Eduardo Ruiz Ramírez

Master's Degree in Automatic Control and Robotics

September 2016

Master's Thesis

Master's Degree in Automatic Control and Robotics

Control of a hand prosthesis using mixed electromyography and pressure sensing

MEMORY

Author: Eduardo Ruiz Ramírez

Supervisors: Cecilio Angulo Bahón and Claudio Castellini

Call: September 2016

ETSEIB

Escola Tècnica Superior
d'Enginyeria Industrial de Barcelona

UPC

CONTROL OF A HAND PROSTHESIS USING MIXED ELECTROMYOGRAPHY AND PRESSURE SENSING

EDUARDO RUIZ RAMÍREZ

Master's Degree in Automatic Control and Robotics

Universitat Politècnica de Catalunya
Escola Tècnica Superior d'Enginyers Industrials de Barcelona (ETSEIB)

Deutsche Zentrum für Luft- und Raumfahrt
Robotics and Mechatronics Center

September 2016

supervisors:
Claudio Castellini
Cecilio Angulo Bahón
submitted:
September 2016

ABSTRACT

During the last years, new technologies approaches have helped to develop realistic robotic hands for prosthetic use. Even so, the strategies to control them (input signals, prediction algorithms) are still limiting a complete match between the robotic hand and the real hand movements and behaviors.

On this thesis, two different input signals (FMG and sEMG) were evaluated. From this analysis characteristic properties from each kind of signal were obtained, related with wrist and hand movements. In this way two different learning methods were implemented for the first time on robotic hand research. The goal of these two methods was to combine both kind of input signals, supported by the feature analysis previously done, in order to improve the movements prediction performance. The methods' performance were compared with the separate input signals methods, so the improvement could be measured.

Both mixing methods presented better results than the single input signal ones. These results along with other considerations defined, could lead to a robotic hand performance improvement from different perspectives.

ACKNOWLEDGEMENTS

First of all, I would like to thank God for the experiences lived, for the opportunities given. God's timing is perfect timing.

In the same way I want to thank my parents and my sister for all their support during the two years abroad. In spite of the ten thousand kilometers between us, they were always at my side.

I want to express my gratitude to Dr. Claudio Castellini, not just for giving me the chance to develop this thesis on DLR, but for all the knowledge shared, the advices given and the support provided.

Thanks a lot to all the DLR's people who contributed in the development of this project. Specially to Mathilde Connan and Bernhard Vodermayer for their collaboration on the first experiment and to Markus Nowak for his collaboration on the second experiment.

Furthermore I want to thank Dr. Cecilio Angulo Bahón, for all the help given since the beginning of my Master's until the supervision of this work.

I also want to thank Tamara von Sawitzky for reviewing this thesis even she did not have to do it. It came as a great relief.

Last but not least, my deepest gratitude to CONACYT for giving Mexican students the opportunity to rise above ourselves, having as goal the development of our country.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AVG Average

CV Cross-validation

DLR Deutschen Zentrums für Luft- und Raumfahrt (German Aerospace Center

ELM Extreme Learning Machine

ENS Ensemble

FMG Forcemyography

FSR Force-sensing resistor

GUI Graphical user interface

IMU Inertial measurement unit

iRR Incremental ridge regression

iRRRFF Incremental ridge regression with random Fourier features

NASA National Aeronautics and Space Administration

nRMSE Normalized root-mean-square error

OMG Optical Myography

PC Personal computer

RFF Random Fourier features

RMC Robotics and Mechatronics Center

RMSE Root-mean-square error

SD Standard deviation

sEMG Surface electromyography

SUS System usability scale

TLX Task Load Index

US Ultrasound

1 INTRODUCTION

Around 49 million people are living with limb loss around the world [6, 18]. 35% of them have an upper limb amputation [6], and according to British statistics [15], 2% have a wrist disarticulation (hand amputation).
The most common causes for amputation are vascular diseases and trauma at work or car accidents [6], but also diabetes is taking strength being nowadays one of the main causes too. Furthermore, in war-torn countries losses caused by war effects are a considerable quantity on amputees statistics [32].

The scope of the amputee statistics is most of the times only focused to the United States and, even there, it is limited and sometimes non-existing. Even so, LeBlanc [18] proposes that the global statistics could be scaled for each country according the rate between the country population and the world population. If this approach is applied, then global and country statistics can be estimated (was this the way the very first statistic of this work was calculated).

Through the years, prostheses have been created for replacing the amputated limbs. With the development and improvement of technology, robotic prosthesis appeared. In this way the user can move an automatic device by using signals from the rest of his/her body (bio-signals). Despite this fact, there are not enough users who prefer to use a bio-signal control prosthesis. Just 25% of them are using myoelectric prostheses for instance [16]. The main reason of this gap is the fact that most of the prosthesis control are not robust, they sacrifice robustness for more joint movements accuracy, losing at the same time potential real life applications.

As mentioned in Jiang et al. [16], this gap can disappear by mixing the advantages that both, commercial and academic prostheses, offer: on one hand a robust system, and on the other hand an intuitive control. This gap is the starting point for upper limb prosthetics control research, based on the experimentation with different kind of sensors or acquisition methods.

The replacement of a human limb with an artificial one implies several challenges. The human involved needs to interact with the prosthesis the same way with a real limb. The main key for a successful interaction is communication [9]. Communication can be understood as the link between what the human wants to do and what the prosthesis can do, in other words how similar the artificial limb compared with a human one is. Although in the last years there have been significant improvements in the electronics, mechanics and materials of the prosthetics limbs, managing to build for instance robotic hands with 22 joints that approximates to the human hand dexterity [9], if this communication feature is not the proper one, then the patient would have a limited performance using the robotic hand, reducing completely the real life application of the prosthesis.

This is the main problematic in prosthetics field, great mechanical designs close to a real hand, with innovative materials built in innovative ways (3D printed for instance), but they cannot be controlled properly by users who need more robust and real behavior from their new limbs. Hence, instead of really working as physiologically replacement of the human hand, most of the prosthesis work only as tools, to provide some functions that were lost (like grasping), but then, after the desired task performance, they could be ignored, as they are not working for something else. Despite those facts actual research is doing a hard work to focus the field progress toward the creation of true hand replacements [35].

The main goal of robotic prostheses development is to help people with limb loss to recover at least some of their usual tasks and activities that were used to do before their loss. Just as an example, 36% of people living with limb loss experience depression [6]. Here lies one of the main motivations for this project.

The German Aerospace Center (DLR) in Oberpfaffenhofen, by its Robotic and Mechatronics Center (RMC), has been working in the research of upper limbs improvement, has been working to reduce the gaps, to overcome the challenges and limitations previously defined. The development of this project is just a small part of the entire research.

1.1 OBJECTIVES AND HYPOTHESIS

Following the motivation line, and based on the work previously done on DLR, the next objectives were established for being accomplished by the current project:

Test and design a software suite for acquisition and control devices.

Design two experiments:The first one to make a comparison between two different bio-signals (FMG and sEMG). The second experiment has as objective the effectiveness test of a multi-sensor control, as opposed to the single-sensor one used on the first experiment.

Identify from the first experiment, what type of sensor (sEMG or FSR) has better response to different hand and wrist movements.

Apply the results of the second experiment to predict movements and test the results in a 3D hand model.

Improve the user interface, adapting it to the needs of the experiment.

Finally, improve the machine learning algorithm for better prediction.

All these goals are focused to deal with the limitations exposed previously in this section. The main hypothesis of this porject is that the hand movements prediction performance could be improved by combining the defined bio-signals from the two kind of sensors, as those signals are presenting different features and behaviors.

It is important to remark that due to the unavailability of the robotic prosthesis hand, this project is limited by the tests on a 3D hand model, however the project is still focused on the goals appliance on a robotic hand. In fact it is relatively simple to move from the 3D hand model control to a real prosthesis one.

1.2 thesis structure

This thesis is composed by seven chapters.Chapter 1 is the current introduction.

Chapter 2 is a literature review showing the background and state of art of the hand anatomy, robotic hand prostheses, sensing methods and machine learning algorithms.

Chapter 3 presents the state of art of the work previously done on DLR, as much of the work used on this project was already used in other projects. This chapter is important as it is a direct background of the project, definitely needed for a proper understanding of it.

On Chapter 4, the first experiment is executed and evaluated. This experiment is considering the use of sEMG sensors and FSR in order to compare their signals and identify special characteristics and properties for each of the sensors. This chapter also contains a study about the usability and user experience of the device.

On Chapter 5, on a similar way as in the previous chapter, the second experiment is exposed, showing first the work done for a better experiment design, then is presented the experiment definition, the results and the discussion.

Chapter 6 exposes the social, environmental and economical commitments this project has. Additionally a cost analysis is evaluated.

Finally on Chapter 7 the conclusions and recommended future work are presented.

Going further, the project work was planned to be done on six months, dividing it into four stages:

Familiarization with the work previously done on DLR, which corresponds to Chapters 2 and 3, this stage was performed on two months.

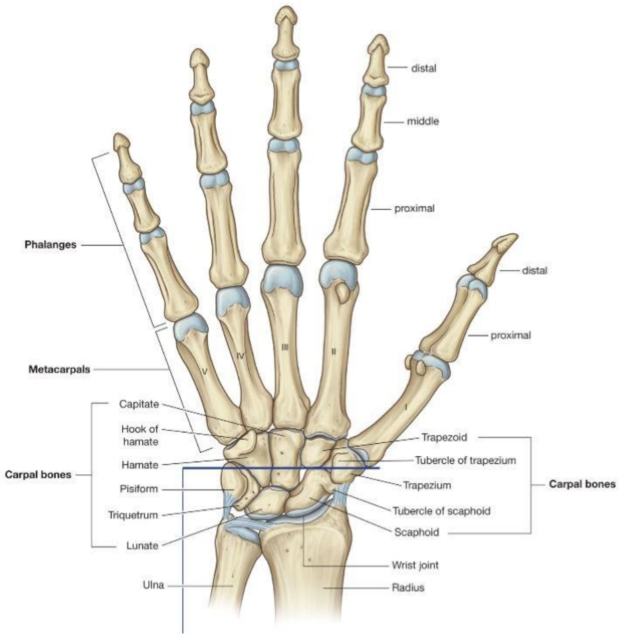
Feature analysis of the two types of sensors (sEMG and FSR), the goal of this stage was the development of Chapter 4, along with it a publication was written with the research findings. This stage took one and a half month.

Machine learning development for prediction performance improvement by combining sensors, the work done in this stage is documented on Chapter 5. It was performed on two months.

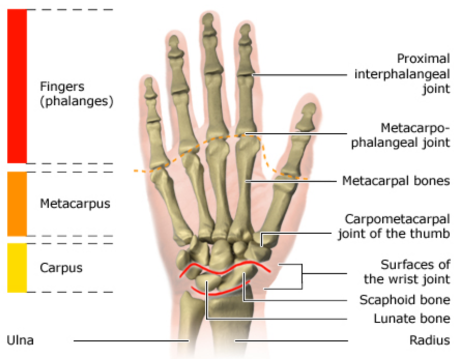
Thesis Documentation, even though the thesis writing was done during the 6 months, supported also by the publication of the second stage results, the last two weeks of the project were used for a final arrangement.

2
STATE OF ART

This chapter is a literature review showing the background and state of art of the hand anatomy, robotic hand prostheses, sensing methods and machine learning algorithms. All this information is needed for the project development.



(a) Hand bones [11]



(b) Hand joints [27]

Figure 1: Hand bones and joints

2.1 HUMAN HAND STRUCTURE AND MOVEMENTS

The anatomical explanation is based on the work completed by Chen [5]. In prosthetics field, the system of the hand considered, at least in most of the prosthesis, is the skeletal system, considering that there are also integumentary, muscular, lymphatic, nervous and cardiovascular systems. The skeletal system includes bones and tendons. For the purposes of this project, the main goal is the performance of wrist movements, so there will be considered also the forearm and wrist skeletal systems. The interested skeletal region goes from the elbow passing through the forearm, where the Ulna and Radius bones are located, then goes across the wrist, which works as a connection for the hand and forearm (the Scaphoid, Lunate, Triquetrum, Pisiform, Hamate, Capitate, Trapezoid and Trapezium bones are located in this part). Finally there is the hand, which is first composed of five Metacarpals (palm bones) and five Digits. Each Digit is built by three bones, the Proximal, Distal and Middle phalanxes; except for the thumb that has just two bones.

If the anatomical terms are translated to robot kinematics, then each bone can be considered as a link. Connecting those links, as always, there are joints. Starting with the Radiocarpal one at the wrist, then the Carpo-metacarpal that is connecting the Hamate, Capitate, Trapezoid and Trapezium with the Metacarpals; after that, there is the Metacarpal-phalangeal, that is located between metacarpal and the proximal phalange bones, and finally the proximal and distal interphalangeal joints that separates the phalangeal bones.

This hand anatomy explanation is important for a 3D Hand Model construction, and, in a closer way to this project, with the wrist and finger movements explained in the next paragraph. For a better understanding, Figure 1 shows in a detailed and graphical way the human hand bones and joints. The joints can perform gliding movements, angular movements, circular movements and special movements (like inversion, eversion, protraction, and retraction). This thesis deals with the angular (flexion/extension) and circular (rotation and pronation/supination) movements.

These movements are shown on Figure 2 for both, fingers and wrist joints. Each joint can be able to perform one or more movements, but the joints and movements which this project focuses on are the following:

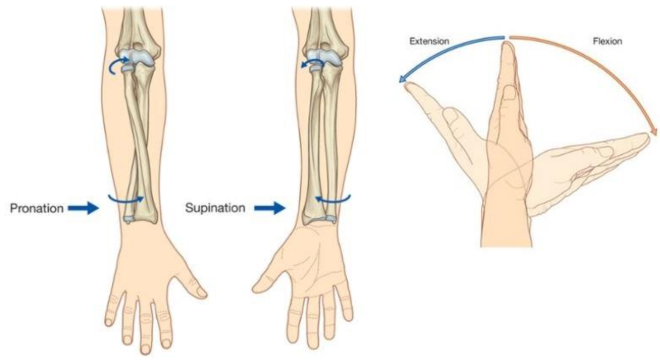
Wrist Flexion/Extension, wrist (Forearm) Pronation/Supination and grasping movement (which includes fingers and thumb flexion and thumb rotation).

Summarizing, the human hand is composed by 27 bones and 33 muscles, which result in a total of 22 DOFs [8]. The hand system is able to perform basically the following tasks, which are the ones that the robotic prosthesis hand tries to emulate [39]: Reaching and pre-shaping, grasping, manipulation with stable grasp, exploration with sensori-motor coordination and gesture expressiveness.

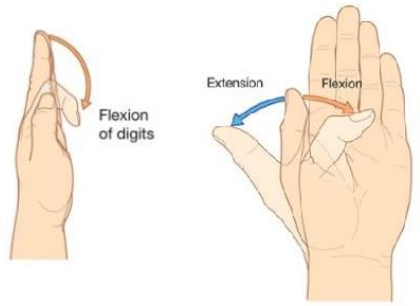
2.2 ROBOTIC HAND PROSTHESIS

In order to control let a user control a prosthesis, Craelius [9] proposes two basic requirements.The first requirement prescribes, that the user must have control over the motor function somewhere in the residual limb, such that he could sense and imagine the manipulation of the lost limb. If the residual muscles are damaged, there is the option of surgery to restore the nerves or reroute them to healthy areas, for example, the hand residual signals could be rerouted to the pectoral muscles, so then the muscular activity is easier to access.

The second requirement implies that a device can acquire and decipher the information generated in the previous requirement. This requirement is more challenging because of the complexity of the human movement control. Our smooth and precise movements are the result of functioning sensor-monitor system, that transforms the result of all senses (tactile, positional, visual) into fine movements via millions of electrical impulses. The



(a) Wrist movements



(b) Finger movements

Figure 2: Hand and wrist joints movements [11]

challenge for the prostheses is to restore the lost sensor-monitor function by the assistance of an artificial limb.

There are four functional issues that need to be considered for the design and construction of a prosthetic hand: grasping, sensing, actuation and control strategy. These four issues should be combined in order to fulfill some of the tasks listed section 2.1. For instance, the hand should provide a natural grasping capability, a natural appearance, secure grasping, and a natural and friendly command interface [4].

GRASPING: Grasping, holding an object, is a complex task because does not involve just kinematic motion, but also force control. This action depends on the goal object geometry, strength, stiffness, and surface finish; as well as on the fingertips features. For a successful grasping, the robotic hand should be robust and compliant, must have control of the force and slippage of the object and also needs to be manipulable while grasping [17].

SENSING: The sensing in a prosthetic hand could be divided in two types: The first one are the sensors used for acquiring body signals for the hand joints movement, this kind of sensors ideally should be non invasive, and actually are present din the next section.
The second type are the internal sensors of the , not just the ones for used in the joints for hard control, but also other sensors that could give feedback

to the user so he or she could feel a more real sensation with the hand, this kind of sensing is called exterioceptive sensing.[17]

ACTUATION: The actuators should occupy as least amount of space inside the hand structure. The most common actuators used are DC motors and minimotors coupled with gear reductions, both of them are used for joints movements. Nevertheless, other actuators have been tried, artificial muscles for example. These artificial muscles are built using pneumatic devices or different materials like shape memory alloy, or ionic polymer metal composite. The problem with the artificial muscles actuators is that they are still under development, facing problems in the size of auxiliary elements and the power of actuation. Finally need to be considered the power storage and supply issues.[17]

CONTROL STRATEGY: The main goal of the hand prosthesis, is to move and react as a human hand. In order to reach that goal the robotic hand should apply control strategies that convert the sensed signals into output signals that are able to move the joints and links and produce natural flow and forces.

According to Naidu et al. [23], there are two kind of control strategies: the hard and soft ones. The hard computing includes: multivariable feedback, optimal, nonlinear, adaptive and robust control techniques; applied, for instance, in impedance control (moment of inertia, joint stiffness or viscosity of a muscle model). The soft computing involves: artificial intelligence, neural networks, machine learning, fuzzy logic, genetic algorithms. Some applications of this techniques are the prediction of the joints movements depending the sensing of a body signal. After all, the final goal for robotic hand control is the fusion of both approaches to achieve a more realistic behavior of the hand.

