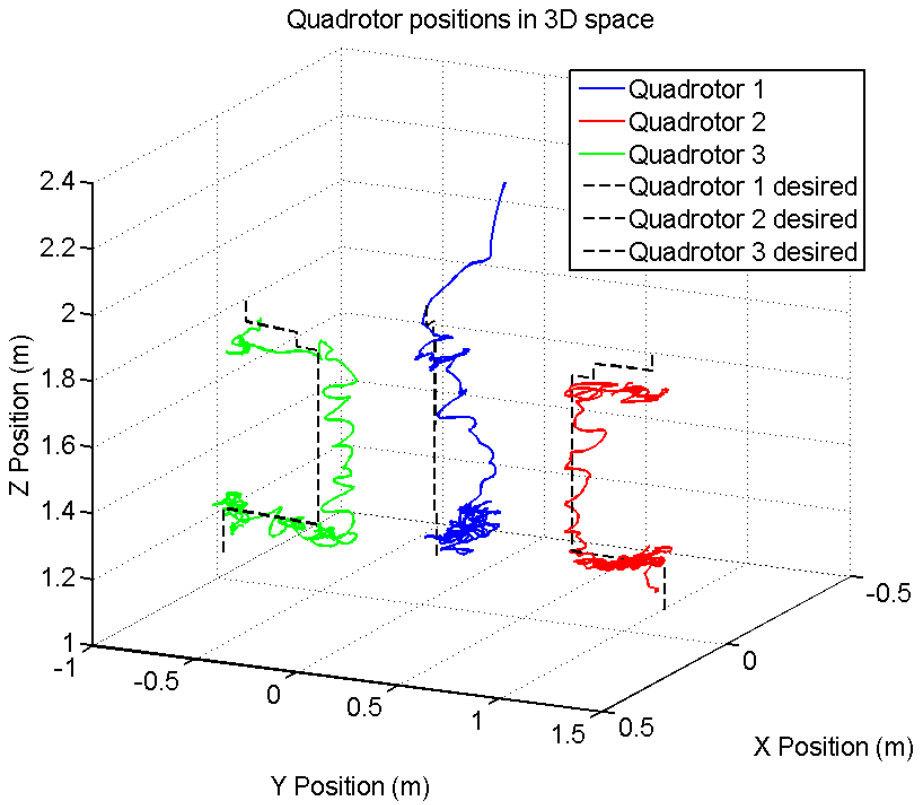

Figure 4.19: Position of the 3 quadrotors in 3D-space during the second experiment.

parameter/behavior (i.e. the distance between the agents) was successful proving its feasibility for swarm related applications. It should be mentioned here that although the user was controlling the system in a binary fashion (increase/decrease of size or distance), the classification algorithm differentiated between three different states. The purpose of this approach was to explicitly recognize the use of the joystick and deal with it appropriately. In the opposite case, where the classification would be done only between the "Left Hand" and "Rest" states, the algorithm could mix the motion of the right hand due to joystick movement with that of left hand imagination and could provide false output. Furthermore, the thresholding procedure provided robustness to misclassification errors which would potentially jeopardize the smooth control of the robotic platform. In addition, allowing the users to control the activa-

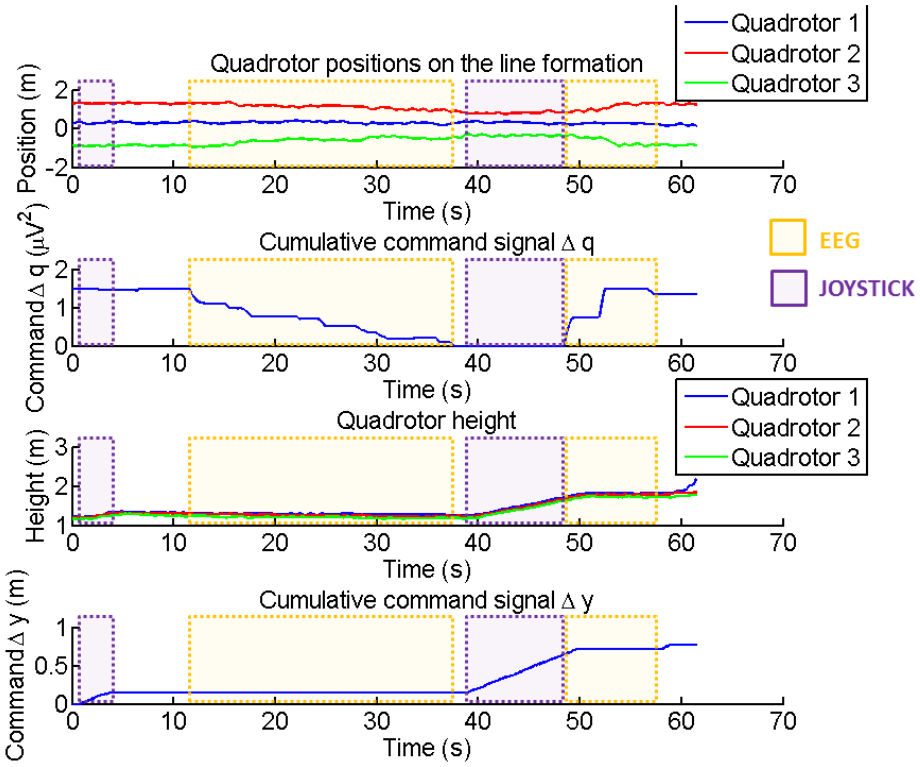

Figure 4.20: Position of the 3 quadrotors and EEG and joystick input during the second experiment versus time. (Karavas *et al.* (2017), ©2017 IEEE)

tion variable Δq_k in a continuous manner provides more flexibility to the system as they can control not only the direction of the desired change but also its rate.

On the other hand, one possible drawback of this approach was that the mental tasks that the users were applying in order to control the system were not directly related to the task at hand, since they were trying to change the size of the disk or the distance between the agents by imagining limb movement. In the author's opinion, this may have provided a less intuitive solution with regard to swarm applications. In addition, the use of motor imagery as a mental task for the BCI put certain constraints on the resulting system, since the use of the joystick did not allow right hand movement imagination as an extra input from the subject. Chapter 5 proposes a different type of mental imagery, namely speech imagery, which has the potential to provide an elegant solution for the aforementioned problems.

Chapter 5

SPEECH IMAGERY AS AN ALTERNATIVE MENTAL TASK FOR BCI

As a continuation of the previous chapter, this one proposes a different type of mental task that can be used to interact with a swarm of robots. This task relies on speech imagery, i.e. the act of pronouncing a vowel or a word inside someone's mind without actual vocalization of the sounds. This results in activations at specific brain areas that are related to speech and hearing. In this chapter a novel method of feature extraction is described, while some insights about the proper use of speech imagery are also presented. An additional analysis about the brain activations during the execution of the mental task is also given. The chapter concludes by discussing potential implications of the method.

