

Towards Improvement of Human-Machine Interaction:

Design of Multimodal Human Intent Recognition System

Master's thesis in Biomedical Engineering

ASTA DANAUSKIENE, MAURICIO MACHADO

MASTER'S THESIS 2019: 94

Towards Improvement of Human - Machine Interaction

Design of Multimodal Human Intent Recognition System

ASTA DANAUSKIENE
MAURICIO MACHADO



CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Mechanics and Maritime Science
Division of Vehicle Safety
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2019

Towards Improvement of Human-Machine Interaction
Design of Multimodal Human Intent Recognition System
ASTA DANAUSKIENE, MAURICIO MACHADO

© ASTA DANAUSKIENE, MAURICIO MACHADO, 2019.

Supervisor: Pinar Boyraz Baykas, Department of Mechanics and Maritime Sciences
Examiner: Pinar Boyraz Baykas, Department of Mechanics and Maritime Sciences
Master's Thesis 2019: 94
Department of Mechanics and Maritime Science
Division of Vehicle Safety
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Experiment setup - test subject performing an experiment with a robotic manipulator

Typeset in L^AT_EX
Gothenburg, Sweden 2019

Towards Improvement of Human-Machine Interaction
Design of Multimodal Human Intent Recognition System
ASTA DANAUSKIENE, MAURICIO MACHADO
Department of Mechanics and Maritime Science
Chalmers University of Technology

Abstract

This master thesis focuses on investigating the electrical brain activity, eye gaze and pupil behaviour in the scope of goal-directed movement intention recognition for human-machine interaction applications. Previous studies support that the electroencephalography (EEG) data is suitable for early motion recognition and prediction and the pupil size changes correlate with the difficulty of the task. However few studies have looked into neural correlates of goal-directed and no-goal movements as well as the correlation between the pupil changes, EEG data and hand motion. We explore these questions through a set of cue-based movement experiments that include changing goal, repeating goal and no-goal scenarios and are performed in collaboration with a robot. The results were analysed with regard to movement related cortical potentials (MRCP) and event related spectral perturbation (ERSP) of EEG data, evoked pupil response, gaze patterns as well as binary goal\ no-goal classification of the data and correlation between different biosignals. Our results indicate that changing goal-directed movements are distinguishable from no-goal movements in EEG data in both temporal and time-frequency domains, when performing the task with a passive robot. Collaborative robot experiments showed great intersubject variability, therefore need to be further investigated. No correlation between evoked pupil response and MRCP was found in this study, however results suggest a correlation between MRCP and motion velocity profile.

Keywords: Human-Machine Interaction, Human-Robot Interaction, Human Intent Recognition, Goal-Directed Movement, Movement Prediction, Gaze Tracking, Pupilometry, BCI.

Acknowledgements

First and foremost, we would like to thank our supervisor, Pinar Boyraz Baykas at SAFER, Chalmers University of Technology, for her supervision and assistance regarding the creation of this master's thesis.

Furthermore, we would like to thank our colleagues Abdulrahman Alsaggaf, Alberto Nieto Sandino, Alexander Doan, Berglind Hlíðkvist Þorgeirsdóttir, Isabelle Montoya, Laura Guerrero, Ryan Thomas Sebastian, and Tryggvi Kaspersen for their invaluable support and unbiased insights into our work throughout the course of this thesis.

Asta Danauskiene, Mauricio Machado, Gothenburg, June 2019

Contents

List of Figures	x
List of Tables	1
1 Introduction	2
1.1 Background	2
1.2 Project aim	2
1.3 Limitations	3
2 Theory	4
2.1 Motor planning activity in the brain	4
2.2 Brain - Computer Interface	4
2.2.1 EEG data acquisition	5
2.2.2 EEG signal preprocessing	6
2.2.3 Feature extraction	7
2.2.4 Classification: Support Vector Machines	9
2.3 Gaze tracking	10
2.4 Pupillometry	12
3 Methods	13
3.1 Methods and equipment	13
3.1.1 EEG recordings	13
3.1.2 Gaze Tracking & Pupillometry	14
3.1.3 Synchronization	16
3.1.4 Robot control	17
3.2 Experiments	19
3.2.1 Participants	19
3.2.2 Experiment setup	19
3.2.3 Experiment session one	21
3.2.4 Experiment session two	21
4 Results	22
4.1 Behavioural analysis	22
4.2 EEG result analysis	24
4.2.1 Temporal analysis	24
4.2.2 Time-frequency analysis	27
4.3 Gaze tracking & Pupillometry	28

4.3.1	Gaze Tracking	28
4.3.2	Pupillometry	34
4.4	Classification	34
4.5	Correlation	36
4.6	Discussion	36
4.6.1	EEG	36
4.6.2	Gaze tracking & Pupillometry	38
4.6.3	Classification	39
4.6.4	Correlation	39
5	Conclusion	40
5.1	Summary of results	40
5.2	Future work	41
	Bibliography	42
A	Appendix: Subject specific EEG data	I
B	Appendix: Pseudo code for robot control algorithm	V

List of Figures

2.1	International 10-20 electrode placement system.	6
2.2	Binary SVM classifier	9
2.3	Pupil detection methods	11
3.1	EEG setup	14
3.2	Gaze tracking headset configuration	15
3.3	World camera view of gaze tracking headset	16
3.4	Lab Streaming Layer architecture	17
3.5	Block diagram of control algorithm.	18
3.6	Experiment setup	19
3.7	Experiment tasks cycles	20
3.8	Example task: changing or repeating and free movement	20
4.1	Reaction times and movement duration: passive robot	22
4.2	Reaction times and movement duration: active robot	23
4.3	Velocity profiles: passive robot	23
4.4	Velocity profiles: active robot	24
4.5	Grand average participant MRCPs: passive robot	25
4.6	MRCP image plots with corresponding ICs	25
4.7	MRCPs of experiment set two: collaborative robot	26
4.8	Average participant ERSPs: passive robot	27
4.9	ERSPs of experiment set two: active robot	28
4.10	Gaze heat maps of test subject S3	29
4.11	Subject S3: Gaze behavior through time for the changing task with passive robot	31
4.12	Subject S3: Gaze behavior through time for the repeating task with passive robot	32
4.13	Subject S3: Gaze behavior through time for the random task with passive robot	33
4.14	Baseline mean pupil sizes for all tasks: passive robot	34
4.15	Baseline mean pupil sizes for all tasks: active robot	35
4.16	Classification confusion matrices	35
4.17	Eye data classification	36
4.18	Temporal correlation of subject S3 MRCP, pupil diameter and veloc- ity profiles	37
A.1	MRCPs of test subject S2: passive robot	I

A.2	MRCPs of test subject S1: passive robot	II
A.3	MRCPs of test subject S3: passive robot	III
A.4	ERSP images of test subjects S1 and S3: passive robot	IV

List of Abbreviations

BCI	Brain-Computer Interface
BP	Bereitschaftspotential (Readiness Potential)
ECG	Electrocardiogram
EEG	Electroencephalography
EMG	Electromyography
EOG	Electrooculography
ERP	Event Related Potential
ERSP	Event Related Spectral Perturbation
fMRI	Functional Magnetic Resonance Imaging
HEOG	Horizontal EOG
HITL	Human-In-The-loop
HMI	Human-Machine Interaction
IC	Independent Component
ICA	Independent Component Analysis
IR	Infra-Red
LC	Locus Ceruleus
LSL	Lab Streaming Layer
MD	Motion Duration
MMP	Motor Monitoring Potential
MP	Motor Potential
MRCP	Motor Related Cortical Potential
RT	Reaction Time
SCARA	Selective Compliance Assembly Robot Arm
SNR	Signal-to-Noise Ratio
SVM	Support Vector Machine
VEOG	Vertical EOG

1 Introduction

1.1 Background

Modern robotic systems are able to perform a myriad of tasks better than humans, but there are cases, mostly those requiring more dexterity and complex decision-making, where humans still perform better. Therefore by implementing a human-in-the-loop (HITL) control strategy, it should be possible for a mechatronic system to improve its performance by actively recognizing human intention. The idea is not new and there are collaborative robots (cobots) that are already working together with humans, however the technology still has some noteworthy limitations such as safety concerns and consequent restrictions on robot performance and level of integration [24]. Cobots are designed to prevent inflicting harm to humans, however these machines are not inherently safe. This means that safety needs to be assessed for every installation case separately and often a robot is working with restricted power, force and speed or even completely stopping when human approaches collaborative workspace. Although it assures the safety, it also prevents the human-robot system from working in its full potential. Development of better HITL algorithms is a next step towards a more seamless, intuitive collaboration. Such synergy, if achieved, can have many applications such as in assisted and automated driving, assembly lines where more intricate steps are needed or exoskeletons for augmentation of humans with motion impairment.

1.2 Project aim

The aim of this Master's project is to design a human sensory-motor intent recognition system to improve robot and human collaboration when performing a shared control task. Through the implementation of such a system on Human-Machine Interaction (HMI) application, we expect to see an efficient and safe accomplishment of a given task, pushing further for human and robots working close together. Previous studies in this field have mostly focused on intent recognition as motion prediction based on kinematics and dynamics of a human collaborator. Meanwhile other ones, especially in the field of prosthetics and rehabilitation, have investigated human cognitive processes and bio-signals, such as electrical muscle activity through electromyography (EMG) and electrical brain activity through electroencephalography (EEG) to better understand motion volition. The authors of this report suggest that a combined approach could be beneficial for improvement of HMI applications.

This thesis work will be focused on the following hypotheses:

- The human motor intent can be predicted based on the EEG and eye tracking measurements. Furthermore, the prediction can be supported by motion tracking data.
- There exists a correlation between eye movements and sensory-motor human intent related changes in EEG signals.
- Implementing the multimodal human intent recognition algorithm as part of the robot control system will improve the performance of the task as compared to a single mode HITL system.

Besides investigation of these hypotheses and testing them through a set of carefully designed experiments, this project also intends to explore the following research questions:

- Is the correlation between EEG and eye-gaze tracking sufficient enough to be able to find "surrogate" measures from gaze-tracking data to replace by-nature-very-intrusive EEG signals for future applications?
- Is it possible to fuse EEG and eye-tracking data to obtain a combined cognitive indicator that could be used to identify different pre-movement stages such as onset decision making, anticipatory behaviour or intention?

1.3 Limitations

Due to time restrictions, the scope of this project will be limited to a narrow part of recognition of human intent, specifically, the prediction of human sensory-motor behaviour.

The experiments will be conducted in a controlled environment where human will have to perform a defined task together with a 2 degree of freedom robotic arm. The number of the test subjects is restricted to 3, which is sufficient for a proof of concept, however a larger study involving more human subjects would be needed for validation of the findings of this project.

No online processing of EEG signals will be performed in this study and no comparison of different EEG processing and classification techniques will be made. Pupil measurements were made through an off-the-shelf solution comprising both hardware and software. Therefore, it is not within the scope of this work to produce the mathematical background for how the data is defined, e.g. pupil size estimation or gaze calculation. When necessary, a short theoretical description will be provided to support the work.

2 Theory

2.1 Motor planning activity in the brain

There have been a number of studies on motor activity in the brain and possibility to employ it with Brain Computer Interface (BCI). Many studies focus recognizing subject's intention by analyzing motor imagery as a tool for rehabilitation and functionality restoration [26, 46]. The recognized intent can be coded into commands in order to augment or assist a human user in performing the intended or predefined task. Similarly, the same technology could be applied to work with collaborative robots. However, the task is not trivial.

It has been proven that significant changes in both frequency and time domains of brain activity signal occur in motor cortex (C3, Cz, C4 locations) when performing a motor task [35, 39]. However, studies have shown that the brain activity changes with the task. Different activation patterns were found in the functional magnetic resonance imaging (fMRI) signal when given shared motor imagery task compared to a single subject motor imagery task. Single action primarily activates the middle occipital lobe, the cerebellum and the precuneus, meanwhile joint action shows significant activation in the bilateral inferior frontal gyrus and the middle frontal gyrus [44]. Fluctuations in attention have also been proven detectable during motor task. These differences occur most prominently when analysing time-frequency signals of motor cortex channels [7]. Moreover, the differences in the presence of the goal is suggested to have an effect on the brain activity. When considering goal-directed movement and no goal movement, there are noticeable changes in premotor, primary motor areas and the posterior parietal cortex and the average amplitude of the brain waves over the central electrodes [31]. This suggests areas for improvements in movement classification for BCI applications.

2.2 Brain - Computer Interface

Brain-Computer Interface (BCI) is a system for direct communication between the brain and external environment. There are three essential parts to the system:

- **Signal acquisition** through sensors placed on the scalp (non-invasive BCI) or brain directly (invasive BCI). In this thesis electroencephalography (EEG) is used to acquire the signals.

- **Signal preprocessing** to amplify the signal and improve signal to noise ratio (SNR). This step also includes filtering to remove DC bias and high frequency noise as well as removal of artifacts.
- **Extraction of meaningful, interpretable features** that can be translated into commands for a further use with various devices. The commands are most often obtained through classification of the features.

One challenge for BCI is its implementation in real-time applications. Even though online processing and classification of the data are computationally challenging, there are other, more challenging factors that make classification task complicated. Firstly, a lot of research mentioned above is conducted in a synchronized manner- the experiments are conducted in a well defined environment, performing a single movement at a time that is synchronized to some onset. In reality, a person executes movements continuously, making the classification less accurate. A recent study showed that a multi-modal approach using electrooculography (EOG), EMG and EEG can significantly improve classification results, thus improving the BCI performance [47].

2.2.1 EEG data acquisition

EEG is a method for monitoring and recording electrical brain activity. This method has many applications in medical diagnostics and research. One of many well studied research applications for EEG is sensory - motor tasks and its potential use for HMI through BCI.

The electrical activity in the brain is created by neuronal communication. A single neuron activity is too weak to be detected using non-invasive EEG methods, hence the signals picked up by electrodes on the scalp are cumulative electric field potentials generated by numerous neurons. The differences in these fields create characteristic representations of brain activity [10]. The signal is typically recorded using non-invasive electrodes placed at fixed locations on the scalp. The standard locations are defined by international 10-20 system, where the distance between adjacent electrodes is either 10% or 20% of the entire distance from nasion toinion and from one preauricular point to the other, see figure 2.1. In studies requiring more spatial resolution, additional electrode locations in between the ones presented in the figure 2.1 are used.

A typical EEG signal is measured in range of microvolts and can easily be masked by the noise that may be as high as 100 times the value of the EEG signal, making it challenging to filter the noise without removing the main signal characteristics. The artifacts may come from different sources such as internal noise due to other processes in the brain (e.g. activity in alpha (8 - 15 Hz) frequency band), biological sources (ocular, cardiac or muscle activity), noise from external devices (e.g. 50 Hz power line noise), noise due to movement or electrode artifacts due to high impedance between the skin and electrodes. To reduce the latter, the choice of electrodes is important.

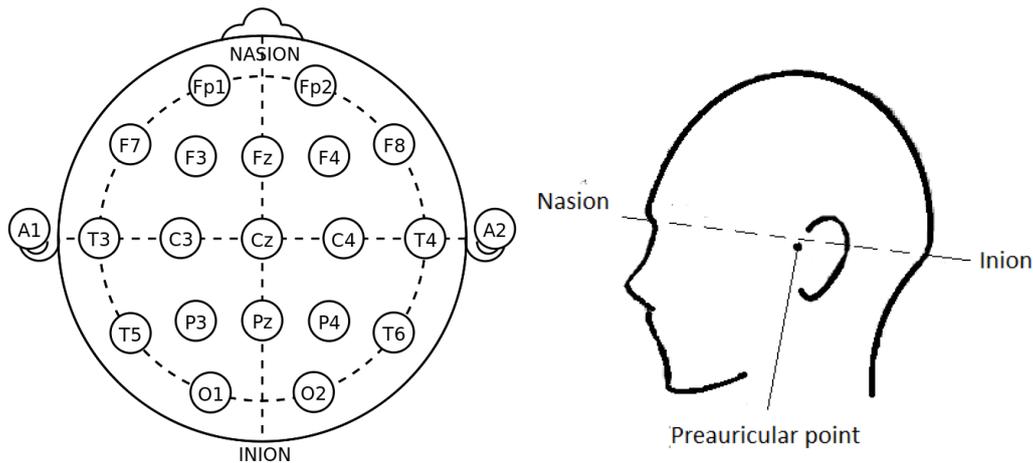


Figure 2.1: International 10-20 electrode placement system.

There are many different types of electrodes and different ways of classification. The electrode categories can be loosely separated into:

- **Surface electrodes or implantable electrodes.** Surface electrodes are placed on the skin and are non-intrusive, however the SNR is often poor due to natural skin oils, hair or dead skin cells that reduce skin conductance. Implantable electrodes give a much cleaner signal than surface electrodes, however are highly intrusive and therefore are rarely used.
- **Active or passive electrodes.** Active electrodes have an in-built preamplifier which significantly increases the quality of the signal and allows the electrodes to be used without special skin preparation. Passive electrodes require special skin preparation: cleaning, abrasion and most often a use of conductive paste. However, passive electrodes are much cheaper.
- **Dry or gel electrodes.** Dry electrodes are equipped to work without skin preparation or conductive paste and often have a brush like shape to bypass the hair. Gel electrodes need a considerable amount of gel to ensure good impedance, which is not preferred by users for short time use as the paste leaves residuals in the hair.

2.2.2 EEG signal preprocessing

As discussed in the previous section, raw EEG data contains a considerable amount of artifacts and noise, hence preprocessing is needed before features can be computed. There are many paradigms and tools for EEG preprocessing. Here, only the steps used in this thesis are presented:

- **Referencing.** In EEG recordings, the voltages recorded in one electrode are relative to other electrodes, reference electrode or a calculated average. Often reference electrodes are chosen on the mastoid bone, earlobes or one of the central electrodes (e.g. Cz electrode) as these locations are close to the measurement area and pick up the same noise as all other channels.
- **Uploading the data.** Before preprocessing can be started, raw EEG data needs to be uploaded to software that will be used for preprocessing. Channel

locations in three dimensional space and events (e.g. cues for movement) should be imported together with the data in order to visualise the scalp maps and estimate source locations for data components.