2.3 NON-INVASIVE SENSING METHODS

Different sensing methods are described in this chapter apart of the chosen sEMG and FMG. These last two methods are compared with each other in subsection 2.3.6.

2.3.1 *sEMG*

Electromyography is evaluating and recording the electrical manifestation of the human muscle movements, it is possible to get information about the neural signals that control the muscles from the central nervous system [16]. In prosthetics the sEMG signals are used due to the relation between the neural drive to the desired movement and the power intensity of the signal. The sEMG signals are acquired by surface sensors located in the remnant muscles of the stump, they are processed and after that activate certain prosthetic functions or movements. Although there are researches with other bio-signals for the prosthesis control, the surface sEMG signals have been the only one that have worked with practical uses of multifunction upper limb prosthesis since the middle of the last century, consequence of the easy acquisition of the signal (non invasive, commercial sensors), and it is likely to be a trend in the near future [16].

2.3.2 *FMG*

Forcemyography is a relatively unexplored technique that captures, by using Force Sensor Resistors(FSR), the expansion/contraction of the muscle surface by registering its pressure. Before they were applied in the prosthetics area, FMG signals were used to monitor physical activity. They were able to distinguish different limb movements and predict them. Although this analysis was done in a non real time mode, those applications showed the feasibility of using FMG for muscle movement capture.[38]
The main difference between FMG and sEMG is that the first one is sensing mechanical signals, which represents a low-bandwidth and degraded version of the neuroelectrical signals, due to this fact, it cannot define activities for individual muscles [36].

A force sensor resistor (FSR) is made of a polymer film that decreases its electrical resistance when a force is applied over its sensing area [38]. That means that the output voltage is inversely proportional to the force applied, additionally, the output response is not linear, presenting different curves shape depending the resistor used in the voltage divisor of the signal conditioning circuit.

Until the date of the publication of [36], FMG was the only neuromuscular imaging method that had worked with simultaneous, multifunctional and multi-DOF control of prosthetic hands. But what is really important about the FMG is that it needs to fulfill practical needs. For example the sensors and wearing devices must be reliable and durable, easy to calibrate and compatible with control and learning interfaces so earlier works with other sensors could be applied for FMG and also the new work done with FMG could be used in other devices. Preliminary trials have demonstrated that the FMG devices are easy to wear and calibrate and have had a good performance with several damaged amputees residual limbs [36].

2.3.3 *Ultrasound imaging*

The ultrasound waves have a frequency over 20kHz. They can penetrate soft tissues without harming them, due to this fact, they are used a diagnostic tool to visualize innards of the human body.

Ultrasound Imaging (US imaging) builds 2D- or 3D-live images from the interest body parts by the partial reflection of the waves at the tissues with different acoustic impedance. Until now US imaging does not have any known side effect, so it is used in most hospitals.

US imaging detects conditions of the muscoloskeletal system, the information obtained is starting to be exploited to build a Human-Machine Interface with clear and potential future applications, like advanced hand prosthetics control, becoming a great competitor of the already established non invasive HMIs, such as sEMG.

There is a relation between the fingers joints and the spatial features extracted from US images of the forearm, so it is possible to reconstruct the hand configuration using those US images. The information extracted is based on the forearm position so it does not depend on the subject's speed movement, having its only limitation in the hardware and software features [31].

2.3.4 *OMG*

Thanks to the high quality and cheap cameras that exist nowadays, the optical motion tracking and image processing are applied in more complex tasks, for instance rehabilitation devices for disabled.

Optical Myography (OMG) is based on the idea that, by using optical tracking and recognition, an amputee's intended movement can be reconstructed, just by looking at the muscle activity deformations of the user's stump and associated them with movements the subject is trying to perform. The idea stems from the principles of pressure and tactile sensors.

The first drawback is the fact that human skin provides very little texture features for computer vision applications, so an artificial tag attached to the human limb that offers robust and reliable texture fixtures is needed for the detection and posterior tracking. Even so the tracking of precise features problem still exists.

At the end the use of plain optical recognition to track the forearm could improve the interaction between an amputee and a virtual world (virtual environments, robotic prosthesis) without using any wearable device [24].

2.3.5 *IMU and accelerometers*

Another option to measure arm movements is the use of accelerometers, there are commercial wrist accelerometers that can give measures of the arm movements reading their multidirectional acceleration data. The disadvantages of this type of sensors are, first of all, that they cannot provide real-time feedback to the user, nor are able to capture information about the hand [38]. So for the purposes of this project the use of accelerometers is limited. These sensors can be used only for the arm tracking and to improve the prediction response [14].

2.3.6 *Discussion*

Among the non-invasive methods, other than the ones here presented, can be found electroencepahlography (EEG), mechanomyography (MMG), functional magnetic resonance imaging, etc [12, 19]. Anyway, recently the research community is pushing the development and research of multi-sensing methods.

Even though the sEMG sensors have been the most used input in the research of prostheses control methods, it is difficult to preform a robust control solely using this kind of sensors. The point here is that several studies and works have demonstrated the potential of the sEMG signal processing to predict limb movements, but none of them has shown that this features have been implemented in a real and practical prosthetics application [36]. It is necessary to mix different kind of inputs to control complex prosthetic devices and also give them more intelligence and autonomy, so the the operation can be simple to the user [16].

Additionally it is important to keep the device in an accessible price for the user. On one hand there are the sEMG sensors which can reach, at least the ones used for this project, 150€each sensor, on the other hand there are the FSRs, which each sensor can be purchased for less than 5€each [29].

Furthermore, on the same work done by Ravindra and Castellini [29], a comparison between FMG, sEMG and ultrasound imaging is done, discarding the ultrasound imaging due to its bad performance wearability and cost, but putting sEMG and FMG in similar levels, giving some advantage to the FMG cause its stability and the already mentioned low cost.

Considering these facts, FMG and sEMG were the signals chosen in the previously DLR's projects and hence in this one. Afterwards it is expected to go further through the signals comparison already done.

2.4 MACHINE LEARNING ALGORITHMS

The algorithms presented in this section have been used for prosthetic hand movements prediction, except for the ensemble learning, which is firstly proposed in this field as a method for mixing different models. Further information about the algorithms can be found in the references pointed in each subsection.

2.4.1 *Pattern classification*

The Pattern Classification applied to the myoelectric control is based on the assumption that the obtained signals are distinguishable and repeatable patterns depending the muscular activation performed. The drawback of classification is that the accuracy and usability are inversely proportional. The reason of this is that while the human limb usually has simultaneous and proportional movements, given by the articulations of the limb's DOFs, the number of patterns is limited, so what is obtained from the pattern classification is just an approximation of the real parameter space. This approximation produces two problems, the first one is that only one class can be selected as a decision, the second one is that proportional control is not obtained directly from classification, actually to try to perform a proportional control affects the accuracy of classification [16].

2.4.2 *Extreme Learning Machine*

The Extreme Learning Machine (ELM) is a batch regression learning algorithm for training one kind of Artificial Neural Network, the Single-Hidden Layer Feedforward Network (one input, one hidden and one output layer). The algorithm is similar to the non linear regression explained in the next subsections, it performs linear least squares on a projected feature space [33].

ELM has been used for FSR signals processing and learning [38]. This algorithm has similar and sometimes better performance compared to the Support Vector Machine algorithm and Artificial Neural Networks, but with the advantage of being simpler and having a faster learning speed. This features are important if the learning is desired to be applied in portable devices, so it could be programmed using a microcontroller. If the sensor's position changes then the learning can be redone fast without any problem [38].

2.4.3 *Linear Least Squares Regression*

Linear Regression is a supervised learning algorithm, used to train relationships between an input variable and continuous outputs values from defined data sets, so future output predictions can be performed using new inputs. This relationship is represented as a function $f : X \to Y$, that should be linear, so $f(x)$ is represented as a line (for one dimension), as a plane (for two dimensions) or hyperplane (for three or more dimensions).

The main goal of this learning approach is to find a vector that contains the slopes of the linear function. The method for this vector search is the least squares algorithm. It works based on the idea of minimizing the sum of the squared distance between observed values, ground truth, and the prediction, $f(x)$. This algorithm is explained as a background for the Random Fourier Features Regularized Least Squares method, that actually is the one that has been used in research for hand movements prediction [33].

2.4.4 *Random Fourier Features Regularized Least Squares*

The Linear Least Squares Regression algorithm can be translated to a non-linear regression approach in two different ways. The first one by building a new model product of the mixture of linear models (each one using distinct input-dependent weighting functions). The second way to reach the non linearity consists in project the input space into a feature space using a non linear kernel function, then a Linear Least Squares Regression is performed in this feature space. For prediction is used a model built by a weighted sum of the kernel functions.

A way to build those kernels is by using randomly sampled cosine features, as defined in the Fourier Transform Theory. This method called Random Features Regularized Least Squares was first purpose by Rahimi and Recht [28] and improved adding an incremental feature by Gijsberts and Metta [13] [33].

This approach has been used in several papers for robotic prostheses control using biosignals as input data, for instance the one done by Gijsberts et al. [14], being this algorithm the one previously used in DLR research, which was the base of this project.vHow this method has been applied in DLR's projects is explained in deeper way on Chapter 3.

2.4.5 *Ensemble Learning*

As was defined in the Introduction, one of the goals for this project is to mix both kind of sensors (sEMG and FSR) to improve the prediction performance. In the Machine Learning field, this improvement could be performed by the appliance of the Ensemble Learning approaches.

In a normal learning algorithm just a single hypothesis is proposed, which explains the data in the best way. Ensemble Learning on the other hand construct a set of hypotheses, this hypotheses are voted to predict the new data samples. Experimental tests have proved that ensemble methods are often much more accurate than single models.[10].

Ensemble Learning includes three methods that are applied for different purposes. The one interested, as will be explained on Chapter 5, is the stacking one. Anyway, all the three algorithms are presented in the next paragraphs.

2.4.5.1 *Bagging*

Was the first effective ensemble learning method and it is one of the simplest ones. The method uses different "versions" of the training data set by using bootstrap. Each of this versions are used to train a different model (usually are classification models, but also could be used with regression). The outputs of the models are averaged (for regression models) or voted (for classification models) in order to combine them. This method is only useful when unstable non-linear models are used [30].

2.4.5.2 *Boosting*

Boosting is the most used ensemble method, and actually one of the most important findings in the learning topic in the last twenty years. Its main idea is to create a weak classifier, which means that the accuracy on the training set is just a little bit better than a random prediction. After that, iteratively, the model is trained using a weighted training set, where the poorly predicted values from the previous iteration are more weighted. After all, the successive models are weighted according their success and the outputs are voted/averaged and combined, creating in this way the final model. The most popular boosting algorithm is the AdaBoost [30].

2.4.5.3 *Stacking*

Stacking method is most of the times used for combining two or more different machine learning methods. It works by dividing the training set in folds, so a cross-validation can be performed, training in a certain number of folds and validating in the rest, then change the folds involved, so after all the training all the folds are trained and validated the same number of times. For each repetition, the prediction performed with the validation set is stacked, creating a new input space for a new machine learning appliance, usually a higher level learner [30].

3 PREVIOUS WORK ON DLR

As was mentioned in the Introduction, this project is the continuation of the work previously done in the DLR related with upper limb prostheses control. Most of the devices and software used for the project's experiments were already developed. This chapter presents a more specific background in which this work was directly based.

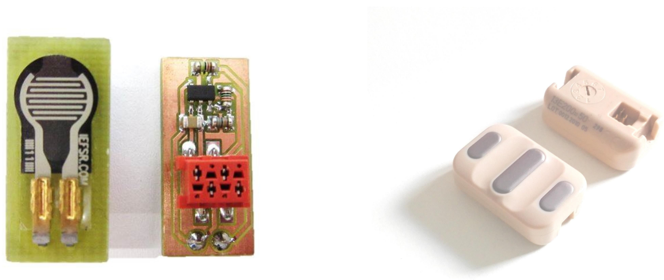
This chapter is structured in 4 parts, the first one explains the acquisition device used, a second part explaining the 3D Hand Models used , then a section that shows the GUI applied and finally a fourth part that presents the learning approach applied.

3.1 ACQUISITION DEVICE

3.1.1 *Sensors*

For the previous projects done in DLR, sEMG sensors and FSRs were used.

For the sEMG acquisition signal, the Ottobock *MyoBock 13E200* sensors were used (Figure 3). This sensors already provide an on-board amplification, rectification and filtering, in order to give a high-quality signal. Additionally this sensors are the standard ones used for clinical prosthetic applications, furthermore they are commercially available.



(a) FSR [8]

(b) sEMG sensor

Figure 3: FSR and sEMG sensor

For the FMG signal acquisition, the FSRs 400 Short (Figure 3) manufactured by Interlink Electronics were used . They are made of a polymer thick film, having a 5.6 mm diameter sensitive area. They have a large force sensitivity range (0.2N-40N), despite of these facts, FMG also presents some drawbacks: a non-linear behavior, a non-negligible hysteresis at high force appliance, and no guarantee of repeatability through sensors. Even so, it has been shown that this drawbacks are reduced for small forces (0N- 15N), so the sensors behavior is comparable and almost linear [29].

In order to condition and process the FSR signal, a circuit board was made. It consists in a voltage amplifier which voltage output corresponds to: $V_{out} = \frac{R_2 V_{CC}}{R_1+R_2} - \frac{R_1 R_4 V_{CC}}{R_1+R_2} \times \frac{1}{R_{FSR}}$, where R_{FSR} is the lowest admissible resistance and its equal to $R_{FSR} = \frac{R_1 R_4}{R_2} = 6k\Omega$, this value corresponds to a theoretical maximum force applied to the FSR surface of 3.33N.(Figure 4)

Both kind of sensors are arranged in Velcro straps, so can be created sensors bracelets that are tightened to the user's forearm. In order to allow the attachment of the sensors to the strap, 3D-printed housings have been designed. These housings are made of flexible thermoplastic polyurethane (Figure 5). They are designed with braces to allow them sliding on the strap, so the sensor position can be individually adjusted, changed and maintained,

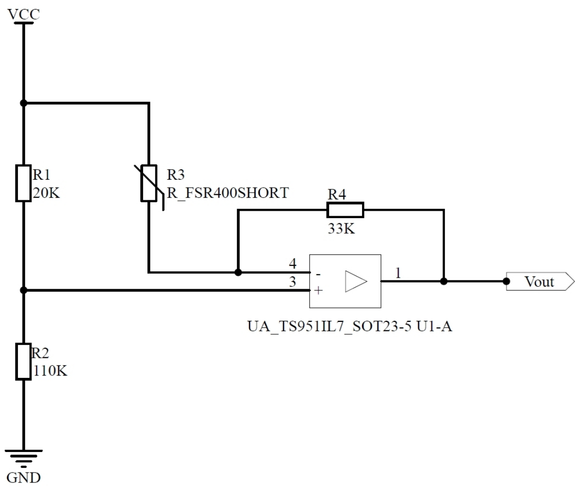
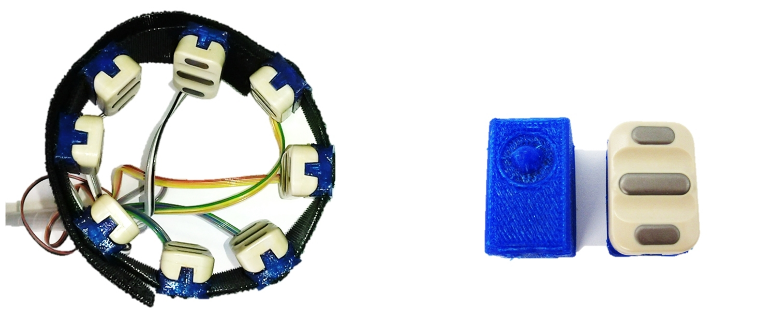


Figure 4: FSR amplifier circuit diagram [8]

having a big range of possibilities for the sensor arrangement. This feature allows any combination of FSR and sEMG, so a complete forearm circumference signal acquisition could be performed. For the experiments done, the sensors were arranged in a *low-density surface electrode layout* or *uniform electrode positioning* approach, that basically puts the sensors in an evenly spaced placement around the forearm (Figure 5). This approach has already been proved to be effective and useful for robotic hand prosthesis in a number of previous publications.

The FSR housings are not only useful as a retainer, but also have a structure that concentrate the force produced by the muscles on the FSR sensitive area. This concentration is possible thanks to a cone shape in the housing that points inwards to the sensor sensitive area. The real impact of this structure in the force signal has not been object of study, so it is assumed that it increases the signal stability, and hypothetically could add some mechanical filtering to the signal. A detailed information about the sensors and housing design can be found in [8]



(a) Bracelet with sensor arrangement

(b) Sensors' housings (Left FSR, right sEMG)

Figure 5: Sensors' housings and bracelet arrangement [8]

3.1.2 *Data acquisition wireless device*

The device used for the sensors' signals acquisition(Figure 6) was designed and built in DLR and is presented in [8]. It is sending the sensors' data to a GUI in a wireless way, by using Bluetooth communication. Going deeper in its structure, the device consists, first, in a multiplexing stage, so a sensor number higher than 15 (microcontroller analog inputs) could be used. The multiplexed signals are then sent to a microcontroller, that converts the analog signals into digital ones. In this way, the signals could be sent via serial communication to the Bluetooth chipset, which in turn sends the signal to the device where the GUI is running. This chipset allows data emission and reception.

In fact this device can also send the prediction obtained in the machine learning to a robotic hand (iLimb[1]), so its Bluetooth chipset is working in a two communication channels way (one for the iLimb, and the other for the GUI). However as will be later explained, for this project purposes, the device is used just for data acquisition, so it is working just in with the GUI communication channel. It is important to remark that the device is just a communication/acquisition device, the training and prediction tasks, the data saving, and other signal condition features are performed by the device in charge of the GUI.

Figure 6: Wireless device

3.2 3D HAND MODEL

There are two ways to test the prediction and have visual feedback of the performance. One way is to apply the predicted output to a real prosthetic hand, the other one to use a 3D hand model. The latter option is also used for having a visual feedback during training.