5.1 Experimental Procedure

In this investigation, 15 healthy subjects (S1-15, 11 males and 4 females, ages 22-32) performed three different types of imagined speech, namely imagined speech of *short words*, *long words* and *vowels*. The aim was to compare these types of imagined speech to each other and evaluate the suitability of each group for BCI applications. The group of short words included the words "in", "out" and "up", while the group of long words consisted of "cooperate" and "independent". These words were chosen in order to evaluate the effect of the meaning and the complexity of the words. In order to evaluate the effect of the sound, three phonemes were used, namely /a/, /i/ and /u/. Finally, in order to further analyze the effect of the complexity, i.e. the length and the different sounds, an additional experiment was performed where the subjects had to imagine either one of the short words ("in") or one of the long words ("cooperate").

Figure 5.1: Experimental setup. The subject is wearing the cap with EEG electrodes and looks at a monitor a few inches away. The monitor shows the task that the subject must execute. In this illustrative figure, the subject was imagining pronouncing the word "in".

During the experiments, the subjects were instructed to pronounce these words internally in their minds and avoid any overt vocalization or muscle movements. The subjects were receiving instructions about the desired word/phoneme based on visual cues from a computer monitor. The experimental setup is shown in Fig. 5.1, while 5.2 shows the experimental procedure which is described in more detail below. The experimental protocol was approved by the ASU IRB (Protocols: 1309009601, STUDY00001345) and each participant signed an informed consent form before the experiment.

Each subject performed one to three sessions of imagined speech. Each session corresponded to one of the mentioned groups of speech imagery, e.g. 3 short words, and was conducted approximately in 1 hour. A single experimental session was comprised of 100 trials per word or sound, which were shown randomly. During each trial, the subject would hear a beep sound that was repeated at period T. This helped create the rhythm that subjects should imagine pronouncing the words or phonemes. In more detail, the beep sound appeared when the trial started and was repeated 4

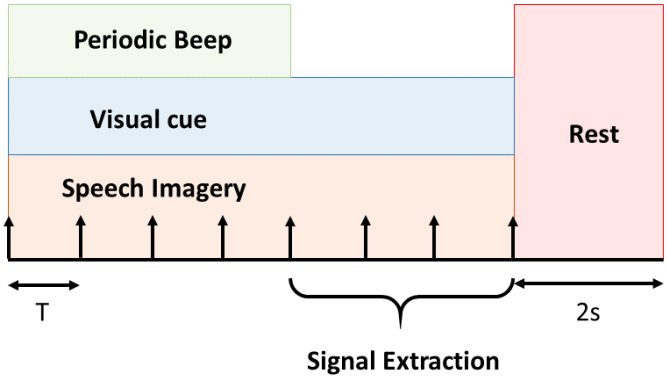

Figure 5.2: Experimental procedure. The vertical arrows represent the time instants where the subject was expected to perform speech imagery, and T denotes the rhythm period.

times. At the beginning of the trial, the subject was also prompted with a visual cue indicating the desired word to be imagined. The cue lasted for $7 \times T$s. The subject was instructed to perform speech imagery at each beep sound and continue at the same rhythm until the visual cue disappeared. This resulted in the subject performing speech imagery for an additional 3 periods after the last beep sound. Finally, the trial ended with a rest period of approximately 2s where no cue and no sounds were present. For short words and vowels, the period was $T = 1s$, while for long words the period was $T = 1.4s$. In the case of comparing between a short and a long word, $T = 1.4s$ was also chosen to render the comparison more accurate. The values for T were chosen empirically based on how long it would take the subjects to pronounce the words overtly.

This experimental procedure is similar to the ones conducted by D 'Zmura *et al.* (2009) and Brigham and Kumar (2010) where a rhythm is used to help the subjects perform the task more efficiently. The signal that is used for classification of the task though is extracted from a segment of each trial, where no auditory stimulus is present. This guarantees that the classification results rely only on the task of speech imagery and not on any external stimuli.

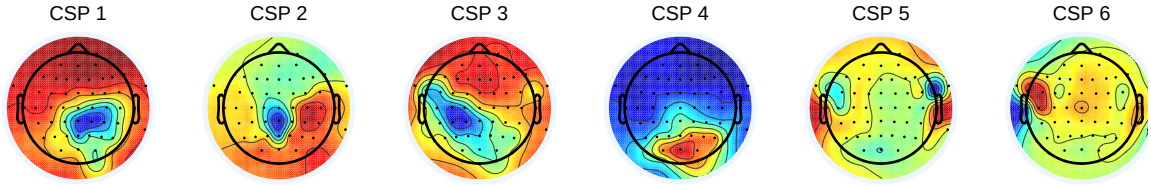

Figure 5.3: The first 6 CSP patterns for subject S3 during speech imagery of short words. The CSP analysis is based on the comparison of speech imagery vs. a rest condition.

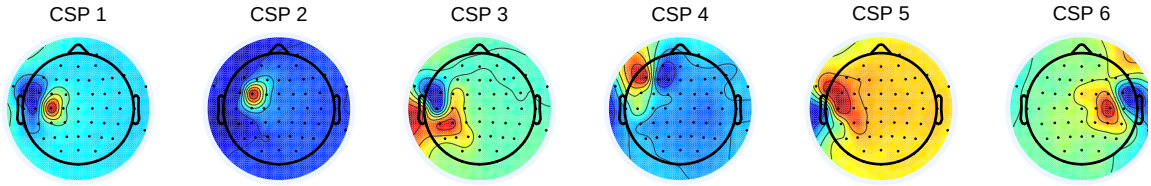

Figure 5.4: The first 6 CSP patterns for subject S3 during speech imagery of short words. The CSP analysis is based on the comparison of brain activity between speech imagery of 3 different short words.

5.2 Data Acquisition and Preprocessing

The EEG signals were collected using the BrainProducts ActiCHamp amplifier system from 64 electrodes placed on the subjects according to the 10/20 International system (Klem *et al.*, 1999) with the help of a BrainProducts ActiCAP electrode cap. The data were recorded at 1000Hz and later downsampled at 256Hz for faster processing and ease of storage. During preprocessing, a 5^{th} order Butterworth bandpass filter between 8-70 Hz was applied to the data in order to remove any low-frequency trends as well as possible artifacts related to ElectroMyoGraphic (EMG) activity. A notch filter at 60Hz was also applied in order to remove line noise. Finally, an ElectroOculoGram (EOG) artifact removal algorithm (He *et al.*, 2004b) was applied on the data to eliminate any eye blinking or eye movement artifacts. After this procedure, the data were ready for analysis and classification.

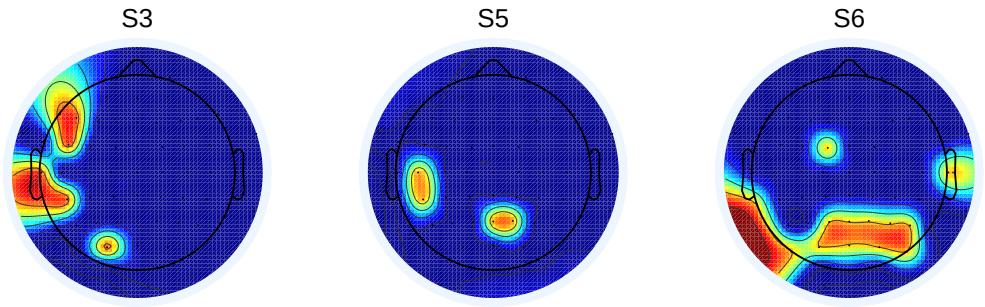

Figure 5.5: Scalp map of the thresholded autocorrelation score of the channels for subjects S3, S5, S6 performing short words imagery.