- **Filtering.** Generally the signal should be filtered with a notch filter to remove the line noise and a high-pass filter to remove the low frequency drift. The cut-off frequency of the high-pass filter depends on the application and can be anywhere from 0.1 Hz up to 1 Hz, a higher limit is not recommended since it starts to interfere with the data. Additionally, a low-pass filter can be applied to attenuate the high frequency noise above 40 - 50 Hz. However, several studies have shown that both low-pass [38,41] and high-pass [34] filtering can significantly alter the temporal structure of the data, therefore, if possible, filtering should be avoided.
- **Artifact and trial rejection.** Some of these artifacts are very strong and not systematic, hence they cannot be removed by filters and have to be removed manually by inspection. It can include noisy channels or time windows when the noise appears in the majority of the channels and sometimes the whole epoch will have to be removed. If some channels need to be removed, then interpolation of removed channels needs to be performed. The purpose of interpolation is to avoid the bias when calculating the average reference (explained in the next step). Other types of noise that is systematic, e.g. heartbeat or eye blinks, can be filtered out using simultaneous and cross-checking filtering techniques employing both time and frequency domains.
- **Division of data into epochs.** When EEG is recorded, the data is represented as a continuous time series for the whole recording time. However, it is often useful to divide it into time windows that are locked to a specific event, e.g. the time point when stimulus was applied.
- **Baseline removal** is used to remove the mean of the recorded baseline for each electrode.
- **Independent Component Analysis (ICA)** is a method to separate Independent Components (IC) of linearly mixed signals in different sensors. Since EEG recorded in one channel is a mixture of all neuron potentials in an area as well as other biological signals, the recording between electrodes can be highly correlated (an effect known as cross-talk) and ICA is a tool to separate these signals. If the EOG or electrocardiogram (ECG) was recorded in a separate channel, ICA method can be used to remove the EOG or ECG artifacts from the data without losing the useful information as well as separate the distinct otherwise mixed EEG signals in different channels [11]. It is important to have clean data before running ICA, filtering with 1 Hz high-pass filter before running ICA is recommended in order to get good results [43].

2.2.3 Feature extraction

Preprocessed data can already provide some information to interpret by observation; however, due to its high dimensionality and variability it is unlikely to produce good results in classification. For this reason, feature vectors are used to extract the meaningful information that can easily be interpreted by a classifier, hence good

selection of features can give good results even with a simple classifier.

There are some common properties that need to be taken into account when selecting the features and a classifier. Lotte et al. [27] points the following feature properties:

- **noise and outliers:** due to poor SNR, EEG signals are contaminated with a lot of noise;
- **high dimensionality:** feature vectors often contain information from different features that in themselves contain information from several time windows, frequency spectra, etc.;
- **time information:** the dynamic properties of EEG signal in time domain hold important information about cognitive processes and need to be accounted for;
- **non-stationarity:** EEG signals change over subjects and even over time in the same subject;
- **small training sets:** the training sets in EEG are usually small because the data collection process is time consuming.

There is a great variety of techniques used for feature extraction that study EEG signals in both time and frequency domains. For the purpose of this study, we will only use two types of features, namely, Event Related Potentials (ERP) and Event Related Spectral Perturbations(ERSP).

ERP is a low frequency electrophysiological brain response to a stimulus. ERP is one of the classical methods of feature extraction and is well studied in the literature [28]. ERPs are acquired by averaging a time locked signal from multiple trials of the same task. It is defined by slow positive and negative deflections in the amplitude of a signal in the time domain. When associated with a motor task, ERPs are referred to as Motor Related Cortical Potentials (MRCPs). MRCPs are characterized by three components: Readiness Potential (Bereitschaftspotential or BP) - a slow negative deflection in amplitude that starts at around 1.5 seconds before movement, reaching its peak negativity at the time of an onset where Motor Potential (MP) occurs and followed by a Movement Monitoring Potential (MMP) - an upturn before returning to the baseline level [30]. BP is associated with intentional involvement in the task [18] and the peak negativity is an important parameter determining such involvement. One study investigated MRCPs in context of presence of a goal and suggested that there is a statistically significant difference between a goal movement versus no-goal movement when performing a classification task [31]. The same study found that maximal negativity over central electrodes and the reafferent potential after onset were more pronounced in the goal directed task rather than a no-goal task.

ERP is unable to capture all brain dynamics as it does not hold information about event related changes in the frequency domain, which are important since evoked responses are not stable across trials and are dependent on ongoing EEG activity. ERSP was introduced as a measure to improve ERP. ERSP measures average dynamic event related changes in EEG power spectrum as a function of time relative to an experimental event [29]. ERSPs are computed by averaging the amplitude

spectra on a sliding time window and normalizing to mean baseline for each epoch. After that, the results are averaged over many trials and plotted in a time - frequency domain.

2.2.4 Classification: Support Vector Machines

Many classification algorithms have been used in EEG applications over the years. In this work, we will not discuss the differences between algorithms, but introduce an algorithm that was used in this study for binary goal\no-goal movement classification of EEG data epochs. A review on several algorithms [27] suggests that Support Vector Machines (SVM) is a reliable and robust option because of regularization and simplicity. Moreover, SVM is effective when it is applied on small datasets, which is the case in this study. To preserve the continuity, the same classifier will be used for eye gaze and pupillometry data classification.

SVM is a discriminative classifier that separates the data using hyperplane as a decision boundary. In a two dimensional binary classifier, the hyperplane is a line separating the data at each side, with least possible amount of outliers 2.2. The position and orientation of a hyperplane is determined by data points called support vectors that are located nearest to decision boundary. The hyperplane is selected to maximize the margin between two classes and minimize the classification error [9].

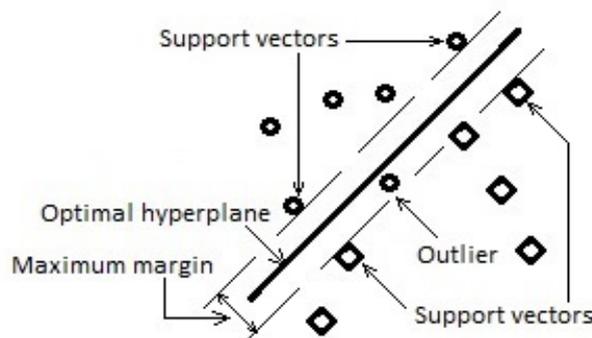


Figure 2.2: An illustration of a two dimensional binary SVM classifier and its main components

Since SVM is a supervised machine learning algorithm, it needs labelled data to be trained and a separate dataset for validation. The training data is a set of n-dimensional vectors $x_j \in R^n$ and classes $y_j = \pm 1$. The hyperplane is defined by equation:

$$f(x) = x^t\omega + b = 0 \quad (2.1)$$

where $\omega \in R^n$ and b is a real number.

Essentially, SVM aims to find the best separating plane by finding ω and b such that $\|\omega\|$ is minimum and $y_j f(x_j) \geq 1$ for all data points (x_j, y_j) . If $y_j f(x_j) = 1$, the

vector is a support vector. For mathematical convenience, the problem is substituted to minimizing $\|\omega\|$. The optimal solution $(\hat{\omega}, \hat{b})$ allows to classify the vector z :

$$\text{class}(z) = \text{sign}(z^t \hat{\omega} + \hat{b}) = \text{sign}(\hat{f}(z)). \quad (2.2)$$

$\hat{f}(z)$ is the classification score and z is a vector distance from the hyperplane. In case when the data is not separable, SVM can be used with a soft margin that separates most of the points, but allows some outliers. The formulation of the soft margin may differ, but generally it is a similar minimization problem that aims to minimize ω with added penalty parameter C and slack variable ξ :

$$\begin{aligned} \min_{\omega, b, \xi} \quad & \frac{1}{2} \omega^t \omega + C \sum_j \xi_j \\ \text{s.t.} \quad & y_j f(x_j) \geq 1 - \xi_j \\ & \xi \geq 0 \end{aligned} \quad (2.3)$$

Penalty parameter and slack variable define how soft the decision boundary is – increasing C value puts more weight on ξ , making the boundary less lenient [17], [4]. In cases when classes cannot be separated by a linear hyperplane, SVM hyperplane linear function can be replaced with a non-linear kernel function, such as polynomial, Gaussian or radial basis function.

2.3 Gaze tracking

It has been studied that eye tracking can be of support for assessing brain function and attention [14]. Moreover, eye tracking can be further divided into two main components, gaze tracking and pupillometry. Since the pioneering research made by Yarbus [45], which linked gaze fixations and thinking, many others followed, further exploring this concept. With the recent improvements in hardware and algorithms for eye tracking, a new window of possibilities was open for this method. In addition, higher sampling rates used now (between 25 to 2000Hz) enabled increased resolution for the measurements, thus improving data quality.

The importance of gaze tracking for this project lies on how it can relate to human attention and intention. Several studies focused on that relationship, in one of them it was investigated, with the use of heat-maps, how human attention changes when a person is looking for something specific in a scene or just scanning through it. The results showed a less scattered and highly focused gaze fixations for the specific task, labeled as "informational intention" [5].

A study assessing steering performance in drivers and their gaze behavior showed strong correlation between them. In that case, the horizontal gaze angle and the steering angle varied in the same proportion with the gaze changing earlier than the steering, therefore giving a previous indication about the action [40]. Another research investigated the relation between gaze behaviour and way finding for subjects navigating in a virtual maze. Again it was found a strong relation between

gaze fixation patterns and spatial decision making, with gaze marks peaking at the chosen path close to a second before the actual decision [42]. Finally, it was theorized in [25] that gaze tracking can be of use when determining attention from a human when interacting with a robot.

Gaze tracking is a method that collects metrics from the eye through time such as, eye angles and pupil position, in its own reference frame, and translates them into the scene world. A camera catches the light reflections from the eye by a light source and an algorithm process the image and calculates the gaze parameters, outputting that as a focus point into the 2D space. It is mainly used to asses attention from a person, through fixations and saccades, while executing a given task.

Eye detection can be achieved in two ways: bright pupil or dark pupil method. Both of them use the same principle described earlier with a light source and a camera. The main difference is that for first one, the camera and the light source are close together and have their axes parallel to each other (on-axis set-up). Because of that, the light hits the retina and bounces straight back to the source, where the camera also is. By getting the reflection straight from the retina, the pupil becomes bright (the same can be observed in camera shots when people get red eyes). The second method places the camera a bit far from the light source, in a non coincident axis (off-axis set-up) so the camera cannot get the reflections from the retina, but only from its surface, therefore showing the pupil as black. Both methods are show in figure2.3. The light source used is normally infra-red (IR) because it falls out of the visible spectrum, therefore preventing distraction or discomfort and also for being harmless to the human eye [16]. For our solution the full algorithm is described in [23].

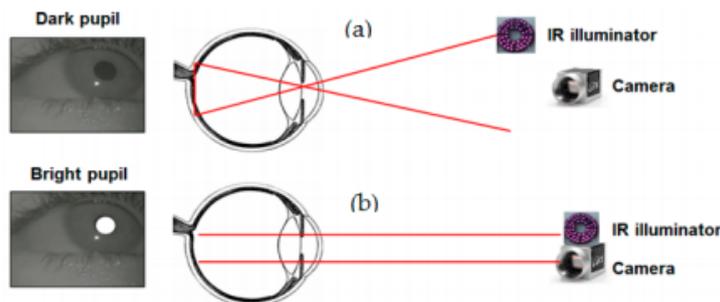


Figure 2.3: Two methods of pupil detection. a) Dark Pupil, b) Bright Pupil [13]

After detecting the pupil, the next step is to determine the gaze position. The literature refers to many different ways to approach this problem [16]. For this work, the gaze mapping was provided from the software used together with the selected hardware and therefore the authors did not take part in that step. The manufacturer states that the mapping "is implemented with a transfer function consisting of two bi-variate polynomials of adjustable degree. The user specific polynomial parameters are obtained by running one of the calibration routines" [23].

2.4 Pupillometry

Pupillometry is the study of the changes in pupil size by the means of cognitive process. Here, changes related to the eye's natural accommodation to light are discarded since they are regarded as a mere reflexive response. Pupil size is controlled by two types of nerve fibers: sympathetic and parasympathetic. Sympathetic fibers respond to the neurotransmitter norepinephrine, which has an effect of relaxing the iris therefore dilating the pupil, while the parasympathetic fibers act the other way with the neurotransmitter acetylcholine, constricting it [14].

The brain structure mainly responsible for noradrenergic pathways that influences pupil dilation is the Locus Ceruleus (LC). Located in the brain stem, the LC is a small structure comprising a bundle of neurons that have a direct relation in regulating "arousal and cognitive functioning". In short, when LC is more active, due to increased cognitive load for example, it inhibits the parasympathetic centre while increasing the sympathetic activity too, which will make the pupil dilate [14], [21].

Studies with monkeys supported this pupil/LC relation by monitoring the activity in the LC and pupillary responses while the subjects performed different types of cognitive tasks. In those, it was observed that the pattern of activation in the LC was replicated in the pupil in the form of changes in its size, [33], [20].

In humans, minimum and maximum pupil range can vary from 1.5 to 9mm and for people in rest condition and in a environment with "standard light conditions", the pupil has a 3mm diameter in average [36]. Changes in pupil size due to cognitive load can be at most of 0.5mm but those changes can also get highly impacted from a person's condition, for example, fatigue can raise the pupils diameter baseline while introducing fluctuations [8]. With that in mind, it is important to assure constant illumination conditions in the test environment while also checking if the subjects are well rested before starting the trials.

3 Methods

3.1 Methods and equipment

3.1.1 EEG recordings

The EEG was recorded with an open source BCI device at 125Hz sampling rate in 11 EEG channels and two EOG channels - vertical (VEOG) and horizontal (HEOG). OpenBCI Cyton board with v3 Daisy Module is used with dry passive EEG electrodes mounted on an electrode cap (see figure 3.1). The electrode impedances were kept below $5k\Omega$. Electrode locations are based on international 10-20 electrode placement system 2.1 and are covering inferior frontal cortex (F7, F8), supplementary motor area and premotor cortex (F3, Fz, F4), motor cortex (C3, Cz, C4) and superior parietal lobule (P3, Pz, P4) as these regions have been proven to be the largest contributors to motor action planning and movement onset, as discussed in chapter 2.1. Four electrodes were located above and below the center of left eye (VEOG) as well as on the sides of both eyes (HEOG), just below temples. These electrodes record the eye movement related muscle activity in order to filter out EOG artifacts from other channels in the preprocessing step. The ground and reference electrodes were placed on right and left earlobes respectively.

Offline processing of the data was performed using MATLAB R2018B (The MathWorks, Massachusetts, USA) with an open source EEGLAB toolbox [12]. The data was bandpass filtered with zero-phase non-causal finite impulse response filter with low cut-off frequency at 1 and high cut off frequency at 45 Hz. The baseline was removed, ICA was performed using an infomax ICA algorithm and then the data was divided into epochs [-2 3] seconds from the "Go" cue. ICA weights were imported to the dataset filtered between at 0.1 Hz low cut-off and 45 Hz high cut-off frequencies and independent components (IC) corresponding to the EOG artifacts or other muscle activity were removed. Then, MRCP and ERSP were calculated.

Binary goal movement vs. no-goal movement classification of EEG data was made based on temporal and time-frequency features that were chosen based methods and findings of previous similar studies [6, 31]. Temporal features include minimum and maximum MRCP values as well as the slopes in time windows [-2 -1], [-1 0], [0 0.5], [0.5 1] and [1 2] seconds as in. Features in time-frequency domain were mean ERSP values in Delta (0–4 Hz), Theta (4–8 Hz), Alpha (8–15 Hz), Beta (15–30 Hz) and Gamma (30–60 Hz) frequency bands for time windows at [-1 -0.5], [-0.5 0], [0 0.5]

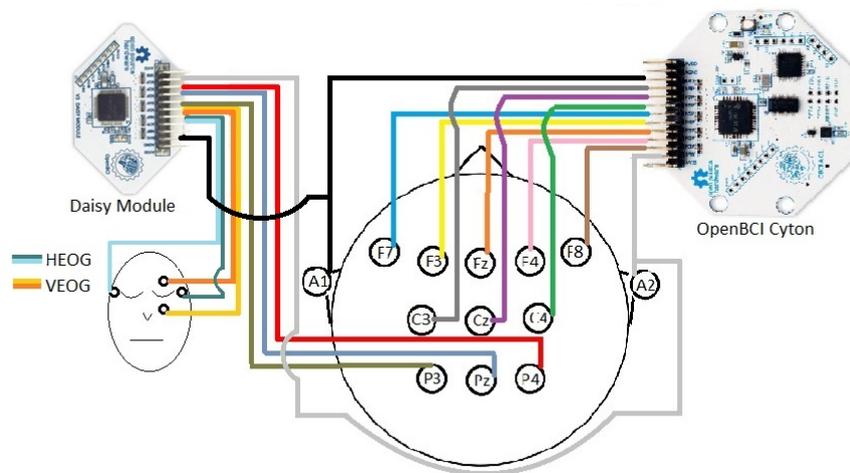


Figure 3.1: EEG setup. Every electrode is connected to a separate channel on either OpenBCI Cyton board or Daisy module and referenced to the A2 channel. Both boards share the ground and reference pins, both, horizontal (HEOG) and vertical (VEOG), EOG channels are connected such that both electrodes of one channel are referenced to one another.

and [0.5 1] seconds. The features were divided into the training and test datasets. Training and cross-validation of the SVM was performed on the training dataset using MATLAB function *fitcsvm()* using the hyperparameter optimization option to find the best classifier based on Sequential Minimal Optimization [15].