1 Robotic hand manufactured and distributed by Touch Bionics, for more information please consult the product datasheet [**Bionics2016**]

The values representing the DOF of the training 3D hand model, which will be called stimulus in this thesis, are acting as the ground truth for the machine training and for performance calculation.

The hand model was designed in Blender. To control it, the degrees of freedom from the stimulus (9 or 12) are coupled with the real hand degrees of freedom (26, now including also wrist DOF). Each of the real hand DOF is matched also with a model bone, so for instance the wrist pronation/supination DOF is matched with the forearm bone, or both the finger metacarpophalangeal adduction and the metacarpophalangeal flexion DOF are matched to the proximal phalange bone.

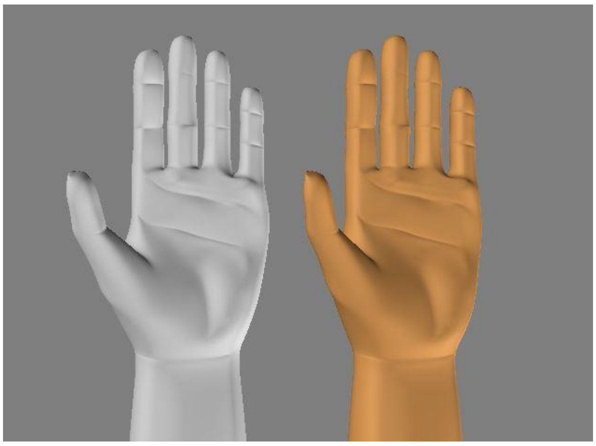


Figure 7: 3D hand model (Left:Training/performance. Right:Prediction)

On the screen there are two hands displayed. The purpose of displaying two hands is to use one for a stimulus ground truth for training or for performance measurement, while the other hand is used for prediction, so the user can have a visual feedback about what the machine learning is performing. The communication between the model and the GUI is done using a UDP protocol, so for each hand model (training / performance and prediction) a UDP Port is assigned.

3.3 GUI

In this approach, a mobile phone or a PC can be used for visualizing the GUI. For each of the options was developed a different GUI, with common and specific tasks.

3.3.1 *PC GUI*

Basically the GUIs previously programmed on DLR, offered a training and prediction stage, fingers and wrist movement selection, data storage, and communication to different applications (iLimb prosthesis, the 3D hand model), it allows also to use not only the sEMG sensor or FSRs, but additionally other types of sensors for different kind of measurements, such as fingertip forces or position tracking. The language used for the GUI programming was C#

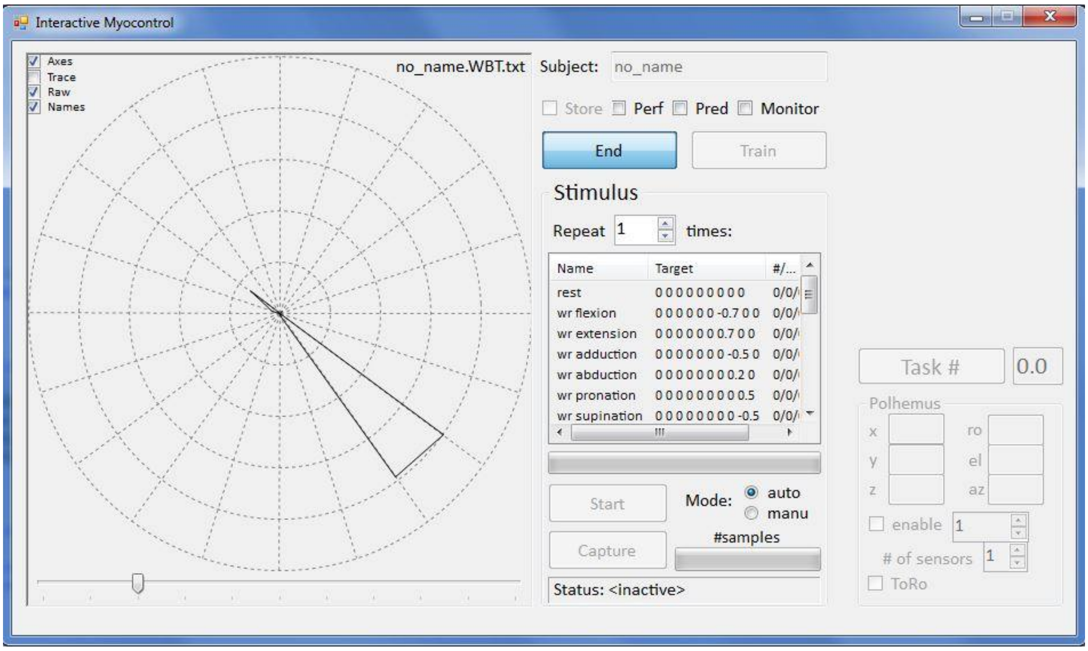


Figure 8: PC GUI

The GUI is shown on Figure 8. As could be appreciated, it has a Radial Plot that shows all the sensors' values, the name of the subject can be assigned, and also has options for storing the sensors data and the stimulus one. The GUI provides information about the data capture rates, so the user can know if the stimulus or the acquisition device is having any delay or lag. It is important to say, that this code was designed for working not only with the wireless device already presented, but with the sEMG bracelet *Myo* and with a PC signal generator (so the GUI could be tested without any acquisition device).

Continuing with the GUI, the user can select the actions to train from a list of possible ones, and the number of repetitions the training will perform. This training could be done in automatic mode (the data capture for training is done in a previously established time) or manual mode (the user decides how long the capture stage will last). The GUI is user-friendly, in the way that it displays in the radial plot also the representations of previous repetitions, so the user can have a reference when he or she is performing the movements. Additionally the GUI shows two progress bars, one showing the progress of the entire learning stage, and a second one showing the progress of the data capture stage. Finally the GUI allows the use of other sensors for hand position tracking (not used for this project).

Internally the GUI code performs the training and prediction of finger and wrist movements considering all the sensors of the device as input data and using the incremental Ridge Regression (iRR) algorithm and its non linear representation, the using Random Fourier Features (iRRRFF) algorithm, as machine learning approaches (explained in the last section of this chapter). Apart of that, the code creates and manage the communication between the Hand Model and the PC using the already mentioned UDP protocol. In this way the data used by the code could be sent for visual feedback to the hand model.

This GUI was used for the second part of the thesis that is presented on Chapter 5.

3.3.2 *Mobile phone GUI*

The code for this approach was developed in the Xamarin Studio software, which allows the development of Android applications using C# (and other languages) syntax, so the code structure previously done could be used for the mobile phone application.

This GUI, shown on Figure 9, is a simplified version of the PC one, mainly because of the mobile phone limitations (processor speed, screen size). Its advantage over some previous C# codes is that this GUI was designed specially for the wireless device, which main purpose was the use not only of the sEMG sensors but also of FSRs, that is why the interface included some features for the FSR management.

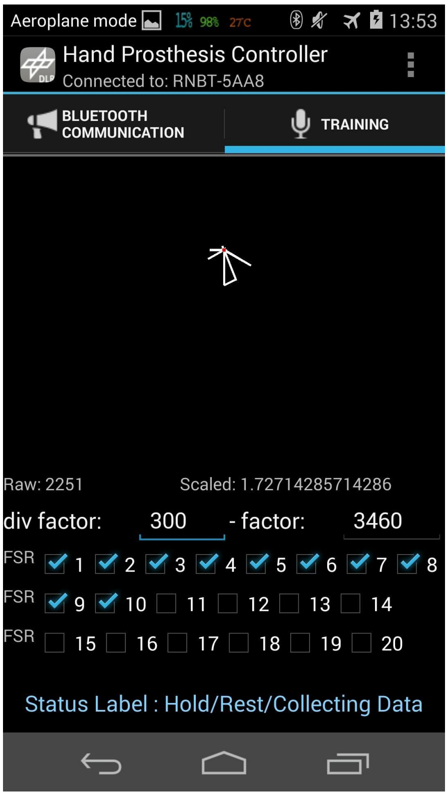


Figure 9: Mobile phone GUI

The GUI has, as the PC one, a radial plot for sensor readings, a list for the actions selection, the modifiable number of repetitions, the storing feature, and some text labels showing instructions for the user. Internally the training and prediction were performed using the same iRRRFF approach.

The difference compared with the PC GUI, as was mentioned, is the possibility to manage in different ways both kind the sensors. First of all the user can select the number of sensors and which of them are FSR. It is important to differentiate them due to the inverse behavior compared with sEMG, so then can be applied some kind of signal conditioning to convert the FSR signals to a similar behavior compared with the sEMG ones. Thanks to the sensors selection features, the bracelets can be customized, changing the number of sensors on them or their arrangement. While the system is used in the training stage, there is a stimulus signal generated, that represents the movement that is followed by the user. This signal has as elements as DOF (12 in this case), and the element's values go from 0 to 1.

The drawback of this interface, at least when it was used for this project experiments, was the lack of visual feedback for the user. First of all there was not a display of previous repetitions (as in the PC GUI), also there was not possible the connection with the PC to move and work with the 3D Hand Models. The only visual feedback was the normal sensor data polar plot, and the text labels indicating the movement to perform. This fact could produce some uncertainty in the measures, nonetheless this uncertainty is reduced depending the user performance. After all this lack of visual feedback did not affect the experiment performance, as will be explained in the next chapter.

This GUI was used for the development of the experiment presented on Chapter 4.

3.4 MACHINE LEARNING

There has been used a machine learning approach to predict the prosthesis movements for a number of hand prosthesis works, experiments and papers in DLR, this approach is the incremental Ridge Regression Random Fourier Features (iRRRFF), already introduced on Chapter 2. In order to understand how it is applied to hand prosthesis control, it is necessary to start with the Ridge Regression(RR) approach.

RR is a regularized variant of least squares regression. It builds a linear output model of the form:

$$f(x) = w^T x \tag{1}$$

where x represents the input space (sensors values) and w is a weighting vector, trained from a set of input/target pairs (training set) previously collected. The way to train w follows the next equation:

$$w = (X^T X + \lambda I)^{-1} X^T y) \tag{2}$$

Here X denotes the input training set, and y is a vector collecting the target values for each collected sample, this vector is considered as the ground truth and is obtained by the stimulus values gotten in the training stage. The regularization coefficient λ is normally set at a standard value of 1.[29]

One of the motivations found to use RR algorithm is that it allows incremental updates (from this fact the incremental name of the algorithm) of the model without the storage of any training samples. A more detailed explanation of the RR algorithm (specially the incremental part) could be found in Gijsberts et al. [14].

For sEMG and FSR prediction, a linear approach is not enough, so RR is limited due its linearity, this limitation is overcome by using the Random Fourier Features applied to least squares regression (in this case the regularized RR)also already explained on Chapter 2.

The goal of this improvement is to perform the algorithm implicitly in a n-Dimension feature space, using for that kernel functions. As standard kernel functions(Radial Basis Function for instance) requires high computational time due to its potentially infinite dimension, a solution found was to perform a finite dimensional feature mapping.Rahimi and Recht [28] applied this strategy and proposed to take a finite number of random samples in the Fourier domain.

The Random Fourier Feature is represented by:

$$\phi(x) = \cos(\omega^T x + \beta) \tag{3}$$

It produces an unbiased estimation of a kernel if ω is drawn from a normal distribution with mean 1 and if β is drawn from a uniform distribution from 0 to 2π. Furthermore this ω value could be scaled by a learning architecture parameter σ, that can be tuned for a better algorithm performance. Anyway, such as the λ parameter, normally σ is set at a value of 1.

To make RR non-linear, should be replaced the input vector (x, X from equations 1 and2) with its projection $\phi(x)$, so the new, and final equations for RR involving RFF are:

$$f(x) = w^T \phi(x) \tag{4}$$

$$w = (\phi(X)^T \phi(X) + \lambda I)^{-1} \phi(X)^T y \tag{5}$$

4 QUALITATIVE ANALYSIS OF FMG AND SEMG SIGNALS AND DEVICE USER EXPERIENCE

This chapter presents the first experiment performed, which main goal was to analyze the properties and features of each of the sEMG and FMG signals. The chapter shows a brief introduction, then the experiment definition, followed by the experiment results and a discussion of the results.

This first experiment was designed to perform a qualitative analysis of each kind of sensor signal, so the next goals can be accomplished: Identify specific characteristics for each kind of signal, analyze the separability of different hand and wrist movements according to each kind of sensor and define pros and cons of each signal for future prosthetic control applications.

Along with the qualitative analysis, another goal of the experiment was the analysis of the subject experience using the wireless device presented in the previous chapter. This experiment was the first time the device was tested on subjects. An important point was to get to know about how the subjects feel during the experiment and how they evaluate their experience. The output of this user study provides information, if further experiments on this topic can be performed using the wireless device.

For the posterior analysis, the most important data were the raw sensor values and the stimulus ones (the ground truth of the movements performed by the user). The reasons, and how were used this values, are explained later in this chapter. The experiment was applied and guided by the author of this thesis (experimenter). Furthermore the work presented on this chapter is part of a paper submitted for publication [8].

4.1 EXPERIMENT 1 DEFINITION

The experiment was applied to ten intact subjects, nine of which were right-handed. Summarizing, the subjects were 3 females and 7 males, with an average age of 28 ± 7 years old, an average weight of 72.4 ± 9.91 kg, and an average height of 177.8 ± 12.14 cm tall. Each subject received a detailed description of the experiment, both in oral and written form. In order to support the experiment for publication, an informed consent was signed from all participants. Also for the experiment support, as the experiment is using sensors for bio-signals acquisition, it must be approved by the DLR's Ethical Committee.

The experiment consisted on performing a sequence of wrist and hand movements ten times: wrist flexion, wrist extension, wrist pronation, wrist supination and power grasp (already explained on Chapter 2), with a rest stage at the end of each repetition task set. Each movement could be followed by the subject using the polar plot from the cell phone GUI, additionally the experimenter was always paying attention, in case the subject was not performing correctly the tasks. For this experiment, the sensors were separated in two different bracelets, the first one with ten sEMG sensors, positioned on the left forearm and the second one with 10 FSRs, positioned on the right forearm. The bracelets were located approximately 10cm below the subject's elbows. After the bracelets attachment, the subject was asked to sit in a relaxed position with their forearms over their thighs and the hands in a lateral position (with the palms looking towards each other); the experiment movements had to be performed in a bilateral way (Figure 10).

The user experience was measured using: The System Usability Scale (SUS) [3], the NASA Task Load Index [22] and a modified version of the rework of the Microsoft Desirability Toolkit by David Travis [34].

Figure 10: Device setup for experiment performance.

The SUS consists of ten questions (Appendix) with answers represented on a 5 point Likert scale (1 - *strongly disagree* to 5 - *strongly agree*). For this experiment were used the statements just as Brooke proposes, except for one. The *I think that I would like to use this system frequently* statement was changed by *I felt comfortable with the device*.

As can be noticed, the even statements are negative focused and the odd ones are focused in a positive way. Due to this fact, the scoring is performed in a different way depending the statement. For the positive statements, from the original score, there must be subtracted one point. For the negative ones, the original score must be subtracted from 5. With this score method, the scoring in this survey is such that the answers to the *strongly agree* positive questions and to the *strongly disagree* negative questions generate a higher impact over the final score.

The NASA Task Load Index provides an overall task workload score on six subscales: Mental, Physical and Temporal Demands; Own Performance, Effort and Frustration. For each subscale, one question was asked in a way, such that the answer could be in a range of 21 points, reaching from *very low* to *very high* (except in the 4th question, Own Performance, that goes from *perfect* to *failure*). The questions are showed on the Appendix. In the experiment results the scores are shown in a percentage way, so the higher the value is, the higher the workload is. All the subscales were considered with the same importance weight.

Travis's survey is based on the idea that, according to him, the surveys presented previously are not reliable, mainly because of the user predisposition to answer in a positive way to help the experiment results or to show that he or she had a good experiment performance. The survey is based on the Microsoft Desirabilty Toolikt, which consists in a series of "reaction cards" with adjectives that could be applied to the system tested; the user is asked to select the five cards that most closely match their personal reactions to the system.

Travis applied a similar questionnaire and saw that the methodology seems to allow the participants to be critical with the system, selecting both positive and negative adjectives.

For this experiment in specific, there were used a list of 75 adjectives instead of the cards (most of them based on Travis's questionnaire), then the subject was asked to choose all the adjectives he or she felt more related with the device. After that, in a more precise selection, the user had to choose only the 5 most important words and try to give a simple reason about his or her decision. For a complete view of this survey, please consult the Appendix.

4.2 EXPERIMENT 1 RESULTS

This section displays the results of the first experiment.

4.2.1 *Qualitative signal analysis*

Using the signals recorded during the experiment, an off-line evaluation was performed after the experiment. The first thing done with the signals, was to apply the same filtering in both of them (3rd order low-pass digital Butterworth filter with a cutoff frequency of 1Hz). This off-line assessment was done using *MATLAB* software. Figure 11 shows typical FSR and sEMG signals obtained while a subject was doing two repetitions of the already explained task set.

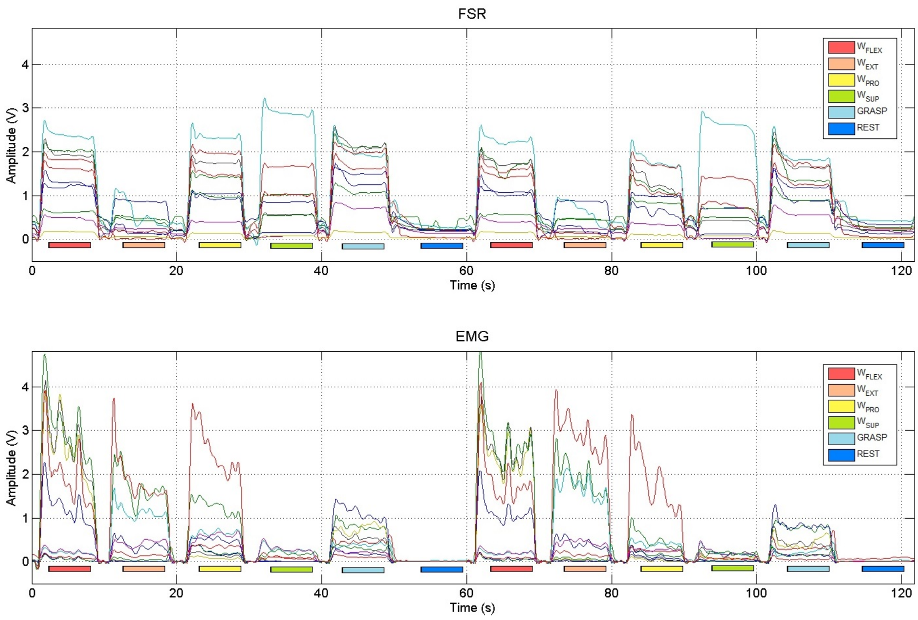


Figure 11: Typical FSR and sEMG signals obtained from one subject two repetitions of the instructed hand and wrist movements (wrist flexion, wrist extension, wrist pronation, wrist supination, power grasp and rest).