5.3 EEG Data Analysis

The aim of this section is to analyze the collected data and assess the level at which they reflect brain activity corresponding to actual speech imagery and not some other type of mental task or artifacts. In order to evaluate the data, three different approaches were used. The first two were based on Common Spatial Patterns (CSP) (Lotte *et al.*, 2010), which have been used extensively in the literature in cases of limb movement imagination (motor imagery). The last approach was based on the examination of the autocorrelation of the signals and its characteristics.

In the first CSP approach (CSP Method 1), a classical binary CSP analysis was performed. The first class was related to the general speech imagery task, i.e. there was no distinction between the imagined sounds, while the second class was related to a resting state. The resting state signals corresponded to the last 1.5s of each trial (pause period) where the subjects were not performing any mental task. The first 6 CSP patterns from this analysis for subject S3 for the case of short words imagination are presented in Fig. 5.3. In this binary analysis, the first few and the last few patterns correspond to the filters that increase the variance of the speech imagery task and the resting state, respectively. Since the interest lies exclusively on the speech imagery task (there is no merit in showing what patterns are elicited

during rest), only the first patterns of the CSP analysis are shown in Fig. 5.3. The first 6 of them were deemed enough to examine if the subjects activated the correct brain regions during the trials.

In the second CSP approach (CSP Method 2), a multi-class CSP algorithm (Grosse-Wentrup and Buss, 2008) was applied in order to distinguish explicitly between the different sounds during the speech imagery task. In this algorithm, an Independent Component Analysis (ICA) was first performed in order to obtain orthogonal channels. Then, the channels were ranked based on their Mutual Information with the corresponding labels from highest to lowest. In Fig. 5.4, the first 6 CSP patterns for subject S3 during imagination of short words are shown. These patterns were chosen because, according to the previous description of the applied method, they contained the most information related to the correct discrimination of the different speech imagery classes.

Based on the CSP patterns shown in Fig. 5.3 and 5.4, it can be concluded that the brain activity during the experiments was located mainly on the left frontal, middle and parietal sides of the brain which correspond to Broca's area, the motor cortex and Wernicke's area. According to the related literature (Kim *et al.*, 2014; Martin *et al.*, 2014, 2016), these areas are involved in speech production and recognition and, thus, it can be proven that the detected activity is indeed related to speech imagery and is not an exclusive product of an irrelevant mental task. This observation though is more evident when applying the multi-class CSP approach (CSP1 - CSP5 in Fig. 5.4) compared to the binary CSP method (only evident in CSP3 and CSP6 in Fig. 5.3). In particular, the patterns CSP2 and CSP4 in Fig. 5.3 seem to relate more to the visual cortex rather than any speech related areas. Nevertheless, this is to be expected, since during the main trial a visual cue was presented to the users while in the rest period the cue was absent. Thus, the CSP analysis just detected and showed

that particular difference between the two phases of the experiment. However, this visual stimulus is present in all trials across all classes. Therefore, its effect on the actual classification procedure is negligible.

In order to further verify the above conclusions, an additional analysis based on the autocorrelation of the signals was performed on each channel. In the experiments the subjects were instructed to perform speech imagery at specific time intervals defined by auditory cues at a certain rhythm. The basic premise is that this rhythm should also be visible in the autocorrelation function of the signals. The channels that exhibited such periodicity repeatedly across trials were detected by analyzing the Fast Fourier Transform (FFT) of the autocorrelation function of each channel and scoring those channels whose highest peak in the FFT was close to the frequency of the auditory cue. The ones that were exhibiting such periodicity in $70 - 80\%$ or more of the trials (depending on the subject) were finally chosen. The chosen channels for subjects S3, S5 and S6 are shown in 5.5. As it can be seen, this analysis also points more towards the same areas of the brain as in the CSP methods discussed previously, which solidifies the validity of the data and their relation to speech imagery.

It is worth mentioning that, among the three methods presented here, the binary CSP provided the highest classification accuracy and was quite consistent in most of the cases. It was also the simplest method as it was unsupervised in the implementation described in this work, i.e. all mental tasks were considered as one class, e.g. involving speech imagery, instead of separated classes. The multi-class CSP occasionally yielded better results than the first method, depending on the subjects and conditions. Finally, selecting channels based on the autocorrelation method yielded the least classification accuracy in our experiment. Hence, only the two CSP methods were used in the subsequent analysis for the reduction of the EEG data dimensions.

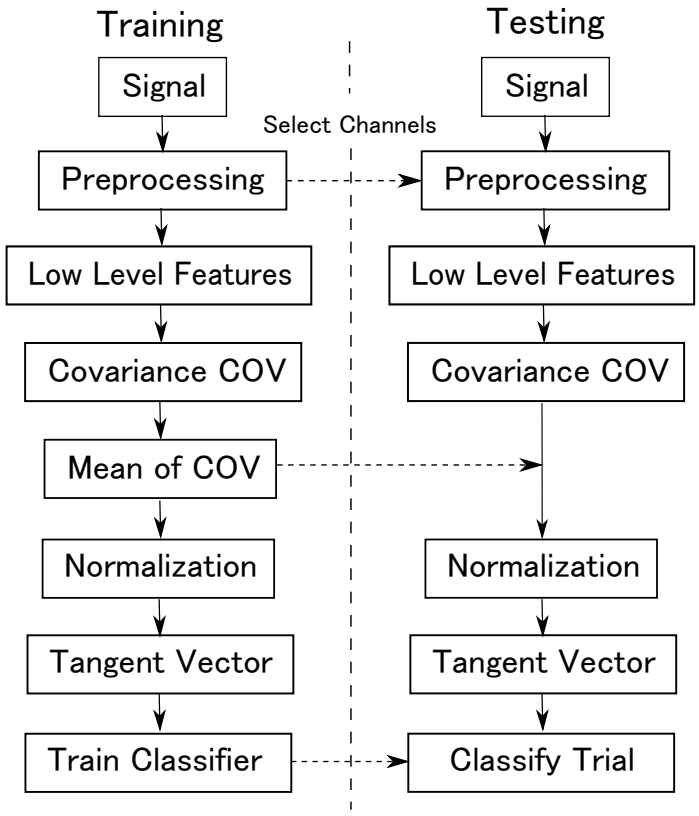

Figure 5.6: A diagram that describes the feature extraction and the classification approach.

5.4 Feature Extraction and Classification Procedure

This section describes the feature extraction and classification method that is applied on the clean data after the initial preprocessing. The procedure can be separated into four main steps. In the first one, the data were further processed in order to extract spatial and/or frequency information related to the speech imagery task. The second step involved the extraction of low-level features that described the underlying signals. Next, the high-level features were extracted from the signals by computing their covariance and mapping them on a lower dimensional space. In the final step, a classifier was trained and applied on the data through a number of cross-validations. The method is shown in Fig. 5.6 and its details are given below.