3.1.2 Gaze Tracking & Pupillometry

For implementing the gaze tracking we picked a solution from Pupil Labs, which consists of a 3D printed headset with two embedded cameras, figure 3.2, one of them, the world camera, always faces forward and displays the perspective from the user on the computer screen. The second camera faces the user's eye with an off-axis IR LED of 860nm of wavelength to illuminate it and feeds that image into the software. Then, an open-source algorithm segments the image using dark pupil method, monitors the pupil size and tracks its position (gaze data) [23]. A quality measure is also provided, ranging from 0 to 1, and indicates the confidence that that circle found is indeed the pupil. The information from both cameras is finally used to map and overlay the gaze point onto the 2D scene view feed as a red dot which moves according to the user's eye movement, therefore providing us with real time eye tracking, figure 3.3. There are also settings available on the menus to improve the eye image capture quality that may need adjustments due to lighting conditions of the environment, such as absolute exposure time, thus making pupil detection more accurate and less noisy.

The software comprises two applications. One is for real time data collection, named Pupil Capture, with the aforementioned configurations together with other plugins that may be of interest, such as surface tracker, blink detection, fixation detector and others. Here, it is important to highlight its calibration process. In order to



Figure 3.2: Headset configuration used for gaze tracking [2]

proper map the gaze, seen from the eye camera, onto the 2D scene, viewed from the world camera, a series of at least 5 markers are shown on the screen when activating the function. The user look each marker and the algorithm asserts how the gaze position translates onto the 2D scene, also the estimated distance from the user to the screen plane should be entered. Other important feature for the experiments is the record button, which enables us to store all that data acquired during the runs and use it for post processing in offline mode.

In order to obtain the gaze pattern for the subjects for every trial, Pupil Capture provides a feature where it detects and stores fixation parameters defined by the user. Fixations occur when an eye position remains fixed in a point, with minor deviations, for a minimum period of time, between 200 and 400ms [19] whereas saccades are normally much shorter than that and relate as transitions between fixations [14]. Therefore we select the middle of this interval, 300ms, as our fixation time threshold. In addition, since eye movement is controlled by muscles which have variability, dispersion in eye movement may happen, therefore a limit of 3 degrees of dispersion for fixations is set. Lastly, to ensure better data quality, only fixation data with confidence above 0.75 is taken.

The other application, called Pupil Player, loads the recorded sessions and provides all the recorded video and gaze data along with other plugins for visualizing the data, such as the Offline Fixation Detector and the Surface Tracker, apart from

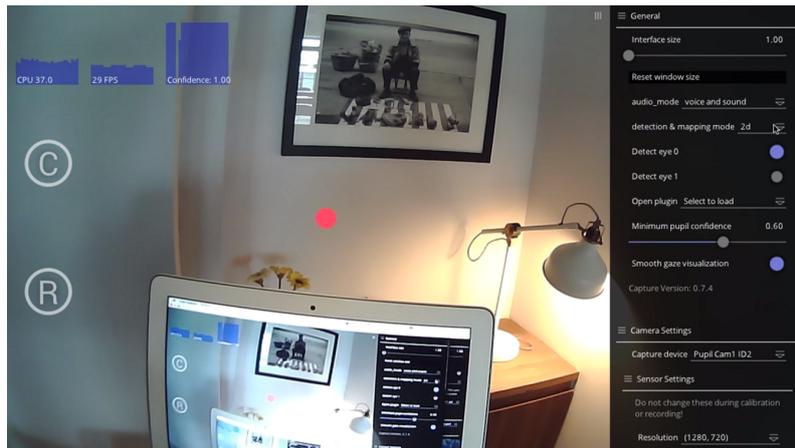


Figure 3.3: World camera view with overlaid red dot marking gaze position [2]

also generating reports in spreadsheet format that can be used by any other data processing tool. In order to analyse the eye gaze data and its associated patterns, a heat-map function will be applied for each round of experiments. Here we expect to see a correlation between the mapped pattern and the type of movement performed (goal/no-goal).

For pupillometric studies, the series of pupil diameter acquired through LSL via plugin in Pupil Capture is processed through MATLAB. The first step was signal filtering. Pupil reactions are referred as slowly changing, around 2Hz, therefore a low-pass filter with cut-off frequency of 4Hz was applied [19]. The data often contained points far off the main trend, mostly due to blinking, since reading the pupil is not possible in that condition. Therefore outliers were removed and after that linear interpolation took place to fill in the gaps.

After the cleaning process, each trial data is marked with the "Go" cue from the event marks collected. Since we are working with four cues per trial, four events are marked. The epochs are defined to comprise the time in between the marks (6s) while also accounting for the previous behavior of the eye. To achieve that the epoch limits span from $-2s$ before the cue, and $4s$ after [19]. Baseline removal is also applied, for that, samples from the first second of the epoch are averaged and the value is removed from the set.

After setting the epochs we average then within each trial and for all trials in the same task, so twenty epochs are averaged (5 trials per task, 4 epochs per trial). The standard deviation is also calculated for the stacked epochs and later the average pupil size variation through time is plotted together with its deviation.

3.1.3 Synchronization

As it was shown, for this project sensors from different parties will be combined, where each one has its own software, therefore, in order to properly analyze the data from our experimental runs, synchronization must be guaranteed. If the correlation between gaze data and EEG has to be demonstrated on each run, it is very important to ensure that all the signals are within the same time reference. For this scenario we use Lab Streaming Layer (LSL). That is an open-source platform from MIT

(Massachusetts Institute of Technology) that works as a cross-platform protocol, enabling device networking alongside with time-synchronization and near real-time data access, figure 3.4. The "core transport library" can deal with many different programming languages, e.g. C++, Python and MATLAB. It is also possible to create one's own 'Application Programming Interface' (API) that feeds data into the network through an outlet regardless of how many channels it uses or its sample rate. The time series are then received through the inlets, for example, a recorder application, that stores and provides access to all the connected streams later on [1].

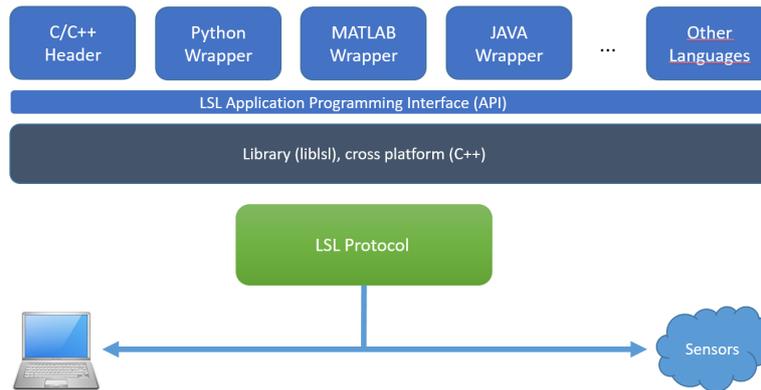


Figure 3.4: Lab Streaming Layer architecture [1].

The data synchronization is handled within the LSL library by assigning timestamps to each sample from the clock source. This source is normally the clock of the computer unless other source is specified. The program used, the Lab Recorder, is in-built with the application, it stores both timestamps and clock offsets for each sample from each of the streams. Later, another application called Importer "performs a linear fit through the clock offsets for each remote stream (thereby assuming a linear clock drift) and uses this linear mapping to remap given remotely collected timestamps" [3].

3.1.4 Robot control

A simple 2D motion detection algorithm based on color recognition was used to track a marker on the robot's end effector. The recordings were done with a simple web camera, seeing the experiment mount from the top (Z axis), since the robot can only move in the X , Y direction. The algorithm for marker recognition is implemented using Python OpenCV library [37]. It provides a live feed from the camera view and adds two overlaid circles, one on the top of the mark indicating the actual position, and the other one indicating the end effector's predicted position. For the former, a Kalman filter was used as a one-step ahead predictor of expected motion that is defined as a discrete time linear dynamic system:

$$x(k+1) = F(k)x(k) + \zeta(k)w(k) \quad (3.1)$$

where $x(k) \in R^{n_x}$ is a current state vector, $F(k) \in R^{n_x} \times R^{n_x}$ is the state transition matrix, $\zeta(k) \in R^{n_x} \times R^{n_x}$ is the process noise covariance matrix and $w(k)$ is the noise.

The system is observed as:

$$z(k) = H(k)x(k) + \epsilon(k)w(k) \quad (3.2)$$

where $z(k) \in R^{n_z}$ is the measurement vector, $K(k) \in R^{n_z} \times R^{n_x}$ is the measurement matrix and $\epsilon(k) \in R^{n_z} \times R^{n_z}$ is the measurement noise covariance matrix. The Kalman filter was defined under constant target velocity assumption with parameters:

$$x = \begin{bmatrix} x \\ y \\ \dot{x} \\ \dot{y} \end{bmatrix}, F = \begin{bmatrix} 1 & 0 & \Delta T & 0 \\ 0 & 1 & 0 & \Delta T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \zeta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 10^{-19} & 0 & 0 \\ 0 & 0 & 100 & 0 \\ 0 & 0 & 0 & 100 \end{bmatrix}, \quad (3.3)$$

$$z = \begin{bmatrix} x \\ y \end{bmatrix}, H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \epsilon = \begin{bmatrix} 0.01 & 0 \\ 0 & 1 \end{bmatrix}. \quad (3.4)$$

where ΔT is the duration of one camera frame in seconds. The process and measurement noise covariance matrices were chosen empirically.

The control algorithm for collaborative behaviour of the robot was implemented based on this prediction. The block diagram in figure 3.5 shows the implemented control loop. Kalman filter takes observed coordinates from the camera as an input and predicts the next step. These predicted coordinates are translated into robot joint angles using inverse kinematics and these angles are written to robot motors through Arduino controller. The algorithm is implemented in Python programming language and the pseudo code for this algorithm is presented in Appendix B.

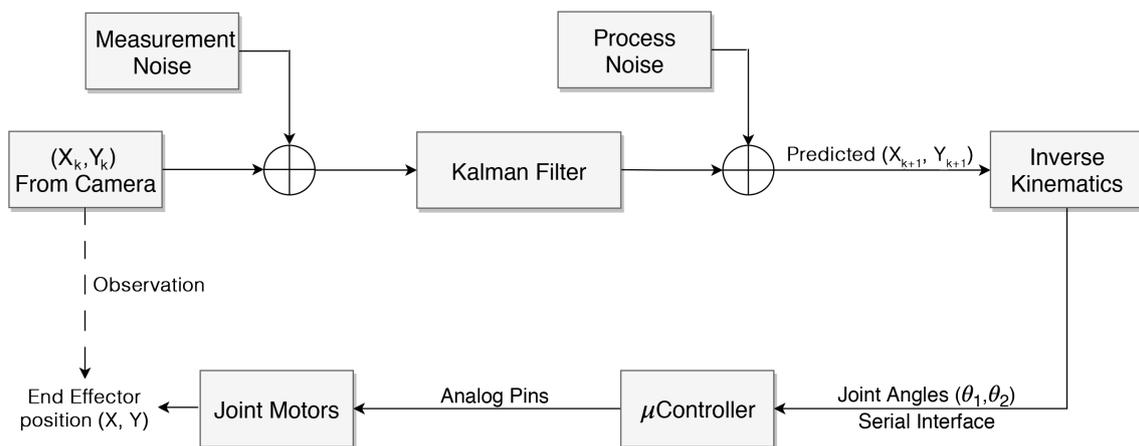


Figure 3.5: Block diagram of control algorithm.

3.2 Experiments

3.2.1 Participants

Due to the lead time for regulatory ethical approval of the procedure, it was not possible to perform the experiments on invited subjects. Therefore, the test subjects were restricted to the authors of this project and their supervisor. The group is homogeneous with regards to some key parameters thus less variability would be introduced into the data. Also, all subjects are right-handed, with no neurological disorders and normal or corrected to normal vision without the use of glasses. Subjects were instructed to stay relaxed at all times but also oriented to minimize unnecessary arm and eye movements. Before the experiment participants were briefed about the procedure and all questions from them were addressed.

3.2.2 Experiment setup

The experiments were carried out in two separate sessions. The first session was focused on data collection for later offline processing and training of the classifier. The second experiment is based upon the findings of the first session and designed to collect the data for evaluation of collaborative behaviour as well as investigating a possible correlation between EEG and eye tracking data.

The robot used for experiments was a simple SCARA robot with two servo motors. A plate with eleven LED lights corresponding to possible focus points defines the workspace that was located in the robot configuration space (see figure 3.6). A camera above the workspace and a marker on the robot's end effector were mounted to facilitate the motion tracking.

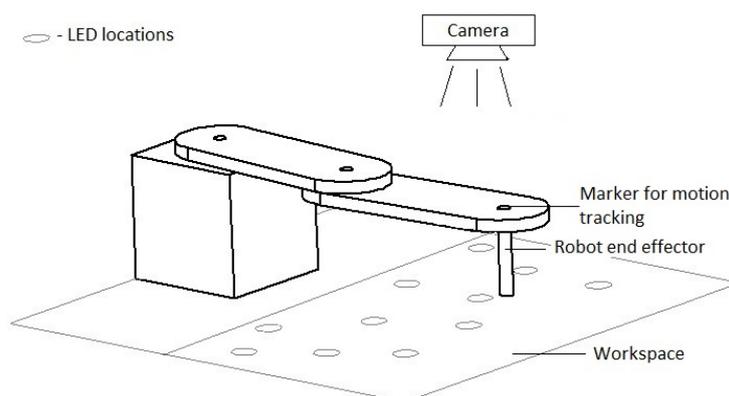


Figure 3.6: Experiment setup. The participant navigates the robot in the workspace (either freely or to one of the LEDs) by holding it on its end effector, the camera mounted above the workspace records the video for tracking of the marker motion.

In both sessions, two types of experiments were carried out: goal movement task

and no-goal (free) movement task illustrated in figure 3.7. In a goal movement task, the experiment starts with a three second period of rest, then a randomly selected LED lights up and stays for a period of three seconds, also a short auditory cue is used. Lighting up of an LED with the sound is a "Go" cue for the participant and it is also used as an onset in both EEG and Pupil measurements. During this period the participant makes a movement to reach the point where LED is located. This period is followed by a period of rest before another LED lights up with another auditory cue. The cycle is repeated four times. A free movement task starts with a six second period of rest and is followed by an auditory cue together with a short blink (100 ms) of all eleven LEDs (a "Go" cue and an onset). After this, there is another six second period during which the participant has to make a random movement of any length, stop and wait for the next "Go" cue. The cycle is repeated until four movements were made. An example of the tasks seen from the camera view is shown in figure 3.8.

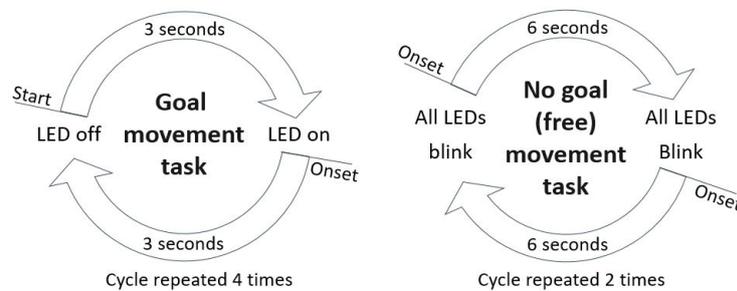


Figure 3.7: Cycles of two types of tasks used in experiments.

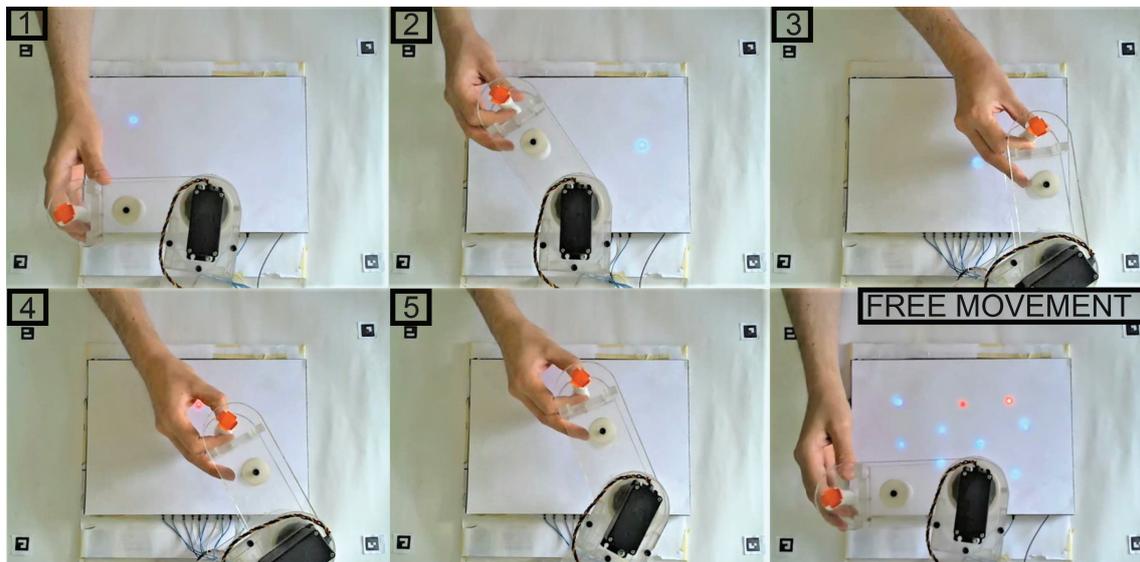


Figure 3.8: Example task. Frames one to five illustrate a changing or repeating task scenario. A bottom right frame illustrates the free movement task scenario.