From Figure 11, can be appreciated that the amplitudes obtained for each kind of sensor are comparable, each movement is well separated in time from each other and produces a distinguishable pattern.
It is here where the first big difference between FSR and sEMG signals appears. FSR signals display a better stability over time compared with sEMG

signals, could be said that FSR sginals are presenting a "plateau" shape while each movement was enforced, on the other hand, sEMG signals exhibit the typical oscillating down-ramp pattern due to muscular motor-unit recruitment (in other words, sEMG is measuring electrical impulses) [20, 21]. To show in a graphic way this signal stability behavior, in Figure 12 the difference between consecutive samples for 5 repetitions of the wrist flexion movement of one of the subjects is plotted. For this plot are just considered the mean of the 5 sensors with the highest amplitude, in this way the stability effect could be better appreciated because there are some sensors that do not have a significant impact compared with others.
As expected, the sEMG sensors show larger oscillations, while FSRs have a more stable behavior, with an almost zero difference.

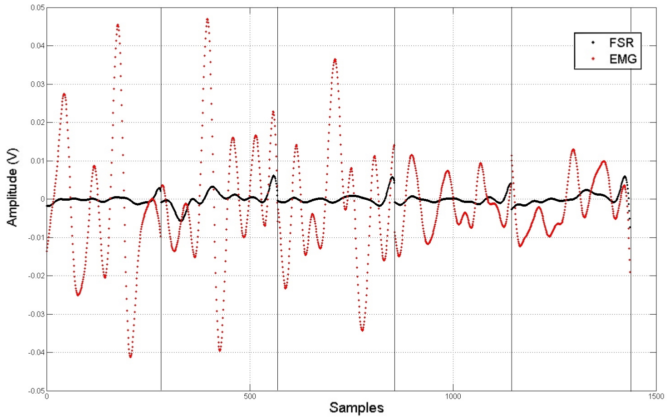


Figure 12: Sensor's samples difference plot. Five repetitions of the wrist flexion movement.

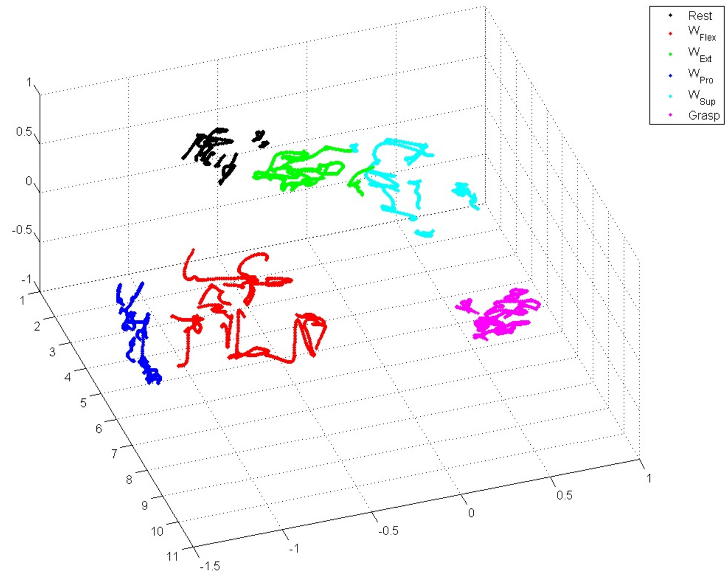
Table 1 shows the mean values of the difference signals explained in the last paragraph, those values were calculated for each user, considering only 3 repetitions and the 3 sensors with the highest amplitude of the wrist flexion movement.

Table 1: Means and SD of the sensor's samples difference. Three repetitions of the wrist flexion movement.

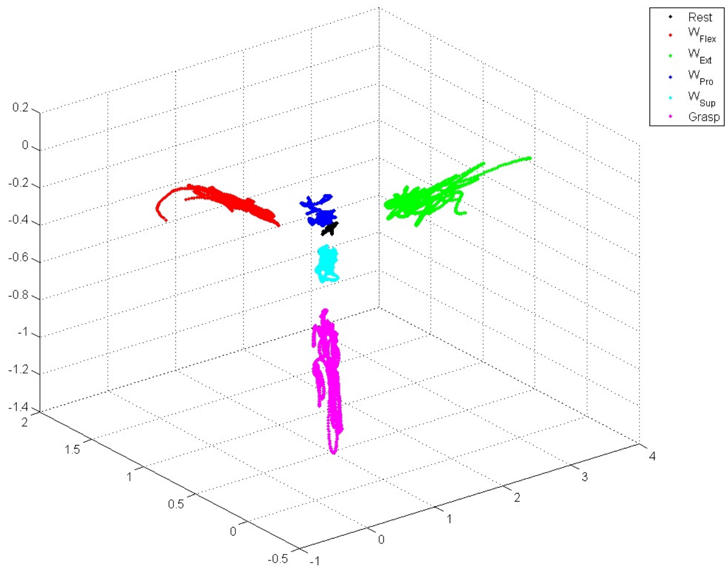
	1	2	3	4	5
FSR	-3.22E-5±0.0018	0.0002±0.0026	-0.00027±0.0006	-0.001±0.003	0.0013±0.0033
sEMG	-0.0018±0.0202	-0.0018±0.0111	-0.0003±0.0019	-1.4E-5±0.0038	-0.0074±0.0144

	6	7	8	9	10	**AVD±SD**
FSR	6.2E-6 ±0.0042	-0.0005±0.0018	-0.0004±0.0016	-0.0012±0.0028	-0.0009±0.0031	**0.0024±0.001**
sEMG	-0.0018±0.0082	-0.0021±0.0037	0.0016±0.0065	-0.0026±0.0082	-0.0039±0.0091	**0.0087±0.0054**

After this first analysis, was checked the separability between movements signal values for each type of sensor. This feature helps to know how easy



(a) FSR Clusters



(b) sEMG Clusters

Figure 13: PCA projection of typical data for each type of sensor.

a machine learning method could classify or predict, as the more separated the movement signals are, the better the performance of the learning method will be.

In Figure 13 is shown a reduced representation of the sensor values by using Principal Component Analysis (were considered the three most significant sensors from each type, and projected considering each axis as a different

sensor), the values are clustered depending the movement performed and was assigned a different color for each cluster.
It is important to remark that , in fact, the clusters are 10-dimensional (number of sensors), so even though in the PCA projection the clusters could be seen separable, in the complete dimension representation they are not. Furthermore, this PCA projection was just performed for visualization purposes, for the analysis that will be presented in the next paragraphs, there was used the complete dimensional data.

With the idea of separability already explained, there was defined a method for evaluating it and obtain a more quantitative approach. The Safety Index value [31], is based on the idea that the separation between the movement clusters could be defined as the relation between the maximum standard deviation of each cluster and the Euclidean distance between each them. The Safety Index s_{ij} was obtained for each subject and each pair of clusters (C_i, C_j), a precise definition of it is given by the Equation 6:

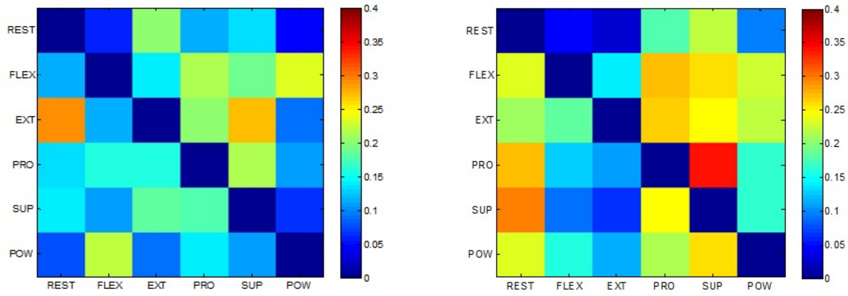
$$s_{ij} = \frac{max(\sigma_i)}{||\overline{C_i} - \overline{C_j}||} \tag{6}$$

where σ_i is the standard deviation of cluster C_i and $\overline{C}$ is the mean of cluster C.
With Eequation 6 it is implied that the lower the Euclidean distance or the higher the standard deviation, then the closer the clusters could be; the same the other way.

After the Safety Index calculation, the values were arranged in a matrix where the rows represented the C_i clusters and the columns the C_j ones, of course if C_i and C_j are the same movement, the Euclidean distance between them is 0 so the separation index is undefined. In order to avoid that fact, the Safety Index for same movement pairs was considered as zero. On Table 2, the means of all Safety Index matrix values for each subject are shown (the diagonal zero values are not considered for the mean).

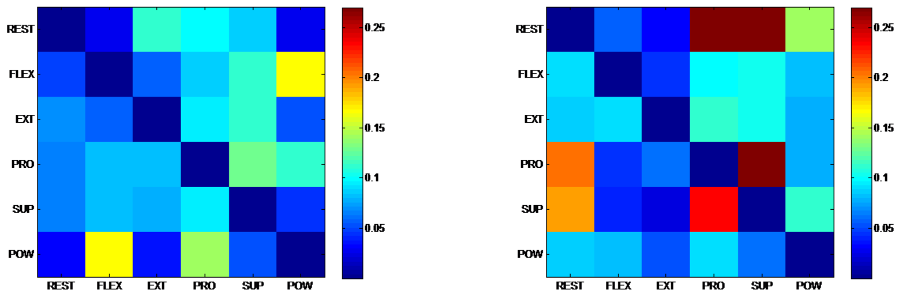
All the subjects Safety Index matrices were averaged, so a mean Safety Index matrix was obtained for both, sEMG and FSR sensors. Additionally was obtained the standard deviations of those movement pairs means. To visualize those matrices was used a color-map image, so the darker red the cell, the closer the movements are; on the other hand, the darker blue the cell, the more separate the clusters are (Figure 14). The standard deviation values were also visualized in a colormap, so in this way could be evaluated which movements were more different performed by the subjects through the experiment (Figure 15).

Finally, the last analysis performed was a prediction performance one. Considering all the sensors data as a data set, it was divided in three sets, the first one just including the FSR data, another one just including the sEMG data, and the third one was the same as the original data set, that means that it was including both kind of signals (called stack approach from now). The learning algorithm chosen for this analysis, was the Incremental Ridge Regression with Random Fourier Features (as was already mentioned in



(a) FSR S-Index (b) sEMG S-Index

Figure 14: Mean Safety Index matrices for each type of sensor. Lower is better.



(a) FSR S-Index SD (b) sEMG S-Index SD

Figure 15: Standard deviation of the Safety Index matrices for each type of sensor.

Table 2: Mean of the Safety Indexes (without counting the diagonal) for each user, and their total means.

	1	2	3	4	5
FSR	0.158±0.14	0.108±0.06	0.133±0.08	0.216±0.15	0.164±0.06
sEMG	0.149±0.1	0.141±0.09	0.152±0.07	0.279±0.15	0.143±0.09

	6	7	8	9	10	AVG±SD
FSR	0.16 ±0.12	0.156±0.09	0.158±0.11	0.115±0.08	0.134±0.08	**0.15±0.03**
sEMG	0.344±0.33	0.182±0.11	0.217±0.15	0.114±0.06	0.176±0.08	**0.189±0.07**

the previous chapter, the algorithm was already successfully used multiple times, for instance in Gijsberts et al. [14] and Ravindra and Castellini [29].

Before continuing, it is needed to explain a concept that will be used not only on this chapter, but also will be an important part of the next one, this concept is K-Fold Cross-Validation. Cross-Validation is a model performance measurement technique, it tries to give a better approach of the performance by training and validating the model through the entire dataset.

In K-Fold Cross-Validation, the interested dataset is divided into K equalized parts (folds). A training set and validation are generated by leaving one of the folds out (validation set) and combining the remaining folds to build the training set. This separation should be done K times, each time keeping out a different fold for the validation. There are two main disadvantages of this method. The first one, in order to keep the training set large enough, the validation set must be small. Second one, the training sets overlap through the repetitions, sharing K - 2 folds. Normally the value of K is 10 or 30, as K increases the estimator is more robust, but as mentioned, the validation becomes smaller . [1]. In spite of this facts, actually cross-validation was applied, just for a general-overview of the prediction, covering statistically all the data set.

Table 3: Prediction accuracy (nRMSE) obtained by each subject for a movement sequence repetition prediction. Were considered the FSRs, the sEMG sensors and both of them together.

	1	2	3	4	5
FSR	0.17±0.016	0.1687±0.0173	0.1495±0.007	0.1636±0.011	0.1762±0.0087
sEMG	0.1494±0.0138	0.1573±0.0175	0.1458±0.0123	0.2056±0.047	0.1673±0.0098
ALL	0.1649±0.0137	0.1608±0.0199	0.1385±0.0083	0.1595±0.0161	0.1636±0.0066

	6	7	8	9	10
FSR	0.173 ±0.0184	0.1728±0.0121	0.1893±0.0352	0.1548±0.0133	0.1803±0.0222
sEMG	0.1857±0.0126	0.1736±0.0193	0.2037±0.0343	0.1346±0.0061	0.1801±0.045
ALL	0.1658±0.0139	0.1736±0.0139	0.2054±0.0267	0.1367±0.0048	0.1526±0.013

For the experiment, a 10-fold “"leave-one-repetition-out”" cross-validation was applied by training each machine on nine of the ten repetition and validating on the remaining one. The input space was chosen from the three data sets already explained (just FSR, just sEMG and both stacked).
The prediction accuracy was measured by the normalized Root Mean-Squared-Error (nRMSE) between the predicted values and the stimulus file generated by the cell phone application, that worked as the ground truth. Notice that the stimulus values were from 0 to 1, as was explained in the Cell Phone GUI section, so the normalization of the RMSE is done considering those values. For general information the RMSE equation is given by:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y - \hat{y})^2}{n}} \tag{7}$$

where y is the ground truth value and $\hat{y}$ is the predicted value. For nRMSE was just divided the RMSE by the difference of the maximum and minimum valued of the stimulus values: $nRMSE = \frac{RMSE}{y_{max} - y_{min}}$.

Table 3 shows the prediction accuracy obtained by each subject measured in the just explained way. Here the nRMSE value showed is a mean of all the nRMSE obtained by the cross-validation. The same values are showed in a bar graph representation in Figure 16.

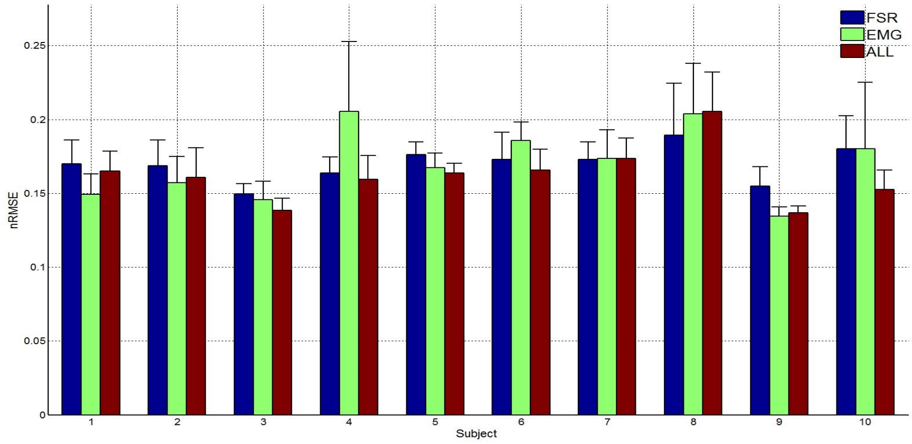


Figure 16: Prediction accuracy plot, separated by FSRs, sEMG sensors and both of them mixed.

4.2.2 *Device user experience*

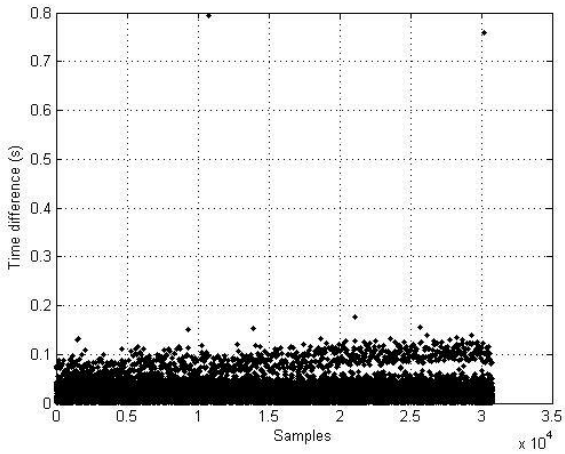
4.2.2.1 *Acquisition time analysis*

Before going through the surveys result, an off-line sample time period analysis was done, looking for any lag problem that could affect the device performance. This analysis was done considering the time stamps printed for each sample. There were two kind of behaviors in the experiment (Figure 17). The first one presents a time increase in the last repetitions , the second one presents also a lag but at the middle of the experiment and also at the end. Despite of this, those lags were so small and were not perceptible during the experiment performance.