During the first step, a CSP algorithm was applied in order to reduce the dimension of the data and choose the most appropriate channels for classification. As mentioned in the previous section, there were two CSP methods that were applied. In the first case, a binary CSP approach was followed (CSP Method 1). The entire dataset was divided into two classes. The first was related to the speech imagery task and the second corresponded to a resting state. The regularized CSP method provided by Lotte *et al.* (2010) was used to perform this binary CSP analysis. The first few and the last few CSP filters were chosen for the subsequent processing. In the second CSP approach, a multi-class CSP algorithm (Grosse-Wentrup and Buss, 2008) was applied (CSP Method 2) in order to distinguish between each individual imagined sound or word. In this case, for reasons that were explained in Section 5.3, only the first few CSP filters were chosen. In both cases, the number of CSP filters varied between 6 to 12 depending on the subject.

Next, the CSP filters were applied to the raw data. Overlapping epochs that corresponded to the speech imagery task were then extracted from each trial. At this point, the low-level features were computed. This was performed in two ways in order to assess which one was the most suitable. In the first case, the features corresponded solely to the time domain data, i.e. the signal $\mathbf{X} \in \Re^{m \times n}$, where m is the number of channels, and n is the number of samples in one epoch. In the second case, a Wavelet Transform (WT) was applied on each channel and the wavelet coefficients of the signals were also extracted using the Morlet wavelet. The feature vector was the concatenation of the channel index i_{Chn}, the time domain data, and the wavelet coefficients of each channel, i.e. $\mathbf{X} \in \Re^{d \times L}$, where $d = n_{wc} + 2$, n_{wc} is the number of wavelet coefficients at one instant time sample, and $L = nm$. It is worth noticing that although the variance of the channel index (diagonal element) was equal for all trials, its correlation to other features (off diagonal elements) would vary among trials. This

approach allowed the combination of both spatial and spectral information into one structure.

After the successful extraction of the low-level features, the covariance matrix $\mathbf{S}$ of the signal $\mathbf{X}$ for each extracted epoch was computed as follows:

$$\mathbf{S} = \frac{1}{n-1}\mathbf{X}\mathbf{X}^T \tag{5.1}$$

This step resulted in the mapping of the EEG signals into the Riemannian space. Then, the resulting signal was remapped into the corresponding tangent space. At first, the mean value of the covariance matrix $\mathbf{C}$ across a testing dataset was computed. The tangent vector $\mathbf{T}$ was then computed through a normalization procedure using $\mathbf{C}$ as:

$$\mathbf{T} = \log_{\mathbf{C}} \mathbf{S} = \log(\mathbf{C}^{-1/2}\mathbf{S}\mathbf{C}^{-1/2}). \tag{5.2}$$

The logarithm of the matrix $\mathbf{A} = \mathbf{C}^{-1/2}\mathbf{S}\mathbf{C}^{-1/2}$ was computed as:

$$\log(\mathbf{A}) = \mathbf{V}\log(\mathbf{\Lambda})\mathbf{V}^T = \mathbf{V}diag([\log(\lambda_1), \ldots, \log(\lambda_n)])\mathbf{V}^T \tag{5.3}$$

where $\mathbf{V}$ contains the eigenvectors of $\mathbf{A}$, $\lambda_1, \ldots, \lambda_n$ are its eigenvalues and $\vec{\Lambda}$ is a matrix that has the eigenvalues of $\mathbf{A}$ on its diagonal. The vector $\mathbf{T}$ represented the high-level features of the signals that were later used for the final classification procedure.

Finally, in order to properly assess the performance and validity of the approach, a cross-validation procedure with 10 cross-validations for each subject and each speech imagery type was adopted. In each case, the dataset containing the final high-level features was randomly separated into a training set and a testing set. A Relevance Vectors Machine (RVM) classifier was trained each time on the training set and later tested on the testing set. For the CSP Method 2 (multi-class CSP algorithm), the CSP filters were computed separately for each different training set. The results of the cross-validation procedure are presented in the next section.

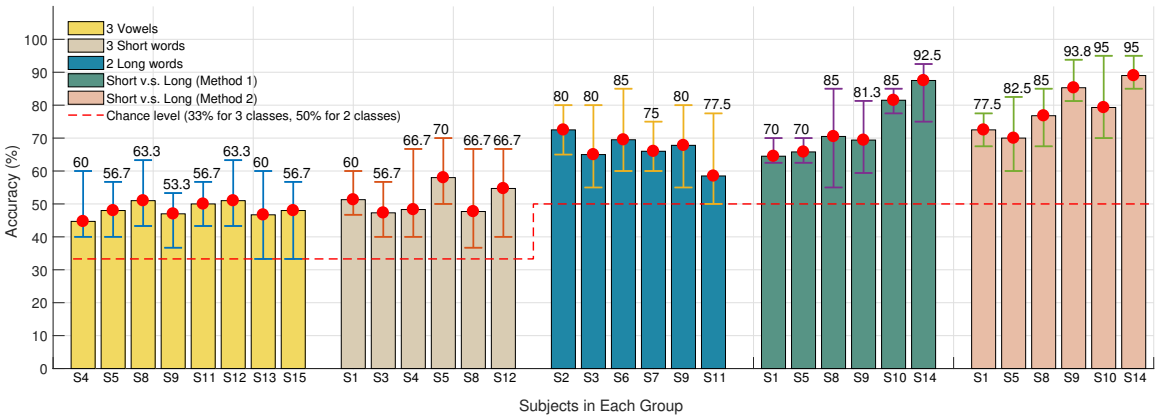

Figure 5.7: Mean, minimum and maximum classification accuracy (%) for different subjects in each mental task.

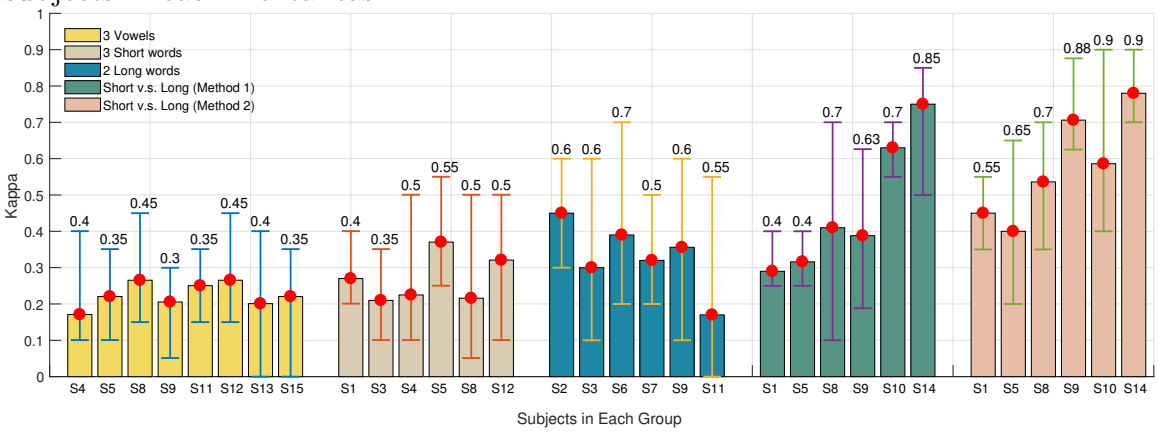

Figure 5.8: Mean, minimum and maximum of Kappa (κ) value for different subjects in each mental task

The aim of this section is a) to compare the different types of speech imagery (i.e. vowels, small words, long words, a short v.s. a long word) with respect to the classification accuracy they can produce and b) to compare the two different types of low-level features (i.e. time domain data v.s. the combination of time and frequency domain information) as applied on the classification of a short v.s. a long word.