3.2.3 Experiment session one

During the first experiment set, all the test subjects were asked to do a set of simple tasks and their EEG and Eye-tracking data was recorded for offline processing and classification algorithm training. In addition to biological signals, the resulting motion tracking of the robot's end effector was recorded. All the recordings were synchronised through LSL platform. There were three tasks in experiment session one, each task was repeated five times with a short break between repetitions. The tasks in the first experiment set were the following:

- **Free movement task** - In this task a test subject was asked to make a four free movements without any goal specified. Each movement was executed after a "Go" cue was given. The task was accomplished together with the passive robot. This is to illustrate the "no goal" movement scenario.
- **Connecting the dots: changing task** - In this task, the test subject was asked to perform the goal movement task described in the previous section. The task is performed by human guiding the passively following robot end effector. For each repetition the order and location of the dots is changed. This task illustrates the "goal" movement scenario without the effects of human learning the task. Therefore, the cognitive effort should remain the same as the subject faces a different set of dots at each try.
- **Connecting the dots: repetitive task** - This task is the same as the task described in the previous point, but this time the location and sequence of the dots is the same for each repetition. This task illustrates the "goal" movement scenario with the effects of human learning the task.

3.2.4 Experiment session two

During the experiment session two, the participant was asked to do the same tasks as in experiment session one. Each one of the tasks was repeated five times for each one of the following scenarios:

- **Passive robot.** In this scenario the participant is manipulating the robot, meanwhile the robot is passively following.
- **Collaborative robot.** In this scenario, the robot and human are sharing the control to achieve the goal. The robot is predicting the movement based on the motion tracking data and moves together with the human.

Throughout each part of experiment session two, the EEG and Eye tracking measurements were collected for later offline processing and classification. The correlation between EEG and eye tracking data as well as the EEG and eye movement classification were assessed based on these measurements.

4 Results

The results of the experiments that were described in the methods section are presented in this chapter. Here, the observations about behavioural analysis, EEG, gaze tracking and classification results will be covered and followed by a brief discussion of the findings.

4.1 Behavioural analysis

The boxplots in the figure 4.1 show Reaction Times (RT) - the times of movement onset relative to the time of the "Go" cue - and the Motion Duration (MD) - times of movement offset relative to the time of movement onset - for all three task types. The median reaction time of the changing task was 0.15s higher than for other tasks. No significant differences were found between reaction times of free movement and repeating tasks nor motion duration of all tasks. The mean RTs are 0.54 (± 0.21 standard deviation), 0.34 (± 0.13) and 0.35 (± 0.20) for the changing, free movement and repeating tasks respectively. The mean MDs are 1.64 (± 0.49), 1.53 (± 0.54) and 1.39 (± 0.46) for the same sequence of tasks.

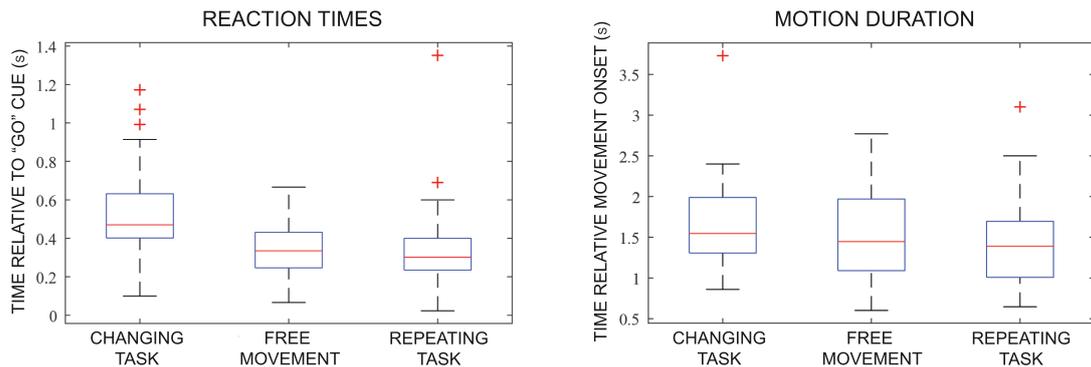


Figure 4.1: Reaction times and movement duration. Boxplots represent the times of movement onset and offset independently for three task types.

No significant changes in RT were observed when the task was performed with the robot (figure 4.2). The mean RTs for changing, free movement and repeating tasks were 0.39 (± 0.51), 0.43 (± 0.17) and 0.25 (± 0.11) respectively. The mean MD during the changing and repeating tasks were 1.19 (± 0.35) and 1.11 (± 0.27) -

slightly lower as compared to the task with passive robot. Meanwhile, the mean free movement MD, $1.79 (\pm 0.73)$, was higher, but within the standard deviation of the first experiment set results.

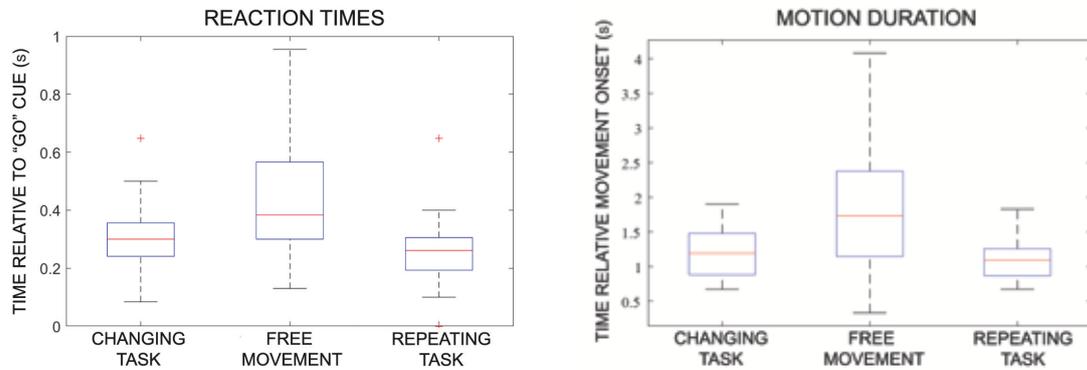


Figure 4.2: Reaction times and movement duration when performing the task with a collaborative robot. Boxplots represent the times of movement onset and offset independently for three task types.

The reaction times and motion duration presented in figure 4.1 correspond with the average velocity profiles of three test subjects in figure, 4.3. RTs vary slightly between test subjects, but are more or less consistent between the tasks of the same subject. In all cases, the reaction time is under 500 ms. The velocity profiles show no significant differences in the peak velocity between different tasks, however the velocity varies between test subjects. The velocity profiles of a collaborative task

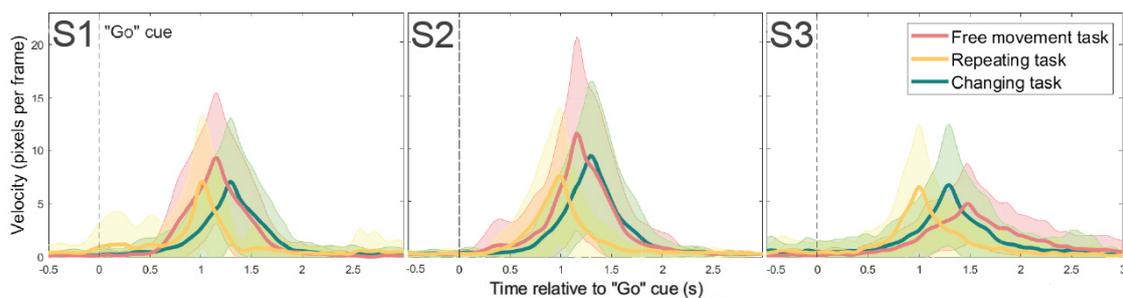


Figure 4.3: Average velocity profiles of three test subjects (S1, S2, S3) for three task types. Standard deviation is shown as a shaded area around the mean of the same colour.

with robot in figure 4.4 show a similar trend as observed previously and correspond to RT and MD results in figure 4.2. There was an exception to the trend in case of the changing task of test subject S1, where the MD appears to be significantly longer than in other tasks, lasting for up to 2 s whereas MDs of other tasks of the same participant only take around 1 s.

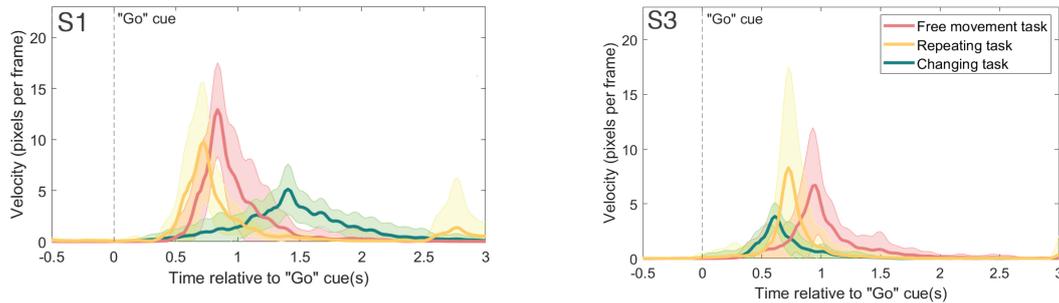


Figure 4.4: Average velocity profiles when performing the task with a robot. Velocity profiles are shown for test subjects S1 (left) and S3 (right) for three task types. Standard deviation is shown as a shaded area around the mean of the same colour.

4.2 EEG result analysis

The EEG data was analyzed in terms of temporal and time-frequency features within $[-2\ 3]$ second windows time-locked to "Go" cues. The results discussed in this section are averaged over all participants, the results for separate participants are presented in the appendix A.

4.2.1 Temporal analysis

The grand average MRCPs at five channels for three different test tasks are presented in figure 4.5. The largest amplitude changes for all three tasks can be observed at electrodes covering the left primary motor cortex (electrodes C3, Cz) and superior parietal lobule (Pz). For the changing task, BP starts before the "Go" cue is shown, indicating a possible anticipation of the task start. The deepest negative deflection over all electrodes can be observed when performing a changing task. During the MMP period, the results of the changing and repeating tasks compared to the free movement task exhibit a smaller positive deviation from the zero level over all electrodes. Overall, free movement and repeating task results show a similar behaviour over all electrodes. It is evident from single participant MRCPs shown in Appendix A figures A.1 and A.2 and A.3 that, with few exceptions, this tendency is recurrent over test subjects and trials.

Figure 4.6 shows MRCP trial images of test subject S1. By investigating the temporal variability of the MRCPs over trials one can observe that there is a clear decrease in the potential that starts around the time of the go cue and continues up to around 600 ms. This period is followed by an increase period that peaks around 1000 ms and lasts for up to 1500 ms after onset. In case of the changing and free movement tasks, this behaviour is consistent through most of the trials. In contrast, the repeating task shows more variability over trials. The scalp plots of the main ICs contributing to MRCPs show that the main activity appears at the motor and premotor cortices and spreads into the prefrontal cortex.

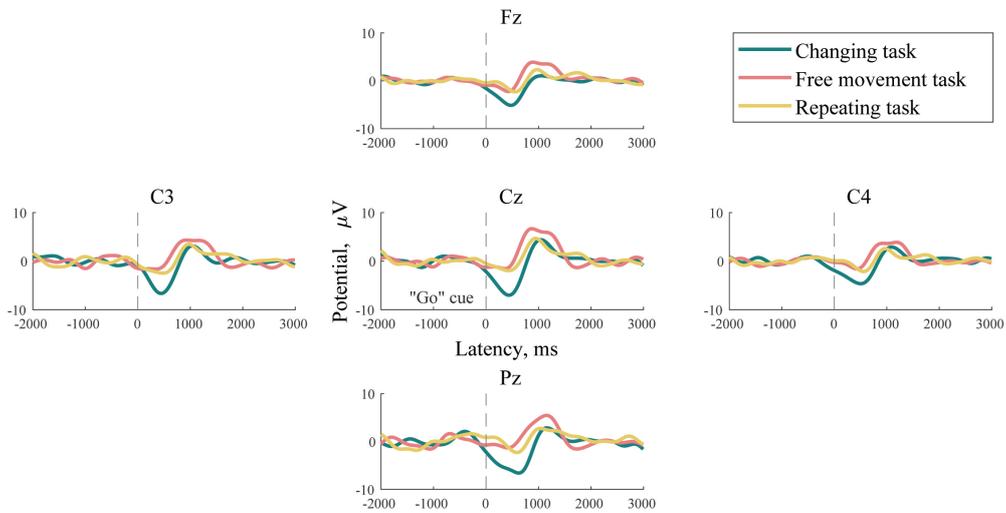


Figure 4.5: MRCPs averaged over all test subjects presented for three different tasks in five electrode locations. The MRCPs were calculated over time window locked to "Go" cue that is marked with a dash line.

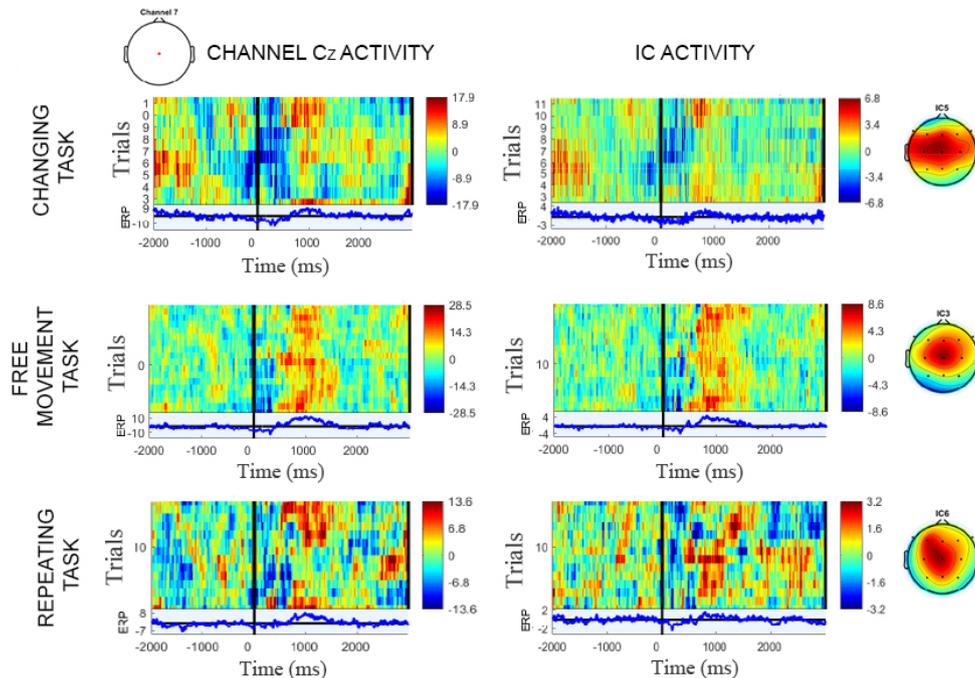
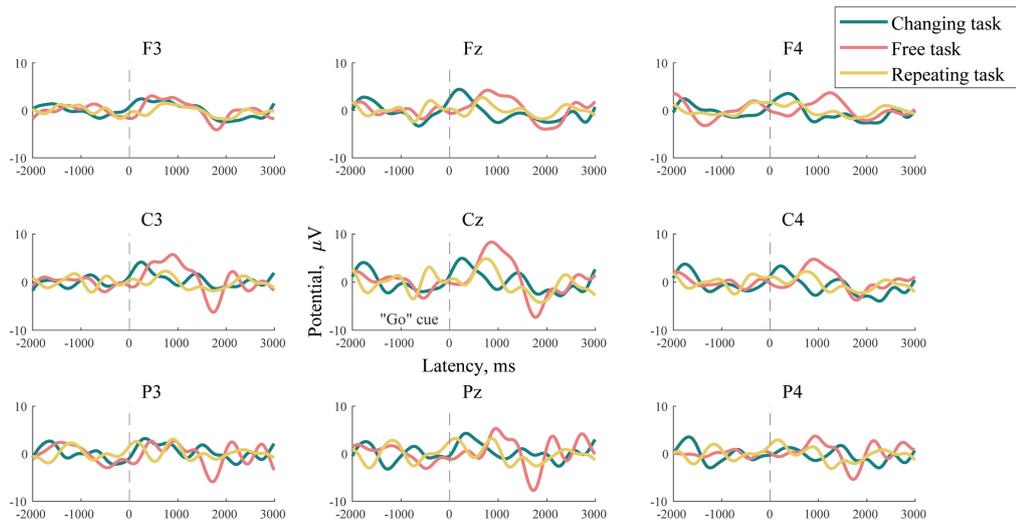


Figure 4.6: MRCP image plots of test subject S1 at electrode Cz (left) and IC (right) that corresponds to motor activity of the same task. The right-most column shows scalp maps corresponding to IC activity presented to the left of it. In each panel, the potentials in μV are color coded (see color bar on the right) and sorted from bottom to top in order of trials. The time window is locked to the "Go" cue. The blue curve at the bottom of each panel represents the MRCP magnitude in μV .

(a) Test subject S1



(b) Test subject S3

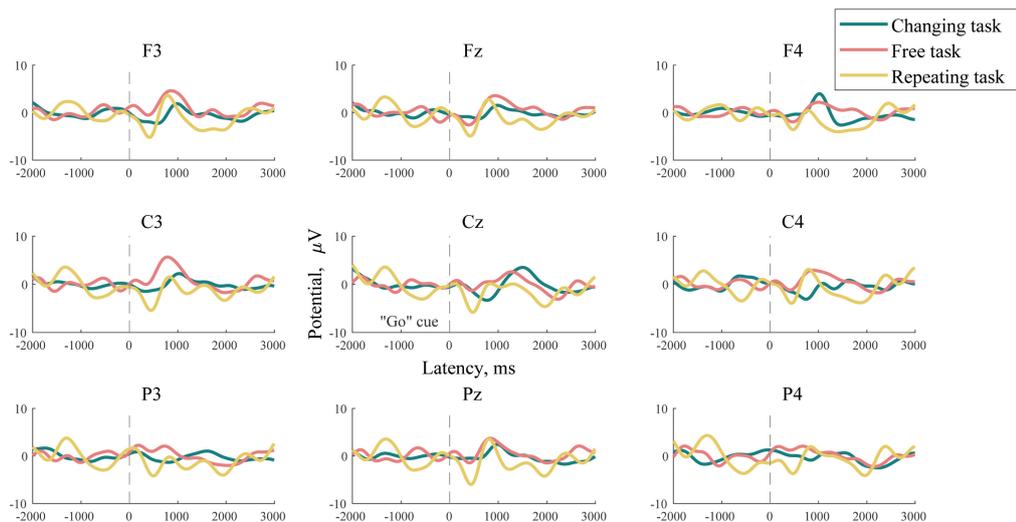


Figure 4.7: MRCPs of test subjects S1 (a) and S3 (b) when performing the task with collaborative robot. MRCPs are presented for three different tasks in nine electrode locations. The MRCPs were calculated over time window locked to "Go" cue that is marked with a vertical dash line.