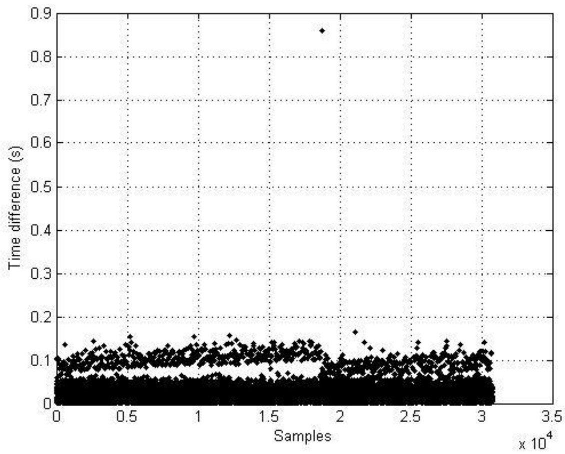
4.2.2.2 *User satisfaction survey results*

The results of the surveys are presented in the same order as they were presented in the previous section.
For the SUS, the total result of each subject and the mean score are presented in Table 4. Notice that the higher the score, the more usable the user judged the device. In this survey the statements that had the worst scores were *I think that I would need the support of a technical person to be able to use this device* and *I felt very confident using the device.*

For the NASA TLX, the total result for each subscale of each subject and overall workload are shown in Table 5. In this case, the highest the score, the more workload the user had when using the device. A plot with average percentages of the workload by subscale is visible in Figure 18.



(a) Incremental lag



(b) Double incremental lag

Figure 17: Samples time period plots.

Table 4: Subject's system usability total scores.

ID	1	2	3	4	5	6	7	8	9	10	AVG
SUS	85	90	100	77.5	92.5	67.5	75	85	100	75	*84.75*

The first two surveys applied threw suitable results with an usability score of almost 85 points and an overall workload of 25%.

Finally, for the desirability survey proposed by Travis, there were obtained two kind of results, the first one using all the words chosen by the user in the first selection, and the second one considering only the 5 final selections. In order to have a different visualization, a word cloud for each result was created (Figure 19), where the bigger and darker the font is, the more often the word was selected.

Table 5: NASA TLX workload percentages.

ID	1	2	3	4	5	6	7	8	9	10	AVG
Mental Demand	66.66	19.04	9.52	14.28	23.8	28.57	9.52	14.28	4.76	9.52	*20*
Physical Demand	33.33	23.8	14.28	14.28	23.8	23.8	9.52	14.28	4.76	76.19	*23.8*
Temporal Demand	71.42	61.9	52.38	14.28	52.38	57.14	14.28	61.9	4.76	33.33	*42.38*
Performance	28.57	28.57	19.04	14.28	14.28	19.04	52.38	14.28	4.76	38.09	*23.33*
Effort	52.38	19.04	23.8	9.52	52.38	19.04	52.38	14.28	4.76	38.09	*28.57*
Frustration	19.04	14.28	4.76	19.04	23.8	9.52	9.52	9.52	4.76	14.28	*12.85*
Overall WL	*45.2381*	*27.77*	*20.63*	*14.28*	*31.74*	*26.19*	*24.6*	*21.42*	*4.76*	*34.92*	***25.15***

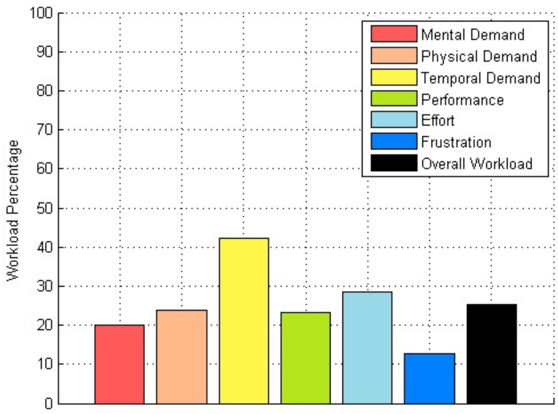


Figure 18: Workload percentage plot.

(a) First selection word cloud

(b) Final 5 words selection cloud

Figure 19: Word clouds from the results of the desirability survey .

Considering just the most common adjectives in the final selection, the reasons given by the subject for choosing them were: The word *simple* was chosen due to the device setup and the experiment performance. The adjective *intuitive* because the cell phone application shows what to do. The subjects considered the word *easy to use* because the device has an easy setup. For the subjects the experiment was *familiar* as they had done previously similar ones, furthermore the device can be used for daily life. The adjective *reliable* was chosen due to the sensor responses which are consistent, additionally the device works all the time. Finally subjects felt the device was *stable* giving similar reasons as the ones given for the reliable adjective, also the system presents none errors.

As could be already noticed, an important thing to mention about the surveys is that even the instructions and the questions were oriented to the device, some answers referred to the experiment performance. So for instance the familiar adjective was chosen in some cases because the subject had already done some experiments with sEMG sensors.

4.3 EXPERIMENT 1 DISCUSSION

4.3.1 *Comparison between sEMG and FMG*

There are marked differences between both kind of signals, proving in that way one part of the hypothesis statement defined in the Introduction. Even though the amplitude and features of the sensors change significantly across all subjects, can be established important remarks about the comparison of sEMG and FMG signals. Focusing first on an amplitude signal analysis and looking again to the Figure 11, can be set that:

FSR measurements present an oscillating and different amplitude behavior in the rest stage, so the signal amplitude was rising if the subject was changing his/her initial rest position (actually the changing of position is natural, as it is really difficult to go back always to the same position after a movement performance), creating in this way a noisy rest measurement. This behavior was expected, due to the stable behavior of the FMG signal and perhaps also due to the inherent hysteresis that the FSRs have.
On the other hand, sEMG resting phase remains almost in the zero value, it is not oscillating, even if the user was changing the initial position, again this behavior was expected, as none considerable impulse muscle activation was performed.

Supported also by the Safety Index, the movements with the lowest amplitude (apart of the rest of course) were: for FSR signals was the wrist extension, having values comparable with the resting stage on some subjects. for the sEMG signals, the weakest movement was the wrist supination.

Thanks to the signal conditioning performed in the cell phone application, the FMG and sEMG produce similar amplitudes inside the same voltage range (0 - 5 V). However, as was already established this amplitude is reduced in the sEMG case across the movement performance, reaching sometimes almost half of the initial movement amplitude.

In a general overview, the FSRs shows more sensibility to the muscle movements than the sEMG sensors, in the way that each FSR is excited by each pattern while with sEMG sensors just some of them are considerable excited. Going deeper, an average of six FSRs present amplitudes over 25% of the maximum sensor amplitude, whereas just an average of three sEMG sensors have this behavior.

The Safety Index is really helpful as it is showing unit-less values, independent of the sensors voltages, so could perform a more confident analysis through the subjects.

Considering the Safety Index mean and standard deviation colormap images, the FSR matrices show an average better performance over the sEMG ones. Some remarks about this comparison are: FSR Safety Index showed a higher number of considerable separated values (blue-dark blue), and a lower number of poor separated values (red-dark red) than the sEMG Safety Index.

Separation between rest position and all the other movements using sEMG is poor.

There is a poor separation for the FSR signals only between extension and rest, and extension and supination.

Looking at the standard deviation matrices, the FSR matrix just had 4 high values, on the other hand sEMG matrix presented 7. From this, can be conclude that the users were performing more similar FSR Safety Index values than sEMG, in other words FSR movements separation were more constant through the subjects, supporting with this fact the idea of high stability of the FSR measurements.

According to all the previous remarks, there can probably be settled that FMG enforces a better overall performance than sEMG, although not uniformly. The defined pros and cons, combined with the fact that there is not apparent significant difference in the prediction accuracy, even considering the combined input space(the nRMSE values range goes from 0.13 to 0.21); support the idea of interleave or fuse the two type of signals, defining as an hypothesis the consideration of the best movements performed by each sensor type.

4.3.2 *User satisfaction*

Even though there was already done an analysis on the survey results, a better way to visualize those results is a radial chart (Figure 20), which separates the results in different categories (Low Workload Demand, Stability, Task Accomplishment, Interface, Easy to Use, Comfort and Setup) that represent the main features appreciated by the subjects. Just as a special remark, the representation of the Low Workload Demand has been inverted to perform a better and easier comparability measure with the other categories.

Additionally, the categories are clustered in two main classes: the usability related features (Interface, Easy to Use, Comfort and Setup) and performance related features (Workload Demand, Stability and Task Accomplishment). Those classes are present in the radial chart, where the categories are arranged in that way, the performance features on the top of the chart, and the usability ones on the bottom.

From the radial chart, can be stated that: The wireless device has a consistent, reliable and stable operation.

The device and interface setup are easy, simple and comfortable, creating with that a low frustration rate;

The user interface is helpful and useful, as it is also friendly, intuitive and well structured;

The experiment performance took a considerable amount of time (from this fact the highest workload demand). In fact this workload demand could become much lower, as in real applications the subject is not expected to perform several number of repetitions for training, avoiding with this a huge amount of repetitive movements.

In general the experiment presented a high successful task accomplishment rate, all the subjects were able to achieve the experiment goal.

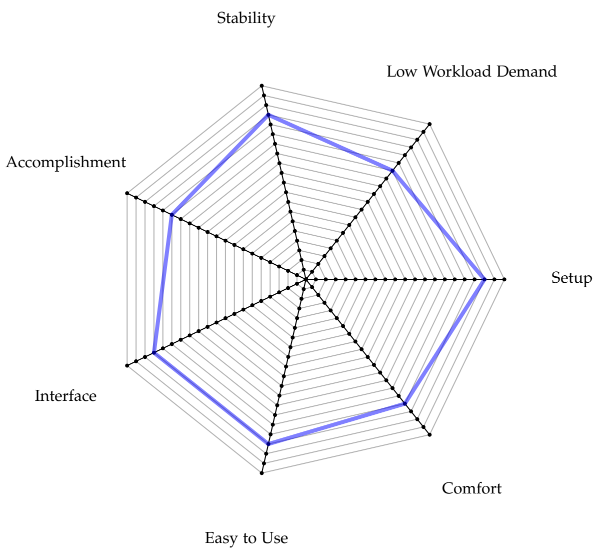


Figure 20: Radial chart showing the user satisfaction surveys' results summarized in different device features.

The survey results supported the idea, form different perspectives, of using this device for future applications and experiments, including the final experiment done for this project, as the user was having positive sensations. However, the "good" results obtained in the surveys, as was already established back into this chapter, could have been influenced by the subject's behaviors explained by Travis.

Although the subjects considered on this experiment are not the potential real application subjects, their feelings and opinion about the device define a starting point for the device conditioning and possible modification to use it on amputees and real robotic hand control.

5 PREDICTION PERFORMANCE IMPROVEMENT BY SENSORS MIXTURE

According to the goals defined in Chapter 1, one of the main purposes of this project is to find a new machine learning method that can mix both kind of sensors and improve in some way the performance of some, or if possible all, the actions tested until now. In this chapter the new machine learning approach is defined. Then a learning algorithm architecture parameter used in iRRRFF is tuned and tested using the previous experiment data. Finally, the last experiment is explained, applied, its results are showed and proper discussions are made.

As it was stated in the previous abstract, this chapter is following the next goals in order to accomplish the main ones defined on the project introduction: First of all, find and apply a method for mixing both FSR and sEMG signals. Perform a further off-line analysis with the previous experiment data, focusing on the prediction performance. Design an experiment which main goal is the measurement of the online prediction performance using four different machine learning methods (including the one just defined). Along with this goal, modify the PC GUI to adapt it to the designed experiment. Furthermore, carry out an off-line analysis with the data obtained in the experiments. Finally, build the proper conclusions focusing on the hypothesis defined in the thesis introduction.

5.1 ENSEMBLE LEARNING

Ensemble Learning has been used for different applications and purposes, it proposes the use of different hypotheses, so the learning has a bigger range of possibilities to build a final better performed prediction.

Although all the ensemble learning approaches explained in Chapter 2 have shown important results, the nature of the problem presented in this project led to the use of the stacking approach.

Notice here that this stacking approach is not the same as the approach used for prediction in the last part of the previous chapter. While the previous approach is *stacking* both sensors signal for the input space, this one, as will be explained in this section, is *stacking* the outputs of different models. In order to avoid any confusion, after this section, the ensemble learning stacking approach will be referred just as ensemble learning while the input stacking method will be called, as in the previous chapter, stack method.

Ensemble learning stacking is used when different machine learning methods are used (ANN, SVM, pattern classification). The way to implement this approach is by training and then obtain a prediction using cross-validation, so the predicted values are stacked creating a new input space that is used for training a final machine learning model, which prediction will be the final system prediction.

Even though the stacking approach is normally used for different learning algorithms, in the case of this project, the methodology was changed to fit with the project's needs, expecting even so the same performance improvement. For this project there were not used two different algorithms, but two different input spaces (sEMG sensors and FSR) trained with the same algorithm, creating in that way a different output for each model. Basically this is why this approach was chosen, as each of the different input space algorithms can be considered as different learning methods.

The algorithm in which the code development was based is the one proposed by Polley and Laan [26] that is based at the same time on the earlier work done by Wolpert [37]. The algorithm was modified so it could be adapted for this project purposes. The steps followed in this project for the ensemble learning stacking approach implementation were:

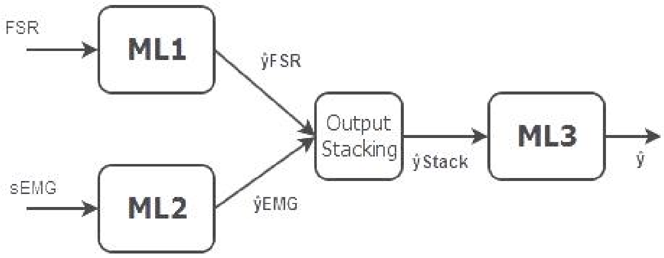


Figure 21: Ensemble learning stacking approach block diagram

1. Fit each data set used in the training stage(EMG and FSR), with n samples, on the learning algorithm, Ridge Regression Random Fourier Features (RRRFF), to estimate the weight matrices W_k, with $k = EMG, FSR$, in order to calculate, by using new data, the predicted output Y_k and then stack them to obtain $\Psi = [Y_{EMG} Y_{FSR}]$

2. Split the training set into a training and validation set, according to a V-fold cross-validation, where V is equal to the number of repetitions performed in the training stage, let the v-th group be the validation set, and the remaining data the training one, v=1,...,V. Define $T(v)_k$ to be the v-th training data set and $V(v)_k$ the corresponding validation set.

3. For the v-th fold, train each $T(v)_k$ using the learning model, predict on the corresponding $V(v)_k$ and save them as in step 1 in $\Psi_{T(v)} = [Y_{T(v)EMG} Y_{T(v)FSR}]$.

4. After the CV, stack the predictions $\Psi_{T(v)}$, creating a n size array with all the predictions.

5. Use the stacked predictions as a new input space for a third machine learning model (Again a RR RFF approach), get a new weight matrix W_{ENS}.

6. Use Ψ obtained in step 1 and W_{ENS} to obtain the final prediction $Y_{ENS} = W_{ENS}^T \Phi(\Psi)$. Ψ should be obtained for each sample of the online prediction stage.

Because of the nature of the algorithm, as it is using the entire training set for the CV, was not possible, at least on this project, to apply the incremental feature of the *RRRFF* algorithm. Even so, the implementation of this characteristic could be considered as future work.

A summarized and graphical view of the algorithm is shown in Figure 21.

5.2 FURTHER FIRST EXPERIMENT ANALYSIS

Before testing the ensemble learning approach on an online experiment, it was important to check first if it was an improvement compared with the

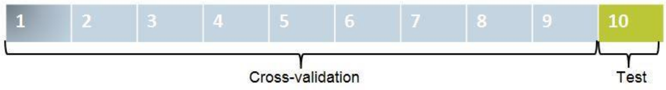


Figure 22: CV diagram for $\sigma tuning$

approaches already designed (Stack, EMG and FSR). In order to accomplish that comparison, all the learning methods, including the ensemble one, were tested on the data obtained in the first experiment (sensors and stimulus data). The main goal was to find which approach had better performance in the prediction of movements, separating for that the acquired data in learning, validation and test sets.

As was mentioned on Chapter 2, there is a Ridge Regression Random Fourier Features architecture parameter σ that is scaling the random distribution ω. Tuning this parameter, the performance of the prediction changes. From this fact, were proposed two hypotheses: First, each learning method has a different *optimal* σ value. Second, this parameter has almost the same value across all subjects, so can be defined a generalized value for future applications.

Next subsection explains how the σ parameter was tuned, and the results of the analysis focused on the hypotheses proof.

5.2.1 *σ parameter tuning*

First of all, as the experiment data would be used not only for the σ tuning, but also for the prediction performance evaluation, then each subject data set should be divided on three sets: the training set (composed by eight repetitions), the validation set (composed by one repetition), and the test set (composed by the repetition left). Training and validation sets were used for the parameter tuning, while test set was used for the prediction performance final test, applying the tuned σ.

The tuning was done by using again the cross-validation method. In this case was a 9-fold cross validation, performed in the first nine repetitions of each subject (one fold for each repetition), the data was split as mentioned in the last paragraph and illustrated in Figure 22. As could be noticed the test set is fixed, while the other nine repetitions are involved in the CV, so for each iteration the validation and training set are changing, keeping always the same length.

The optimal σ was found by applying a grid search method, so a range of possible σ was defined and all its elements were considered for the search. For each CV iteration, all the elements of the σ range were tested, so the training was performed using them and then a prediction was done in the validation set. For each σ value, the nRMSE between the validation set prediction and the stimulus signal was obtained, the lowest nRMSE determined the optimal σ. In order to avoid randomness, but also give the enough statistical support, there were built ten different ω and β representations, so the σ were also tested using those fixed values. After this procedure was applied on each of the CV iterations, then the optimal σ was already defined. The weight matrix obtained in the training stage using the optimal σ value, was used for a final prediction performed in the test set (tenth repetition), this prediction performance is showed in the next subsection.

An optimal σ was obtained, as mentioned, for each of the subjects (Table 6). Looking at the average of those values and the standard deviations can be concluded that each approach presents a different value compared with the others, especially the multi-sensors approaches compared with the single-sensor ones. As the standard deviation was not big enough, could be considered the average values as generalized σ for future applications, proving in this way the hypotheses previously defined in this section.