In order to evaluate the performance of the proposed method and to assess the suitability of each different speech imagery type for BCI applications, a 10-fold cross-validation procedure was performed. The training and testing sets were partitioned

randomly from the dataset that contained the high-level features presented in the previous section. In more detail, for short words and vowels, the data were classified across 3 different classes with 90% (270 trials) of the dataset being used for training and 10% (30 trials) for testing. For long words and the comparison between a short and a long word, 80% (160 trials) of data were used for training and 20% (40 trials) for testing.

For each case of speech imagery, a number of overlapping epochs was extracted from each trial. More specifically, for short words and vowels, a segment of 4s was extracted. The first 3s corresponded to the expected speech imagery. A period of 1s after that was further added in order to capture mental activity that the subject might still be performing after the visual cue disappeared. The total 4s segment was further divided into 3 epochs of 2s with a 1s overlap. For long words and the short-vs-long word comparison, the total time segment was 4.5s per trial. This interval was also divided into 3 epochs of 2s with a 1.25s overlap.

The number of CSP components was chosen empirically and it varied between 6 and 12 depending on the subjects and testing conditions. In most of the cases, the binary CSP method (CSP Method 1) was used to select the channels. The multi-class CSP approach (CSP Method 2) seemed to provide occasionally better results when applied on the comparison between a short and a long word. In all cases, multiple Gaussian kernels were used for the multi-class RVM classifier. The kernel parameters were typically set to $[0.1, 0.05, 0.025, 0.01, 0.005]$ with small adjustments for each subject.

When classifying short words, vowels, and long words, incorporating frequency information did not seem to provide any improvement, so that only the time domain information was used to compute the Covariance matrix (Method 1). In contrast, when classifying short v.s long words, using either one of the low level features pro-

vided quite similarly high accuracy. Combination of the two features by simply concatenating the tangent vectors from two covariance matrices (Method 2) significantly improved the results. This suggests that the difference in the complexity of the words could create discriminative features across frequency bands. To extract the second type of low level features, Morlet wavelet transform using the function *cwt* provided by Matlab 2016b was used. The number of octaves and number of voices for Morlet wavelet transform were set to 8 and 10 respectively, which yielded totally 80 ($= 8 \times 10$) scales. Only the scales in the range $[8, 40]$ were used to construct the low level feature vector, which in turn yielded the low level feature vector in the dimension of 35 ($=$ 33 subands + 1 raw signal + 1 channel index). The highest accuracy in this case was obtained by using the multi-class CSP algorithm with 6 CSP components.

The classification results of all cases are presented in Fig. 5.7 , which shows the mean, minimum and maximum values for each subject participating in each group. The two last groups represent the results of classifying a short v.s. a long word using the first set of low level features (Method 1), and the combination between the two kinds of features mentioned previously (Method 2). To compare the classification accuracy between groups of different number of classes, the Kappa κ value is also computed and is defined as:

$$\kappa = 1 - \frac{1 - P\%}{1 - C\%},$$

where $P\%$ is the prediction accuracy, and $C\%$ is the chance level, i.e. $C\% = 50\%$ for 2 classes. The mean, minimum and maximum κ values for each subject participating in each group are presented in Fig. 5.8. The averaged classification accuracies for all subjects in each group are reported in Table 5.1, which are well above chance level.

Table 5.1: The mean ± standard deviation for the accuracy and the corresponding Kappa Values for all subjects in each group of speech imagery.

Group	Participants	Accuracy (%)	κ
Vowels	S4,S5,S8,S9, S11,S12,S13,S15	48.6 ± 1.8	0.23
Short Words	S1,S3,S4,S5,S8,S12	51.1 ± 4.6	0.27
Long Words	S2,S3,S6,S7,S9,S11	67.1 ± 5.4	0.34
Short v.s. Long (Method 1)	S1,S5,S8,S9,S10,S14	71.9 ± 8.8	0.44
Short v.s. Long (Method 2)	S1,S5,S8,S9,S10,S14	80.1 ± 7.6	0.60

5.6 Conclusions

This chapter presented an investigation about the feasibility of speech imagery in the case of BCI applications. The experimental procedure was comprehensive and involved many different types of speech imagery, namely imagination of vowels, short words, long words and the comparison between a short and a long word. A classification method based on the computation of the covariance of the signals and the Relevance Vector Machines classifier was proposed and it was proven to be very successful achieving classification accuracies that in all cases were above the chance level. In addition, the classified data were proven to represent brain activity that was related to speech production and understanding (Section 5.3), while the subsequent analysis and the classification results for each of the examined cases provided useful insights about the proper use of speech imagery as user input to a BCI system. In particular, there was a strong indication that the performance of the system would greatly depend on the complexity of the used words rather than their meaning, which means that a successful application of speech imagery would include words that have different number of syllables with each other. As a result from the above, it can be

concluded that the mental task of imagined speech has great potential as a tool for BCI applications, while at the same time it can provide a viable alternative to the motor imagery paradigm used in the previously reported results (Chapter 4) as well as in other platforms found in the literature.

Chapter 6

CONCLUSIONS

Robotic swarms can be used in many different applications, while their inherent redundancy provides increased robustness against environmental disturbances and system failures. In addition, the incorporation of the human element in the control of robotic swarms could leverage the human intuition and cognition capabilities in order to find solutions for extreme cases where an automated algorithm might fail to respond in a timely manner. The main goals of this thesis were to investigate the feasibility of BCI as a method of interaction with a robotic swarm and to propose possible approaches and solutions for such a system.

At first, an investigation about the connection between observation of swarm collective behaviors and brain activity was performed and showed that such a connection might indeed exist. This result provided motivation for the next steps of this research. At the same time, it proposed a different type of mental imagery that was related directly to the control of swarm behaviors and could potentially lead to an intuitive solution for this problem. Next, as a preliminary step towards the overall goal, a hybrid BCI system that combined the input from the user's brain and the input from a game controller was proposed for the control of a small team of quadrotors. The system was extensively tested on multiple subjects as well as on an actual quadrotor platform and it was proved to be successful. The subjects were able to get acquainted with the system and achieve high levels of accuracy in a small amount of time and without any issues. This was very important considering that many systems require lengthy training procedures until the users reach an acceptable level of control. Finally, speech imagery was proposed as an alternative to the motor imagery paradigm used in many applications. This type of mental task could also provide a more intuitive approach for issuing commands to a swarm of agents using the brain, while, at

the same time, it could allow the concurrent use of other devices that would require hand motions.

Future research on this topic may include initially a more detailed investigation on the correlation between swarm behaviors and brain activity. This would lead to a better understanding of the mechanisms of human perception in a human-swarm interaction case and possibly use it to define a new type of paradigm that is related directly to control of swarms. Another direction could be the incorporation of speech imagery tasks to the mental task vocabulary that a user can employ for BCI application which would greatly increase the amount of DOF that could be controlled with respect to a robotic swarm. In addition, adaptive methods could be applied in order to expedite the training procedure and guarantee that the performance of the system would not deteriorate over time.