As seen in figure 4.7, the same three tasks performed with collaborative robot show a large intersubject and intertrial variability. Test subject S1 shows a strong positive deviation from the baseline, resembling a P300 waveform, at central and frontal electrodes during the first second of all tasks. In case of the free movement task, this deviation is followed by a negative deflection at 1800 ms. Test subject S3 demonstrates results that are more consistent with a passive robot case - one can observe a clear MRCP over the first second after the "Go" cue that is most prominent over the electrodes covering motor cortex. The lowest MP is reached for the repeating task.

In case of the repeating task, another, smaller, negative deflection can be observed at around 2000 ms after the "Go" cue over all electrodes.

4.2.2 Time-frequency analysis

Time frequency analysis (figure 4.8) shows an increase in power over delta and theta frequency bands during the first second after the "Go" cue of all tasks, peaking at 500 ms. Another, milder, power burst appears in free movement and repeating tasks at 1500 ms after go cue. For the changing task, an increase in power appears over all four frequency bands and is more prominent at 500 ms in lower regions of each band. This period is followed by a power drop below baseline level in the alpha and lower beta regions of changing and repeating tasks. Overall, the ERSP images of repeating and free movement tasks are very similar and differ noticeably from the changing task, which correlates with behaviour of temporal features. However, individual ERSP images (Appendix A figure A.4) of test subjects show that there is a high intersubject variability in time-frequency features, indicating a need to calibrate the system for every user.

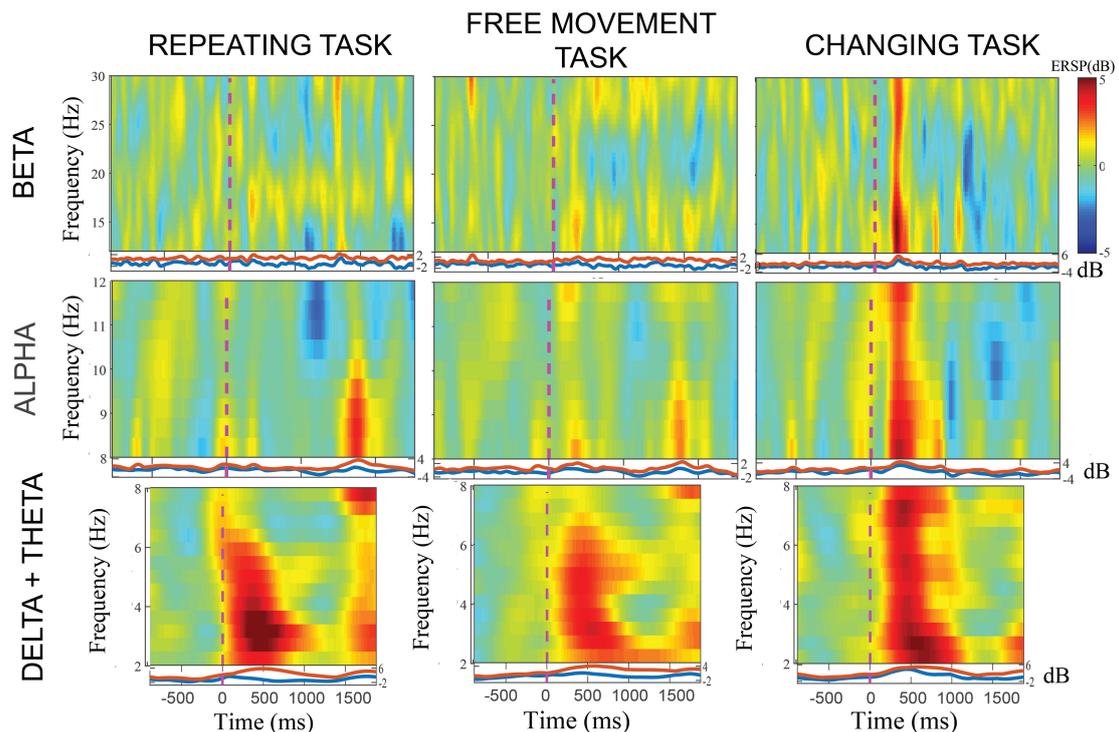
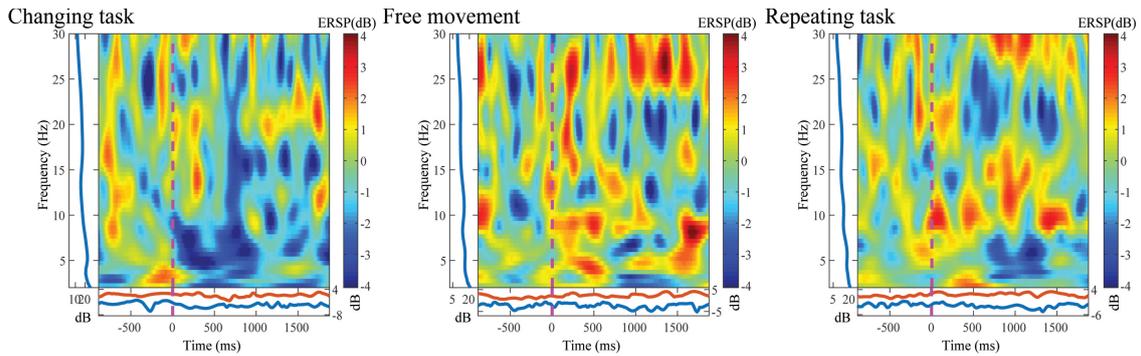


Figure 4.8: Average participant ERSPs at the electrode Cz in delta, theta, alpha and beta frequency bands. The power spectrum is color coded to the bar on the right as well as displayed at the bottom of each panel in a form of an ERSP envelope. The time window is locked at zero to the "Go" cue.

Similar to MRCs, the time-frequency analysis of the collaborative task (figure 4.9) shows a great variability between test subjects and in case of test subject S1, between

the tasks. Subject S3 ERSP images are similar to the ones of experiment one, with strong activity in delta, theta and alpha frequency bands occurring during the first second after the "Go" cue. However in contrast to experiment one, the collaborative case free movement task results show the positive peak in power in all frequency bands, meanwhile the changing task power only peaks in frequencies below 16 Hz and a drop in power can be observed for frequencies between 16 Hz and 25 Hz.

(a) Test subject S1, electrode Cz



(b) Test subject S3, electrode Cz

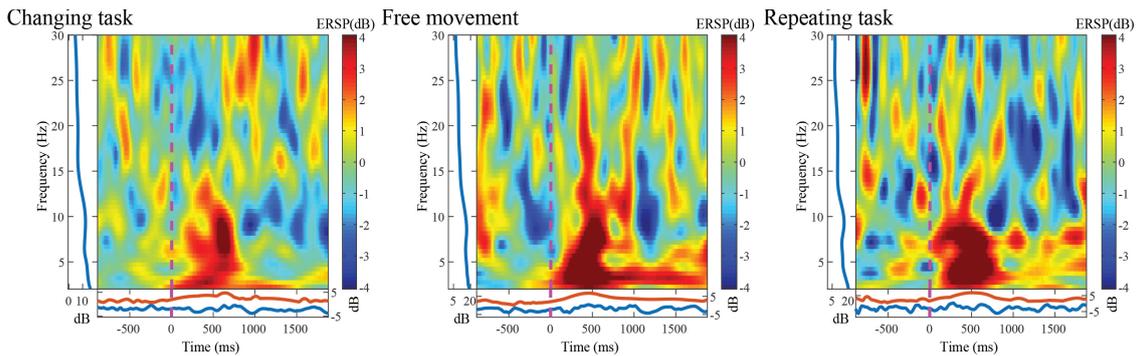


Figure 4.9: ERSPs of test subjects S1 (a) and S3 (b) at the electrode Cz at frequencies [0 30] Hz. Each panel shows the power images of changing (left), free-drawing (middle) and repeating tasks (right). The power spectrum is color coded to the bar on the right as well as displayed at the bottom of each panel in a form of an ERSP envelope. The time window is locked at zero to the "Go" cue.

4.3 Gaze tracking & Pupillometry

For gaze tracking and pupillometry analysis we extract the data from Pupil Labs algorithm through LSL and also use features from the GUI provided for the headset.

4.3.1 Gaze Tracking

The aggregated fixation points for each trial of the subject's runs were gathered as a heat-map overlaid on the recorded scene. For subject S1 it is shown the compiled heat-maps for all the trials performed, five for each task type. For the task type of

"Connecting the dots" with passive robot, either with the changing or repeating set, the gaze matched the dot's location, figure 4.10. For the random task type, it can be seen that the pattern changes either with a very hot spot in the middle of the plane or more spread across the area. Some of the tasks, mostly the changing type ones, were impacted by a reduced frame rate due to a system overload when running the experiment. Sometimes the PC lowered the performance of the applications during the trials due to heavy stream of data, therefore, some of the bright points on the changing task trials should have had more samples. Due to that, the data from subject S2 had to be discarded.

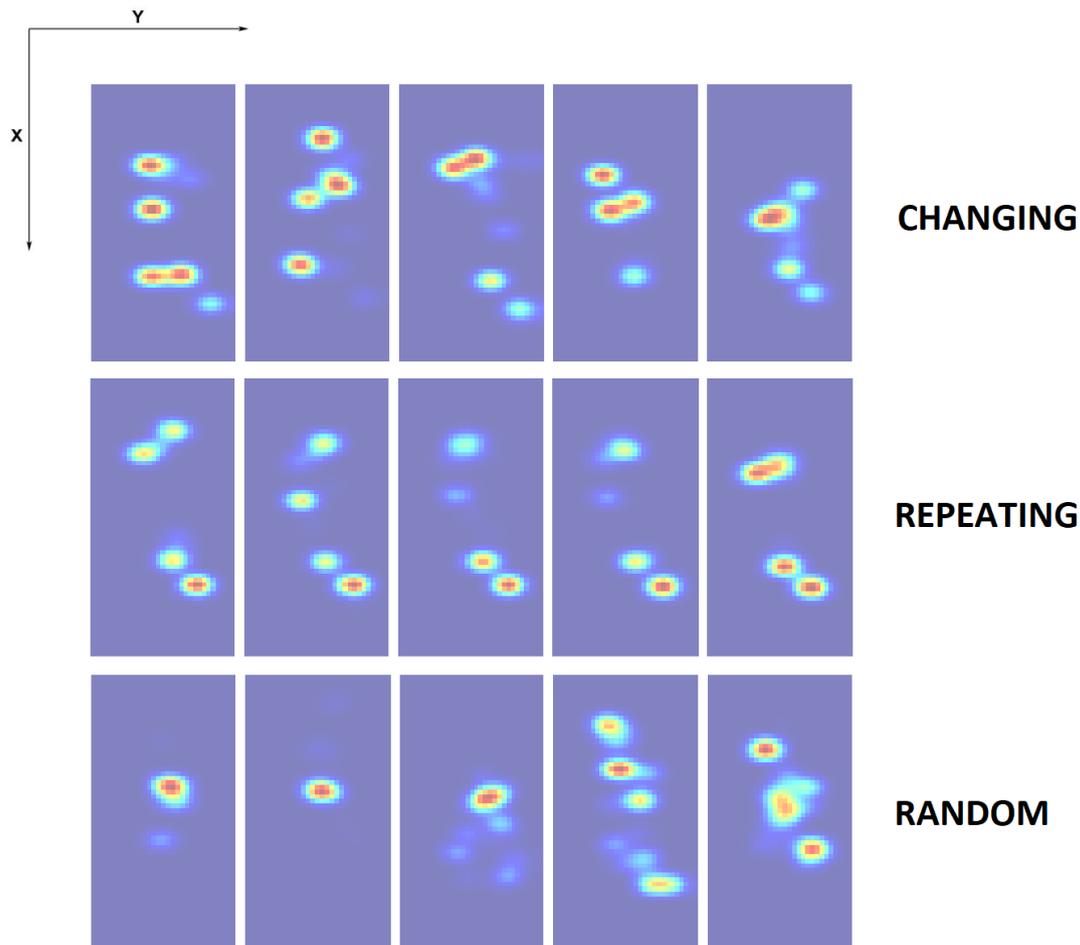


Figure 4.10: Heat maps for all the trials from subject S3 with passive robot, grouped by task type. The vertical axis is Y and the horizontal is X, the figure shows them flipped 90° clockwise for better grouping. The changing task presented the subject with always shifting patterns, while the repeating task lit up always the same LED sequence. The random task did not demand any specific motion from the subject who, therefore, stared at the surface in a non predicted way.

The tasks are also broken down in time to see how the gaze behaved throughout the tasks. It was picked one representative trial for each task type and a 3-D arrow plot was generated in order to better track its development. The synchronized "Go" cues are also marked as a reference to support the study of those behaviors. In

figure 4.11, figure 4.12 and 4.13 it is shown how the heat map, and therefore, the gaze, develops for changing, repeating and random tasks respectively. Both spatial dimensions are highlighted with regards to time to support understanding. It is also worth noting that, specially for the random task, the brightness may vary for the the same points throughout the snapshots. That happens because the brightness is also function of the total number of gaze fixations for a given time window. Since the first two snaps comprise a short interval in time, and therefore less fixation points, a few hits on the same spot end up "heating" more that spot. They later get dimmed down because more different fixation points enter our collection.

Results for subject S1 can be found in the Appendix A.

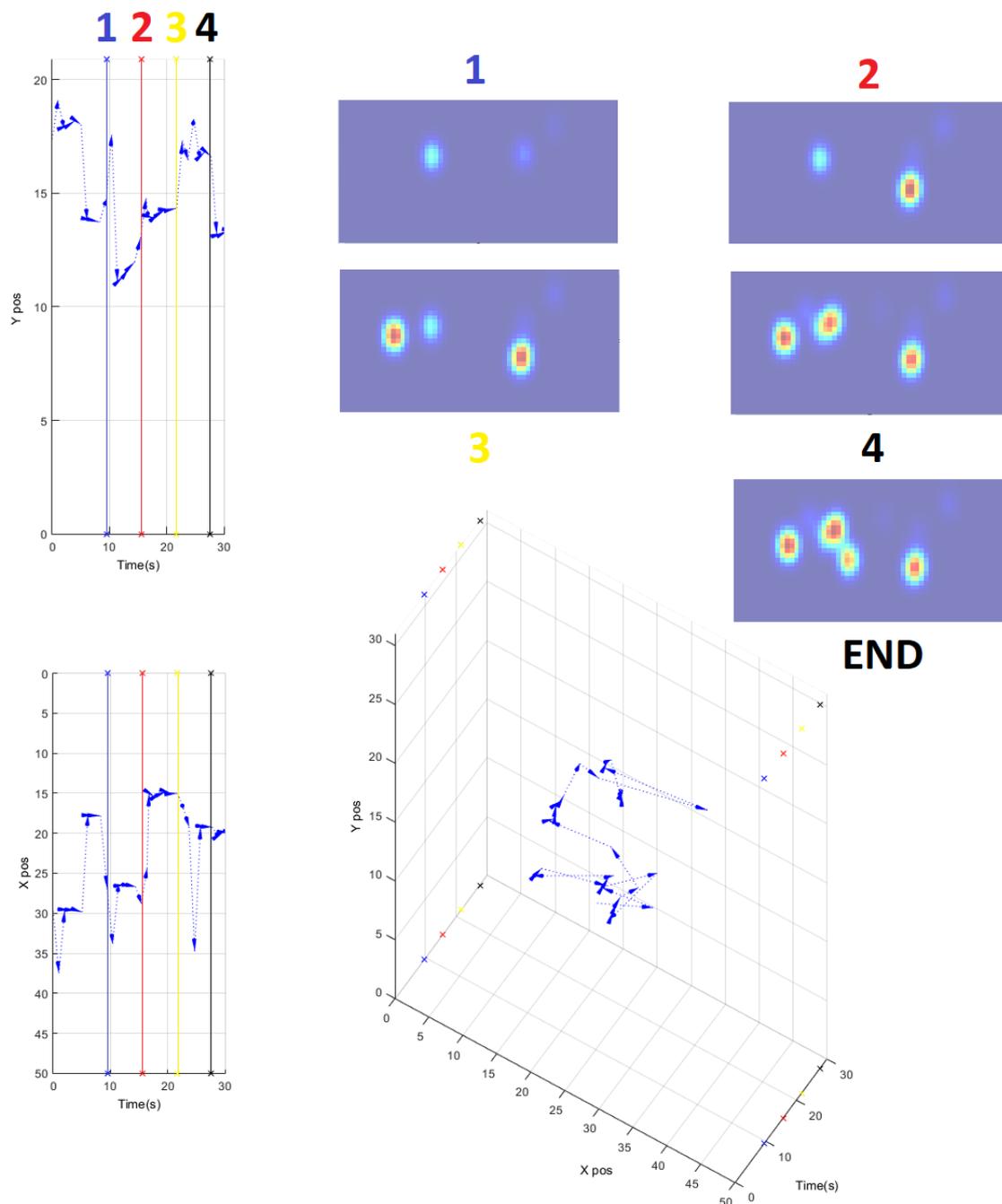


Figure 4.11: Subject S3: Gaze behavior through time for the changing task with passive robot, broken down with cumulative heat maps. The spatial coordinates are isolated into 2D planes with time to track their development. The four "Go" cues (marked with crossed) are marked and assigned to the heat map with gaze points accumulated up to that time, it is also included the map for the end frame, after the last cue. It can be seen that as the cues go off and before the hand movement onset, the gaze is directed to the location where the LED is lit.