Table 6: Optimal σ values obtained by CV for all subjects

Subj.	**1**				**2**				**3**			
Meth.	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**
σOpt	0.1	0.398	0.158	0.398	0.1	0.199	0.126	0.316	0.794	1.259	0.794	1.259

Subj.	**4**				**5**				**6**			
Meth.	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**
σOpt	0.316	0.316	0.501	0.794	0.199	0.251	0.501	0.794	0.158	0.251	0.199	0.631

Subj.	**7**				**8**				**9**			
Meth.	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**
σOpt	0.251	0.794	0.316	0.501	0.199	0.251	0.794	0.316	0.158	0.398	1	0.631

Subj.	**10**				**AVG±SD**			
Meth.	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**
σOpt	0.199	0.398	1	0.501	0.2476±0.192	0.4516±0.313	**0.5391±0.322**	**0.6142±0.270**

5.2.2 *Prediction performance*

The prediction performance analysis can be divided in two parts. The first one consists in the prediction performance of the test set for each user, using the optimal σ obtained for each user. The results of this performance are shown on Table 7 and illustrated on Figure 23. The nRMSE shown on Table 7 were calculated in a similar way than the prediction performance of the first experiment, that means that the values correspond to an average nRMSE of the DOFs considered.

Nevertheless a more robust, and closer to the online application, prediction was performed. A performance analysis was done by applying a 10-fold cross-validation on each subject using the generalized optimal σ values obtained in the previous subsection. A mean of the nRMSE values of each CV iteration is obtained for each subject, these results are presented on Table 8, and again illustrated on Figure 26 Noticeably, as closer the generalized σ was to the subject optimal σ, the performance was better, the same the other way. From this prediction performance analysis and from a separated action one(Table 9), some conclusions were stated. First of all, the ensemble method was the one with better prediction performances, just one subject did not

Table 7: Performance measurement on the test set for each subject

Subj.	**1**				**2**				**3**			
Meth.	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**
nRMSE	0.156	0.156	0.169	0.139	0.191	0.194	0.211	0.169	0.152	0.143	0.169	0.129

Subj.	**4**				**5**				**6**			
Meth.	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**
nRMSE	0.186	0.229	0.196	0.249	0.188	0.188	0.209	0.198	0.174	0.192	0.194	0.165

Subj.	**7**				**8**				**9**			
Meth.	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**
nRMSE	0.176	0.171	0.19	0.169	0.317	0.324	0.325	0.284	0.157	0.155	0.179	0.142

Subj.	**10**				**AVG±SD**			
Meth.	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**
nRMSE	0.157	0.156	0.2080	0.13	**0.1875±0.048**	**0.1908±0.067**	**0.2006±0.061**	**0.1776±0.064**

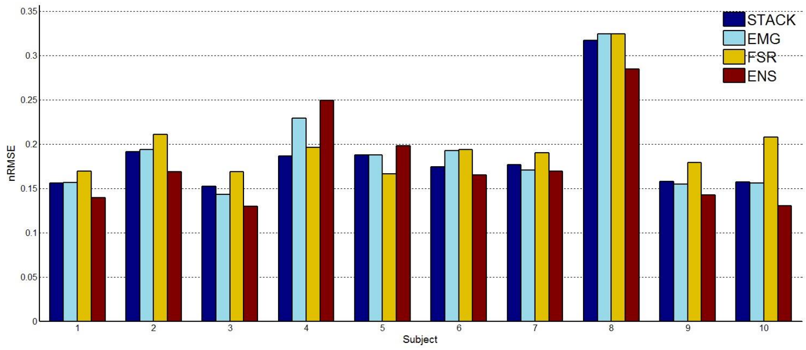
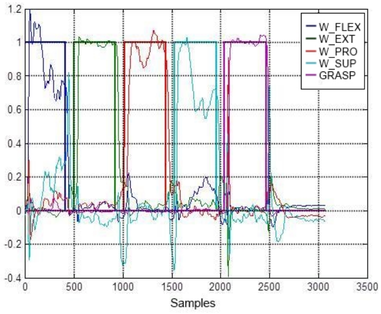


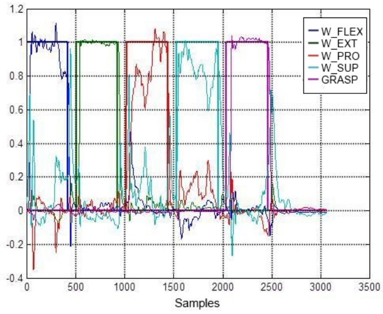
Figure 23: Prediction performance on the test set using the optimal σ for each subject

present that method as the best one. Going through the separate actions analysis, the FSR method presented high error values, especially on the wrist flexion/extension movements. EMG presented a good performance on the flexion and extension actions, but was the worst on the pronation/supination ones. Stack method showed a constant performance, slightly worse than the ensemble method. By plotting the predicted and the stimulus signal for each method on one subject repetition (Figure 25), a qualitative analysis of the prediction signal was done.

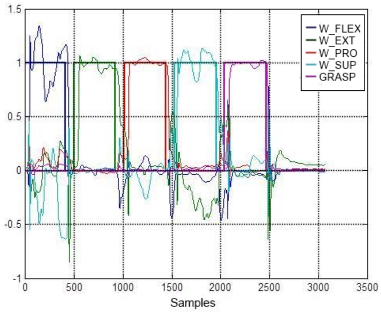
From figure 25 could be defined some features corresponding for each approach. The stack method signals are kind of average of EMG and FSR prediction amplitudes. EMG method is presenting the typical oscillating behavior. FSR approach presented more stable signal but with high noise in the inactive DOF. Finally the ensemble method seems to inherited the FSR amplitude, but also the EMG oscillating behavior, also is improving the performance by reducing the noise of the inactive DOFs and the signal delay respect with the stimulus. The learning approach features obtained in this qualitative analysis are expected to appear also in the online analysis, affecting each learning approach performance in the same way as here.



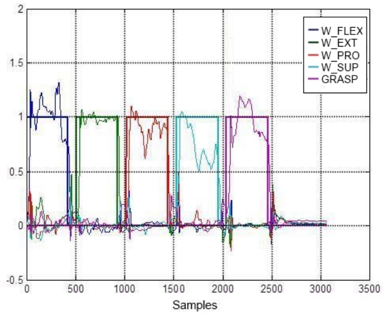
(a) Stack method



(b) EMG method



(c) FSR method



(d) Ensemble method

Figure 24: Prediction performance on one repetition for each method considering the CV approach

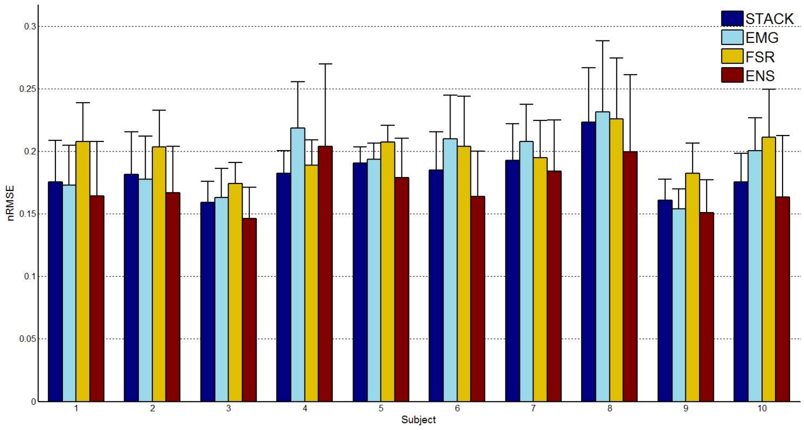


Figure 25: Prediction performance after a 10-fold CV using the generalized σ for each subject

Table 8: Performance measurement using the generalized σ and applying CV for each subject

Subj.	1				2			
Meth.	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**
nRMSE	0.175±0.033	0.173±0.031	0.207±0.030	0.164±0.043	0.181±0.034	0.178±0.034	0.203±0.029	0.167±0.037

Subj.	3				4			
Meth.	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**
nRMSE	0.159±0.016	0.163±0.023	0.174±0.017	0.146±0.025	0.182±0.018	0.218±0.037	0.189±0.02	0.204±0.065

Subj.	5				6			
Meth.	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**
nRMSE	0.19±0.013	0.193±0.013	0.207±0.014	0.179±0.031	0.185±0.03	0.21±0.035	0.204±0.04	0.164±0.036

Subj.	7				8			
Meth.	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**
nRMSE	0.193±0.029	0.195±0.029	0.208±0.03	0.184±0.041	0.223±0.043	0.231±0.057	0.225±0.048	0.199±0.061

Subj.	9				10			
Meth.	**Stack**	**EMG**	**FSR**	**ENS**	**Stack**	**EMG**	**FSR**	**ENS**
nRMSE	0.161±0.017	0.154±0.016	0.182±0.024	0.151±0.026	0.175±0.023	0.2±0.041	0.211±0.038	0.163±0.049

Subj.	AVG±SD			
Meth.	**Stack**	**EMG**	**FSR**	**ENS**
nRMSE	**0.1826±0.026**	**0.1886±0.031**	**0.2026±0.03**	**0.1687±0.039**

Table 9: Average of all the subjects performance measurement (nRMSE) for each action using the generalized σ.

	Stack	**sEMG**	**FSR**	**Ens**
Act.	**Mean±SD**	**Mean±SD**	**Mean±SD**	**Mean±SD**
Wr. Flex.	0.187±0.036	0.182±0.036	0.209±0.036	0.177±0.046
Wr. Ext.	0.187±0.034	0.188±0.039522	0.242±0.049	0.179±0.049
Wr. Pro.	0.191±0.025	0.208±0.03	0.199±0.024	0.173±0.037
Wr. Sup.	0.180±0.018	0.206±0.029	0.186±0.022	0.155±0.031
Grasp	0.168±0.02	0.164±0.020	0.176±0.019	0.159±0.032

5.3 EXPERIMENT 2 DEFINITION

The experiment was applied to twelve intact subjects, ten of which were right-handed. Summarizing, the subjects were 3 females and 9 males, with an average age of 27 ± 6 years old, an average weight of 70.66 ± 12.34 kg, and an average height of 176.91 ± 10.39cm tall. Each subject received a detailed description of the experiment, both in oral and written form. As in the first experiment, an informed consent was signed and the experiment was approved by the DLR's Ethical Committee.

For this experiment, the sensors were separated on two different bracelets, the first one with ten sEMG sensors and the second one with ten FSRs, but this time both bracelets were positioned on the right forearm. The bracelets were located approximately 10cm below the subject's elbows, with the the FSR's bracelet closer to the elbow. After the bracelets attachment, as in the previous experiment, there was asked to the subject to sit in a relaxed posi-

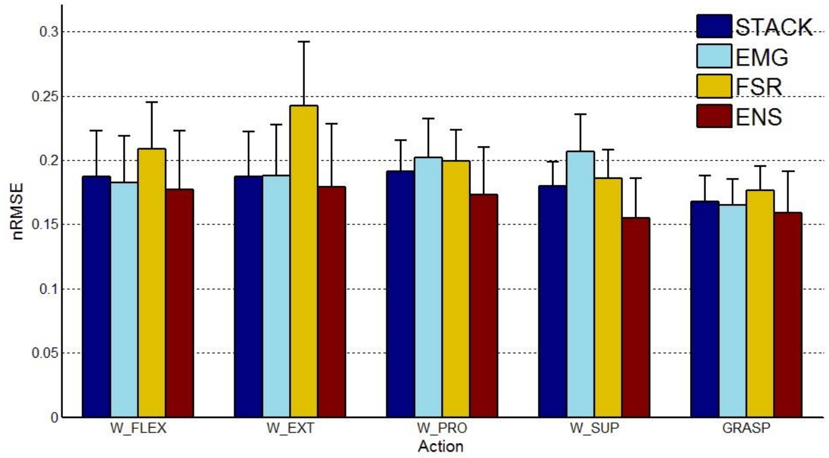


Figure 26: Prediction performance after a 10-fold CV using the generalized σ for each action

Figure 27: Second experiment setup and performance

tion with their forearms over their thighs and the hands in a lateral position (Figure 27).

The experiment consisted on three stages: the training, the performance and the user experience stage (Figure 28) As in the previous experiment, the subject had to perform a sequence of wrist and hand movements: wrist flexion, wrist extension, wrist pronation, wrist supination and power grasp.

On the training stage the subject must perform three repetitions of the movements sequence, in the same order as presented in the last paragraph, with a rest training at the beginning of each repetition. The training data acquisition was done in a manual mode, that means that the data considered for the training of the Stack, EMG and FSR approaches was captured just after a button was pressed by the experimenter. In the case of the ensemble learning, this manual mode is only applied for the training of the first two learning machines (EMG and FSR), but the stacked outputs are obtained by predicting in the entire data set, that means that the final output estimation involves in some way the complete information of the training stage and not

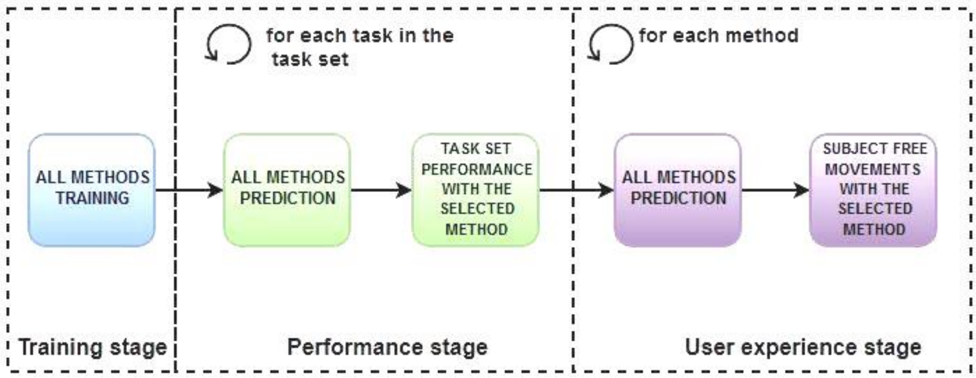


Figure 28: Experiment block diagram

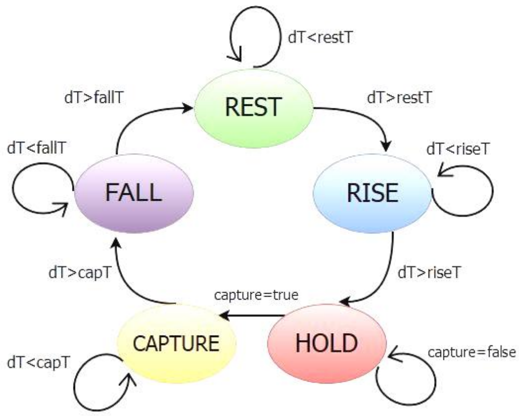


Figure 29: Training stage state diagram

only the one obtained in the capture state, as in the case of the other three approaches.

In order to understand better how each action was trained, Figure 29 presents the state diagram of the training stage, where rise and fall are the transition movements from rest to maximum position and vice versa, hold state represents the movement in a maximum level but without capturing data, event that happens in the capture state. dT is the elapsed time from the beginning of each state. All the states were performed by a predetermined time except for the hold state, which was acting until the capture button was pressed.

For the training stage, the subject could follow the movements to perform using the 3D hand model (Figure 7). In this case the gray hand was the one showing the action and also the level of intensity, so the subject must follow exactly in the same way the 3D model. After the data acquisition, all the four learning methods were trained, obtaining in this way four different weight matrices used later for prediction.

The next stage was the performance stage, which was the main part of the experiment. On this stage the user must perform two repetitions of five actions (wrist flexion, extension, pronation, supination; and grasp) considering

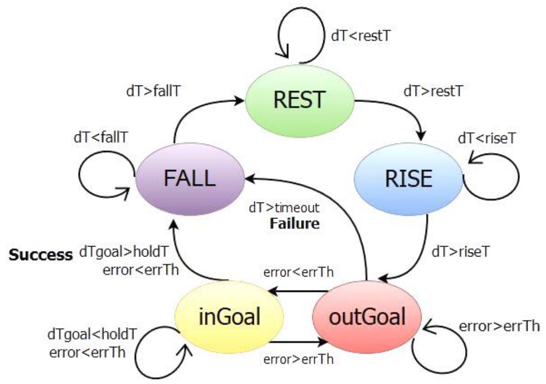


Figure 30: Performance stage state diagram

the four methods (Stack, EMG, FSR and Ensemble) and also three different movement degree levels (a complete movement level like in the training stage, two thirds the level of that movement and one third the level of the complete movement), in this way a total of 120 tasks must be performed to cover all the variables. In order to accomplish those tasks, the prediction was done by all the learning methods, but only was used the one selected for the task.

In this stage the subject must follow the movement of the gray 3D Hand (Figure 7), but now the predicted movement is shown by the colored hand, then the subject must match both of the hands. The way this stage works is illustrated with the state diagram on Figure 30. The rest, rise and fall states worked as in the training stage, but instead of having a hold and capture states, they were called outGoal and inGoal states. The main idea of these new states is based on the absolute error between the stimulus values of the guide hand (gray) and the output values obtained by the prediction using the corresponding learning method. If this error was lower than a threshold (0.15) then the system moved from the outGoal state to the inGoal. The subject must keep the error under the threshold (errTh) for 1.5 seconds (holdT), if he or she could manage to do it, then the task was considered a success, if the subject went over the error threshold before the 1.5 seconds, then the system went back to the outGoal state and the inGoal time counter(dtGoal) was restarted. In total the subject had 30 seconds (timeout) to accomplish the goal, if he or she was not able to do it, then the task was considered as unsuccessful. From this stage were saved important information for posterior analysis, like the learning method used for the prediction of the task, the time elapsed during the task performance, and of course the times the task was successful.

In order to get more reliable results, the learning method, the actions, and the movement sorts must be ordered randomly, but at the same time must have the same number of appearances in a distributed way through the tasks, in this way all the data could be covered by all the possible combinations. In the next subsection is proposed a variable sort design that offers a good enough statistical coverage across all the subjects data.

Finally, the last stage was the user experience. In this stage were obtained again the predictions of all the methods, but was apply just the one chosen by the experimenter using the GUI. In this stage just the colored hand was acting, showing the prediction of the selected approach. The subjects must move their hand free, performing the movements that they wanted with the level they wanted, trying to cover all the possibilities. After some time the subject was asked for some questions about the stability, realism, and smoothness of the prediction 3D Hand, evaluating each factor from 1 to 7, being 1 the lowest value and 7 the highest. In this way could be known the subject impressions about the prediction.