Overall, the BCI field has potential of providing intuitive solutions for all types of applications and the human-swarm interaction could be one of them. The present work shows that there is great potential in this endeavor as well as a plethora of solutions and options that someone can draw ideas from. Hopefully, this work will set the necessary stepping stones towards the realization of Brain-Swarm Interfaces (BSI) but also towards improvements of the existing HSI platforms.

REFERENCES

Amiri, S., R. Fazel-Rezai and V. Asadpour, "A review of hybrid brain-computer interface systems.", Advances in Human-Computer Interaction **2013** (2013). 8

Argunsah, A. O. and M. Cetin, "AR - PCA - HMM approach for sensorimotor task classification in EEG-based brain-computer interfaces.", Proceedings of the 20th International Conference on Pattern Recognition (ICPR). pp. 113–116 (2010). 40

Bashyal, S. and G. Venayagamoorthy, "Human swarm interaction for radiation source search and localization", in "IEEE Swarm Intelligence Symposium, 2008", pp. 1–8 (2008). 4

Bell, C. J., P. Shenoy, R. Chalodhorn and R. P. N. Rao, "Control of a humanoid robot by a noninvasive brain computer interface in humans", Journal of Neural Engneering **5**, 214–220 (2008). 7

Beni, G., *From Swarm Intelligence to Swarm Robotics*, pp. 1–9 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2005). 1

Birbaumer, N., A. Kubler, N. Ghanayim, T. Hinterberger, J. Perelmouter, J. Kaiser, I. Iversen, B. Kotchoubey, N. Neumann and H. Flor, "The thought translation device (TTD) for completely paralyzed patients.", IEEE Transactions of Rehabilitation Engineering **8**, 2, 190–193 (2000). 6

Brigham, K. and B. V. K. V. Kumar, "Imagined speech classification with EEG signals for silent communication: A preliminary investigation into synthetic telepathy", 2010 4th International Conference on Bioinformatics and Biomedical Engineering, iCBBE 2010 pp. 1–4 (2010). 10, 65

Chi, X., J. B. Hagedorn, D. Schoonover and M. D. Zmura, "EEG-Based Discrimination of Imagined Speech Phonemes", International Journal of Bioelectromagnetism **13**, 4, 201–206 (2011). 11

Contreras, S. and V. Sundararajan, "Visual Imagery Classification Using Shapelets of EEG Signals", ASME 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference pp. 1–6 (2016). 9

Coppin, G. and F. Legras, "Autonomy spectrum and performance perception issue in swarm supervisory control", in "Proceedings of the IEEE", vol. 100, pp. 590–603 (2012). 5

Couture-Beil, A., R. T. Vaughan and G. More, "Selecting and commanding individual robots in a vision-based multi-robot system.", in "2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)", (2010). 4

Crandall, J., M. Goodrich, J. Olsen, D.R. and C. Nielsen, "Validating human-robot interaction schemes in multitasking environments", IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans **35**, 4, 438–449 (2005). 5

Cummings, M. L., "Human supervisory control of swarming networks", in "2nd Annual Swarming: Autonomous Intelligent Networked Systems Conference", pp. 1–9 (2004). 5

D 'Zmura, M., S. Deng, T. Lappas, S. Thorpe and R. Srinivasan, "Toward EEG Sensing of Imagined Speech", International Conference on Human-Computer Interaction pp. 40–48 (2009). 10, 65

DaSalla, C. S., H. Kambara, M. Sato and Y. Koike, "Single-trial classification of vowel speech imagery using common spatial patterns", Neural Networks **22**, 9, 1334–1339 (2009a). 10, 11

DaSalla, C. S. C., H. Kambara, Y. Koike and M. Sato, "Spatial filtering and single-trial classification of EEG during vowel speech imagery", International Convention on Rehabilitation Engineering and Assistive Technology (ICREAT) **5**, 27 (2009b). 10

Delorme, A. and S. Makeig, "Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis", Journal of Neuroscience Methods **134**, 9–21 (2004). 18

Deng, S., R. Srinivasan, T. Lappas and M. D'Zmura, "EEG classification of imagined syllable rhythm using Hilbert spectrum methods", Journal of Neural Engineering **7**, 4, 046006 (2010). 10

Esfahani, E. T. and V. Sundararajan, "Classification of primitive shapes using brain-computer interfaces", CAD Computer Aided Design **44**, 10, 1011–1019 (2012). 9

Farwell, L. and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials.", Electroencephalography and Clinical Neurophysiology **70**, 6, 510–523 (1988). 6

Fisher, R. A., "The use of multiple measurements in taxonomic problems.", Annals of Eugenics **7**, 2, 179–188 (1936). 11

González-Castañeda, E. F., A. A. Torres-García, C. A. Reyes-García and L. Villaseñor-Pineda, "Sonification and textification: Proposing methods for classifying unspoken words from EEG signals", Biomedical Signal Processing and Control (2016). 12

Grosse-Wentrup, M. and M. Buss, "Multiclass Common Spatial Patterns and Information Theoretic Feature Extraction", IEEE Transactions on Biomedical Engineering **55**, 8, 1991–2000 (2008). 68, 71

Halkidi, M., Y. Batistakis and M. Vazirgiannis, "Cluster validity methods: Part I.", ACM SIGMOD Record **31**, 2, 40–45 (2002a). 56

Halkidi, M., Y. Batistakis and M. Vazirgiannis, "Cluster validity methods: Part II.", ACM SIGMOD Record **31**, 3, 19–27 (2002b). 56

Harriot, C. E., A. E. Seiffert, S. Hayes and J. A. Adams, "Biologically-inspired human-swarm interaction metrics", in "Proceedings of the Human Factors and Ergonomics Society Annual Meeting", pp. 1471–1475 (2014). 5

Hartigan, J. A., *Clustering algorithms.* (John Wiley & Sons Inc., 1975). 56

He, P., G. Wilson and C. Russel, "Removal of ocular artifacts from electro-encephalogram by adaptive filtering.", Medical and Biological Engineering and Computing **42**, 3, 407–412 (2004a). 37

He, P., G. Wilson, C. Russel and M. Gerschutz, "Removal of ocular artifacts from the EEG: a comparison between time-domain regression method and adaptive filtering method using simulated data.", Medical and Biological Engineering and Computing **45**, 5, 495–503 (2007). 37

He, P., G. Wilson and C. Russell, "Removal of ocular artifacts from electro-encephalogram by adaptive filtering", Medical and Biological Engineering and Computing **42**, 3, 407–412 (2004b). 66

Herff, C. and T. Schultz, "Automatic speech recognition from neural signals: A focused review", Frontiers in Neuroscience **10**, SEP, 1–7 (2016). 9

Herrmann, C. S., "Human EEG responses to 1100 hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena.", Experimental Brain Research **137**, 3–4, 346–353 (2001). 6

Holsheimer, J. and B. Feenstra, "Volume conduction and EEG measurements within the brain: a quantitative approach to the influence of electrical spread on the linear relationship of activity measured at different locations.", Electroencephalography and Clinical Neurophysiology **43**, 1, 52–58 (1977). 36

HORC ASU, "Formation control of robotic swarms using brain interface", URL `https://www.youtube.com/watch?v=BymnXeuSLcY` (2016). 59, 60