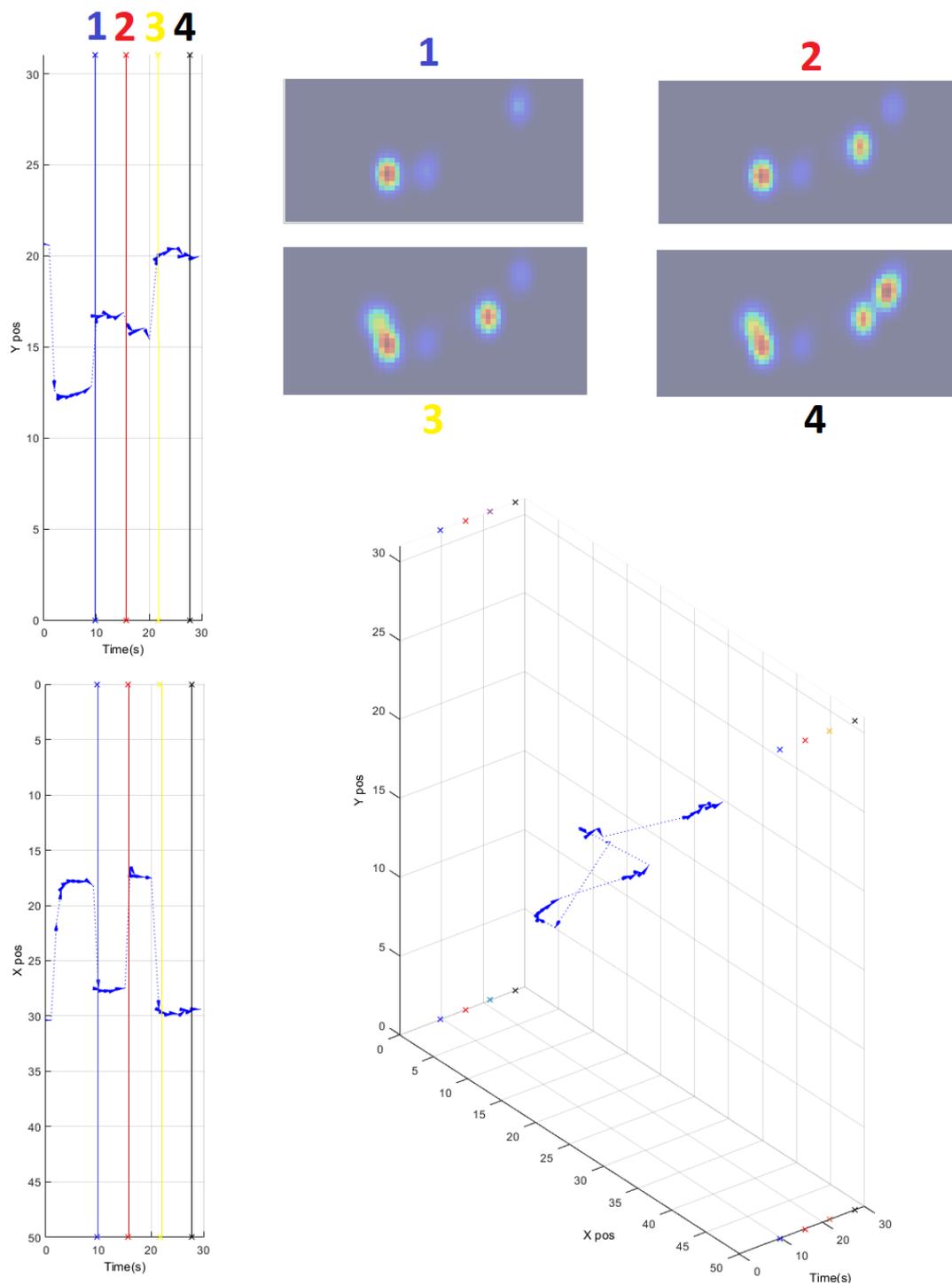


Figure 4.12: Subject S3: Gaze behavior through time for the repeating task with passive robot, broken down with cumulative heat maps. The spatial coordinates are isolated into 2D planes with time to track their development. The four "Go" (marked with crossed) cues are marked and assigned to the heat map with gaze points accumulated up to that time. A careful look shows that for all points, before the cue goes off, the gaze is already located there, this anticipated knowledge can be useful for applications.

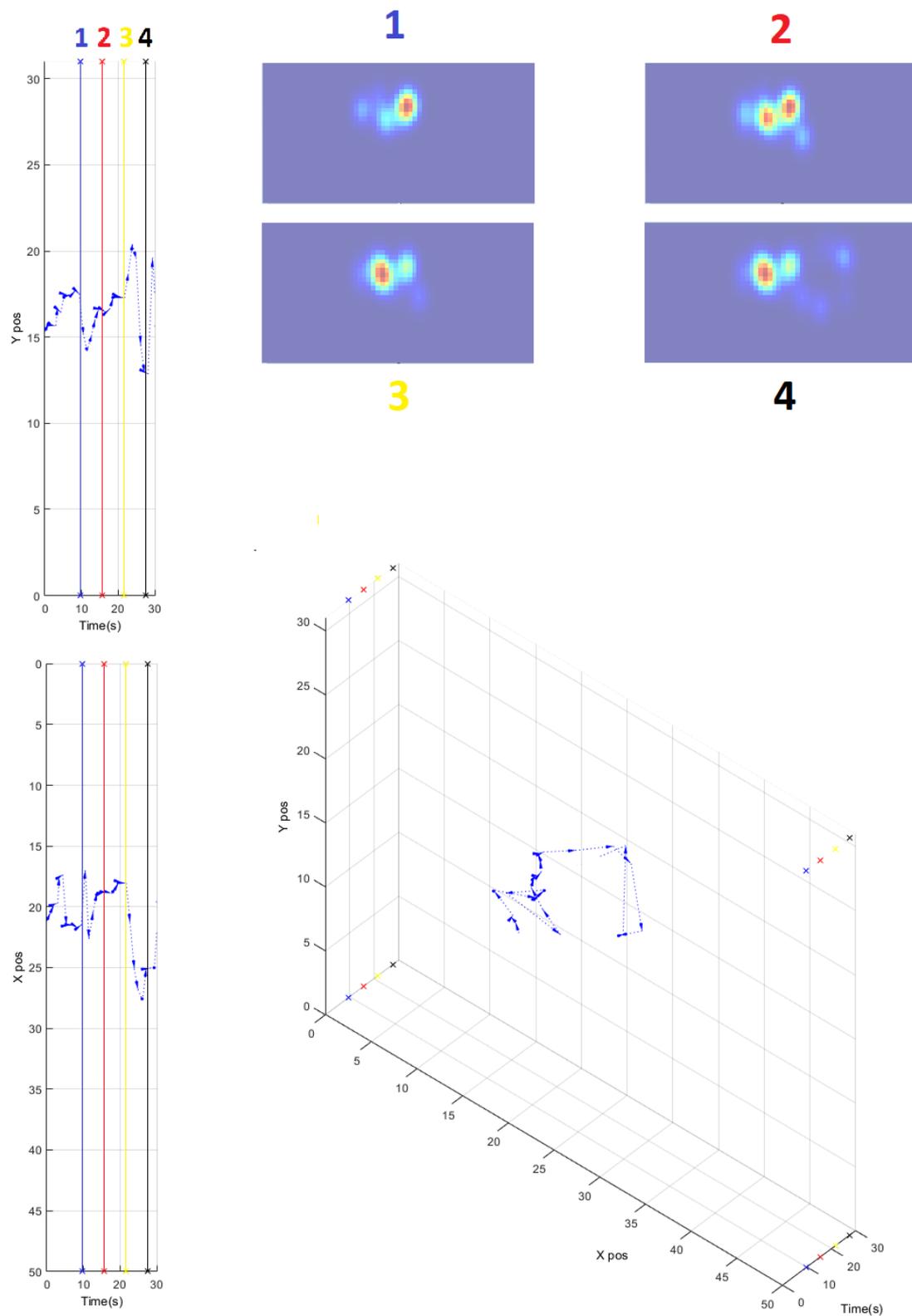


Figure 4.13: Subject S3: Gaze behavior through time for the random task with passive robot, broken down with cumulative heat maps. The spatial coordinates are isolated into 2D planes with time to track their development. The four "Go" (marked with crossed) cues are marked and assigned to the heat map with gaze points accumulated up to that time. Because of the no-goal scenario, gaze behavior is more scattered and develops in a non

4.3.2 Pupillometry

For assessing pupil behavior at all times, pupil diameter (in pixels) was recorded. As stated before, pupil reactions can be a strong indicative of cognitive effort. The data was collected and processed as previously described and the pupil responses for the three task types, with passive robot, are exhibited on the same plot for comparison for subjects S1 and S3. Changes of ± 1 in pupil mean size are already meaningful. For the time windows, the cue is set at time zero and the pupil activity preceding it is set be in negative time, figure 4.14. The standard deviations for all measurements were calculated, but due its high values they are not represented to improve the visualization of the lines. Figures with the deviations are included in Appendix A.

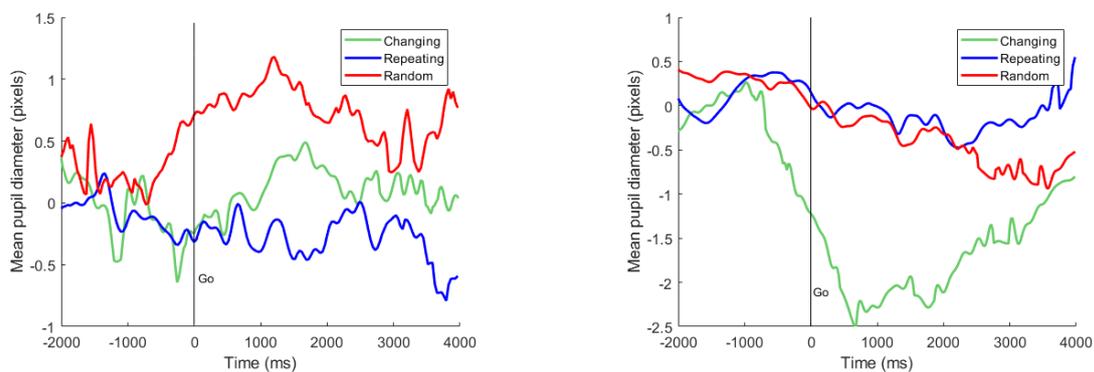


Figure 4.14: Baseline mean pupil sizes for all the tasks performed with the passive robot, for subjects S1 (left) and S3 (right). Negative times relate to pupil activity before the "Go" cue. Results varied between subjects, but for S3 the changing task presents a more distinctive feature, with the pupil size shrinking at first and rising after reaching a minimum around movement onset.

Later, experiments with the subjects performing the tasks with a collaborative robot were carried. Pupillometry for those rounds can be seen in figure 4.15. Again, the data is shown the same way as it was for the non-collaborative task.

4.4 Classification

Classification was performed using an SVM classifier and validated using 5-fold cross-validation on three datasets:

- EEG data with 20 features described in section 3.1.1;
- Eye tracking and pupil data consisting of three features, namely, fixation duration, mean pupil diameter and and slope of pupil diameter change during fixation;
- Combined dataset of both EEG and Eye tracking features.

All datasets are formed from the data collected on experiment set one. The EEG classification accuracy reached 67.2% validation accuracy and 65.3% test accuracy when using a third order polynomial kernel function and choosing the following features: minimum and maximum potentials of the epoch as well as the average

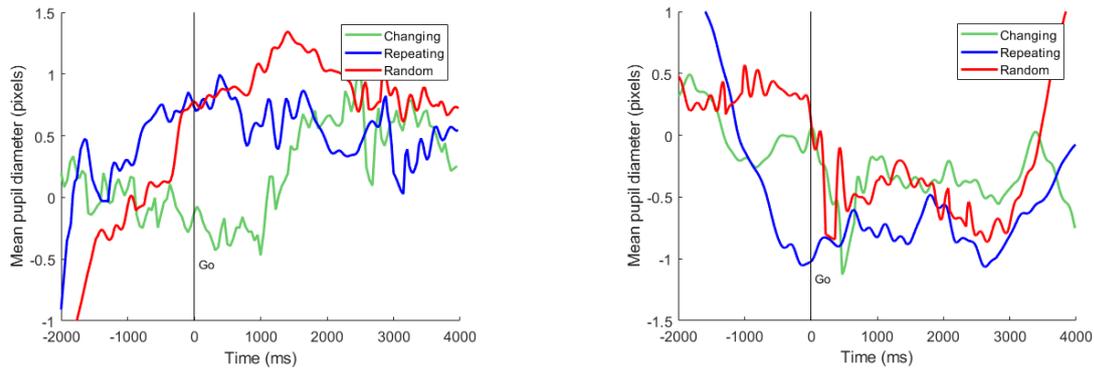


Figure 4.15: Baseline mean pupil sizes for all the tasks performed with the collaborative robot, for subjects S1 (left) and S3 (right). Negative times relate to pupil activity before the "Go" cue. Here the trends for each participant look opposed. For subject S1 all pupil values rise above baseline, where for random and repeating tasks it rises before the cue and after for the changing type. Subject S3 presented below baseline changes in pupil size for all tasks.

ERSP values in alpha frequency band at $[-0.5 \ 0.5]$ seconds from the "Go" cue. The eye tracking classification accuracy reached 85.4% for the validation set and 88.8% for the test set when using a gaussian kernel and fixation duration together with mean pupil diameter during fixation as features. A combination of both sets showed highest accuracy, 93.4%, on validation set, however the test accuracy was 71.2%. The kernel function for this model was a second order polynomial. The confusion matrices for all three test sets are shown in figure 4.16. The same classifiers applied on data from experiment set two, showed significantly lower results, with the best validation accuracy of 53.7%. A classifier trained only on experiment set two data reached 59.1%, however, the results for "No-Goal" class were below chance level.

TRUE CLASS	PREDICTED CLASS		Accuracy
	Goal	No goal	
Goal	45	19	70.3%
No goal	19	35	64.8%

TRUE CLASS	PREDICTED CLASS		Accuracy
	Goal	No goal	
Goal	42	9	82.4%
No goal	6	77	92.8%

TRUE CLASS	PREDICTED CLASS		Accuracy
	Goal	No goal	
Goal	44	20	68.8%
No goal	29	77	72.6%

Figure 4.16: Confusion matrices of three test sets: EEG set (left), eye and pupil tracking (middle), combined dataset (right).

By looking at the feature space of eye data classification (figure 4.17) we can observe that many data points are clustered at fixation duration of 4 seconds, which is limited by the software. Expanding this boundary could potentially increase the accuracy.

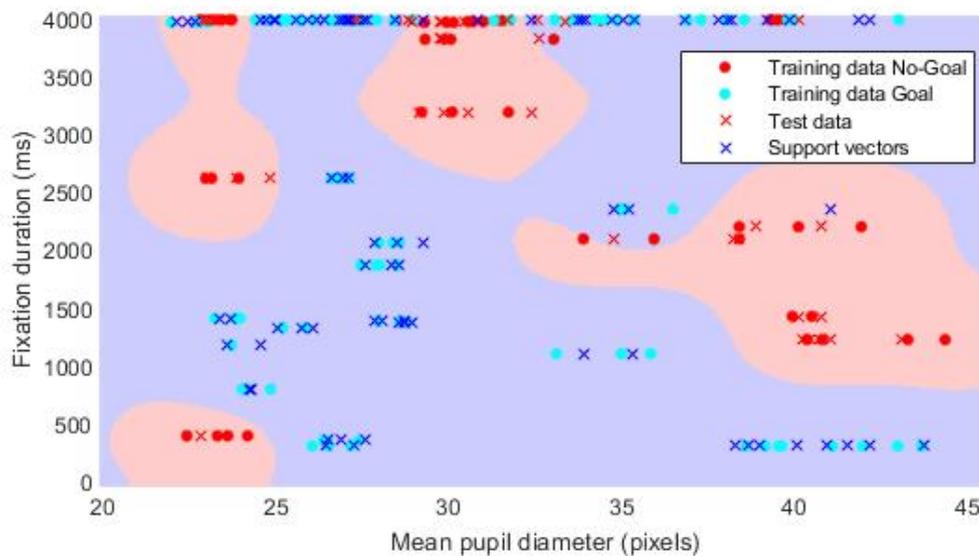


Figure 4.17: Eye data classification feature space and decision boundary

4.5 Correlation

By synchronizing the pupil diameter, MRCP and velocity profile in time, we create the correlation images in figure 4.18. The pupil diameter reaches it's minimum after the movement onset in all cases where significant changes in the pupil diameter can be observed. We can observe no temporal correlation between MRCP and pupil diameter that would be constant throughout the trials. However, there exists a correlation between the MRCP and velocity profile: at the time of movement onset, the motor potential reaches the lowest point and the peak velocity occurs around the same time as the peak MMP.

4.6 Discussion

4.6.1 EEG

Results of the changing task (passive robot case) and free movement task (collaborating robot case) in time-frequency domain directly after the "Go" cue show a strong activity in alpha and beta bands that are linked to readiness and alertness. A strong increase in power in delta and theta frequency bands, that are primarily associated with resting state, occurs over all subjects during all tasks. Some previous studies have also observed an event related increase in delta power during Go or No-Go tasks [22].