After the experiment, the subjects filled the NASA TLX survey (the same as in the first experiment). It was a long experiment, so there was expected a considerable workload present, that actually could affect the experiment performance.

5.3.1 *Variable sort design*

This subsection describes the way all the tasks used in the performance stage were arranged. As there were four learning methods, the best idea would be to have 24 subjects (4! = 24). As this is a large number of subjects, was reduced to twelve, so just half of all the possible methods permutations were considered. Fortunately the number of movement levels was three, so could be fitted to the number of subjects without any problem.

The real problem appeared with the number of actions, as there were five different actions, their permutations did not fit with the subjects involve. In order to overcome this problem, six different blocks including the five actions to perform were built, even was not possible to cover all the possible action combinations, these blocks allowed a close representation of the action's permutations. The blocks were built trying to perform each action in a distributed way, that means that was avoid to repeat for instance the wrist flexion action always in the first position of the repetitions.
The blocks built were:

Table 10: Action blocks including different sorts of the actions to perform (1- Wrist Flexion, 2- Wrist Extension, 3- Wrist Pronation, 4- Wrist Supination, 5- Power Grasp)

Block	**Actions**
B1	1 2 3 4 5
B2	4 5 1 2 3
B3	3 4 2 5 1
B4	2 1 3 5 4
B5	5 4 2 1 3
B6	3 5 1 4 2

Now that the number of blocks (6) allows compatible permutations with the number of subjects, an overall experiment task sort design was purposed (Table 11), the order of each variable (method, action block and level) for the first subject was generated randomly. From there, the rest of the subject sorts are permutations from the first subject sort. According with Table 11, for the first six subjects on the first repetition were performed the action blocks from B1 to B3 related with a certain level sort, on the second repetition were performed the action blocks from B4 to B6 related with another level sort. As there are just 6 possible permutations (3!=6), for the last 6 subjects the action blocks and their assigned level were exchanged between repetitions. In this way the action blocks/levels done on the first 6 subjects first repetition were performed on the second repetition for the last 6 subjects and vice versa.

Table 11: Methods, action blocks and level sort across the subjects (for the methods: 1- Stack, 2- sEMG, 3- FSR, 4- Ensemble)

	Repetitions					
	1	2	1	2	1	2
Subject	Method		Actions		Levels	
1	3 2 4 1	3 2 4 1	B2 B1 B3	B4 B5 B6	1,0.33,0.66	1,0.66,0.33
2	2 4 1 3	2 4 1 3	B1 B3 B2	B5 B6 B4	0.66,1,0.33	0.33,1,0.66
3	4 1 3 2	4 1 3 2	B3 B2 B1	B6 B4 B5	0.33,0.66,1	0.66,0.33,1
4	1 3 2 4	1 3 2 4	B2 B3 B1	B4 B6 B5	1,0.66,0.33	1,0.33,0.66
5	3 4 2 1	3 4 2 1	B3 B1 B2	B6 B5 B4	0.33,1,0.66	0.66,1,0.33
6	4 2 1 3	4 2 1 3	B1 B2 B3	B5 B4 B6	0.66,0.33,1	0.33,0.66,1
7	2 1 3 4	2 1 3 4	B4 B5 B6	B2 B1 B3	1,0.66,0.33	1,0.33,0.66
8	1 3 4 2	1 3 4 2	B5 B6 B4	B1 B3 B2	0.33,1,0.66	0.66,1,0.33
9	3 1 4 2	3 1 4 2	B6 B4 B5	B3 B2 B1	0.66,0.33,1	0.33,0.66,1
10	1 4 2 3	1 4 2 3	B4 B6 B5	B2 B3 B1	1,0.33,0.66	1,0.66,0.33
11	4 2 3 1	4 2 3 1	B6 B5 B4	B3 B1 B2	0.66,1,0.33	0.33,1,0.66
12	2 3 1 4	2 3 1 4	B5 B4 B6	B1 B2 B3	0.33,0.66,1	0.66,0.33,1

This overall arrangement does not show exactly how the tasks are performed inside each subject. That's why Table 12 shows in a precise way how the tasks are sort for instance on subject 1. The two repetitions are performed for each method consecutively and the action blocks and levels are the same for each method, so the subject was doing a total of 30 tasks per method, 6 tasks per action and 10 tasks per level.

As could be noticed was not pretended to cover all the possibilities on one subject, but considering all.

Table 12: Experiment performance for the first subject

Method	Action	Level	Rep.
3	B2, B1, B3	1,0.33,0.66	1
	B4, B5, B6	1,0.66,0.33	2
2	B2, B1, B3	1,0.33,0.66	1
	B4, B5, B6	1,0.66,0.33	2
4	B2, B1, B3	1,0.33,0.66	1
	B4, B5, B6	1,0.66,0.33	2
1	B2, B1, B3	1,0.33,0.66	1
	B4, B5, B6	1,0.66,0.33	2

5.3.2 *Modifications to DLR's previous work*

All the modifications were applied on the PC GUI C# code. The main change was the implementation of the ensemble algorithm, apart of that all the necessary changes for the application of the experiment were done, for instance for the training stage, all the four methods must be applied (the original code just performed the Stack method). For the performance stage a text file with all the task variable sorts was read (generated with MATLAB), also the modifications were done to just apply the prediction of the right method, finally the needed GUI elements for the stage start and the number of task monitoring were created. Additionally just for a better user-interface interaction two face drawings were used (Figure 31), the green face was appearing when the subject could successfully perform a task, the yellow one appeared in case of task failure. These faces were already used on previous works on DLR. Finally for the user experience stage, were created four check boxes, one for each learning method, so the experimenter could check the one to be tested by the subject. The experiment final PC GUI is shown on Figure 32.

(a) Success! face (b) Keep trying face

Figure 31: Visual feedback face drawings

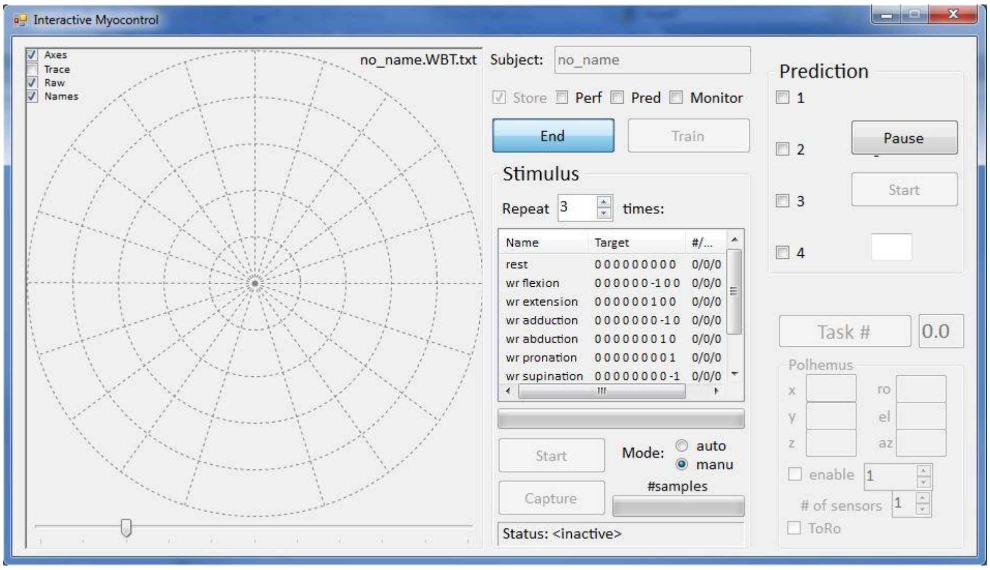


Figure 32: Modified PC GUI

5.4 EXPERIMENT 2 RESULTS

5.4.1 *Learning method performance results*

After data collection, it was evaluated from different approaches. The analyses considered were: Complete data set analysis, level analysis, performance evolution across the complete set and performance evolution across learning method set. There were considered some variables for the performance measurement. These variables are calculated for each learning method and are explained in the next paragraphs.

COMPLETION TIME: It refers to the seconds needed for each method completion. Was just considered the time of the outGoal and InGoal states, in other words, was measured just the time inside the 30 seconds margin, this because some subjects wanted to make a pause during the experiment (the pause was made just on the rest state, if a task was being performed the subject must finish it and then pause), so a total experiment time would not be an objective measurement.

EFFECTIVENESS PERCENTAGE: This value is obtained from the ratio between the number of successful tasks and number of total tasks. The number of total tasks can vary depending the analysis performed, for instance the number of total tasks for an overall approach considering all the actions the total task are 30, while for an action are 6. This value represents the main goal of the experiment: which approach is having better successfulness ratio. But in fact the other two values could give some strength to methods that did not have a good effectiveness percentage.

UNSUCCESSFUL TASKS INGOAL RATE: This value shows the relation between the number of unsuccessful task samples that were inside the inGoal state and the total samples part of an unsuccessful task (again just considering the 30 seconds part). In other words, how many samples of the entire unsuccessful set were inside the goal zone. This value represents a stability measurement, so the higher the value, the more unstable the method. Then a possible solution could be an output filtering to avoid this event and increase the method performance.

5.4.1.1 *Complete data set analysis*

In this analysis, were considered all movement levels. The data was divided by actions, but also there was an overall approach including all of them. For this analysis were considered the three measurement variables: Completion time, effectiveness percentage and unsuccessful tasks InGoal rate.

First of all, in order to see how the subjects were performing the experiment, on Table 13 and on Figure 33 are shown the overall measurement variables. This overall representation will be the only one done by subject, as the information quantity is really big. From now the measurement variables for the other analysis will be done using the average of the twelve subjects. Actually this action is justified by the fact exposed on the previous section, the complete statistical distribution can be seen just from an all subjects point of view. On Table 14 and on Figure 34 are shown the average results of the complete data set analysis for each movement and the overall approach.

Table 13: Overall experiment 2 performance (Experiment completion time, successful tasks percentage and unsuccessful tasks InGoal rate)

Completion Time

Subj.	1	2	3	4	5	6	7	8	9	10	11	12	AVG±SD
STACK	352.40	521.09	347.83	538.26	365.99	417.93	219.66	586.92	452.57	743.20	595.77	432.94	464.55±140.67
EMG	366.52	629.31	660.78	679.49	489.62	654.81	502.54	770.17	567.56	691.91	438.24	429.06	573.33±126.27
FSR	489.87	788.35	644.93	739.55	525.70	599.19	755.01	604.48	630.53	569.04	567.76	667.53	631.83±92.44
ENS	521.24	625.59	345.87	705.66	440.85	578.26	518.23	645.73	699.86	559.84	524.80	651.81	568.14±106.57

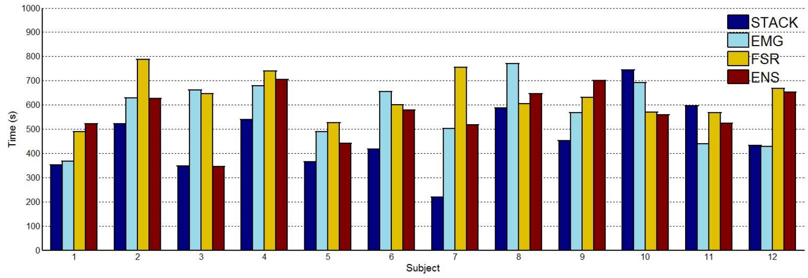
Successful Tasks Percentage

Subj.	1	2	3	4	5	6	7	8	9	10	11	12	AVG±SD
STACK	83.3	60	93.3	63.3	83.3	80	93.3	56.6	73.3	36.6	60	80	71.9±16.9
EMG	80	56.6	43.3	40	66.6	46.6	66.6	26.6	60	36.6	73.3	73.3	55.8±16.9
FSR	63.3	23.3	53.3	30	63.3	46.6	26.6	50	53.3	63.3	66.6	50	49.1±15
ENS	70	53.3	83.3	43.3	80	56.6	63.3	56.6	46.6	60	70	56.6	61.6±12.2

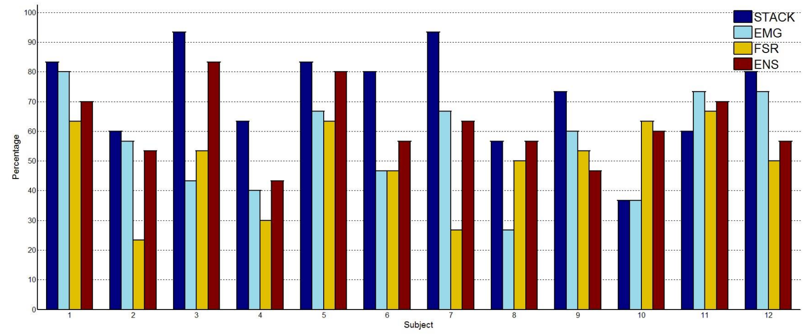
Unsuccessful Tasks InGoal Rate

Subj.	1	2	3	4	5	6	7	8	9	10	11	12	AVG±SD
STACK	0.112	0.044	0.069	0.056	0.034	0.106	0.039	0.027	0.084	0.080	0.049	0.094	0.066±0.029
EMG	0.171	0.060	0.114	0.146	0.163	0.080	0.038	0.092	0.097	0.084	0.123	0.065	0.103±0.042
FSR	0.051	0.007	0.045	0.015	0.004	0.035	0.006	0.016	0.029	0.092	0.023	0.012	0.028±0.025
ENS	0.035	0.031	0.202	0.024	0.089	0.139	0.052	0.060	0.053	0.092	0.150	0.043	0.081±0.056

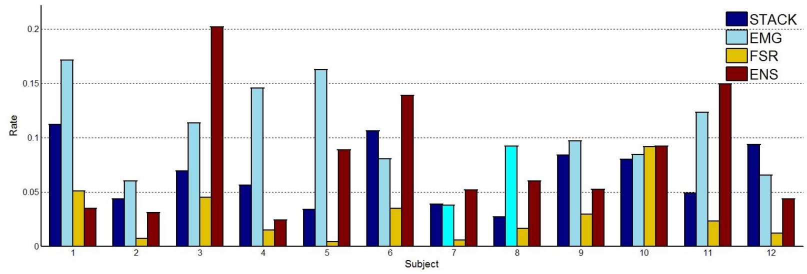
From the results previously presented, the stack method had an average better completion time (just three subjects were not having stack method as the fastest). Actually completion time inside actions was not so different be-



(a) Completion time



(b) Effectiveness percentage



(c) Unsuccessful tasks inGoal rate

Figure 33: Overall measurement variables for all the subjects

tween methods, but when those times were stacked in the overall approach, then the time difference increased.

Talking about the effectiveness percentage the stack method was the one with more successful tasks, on the other hand FSR method was the one having in most of the cases the lower successful tasks. For EMG and ensemble methods, the subjects were having different performances, nevertheless in the average movement's measurements and the overall approach, the ensemble learning had an almost equal or higher number of successful tasks compared with the EMG ones. The most successful action was the wrist supination.

As expected (considering the results of the first experiment), EMG sensors were the most unstable ones, being the method which was more time in the inGoal state but without being able to stay the 1.5 consecutive seconds.

Table 14: Experiment 2 average performance (Experiment completion time, successful tasks percentage and unsuccessful tasks InGoal rate) by action

Completion Time

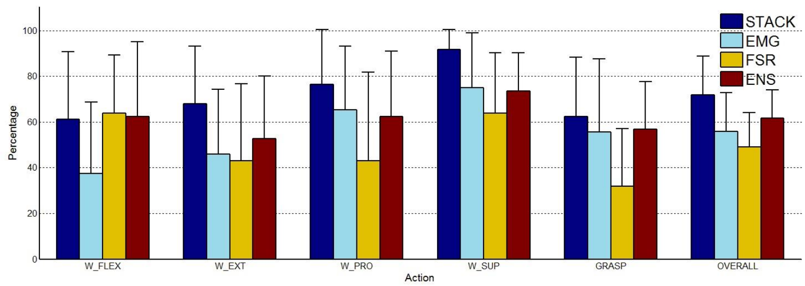
	Actions					
	W.FLEX.	**W.EXT.**	**W.PRO.**	**W.SUP**	**GRASP**	**OVERALL**
STACK	108.7±44.8	99.8±37.2	85.3±44.2	54.7±21.6	116.1±34.0	**464.5±140.67**
EMG	136.6±40.4	126.4±36.1	100.7±43.4	84.1±36.2	125.5±40.3	**573.3±126.27**
FSR	97.5±38.6	142.9±32.8	137.7±40.1	104.5±36.3	149.2±27.1	**631.8±92.44**
ENS	115.9±43.7	122.2±33.6	112.0±44.3	99.7±36.0	118.5±37.8	**568.1±106.57**

Successful Tasks Percentage

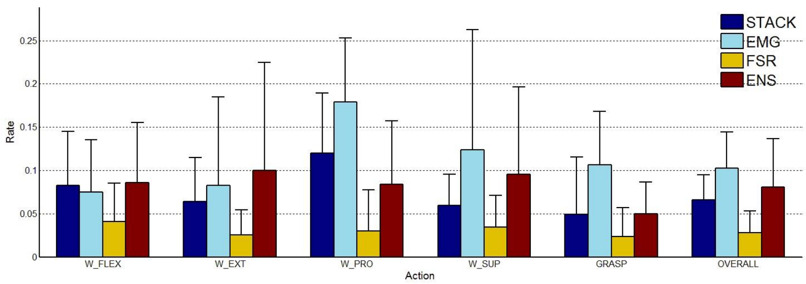
	Actions					
	W.FLEX.	**W.EXT.**	**W.PRO.**	**W.SUP**	**GRASP**	**OVERALL**
STACK	61.1±29.6	68.1±25.1	76.4±24.1	91.7±8.7	62.5±25.7	**71.9±16.9**
EMG	37.5±31.1	45.8±28.5	65.3±27.9	75.0±24.1	55.6±32.0	**55.8±16.9**
FSR	63.9±25.5	43.1±33.7	43.1±38.6	63.9±26.4	31.9±25.1	**49.2±15.1**
ENS	62.5±32.7	52.8±27.4	62.5±28.5	73.6±16.6	56.9±20.7	**61.7±12.3**

Unsuccessful Tasks InGoal Rate

	Actions					
	W.FLEX.	**W.EXT.**	**W.PRO.**	**W.SUP**	**GRASP**	**OVERALL**
STACK	0.083±0.062	0.064±0.051	0.120±0.070	0.060±0.036	0.049±0.066	**0.066±0.029**
EMG	0.075±0.060	0.083±0.102	0.179±0.074	0.124±0.138	0.106±0.062	**0.103±0.042**
FSR	0.041±0.044	0.025±0.029	0.030±0.048	0.035±0.036	0.023±0.034	**0.028±0.025**
ENS	0.086±0.069	0.100±0.125	0.084±0.074	0.096±0.101	0.050±0.037	**0.081±0.056**



(a) Effectiveness percentage



(b) Unsuccessful tasks inGoal rate

Figure 34: Average measurement variables for each action and the overall approach

5.4.1.2 *Level analysis*

For level analysis was just considered the effectiveness percentage value. The data was not just divide as in the previous analysis by action and overall approach, but also by movement degree level. Calling *high* to the complete movement, *medium* to the two thirds level and *low* to the one third level. In this way on Table 15 and Figure 35 are shown the effectiveness percentages by taking into account just the tasks performed by each level.