Huang, N. E. and S. S. P. Shen, *Hilbert-Huang transform and its applications.* (World Scientific, 2005). 10

Huang, N. E., Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis.", in "Proceedings of The Royal Society A Mathematical Physical and Engineering Sciences", pp. 903–995 (1998). 10

Huang, X. D., Y. Ariki and M. Jack, *Hidden Markov models for speech recognition.* (Edinburgh University Press, 1990). 40

Idrees, B. M. and O. Farooq, "Vowel classification using wavelet decomposition during speech imagery", in "2016 3rd International Conference on Signal Processing and Integrated Networks (SPIN)", pp. 636–640 (2016). 10

Karavas, G. K. and P. Artemiadis, "On the effect of swarm collective behavior on human perception: Towards brain-swarm interfaces.", in "IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)", pp. 172–177 (2015). ix, 14, 15, 16, 17, 19, 20, 21, 23, 24, 25, 26, 27, 28, 29

Karavas, G. K., D. T. Larsson and P. Artemiadis, "A hybrid brain-machine interface for control of robotic swarms: Preliminary results", in "2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (*to appear*)", (2017). ix, 8, 32, 34, 38, 39, 44, 46, 50, 51, 53, 54, 55, 56, 60, 62

Kim, J., S.-K. Lee and B. Lee, "EEG classification in a single-trial basis for vowel speech perception using multivariate empirical mode decomposition", Journal of Neural Engineering **11**, 3, 036010 (2014). 11, 68

Klem, G. H., H. O. Luders, H. H. Jasper and C. Elger, "The ten-twenty electrode system of the International Federation.", Recommendations for the Practice of Clinical Neurophysiology: Guidelines of the International Federation of Clinical Physiology **Suppl. 52**, 3–6 (1999). 17, 36, 37, 66

Kolling, A., K. Sycara, S. Nunnally and M. Lewis, "Human swarm interaction: An experimental study of two types of interaction with foraging swarms", Journal of Human-Robot Interaction **2**, 2, 103–128 (2013). 4

LaFleur, K., K. Cassady, A. Doud, K. Shades, E. Rogin and B. He, "Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain-computer interface.", Journal of Neural Engneering **10** (2013). 8, 33, 43

Larsson, D. T., *Dynamics, Modeling, Simulation and Control of Mid-Flight Coupling of Quadrotors*, Master's thesis, Arizona State University (2016). 49

Lebedev, M. A. and M. A. L. Nicolelis, "Brain-machine interfaces: past, present, future.", Trends in Neurosciences **29**, 9, 536–546 (2006). 6

Lee, H. and S. Choi, "PCA + HMM + SVM for EEG pattern classification.", Proceedings of the 7th International Symposium on Signal Processing and its Application. **1**, 541–544 (2003). 40

Leeb, R., H. Sagha, R. Chavarriaga and J. d. R. Millan, "A hybrid brain-computer interface based on the fusion of electroencephalographic and electromyographic activities.", Journal of Neural Engineering **8** (2011). 8

Leonard, N. E., D. A. Paley, R. E. Davis, D. M. Fratantoni, F. Lekien and F. Zhang, "Coordinated control of an underwater glider fleet in an adaptive ocean sampling field experiment in monterey bay", Journal of Field Robotics **27**, 6, 718–740 (2010). 1

Li, Y., J. Long, T. Yu, Z. Yu, C. Wang, H. Zhang and C. Guan, "An EEG-based BCI system for 2-D cursor control by combining Mu/Beta rhythm and P300 potential.", IEEE Transactions on Biomedical Engineering **57** (2010). 8

Lopez-Calderon, J. and S. J. Luck, "Erplab: An open-source toolbox for the analysis of event-related potentials", Frontiers in Human Neuroscience **213**, 8, 9–21 (2004). 18

Lotte, F., M. Congedo, A. Lecuyer, F. Lamarche and B. Arnaldi, "A review of classification algorithms for EEG-based braincomputer interfaces.", Journal of Neural Engineering **4**, 2, R1–R13 (2007). 7

Lotte, F., C. Guan, F. Lotte, C. Guan and S. Member, "Regularizing Common Spatial Patterns to Improve BCI Designs : Theory and Algorithms", IEEE Transactions on biomedical Engineering **58**, 2, 355–362 (2010). 67, 71

Ma, J., Y. Zhang, A. Cichocki and F. Matsuno, "A novel EOG/EEG hybrid humanmachine interface adopting eye movements and ERPs: Application to robot control.", IEEE Transactions on Biomedical Engineering **62** (2015). 8

Martin, S., P. Brunner, C. Holdgraf, H.-J. Heinze, N. E. Crone, J. Rieger, G. Schalk, R. T. Knight and B. N. Pasley, "Decoding spectrotemporal features of overt and covert speech from the human cortex.", Frontiers in neuroengineering **7**, May, 14 (2014). 68

Martin, S., P. Brunner, I. Iturrate, J. d. R. Millán, G. Schalk, R. T. Knight and B. N. Pasley, "Word pair classification during imagined speech using direct brain recordings", Scientific Reports **6**, 25803 (2016). 68

McFarland, D. J., L. M. McCane, S. V. David and J. R. Wolpaw, "Spatial filter selection for EEG-based communication.", Electroencephalography and Clinical Neurophysiology **103**, 3, 386–394 (1997). 36

McLachlan, G. J., *Finite Mixture Models.* (Wiley, 2000). 40

McLachlan, G. J., *Discriminant Analysis and statistical Pattern Recognition.* (Wiley Interscience, 2004). 11

McLurkin, J., J. Smith, J. Frankel, D. Sotkowitz, D. Blau and B. Schmidt, "Speaking swarmish: Human-robot interface design for large swarms of autonomous mobile robots", in "Proceedings of AAAI Spring Symposium", vol. 100 (2006). 2, 3

Mellinger, D., N. Michael and V. Kumar, "Trajectory generation and control for precise aggressive maneuvers with quadrotors.", The International Journal of Robotics Research **31**, 5, 664–674 (2012). 48

Mellinger, D., M. Shomin, N. Michael and V. Kumar, "Cooperative grasping and transport using multiple quadrotors", in "Proceedings of the 10th International Symposium on Distributed Autonomous Robotic Systems (DARS)", pp. 545–558 (2010). 1

Millan, J. d. R., F. Renkens, J. Mourino and W. Gerstner, "Noninvasive brain-actuated control of a mobile robot by human EEG", IEEE Transactions on Biomedical Engineering **51**, 1026–1033 (2004). 7

Millan, J. d. R., R. Rupp, G. R. Mller-Putz, R. Murray-Smith, C. Giugliemma, M. Tangermann, C. Vidaurre, F. Cincotti, A. Kubler, R. Leeb, C. Neuper, K.-R. Muller and D. Mattia, "Combining brain computer interfaces and assistive technologies: state-of-the-art and challenges.", Frontiers in Neuroscience **4** (2010). 8

Mohanchandra, K. and S. Saha, "A communication paradigm using subvocalized speech: translating brain signals into speech", Augmented Human Research **1**, 1, 3 (2016). 12

Monajjemi, V., J. Wawerla, R. Vaughan and G. Mori, "HRI in the Sky: Creating and commanding teams of UAVs with a vision-mediated gestural interface.", in "2013 IEEE/RSJ International Conference on Intellingent Robots and Systems (IROS)", pp. 617–623 (2013). 4