The MRCPs show similar results to previous studies that were investigating the goal directed movement [31], with an exception of subject S3 in collaborative robot scenario, where the behaviour resembles a P300 wave that is associated with attention to changes in the environment and processing of novelty [32]. In the passive robot

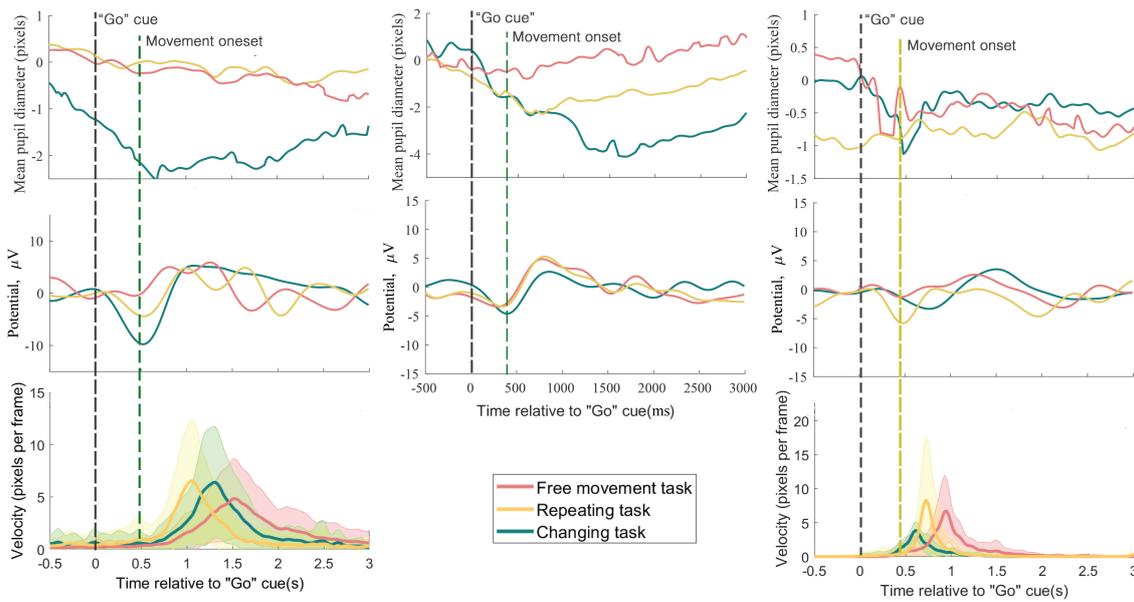


Figure 4.18: Temporal correlation of subject S3 MRCP, pupil diameter and velocity profiles synchronized at the "Go" cue. Each column represents one experiment: experiment set one (left), experiment set two passive robot (middle), experiment set two collaborative robot (right).

case, the changing task demonstrates a lower MP than other two tasks, suggesting that the goal directed movement is distinguishable when the goal is externally presented and changing, however the repeating task shows similar results to the no-goal scenario, suggesting that the effect of learning the task is affecting the level of attention to the task.

Moreover, results change greatly when the robot collaboration is involved. These changes may occur due to human response to robot behaviour, since the test subjects did not have a training period to get acquainted with the robot. Moreover, the robot was behaving based on the Kalman filter prediction, causing the robot to continue moving past the stop point, meaning that the experiment subject had to apply some force to stop the movement of the robot, when the goal destination is reached. We hypothesize that this might have had an effect on the test subject's planning of the movement, shifting the focus from the beginning of the movement to the end of it. Nevertheless, the data is insufficient to determine the true nature of these changes and therefore more tests are needed to investigate how these results generalize to a larger dataset.

Finally, an early BP start before the "Go" cue can be observed in MRCPs of some cases. This might indicate the anticipation linked to experiment design. It is worth noting that the experiments involved an audiovisual stimulus that is known to induce a response in EEG data and the effect of this stimulus was not analysed in this study. Hence, some of the activity might be corresponding to the audiovisual stimulus.

Overall, the results indicate that out of three studied tasks, the highest increase of electrical brain activity is observed in the motor and frontal cortices when performing a changing task with a passive robot. When working with collaborative robot, changes in EEG data occur, which indicates that EEG is sensitive to external disturbances as well as internal distractions from the task and therefore might not be a reliable measure of goal detection in a non-controlled environment.

4.6.2 Gaze tracking & Pupillometry

Results for gaze behavior shown in the pictures are in accordance with the theory. For Changing and Repeating tasks (goal oriented) it can be seen that the majority of the fixations was directed towards the LED positions. The heatmaps showed a predictable progression for fixations as the cues were presented, whereas for the repeating task it was observed less gaze deviation from the LED locations, which may indicate a learning effect for that subject along the trial. In the Random task (No-goal oriented), fixations were irregular, with no characteristic pattern through the trials. Since the subjects were free to gaze around, sometimes the gaze was restricted to a point, sometimes it was more dispersed with a few "warmer" spots, also, from the figure, it can be seen that the progression through the cues was rather not predictable, meaning that there was no distinguishable goal for the gaze. Besides, in the varying sequence scenario, the first fixations, as it can be seen after the cue going off and before movement onset, were mostly directed to the goal point, whereas for the repeating part, for most of the times, fixations were already at the goal point even before the "Go" cue went off. This predictability for the repeating task can be of special use for similar environments, such as assembly lines. Having the previous knowledge from a person's movement could help closing the loop for the control system when the robot is sharing the task.

The results for pupillometry did not behave as expected. Also, the reduced number of subjects prevented us from drawing more general conclusions. For the passive robot set up, Subject S3 showed only a constant negative variation from the baseline for the random task, which could potentially indicate lower activation of the LC, whereas for the repeating task the behavior is mostly around the baseline with a late trend going up, but, again, not significant enough. However for the changing task a steep valley during movement onset (around 500ms) was identified, but with the recovery two seconds later restricted to baseline levels, which are not enough to determine an improved activation of the LC. When S3 performed the tasks with the collaborative robot all the averaged pupil lines stayed below baseline, with the repeating task as the lowest, which may indicate that having the robot helping eased the mental effort. It is important to highlight that steep behaviors (upwards or downwards) for the repeating task at the edges of the time window were mostly due to high data variability within the epochs, with high values for standard deviations.

For Subject S1, with the passive robot, it looks like the random task had more cognitive impact. Also, since the pupil size raises before the cue, it can be hypothesized that maybe an anticipation have happened, but further analysis should be made to confirm it. The repeating behavior stayed mostly below baseline. When interacting with the active robot, S1 showed a similar behavior than before for the random task. For the repeating scenario, the pupil size rapidly raised but stayed around 0.5. Lastly, on the changing task a clear valley followed by a crescent trend away from the baseline was observed, showing that maybe S1 had more mental load, and therefore higher activation of the LC, while performing that task.

4.6.3 Classification

The classification results show that "Goal" and "No-goal" cases are separable for a passive robot case when using all three datasets: EEG, eye tracking together with pupil size and combination of both. Best results are achieved from eye and pupil tracking features and is lower for a combined dataset. Despite this, the combination of both EEG and eye data classification could still perform better as these datasets hold different information in temporal domain. The features used for EEG classification are based on the [-0.5 0.5] seconds of "Go" cue and the features of eye tracking and pupil diameter occur through the whole epoch. We infer that EEG could be used for movement prediction stage and eye behaviour based classification would act as a validation tool, which would improve the reliability of classification. In the case of active robot, the results this not exceed the chance level. We link these results to the behaviour of the robot discussed in subsection 4.6.1.

4.6.4 Correlation

The MRCP related changes are observed in the motor and frontal cortices of the brain, meanwhile the pupil changes correlate to activity in LC. Since no temporal correlation between the pupil diameter and the MRCPs was observed, we conclude that no temporal correlation between motor cortex and the LC regions of the brain was found in this study. The temporal EEG data correlates with velocity profiles and the changes in MRCPs can be observed before the movement onset whereas minimum pupil diameter does not correlate with the velocity profile, however the decrease in the pupil size starts before the movement onset. Hence, EEG can be used for the prediction of the motion, while pupil diameter is not a reliable stand-alone measure for the prediction of movement onset.

5 Conclusion

5.1 Summary of results

In this work, the possibility of improvements in HMI by deriving human intention through the acquisition and further processing of selected biosignals, namely electrical brain activity, gaze tracking and pupillometry was investigated. By studying the pattern of those signals during goal/no-goal tasks a correlation was sought. It was hypothesized that it could help provide a base for intent recognition which in turn could later be fed into the control algorithm of an autonomous system (2DOF robotic arm), improving its responsiveness towards its user, therefore bridging the gap between man and the machine. This "awareness" from its user's goals could be important for motion augmentation, which can be applied in the field of prosthetics and assisted driving, to name a few.

A BCI in conjunction with EEG system was used to record brain activity. The results after processing indicated variability through participants, especially when working together with the robot. Despite that, the results in the passive robot case in both temporal and time-frequency domains indicated a possibility to detect anticipation to movement onset. The trained SVM could correctly classify movement types with 65.3% accuracy. The results of the collaborative robot case demonstrated unexpected behaviour and high variability between subjects and trials and therefore the classification results did not reach above the chance level. However, the subject group size was too small to strongly support or invalidate the hypothesis.

For gaze tracking and pupillometry, data was acquired through two cameras embedded on a head mounted gear. For the first part, the data followed the theory, in goal directed movements, gaze pattern was directed towards the targets before movement onset, regardless to the robot mode, indicating that gaze fixation precedes movement. For pupillometric studies, the results were dissonant with the literature. It was expected to observe higher activation from the LC through higher mean pupil values for goal oriented tasks, which was not consistently verified. High variability between subjects, which might be an effect of having few participants, was also found. In the end, comparisons within the task set and between the two robot's setups are hard to draw.

All in all, working with the passive robot, the results show a distinctive behaviour between "Goal" and "No-Goal" movements in all domains. However, when working

with an active robot, the results vary highly between participants and trials. The most consistent domain through both sets of experiments was the gaze tracking. However, only the gaze tracking data was not sufficient to obtain good classification results. Hence, other modes such as pupil diameter and EEG measurements were used to improve the performance.

5.2 Future work

This work lays a ground for future research into human "Goal" and "No-goal" movement intent recognition. In this work we focused on EEG, eye tracking and pupil measurements for a specific set of experiments, that included an audiovisual cues. A further step towards implementation of such intent recognition platform in the real world devices would be to investigate self-paced movements and goal movement intent recognition in an uncontrolled environment. Moreover, other modes, such as language, movement or haptic feedback could be investigated in order to find the most efficient way for robot to infer human intention.

Due to a small number of test subjects and trials, the variability in the data is high and often subject dependent, therefore generalizing to a larger scale is inexpedient. Hence, a study with more participants is needed to validate the results of this work. In addition, other experiment designs should be explored to investigate the effects of audiovisual cues, task difficulty and robot behaviour on the results.

Moreover, this study did not investigate the effects of different classifiers, which might have a significant effect on the quality of the classification results and consequently, the prediction of human intent and reliability of the robot control. Therefore, in the future work, different classifiers should be investigated to determine if the results generalize well to other classifiers and find the optimal classification strategy that would benefit the implementation of the multi-modal control strategy the most. Lastly, an online classifier should be implemented in order to be able to implement the intent recognition system for real-time communication with the robot.

Bibliography

- [1] Home - sccn/labstreaminglayer wiki - github. <https://github.com/sccn/labstreaminglayer/wiki>. (Accessed on 05/10/2019).
- [2] Pupil docs - master. <https://docs.pupil-labs.com/>. (Accessed on 05/10/2019).
- [3] Timesynchronization - sccn/labstreaminglayer wiki - github. <https://github.com/sccn/labstreaminglayer/wiki/TimeSynchronization>. (Accessed on 05/10/2019).
- [4] *Introduction*, pages 1–20. Springer New York, New York, NY, 2008.
- [5] Human implicit intent recognition based on the phase synchrony of EEG signals. *Pattern Recognition Letters*, 66:144–152, 2015.
- [6] Susan Aliakbaryhosseinabadi, Ernest Nlandu Kamavuako, Ning Jiang, Dario Farina, and Natalie Mrachacz-Kersting. Classification of eeg signals to identify variations in attention during motor task execution. *Journal of neuroscience methods*, 284:27–34, 2017.
- [7] Susan Aliakbaryhosseinabadi, Ernest Nlandu, Ning Jiang, Dario Farina, and Natalie Mrachacz-kersting. Classification of EEG signals to identify variations in attention during motor task execution. *Journal of Neuroscience Methods*, 284:27–34, 2017.
- [8] Jackson Beatty, Brennis Lucero-Wagoner, et al. The pupillary system. *Handbook of psychophysiology*, 2(142-162), 2000.
- [9] Kristin P Bennett and Colin Campbell. Support vector machines: hype or hallelujah? *Acm Sigkdd Explorations Newsletter*, 2(2):1–13, 2000.
- [10] Alex Casson, Mohammed Abdulaal, Meera Dulabh, Siddharth Kohli, Sammy Krachunov, and Eleanor Trimble. *Electroencephalogram*, pages 45–81. 01 2018.
- [11] Stefan Debener, Jeremy Thorne, Till Schneider, and Filipa Viola. *Using ICA for the analysis of EEG data*, volume 1. 04 2010.
- [12] Arnaud Delorme and Scott Makeig. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9 – 21, 2004.
- [13] Yilmaz Durna and Fikret Ari. applied sciences Design of a Binocular Pupil and Gaze Point Detection System Utilizing High Definition Images. 2017.
- [14] Maria K Eckstein, Belén Guerra-carrillo, Alison T Miller, and Silvia A Bunge. Developmental Cognitive Neuroscience Beyond eye gaze : What else can eye-tracking reveal about cognition and cognitive development ? *Accident Analysis and Prevention*, 25:69–91, 2017.

-
- [15] Rong-En Fan, Pai-Hsuen Chen, and Chih-Jen Lin. Working set selection using second order information for training support vector machines. *Journal of machine learning research*, 6, Dec:1889–1918, 2005.
- [16] Dan Witzner Hansen, Qiang Ji, and Senior Member. In the Eye of the Beholder : A Survey of Models for Eyes and Gaze. (March 2010), 2010.
- [17] Trevor Hastie and Robert Tibshirani. J. friedman the elements of statistical learning. chapter 6, 2001.
- [18] Marjan Jahanshahi and Mark Hallett. *The Bereitschaftspotential: Movement-related cortical potentials*. Springer Science & Business Media, 2003.
- [19] Xianta Jiang, M Stella Atkins, Geoffrey Tien, Roman Bednarik, and Bin Zheng. Pupil Responses during Discrete Goal-directed Movements. 2014.
- [20] Siddhartha Joshi, Yin Li, Rishi M Kalwani, and Joshua I Gold. Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus , Colliculi , and Cingulate Cortex Article Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus , Colliculi , and Cingulate Cortex. *Neuron*, 89(1):221–234, 2016.
- [21] Eric R Kandel, James H Schwartz, Thomas M Jessell, Steven Siegelbaum, A James Hudspeth, and Sarah Mack. Principles of neural science. 2013.
- [22] Diana Karamacoska, Robert J. Barry, Genevieve Z. Steiner, Elle P. Coleman, and Emily J. Wilson. Intrinsic eeg and task-related changes in eeg affect go/nogo task performance. *International Journal of Psychophysiology*, 125:17 – 28, 2018.
- [23] Moritz Kassner and William Patera. Pupil : An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction.
- [24] Johan Kildal, Alberto Tellaeché, Izaskun Fernandez, and Inaki Maurtua. Potential users’ key concerns and expectations for the adoption of cobots. *Procedia CIRP*, 72:21 – 26, 2018. 51st CIRP Conference on Manufacturing Systems.
- [25] Dana Kulić and Elizabeth Anne Croft. Estimating Intent for Human Robot Interaction. *IEEE International Conference on Advanced Robotics*, pages 810–815, 2003.
- [26] Ernesto Pablo Lana, Bruno Vilhena Adorno, and Carlos Julio Tierra-Criollo. Detection of movement intention using eeg in a human-robot interaction environment. *Research on Biomedical Engineering*, 31(4):285–294, 2015.
- [27] Fabien Lotte, Marco Congedo, Anatole Lécuyer, Fabrice Lamarche, and Bruno Arnaldi. A review of classification algorithms for eeg-based brain–computer interfaces. *Journal of neural engineering*, 4(2):R1, 2007.
- [28] Steven Luck. An introduction to event-related potentials and their neural origins. *An Introduction to the Event-Related Potential Technique*, 107, 01 2005.
- [29] Scott Makeig. Auditory event-related dynamics of the eeg spectrum and effects of exposure to tones. *Electroencephalography and clinical neurophysiology*, 86(4):283–293, 1993.
- [30] Omar Feix do Nascimento, Kim Dremstrup Nielsen, and Michael Voigt. Movement-related parameters modulate cortical activity during imaginary isometric plantar-flexions. *Experimental Brain Research*, 171(1):78–90, May 2006.