Muller-Putz, G. R., R. Scherer, G. Pfurtscheller and R. Rupp, "EEG-based neuroprosthesis control: A step towards clinical practice.", Neuroscience Letters **382**, 169–174 (2005). 8

Naghsh, A., J. Gancet, A. Tanoto and C. Roast, "Analysis and design of human-robot swarm interaction in firefighting", in "The 17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)", pp. 255–260 (2008). 2, 3

Nagi, J., A. Giusti, L. M. Gambardella and G. A. Di Caro, "Human-swarm interaction using spatial gestures.", in "Proceedings of the 27th IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014)", pp. 3834–3841 (2014). 3

Naseer, N. and K.-S. Hong, "fNIRS-based brain-computer interfaces: a review.", Frontier in Human Neuroscience **9**, 3, 1–15 (2015). 6

Neuper, C. and G. Pfurtscheller, "Event-related dynamics of cortical rhythms: frequency-specific features and functional correlates.", International Journal of Psychophysiology **43**, 41–58 (2001). 6, 8

Nunnally, S., P. Walker, M. Lewis, N. Chakraborty and K. Sycara, "Using haptic feedback in human robotic swarms interaction.", in "Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting", pp. 1047–1051 (2013). 4

Obermaier, B., C. Guger, C. Neuper and G. Pfurtscheller, "Hidden Markov models for online classification of single trial EEG data.", Pattern Recognition Letters **22**, 1299–1309 (2001). 40

Pfurtscheller, G., B. Z. Allison, C. Brunner, G. Bauernfeind, T. Solis-Escalante, R. Scherer, T. O. Zander, G. Mueller-Putz, C. Neuper and N. Birbaumer, "The hybrid BCI.", Frontiers in Neuroscience **4** (2010). 8

Pfurtscheller, G. and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles.", Clinical Neurophysiology **110**, 1842–1857 (1999). 8

Podevijn, G., R. O'Grady, Y. S. Nashed and M. Dorigo, "Gesturing at subswarms: Towards direct human control of robot swarms.", in "Proceedings of Towards Autonomous Robotic Systems-14th Annual Conference (TAROS 2013)", vol. 806, pp. 390–403 (2013). 3

Pollini, L., M. Niccolini, M. Rosselini and M. Innocenti, "Human-Swarm Interface for abstraction based control.", in "Proceedings of the 2009 AIAA Guidance, Navigation, and Control Conference", (2009). 4

Rabiner, L. and B.-H. Juang, *Fundamental of speech recognition.* (Prentice Hall Signal Processing Series, 1993). 40

Rabiner, L. R., "A tutorial on Hidden Markov Models and selected applications in speech recognition.", Proceedings of the IEEE **77**, 2, 257–286 (1989). 39, 40, 41

Riaz, A., S. Akhtar, S. Iftikhar, A. A. Khan and A. Salman, "Inter comparison of classification techniques for vowel speech imagery using EEG sensors", 2014 2nd International Conference on Systems and Informatics, ICSAI 2014 pp. 712–717 (2015). 11

Rousseeuw, P. J., "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis.", Journal of Computational and Applied Mathematics **20**, 53–65 (1987). 57

Royer, A. S., A. J. Doud, M. L. Rose and B. He, "EEG control of a virtual helicopter in 3-dimensional space using intelligent control strategies.", IEEE Transaction of Neural Systems and Rehabilitation Engineering **18**, 581–589 (2010). 9, 33, 43

Salama, M., L. Elsherif, H. Lashin and T. Gamal, "Recognition of Unspoken Words Using Electrode Electroencephalograhic Signals", The Sixth International Conference on Advanced Cognitive Technologies and Applications , c, 51–55 (2014). 11

Soares, J. M., A. Marjovi, J. Giezendanner, A. Kodiyan, A. P. Aguiar, A. M. Pascoal and A. Martinoli, "Towards 3-d distributed odor source localization: An extended graph-based formation control algorithm for plume tracking", in "Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)", pp. 1–8 (2016). 1

Stein, D. T., T. R. Schoen and D. Rus, "Constraint-aware coordinated construction of generic structures", in "IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)", pp. 1–9 (2011). 1

Stoica, A., T. Theodoridis and D. F. Barrero, "Towards human-friendly efficient control of multi-robot teams.", in "Proceedings of the 2013 International Conference on Collaboration Technologies and Systems (CTS)", pp. 226–231 (2013). 4

Suppes, P., B. Han and Z.-L. Lu, "Brain-wave recognition of sentences", Proceedings of the National Academy of Sciences **95**, 26, 15861–15866 (1998). 11

Suppes, P., Z. L. Lu and B. Han, "Brain wave recognition of words.", Proceedings of the National Academy of Sciences of the United States of America **94**, 26, 14965–9 (1997). 11

Walker, P., S. Nunnally, M. Lewis, A. Kolling, N. Chakraborty and K. Sycara, "Neglect benevolence in human control of swarms in the presence of latency", in "2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)", pp. 3009–3014 (2012). 5

Wang, L., X. Zhang, X. Zhong and Y. Zhang, "Analysis and classification of speech imagery EEG for BCI", Biomedical Signal Processing and Control **8**, 6, 901–908 (2013). 11

Wang, Z. and M. Schwager, "Multi-robot manipulation with no communication using only local measurements", in "IEEE 54th Annual Conference on Decision and Control (CDC)", pp. 1–8 (2015). 1

Wester, M. and T. Schultz, *Unspoken speech-speech recognition based on electroencephalography*, Ph.D. thesis, Universität Karlsruhe (TH), Karlsruhe, Germany (2006). 9

Wolpaw, J. R., N. Birbaumer, D. J. McFarland, G. Pfurtscheller and T. M. Vaughan, "Brain-computer interfaces for communications and control.", Clinical Neurophysiology **113**, 2002, 767–791 (2002). 6

Wolpaw, J. R. and D. J. McFarland, "Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans.", in "Proceedings of the National Academy of Sciences of the United States of America", vol. 101, pp. 17849–17854 (2004). 6

Wolpaw, J. R., W. A. Sarnacki and D. J. McFarland, "Electroencephalographic (EEG) control of three-dimensional movement.", Journal of Neural Engineering **7**, 3 (2010). 6

APPENDIX A

LIST OF PUBLICATIONS

The following publications have resulted from research related to this thesis. Portions of several of these papers have been used in the thesis and are correspondingly cited within the text.

Peer-Reviewed Conference Proceedings Articles

1. George K. Karavas, Daniel T. Larsson and Panagiotis Artemiadis, "A Hybrid Brain-Machine Interface for Control of Robotic Swarms: Preliminary Results.", IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2017 (*to appear*).

2. George K. Karavas and Panagiotis Artemiadis, "On the effect of swarm collective behavior on human perception: Towards brain-swarm interfaces.", IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), 2015.

APPENDIX B

COPYRIGHTED MATERIAL

In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of Arizona State University's products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to `http://www.ieee.org/publications_standards/publications/rights/rights_link.html` to learn how to obtain a License from RightsLink.

APPENDIX C

CO-AUTHOR PERMISSION

All co-authors have given permission for inclusion of material from published co-authored papers in this thesis.