-
- [31] Joana Pereira, Patrick Ofner, Andreas Schwarz, Andreea Ioana Sburlea, and Gernot R. Müller-Putz. EEG neural correlates of goal-directed movement intention. *NeuroImage*, 149(January):129–140, 2017.
- [32] John Polich. *Theoretical Overview of P3a and P3b*, pages 83–98. Springer US, Boston, MA, 2003.
- [33] J Rajkowski. Correlations between locus coeruleus (lc) neural activity, pupil diameter and behavior in monkey support a role of lc in attention. *Soc. Neurosc., Abstract, Washington, D. C., 1993*, 1993.
- [34] Guillaume A Rousselet. Does filtering preclude us from studying erp time-courses? *Frontiers in psychology*, 3:131, 2012.
- [35] Aqsa Shakeel, Muhammad Samran Navid, Muhammad Nabeel Anwar, Suleman Mazhar, Mads Jochumsen, and Imran Khan Niazi. A Review of Techniques for Detection of Movement Intention Using Movement-Related Cortical Potentials. *Computational and Mathematical Methods in Medicine*, 2015(vol. 2015, Article ID 346217):13, 2015.
- [36] Sylvain Sirois and Julie Brisson. Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(6):679–692, 2014.
- [37] SriramEmarose. Sriramemarose/motion-prediction-with-kalman-filter. (Accessed on 05/11/2019).
- [38] Rufin VanRullen. Four common conceptual fallacies in mapping the time course of recognition. *Frontiers in psychology*, 2:365, 2011.
- [39] Y. Wang, B. Hong, X. Gao, and S. Gao. Phase synchrony measurement in motor cortex for classifying single-trial eeg during motor imagery. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 75–78, Aug 2006.
- [40] Zheng Wang, Rencheng Zheng, Tsutomu Kaizuka, and Kimihiko Nakano. Relationship between Gaze Behavior and Steering Performance for Driver – Automation Shared Control : A Driving Simulator Study. pages 1–14.
- [41] Andreas Widmann and Erich Schröger. Filter effects and filter artifacts in the analysis of electrophysiological data. *Frontiers in psychology*, 3:233, 2012.
- [42] Jan M Wiener, Christoph Hölscher, Simon Büchner, and Lars Konieczny. Gaze behaviour during space perception and spatial decision making. pages 713–729, 2012.
- [43] I. Winkler, S. Debener, K. Müller, and M. Tangermann. On the influence of high-pass filtering on ica-based artifact reduction in eeg-erp. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4101–4105, Aug 2015.
- [44] S C Wriessnegger, D Steyrl, K Koschutnig, and G R Müller-putz. Cooperation in mind : Motor imagery of joint and single actions is represented in different brain areas. *Brain and Cognition*, 109:19–25, 2016.
- [45] Alfred L Yarbus. *Eye movements and vision*. 1967.
- [46] Han Yuan and Bin He. Brain–computer interfaces using sensorimotor rhythms: current state and future perspectives. *IEEE Transactions on Biomedical Engineering*, 61(5):1425–1435, 2014.

- [47] Jinhua Zhang, Baozeng Wang, Cheng Zhang, Yanqing Xiao, and Michael Yu Wang. An EEG / EMG / EOG-Based Multimodal Human-Machine Interface to Real-Time Control of a Soft Robot Hand. 13(March), 2019.

A Appendix: Subject specific EEG data

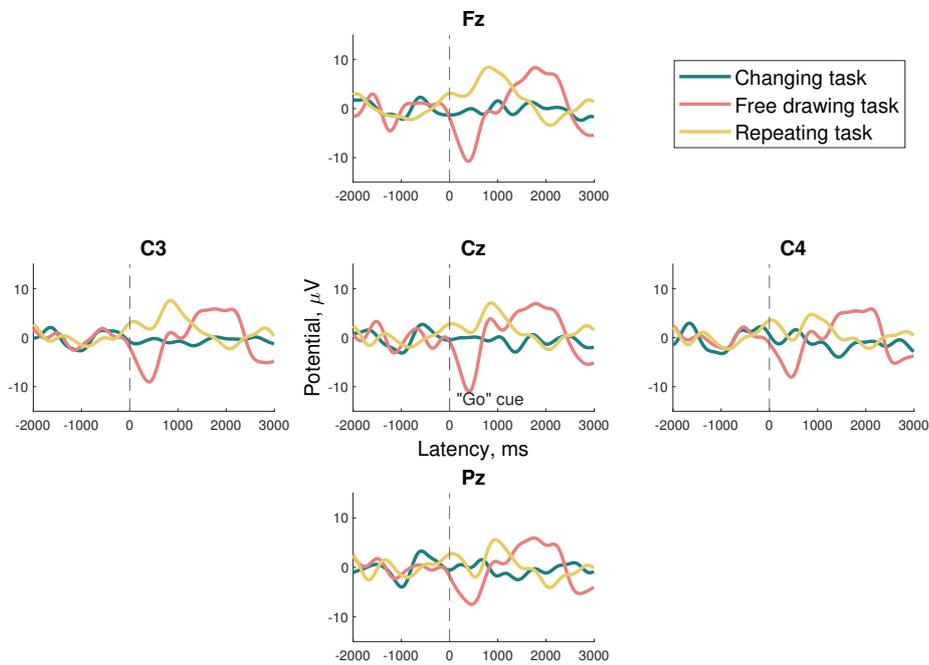
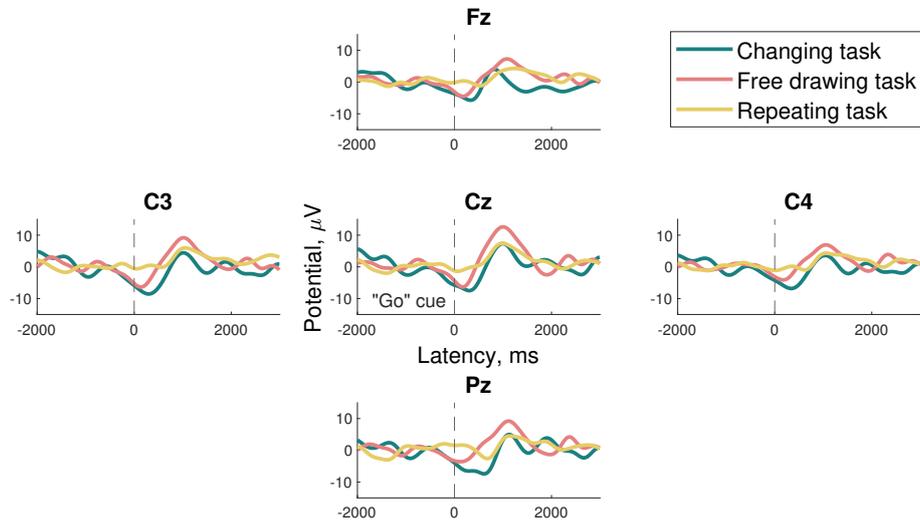


Figure A.1: MRCPs of test subject S2, experiment set one at five electrode locations. The MRCPs were calculated over $[-2\ 3]$ s time window locked to "Go" cue that is marked with a dashed line.

(a) Experiment set one, test subject S1



(b) Experiment set two, test subject S1

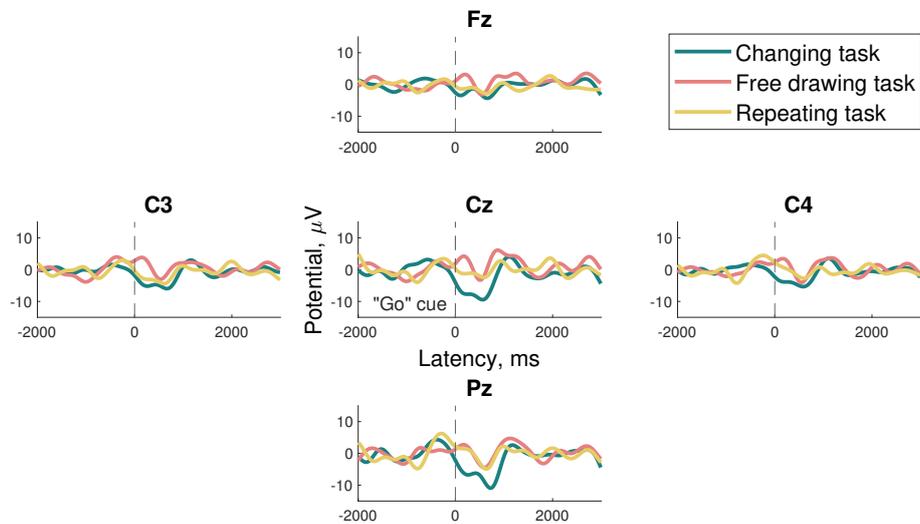
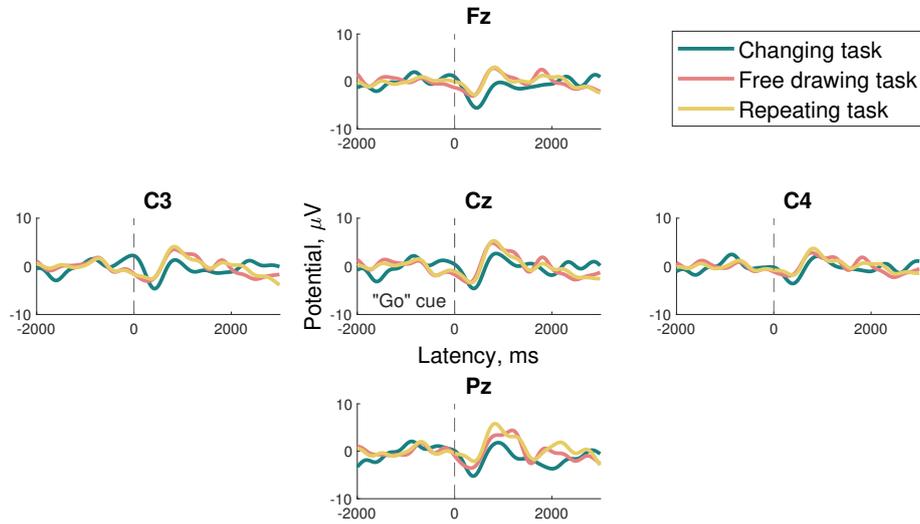


Figure A.2: MRCPs of test subject S1 at five electrode locations when performing the task with passive robot. The results of experiment set one and experiment set two are shown in top and bottom panels respectively. Each panel contains MRCPs from three tasks. The MRCPs were calculated over $[-2 \ 3]$ s time window locked to "Go" cue that is marked with a dashed line.

(a) Experiment set one, test subject S3



(b) Experiment set two, test subject S3

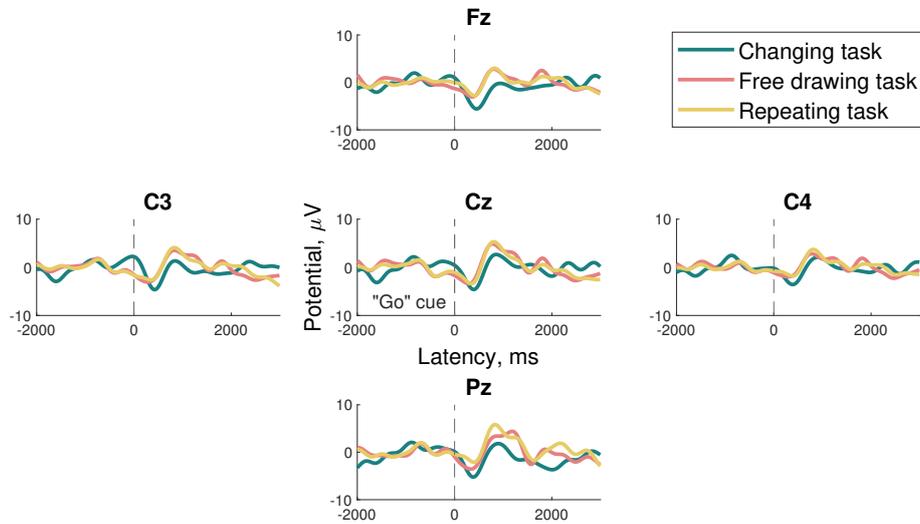
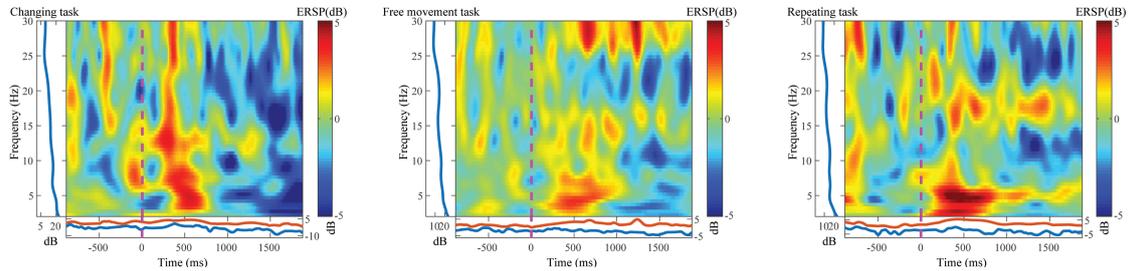
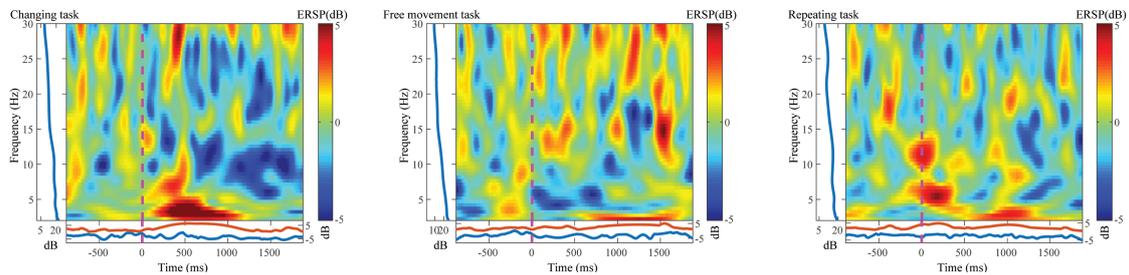


Figure A.3: MRCPs of test subject S3 at five electrode locations when performing the task with passive robot. The results of experiment set one and experiment set two are shown in top and bottom panels respectively. Each panel contains MRCPs from three tasks. The MRCPs were calculated over $[-2\ 3]$ s time window locked to "Go" cue that is marked with a dashed line.

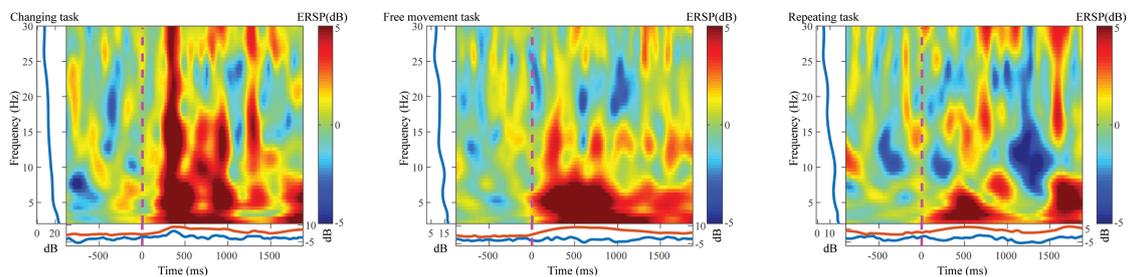
(a) Experiment set one, test subject S1, electrode Cz



(b) Experiment set two, test subject S1, electrode Cz



(c) Experiment set one, test subject S3, electrode Cz



(d) Experiment set two, test subject S3, electrode Cz

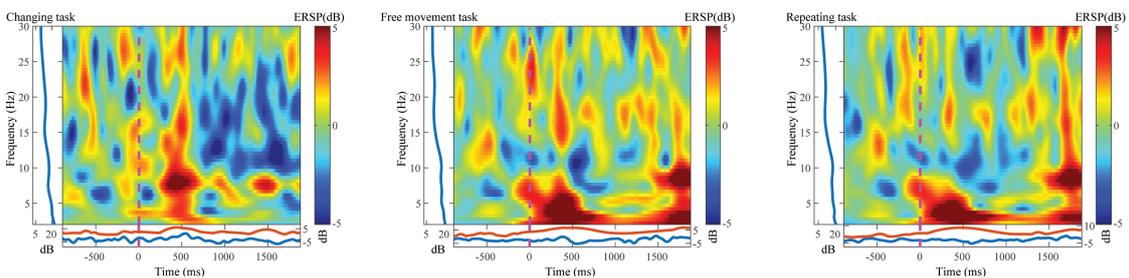


Figure A.4: ERSP images of test subjects S1 and S3 at electrode Cz for three task types. Each panel shows an image of one task power spectrum that is color coded to the bar on the right as well as displayed at the bottom of each panel in a form of an ERSP envelope. The time window is locked at zero to the "Go" cue.

B Appendix: Pseudo code for robot control algorithm

```
1  define Setup():
2      initialize camera
3      initialize ArduinoBoard
4
5  define main():
6      Setup()
7      #initialize multiprocessing to run simultaneously and share a
8      #variable from the parent process to a child process:
9      Process_Detect = parent_process
10     Process_Move = child_process
11     start process_detect
12     start process_move
13
14  define Kalman_Filter():
15     kf = define matrices F, zeta, H, epsilon as in equations
16         (3) and (4) in Chapter 3.1.4.
17     return kf
18
19  define Predict(camera_coordinates, kf):
20     #Update the predicted state from the measurement:
21     kf.correct(camera_coordinates)
22     #One step ahead prediction of the coordinates
23     predicted_coordinates = kf.predict()
24     return predicted_coordinates
25
26  define Detect_Marker()
27     Grab frame;
28     Filter orange color in the image;
29     Dilate and find a contour;
30     Find the center of the contour.
31     return camera coordinates
32
33  define Process_Detect(child):
34     predicted_coordinates = Predict(camera_coordinates)
35     send predicted_coordinates to child process
```

```
36
37 define Inverse_Kinematics(predicted_coordinates)
38     #Get predicted_coordinates, determine shoulder and elbow
39     #angles needed to reach the position and send a command to
40     #the robot.
41     #l1 - length of link between the shoulder and elbow
42     #l2 - length between the elbow and robot end effector
43     elbow_angle = arccos((x_predicted)^2+(y_predicted)^2 -
44         l1^2 - l2^2) / (2 * l1 * l2))
45     shoulder_angle = arctan(y_predicted/x_predicted) - arcsin(
46         l1 *sin(elbow_angle)/(x_predicted)^2+(y_predicted)^2)
47     write elbow_angle to elbow motor
48     write shoulder_angle to shoulder motor
49
50 define Coordinate_Translation(predicted_coordinates)
51     Use affine transformation to translate predicted_coordinates
52     from the camera to robot_coordinates in robot's work space
53
54 define Process_Move(parent_process)
55     predicted_coordinates = receive from parent_process
56     #translate predicted_coordinates to robot coordinate system:
57     robot_coordinates = Coordinate_Translation(predicted_coordinates)
58     Inverse_Kinematics(robot_coordinates)
```