Table 15: Experiment 2 average successfulness percentage by action level intensity for each action (High, medium and low)

High Level

	Actions					
	W.FLEX.	**W.EXT.**	**W.PRO.**	**W.SUP.**	**GRASP**	**OVERALL**
STACK	45.83±45.02	50.00±36.93	75.00±39.89	83.33±24.62	62.50±31.08	**63.33±21.03**
EMG	50.00±52.22	70.83±39.65	62.50±43.30	50.00±47.67	54.17±33.43	**57.50±21.79**
FSR	54.17±45.02	37.50±37.69	50.00±47.67	66.67±38.92	54.17±45.02	**52.50±21.79**
ENS	75.00±45.23	66.67±38.92	54.17±39.65	50.00±42.64	70.83±39.65	**63.33±20.60**

Medium Level

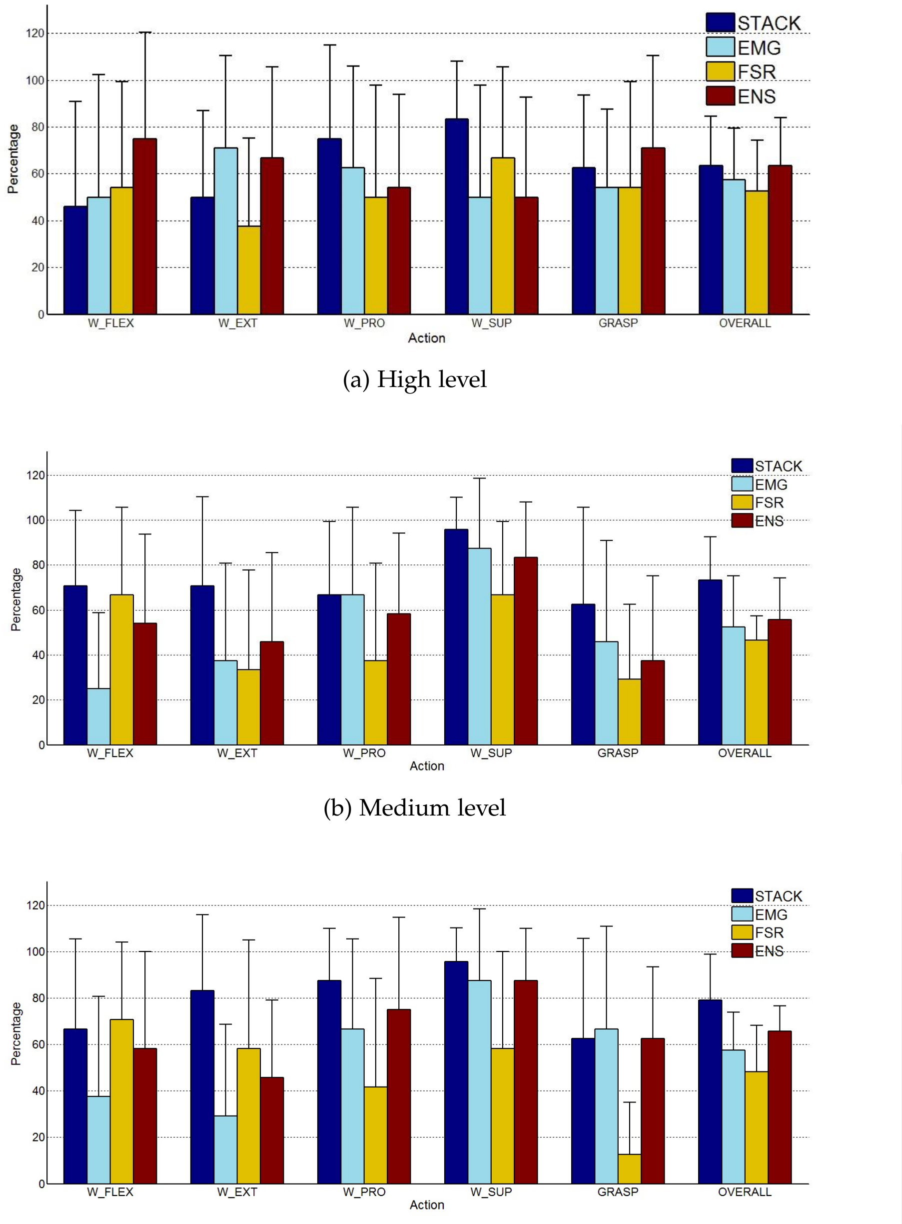
	Actions					
	W.FLEX.	**W.EXT.**	**W.PRO.**	**W.SUP.**	**GRASP**	**OVERALL**
STACK	70.83±33.43	70.83±39.65	66.67±32.57	95.83±14.43	62.50±43.30	**73.33±19.23**
EMG	25.00±33.71	37.50±43.30	66.67±38.92	87.50±31.08	45.83±45.02	**52.50±22.61**
FSR	66.67±38.92	33.33±44.38	37.50±43.30	66.67±32.57	29.17±33.43	**46.67±10.73**
ENS	54.17±39.65	45.83±39.65	58.33±35.89	83.33±24.62	37.50±37.69	**55.83±18.32**

Low Level

	Actions					
	W.FLEX.	**W.EXT.**	**W.PRO.**	**W.SUP.**	**GRASP**	**OVERALL**
STACK	66.67±38.92	83.33±32.57	87.50±22.61	95.83±14.43	62.50±43.30	**79.17±19.75**
EMG	37.50±43.30	29.17±39.65	66.67±38.92	87.50±31.08	66.67±44.38	**57.50±16.58**
FSR	70.83±33.43	58.33±46.87	41.67±46.87	58.33±41.74	12.50±22.61	**48.33±19.92**
ENS	58.33±41.74	45.83±33.43	75.00±39.89	87.50±22.61	62.50±31.08	**65.83±10.84**

In general, the stack method was one more time the one with more effectiveness percentage, however in the high level the best method was shared between stack and ensemble methods. Can be stated that the ensemble method works better with high levels. Additionally FSR method worked better in high level compared with the other two levels, while EMG did it in intermediate levels. This behavior was not expected, actually should be the inverse behavior as the sEMG only works with high muscle activation, while FSRs are more sensible to low intensity actions. This matter is discussed in the next section.

Looking at the unsuccessful tasks inGoal rate, the only considerable remark to do, is the unstable behave of the stack approach in the medium level, where an approximate average of 25% of the unsuccessful tasks time the approach was under the error threshold.



(a) High level

(b) Medium level

(c) Low level

Figure 35: Average effectiveness percentage for each intensity level considering each action and an overall approach

With the analyses already explained, it is evident that the stack approach was the best in all the effectiveness measurements. But an interesting phenomena happened if the data is analyzed in a different way. This analysis is presented on the following paragraphs.

5.4.1.3 *Performance evolution across the tasks set*

Considering that each subject task set can be divided in four sections, performing 30 tasks per section (in other words each learning method was applied on a different section), as there were so many repetitions and the methods were performing 30 consecutive tasks, could be possible that the methods applied in third and fourth section (tasks 61-120), could present a worst performance because of the subject's fatigue. The information presented on Table 16 and illustrated on Figure 36 shows, for each action and the overall approach, how the methods performance were changing depending in which section were applied. Section 1 refers to the tasks 1-30 , section 2 to the tasks 31-60, so on.

Figure 36: Average effectiveness percentage evolution across the experiment sections for each action and the overall approach

From this analysis there are two important remarks to do. Starting with the stack method, when it was performed in the first section it had a really bad performance compared with the previous analyses results, the ensemble method was the best on this first section. Nevertheless in the next sections, the stack method recovered its good performance, while the ensemble one performance started to decrease gradually. The other three methods were decreasing their performance across the experiment, but was with the sEMG signals where this behavior was happening in a gradually way too. This fact supports the idea proved in previous works [29], the sEMG signal is affected during stressing conditions (like long and demanding experiments) by the muscular fatigue and the formation of sweat between the sensors and the skin.

5.4.1.4 *Performance evolution across learning method set*

Finally, a similar analysis from the last one was done. In this case is considered the evolution of the performance through each method tasks (30 tasks), it does not matter the task set section where they were performed. The goal of this analysis is to prove the hypothesis of a possible subject learning through the repetitions, so the performance could be higher in the last ones.

From Table 17 and Figure 37 can be appreciated that there were not a considerable change through the repetitions. Then there were not considerable learning applied by the user, also it is important to remark that 30 tasks are not enough for the fatigue behavior in EMG and ensemble methods, so it was not affecting the performance .

Unfortunately, could not be performed an off-line performance analysis based on the nRMSE as with the previous experiment. The problem in this case was that the prediction signals were not uniform. For instance some subjects were going back to the rest position to try over again to perform the movement. As this behavior was not done for all the methods, then a nRMSE calculation over the prediction signals would not be an objective measurement.

Table 16: Average effectiveness percentage evolution across the experiment sections for each action and the overall approach

Wrist Flexion

	Section 1	Section 2	Section 3	Section 4
STACK	33.33±16.67	77.78±25.46	50.00±33.33	83.33±16.67
EMG	44.44±9.62	61.11±53.58	22.22±19.25	22.22±19.25
FSR	72.22±34.69	44.44±34.69	72.22±9.62	66.67±16.67
ENS	66.67±33.33	44.44±19.25	77.78±25.46	61.11±53.58

Wrist Extension

	Section 1	Section 2	Section 3	Section 4
STACK	44.44±38.49	88.89±9.62	66.67±16.67	72.22±9.62
EMG	44.44±19.25	61.11±34.69	44.44±34.69	33.33±33.33
FSR	55.56±48.11	61.11±19.25	33.33±44.10	22.22±9.62
ENS	44.44±34.69	61.11±19.25	38.89±41.94	66.67±0.00

Wrist Pronation

	Section 1	Section 2	Section 3	Section 4
STACK	44.44±9.62	94.44±9.62	83.33±0.00	83.33±28.87
EMG	77.78±38.49	72.22±19.25	44.44±25.46	66.67±28.87
FSR	83.33±16.67	22.22±38.49	22.22±19.25	44.44±48.11
ENS	72.22±19.25	88.89±9.62	55.56±19.25	33.33±33.33

Wrist Supination

	Section 1	Section 2	Section 3	Section 4
STACK	94.44±9.62	94.44±9.62	88.89±9.62	88.89±9.62
EMG	88.89±19.25	66.67±0.00	72.22±34.69	72.22±34.69
FSR	55.56±25.46	66.67±28.87	72.22±19.25	61.11±41.94
ENS	88.89±9.62	72.22±25.46	66.67±0.00	66.67±16.67

Grasp

	Section 1	Section 2	Section 3	Section 4
STACK	44.44±34.69	77.78±25.46	77.78±9.62	50.00±16.67
EMG	72.22±25.46	72.22±25.46	55.56±41.94	22.22±9.62
FSR	33.33±16.67	22.22±25.46	44.44±38.49	27.78±25.46
ENS	77.78±25.46	55.56±9.62	50.00±16.67	44.44±19.25

Overall

	Section 1	Section 2	Section 3	Section 4
STACK	**52.22±13.88**	**86.67±11.55**	**73.33±11.55**	**75.56±13.47**
EMG	**65.56±8.39**	**66.67±17.64**	**47.78±16.44**	**43.33±16.67**
FSR	**60.00±5.77**	**43.33±11.55**	**48.89±20.37**	**44.44±20.09**
ENS	**70.00±13.33**	**64.44±13.88**	**57.78±11.71**	**54.44±10.18**

5.4.2 *User experience surveys results*

From survey applied during the user experience stage, the results are presented on Table 18 for each subject. Although all the methods were close the moderate feeling for all the features, the stack and FSR methods were having

Table 17: Overall average effectiveness percentage evolution across the experiment tasks

	Task repetitions			
	1-10	11-20	21-30	26-30
STACK	72.50±18.15	73.33±24.25	70.00±21.74	66.67±27.41
EMG	55.00±23.16	55.00±16.79	57.50±19.60	60.00±24.12
FSR	47.50±18.15	48.33±14.67	51.67±23.68	56.67±25.35
ENS	59.17±9.00	64.17±17.82	61.67±18.50	61.67±21.67

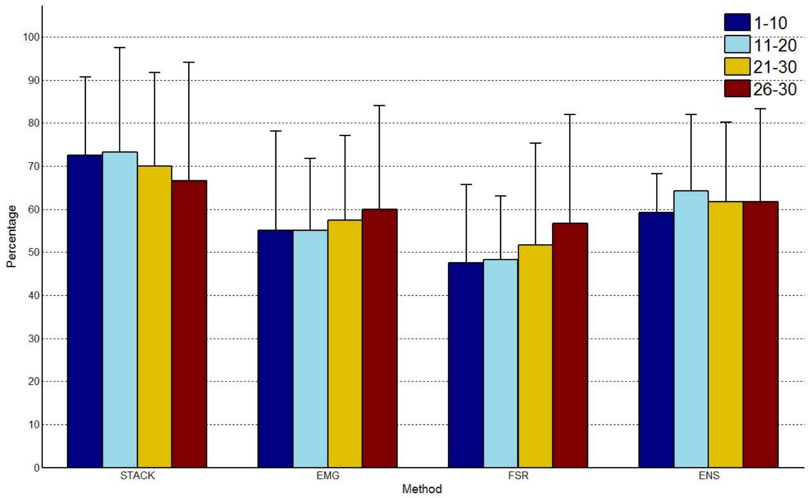


Figure 37: Overall average effectiveness percentage evolution across the experiment tasks

better impressions on the subjects, while the ensemble learning method had the lowest values of the survey.

Meanwhile on the post-experiment NASA TLX survey, as can be noticed on Table 19 in a complete subject way or in Figure 38 in an average representation, the experiment was demanding a neutral workload, being the effort one the highest demanding category, this fact is affecting directly to the EMG and ensemble learning methods as will be explained on the next section.

Figure 38: Experiment 2 workload percentage plot

5.5 EXPERIMENT 2 DISCUSSION

5.5.1 *Experiment performance*

Evidently the multi-sensors methods were having a better performance compared with the single-sensors methods, accomplishing in this way the main hypothesis of the project stated on Chapter 1. But was not expected the higher quality in the stack method performance compared with the other

Table 18: User experience survey results

Stack

Subject	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**AVG±SD**
Realism	6	4	6	3	5	5	3	2	3	3	4	6	**4.17±1.40**
Stability	4	5	7	4	6	5	1	3	4	2	5	7	**4.42±1.83**
Smoothness	5	6	7	6	7	3	1	3	2	4	6	5	**4.58±1.98**

EMG

Subject	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**AVG±SD**
Realism	5	5	4	4	4	5	6	2	3	5	2	5	**4.17±1.27**
Stability	2	5	2	3	4	4	5	4	3	3	2	4	**3.42±1.08**
Smoothness	4	6	4	5	7	5	3	3	2	2	3	5	**4.08±1.56**

FSR

Subject	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**AVG±SD**
Realism	3	4	6	2	4	4	2	6	3	6	3	6	**4.08±1.56**
Stability	3	4	6	3	5	4	6	5	4	7	2	5	**4.50±1.45**
Smoothness	5	6	6	4	7	4	6	5	3	6	3	6	**5.08±1.31**

Ensemble

Subject	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**AVG±SD**
Realism	6	3	4	2	5	5	2	3	2	2	3	5	**3.50±1.45**
Stability	6	2	2	1	6	4	1	3	2	1	2	5	**2.92±1.88**
Smoothness	6	5	1	2	7	4	1	3	2	1	2	5	**3.25±2.09**

Table 19: Experiment 2 NASA TLX workload percentages.

ID	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**AVG**
Mental Demand	4.76	57.14	66.66	28.57	85.71	28.57	38.09	52.38	19.04	76.19	52.38	4.76	*42.85*
Physical Demand	19.04	61.9	71.42	52.38	90.47	80.95	71.42	66.66	52.38	52.38	61.9	71.42	*62.69*
Temporal Demand	61.9	57.14	66.66	47.61	52.38	52.38	52.38	52.38	23.8	76.19	38.09	4.76	*48.8*
Performance	33.33	66.66	71.42	71.42	38.09	52.38	47.61	66.66	42.85	71.42	71.42	28.57	*55.15*
Effort	76.19	57.14	80.95	61.9	80.95	85.71	66.66	66.66	71.42	71.42	61.9	76.19	*71.42*
Frustration	52.38	57.14	52.38	61.9	66.66	66.66	66.66	61.9	52.38	71.42	52.38	4.76	*55.55*
Overall WL	*41.27*	*59.52*	*68.25*	*53.96*	*69.04*	*69.04*	*57.14*	*61.11*	*43.65*	*69.84*	*56.34*	*31.74*	***56.08***

three methods, especially when the stack approach was the easiest implemented method from the two multi-sensors ones. The success of the stack method happened because the method was based on a bigger input space. Then the movement patterns have a higher possibility to be different from each other. If the EMG signals are not so different or are bad trained they will be always supported by the different patterns the FSR could generate, and the same the other way, if the FSR intermediate values are not differentiable from other movements, then are supported by the EMG ones. The main difference with ensemble learning is that the last one is training with predictions, so if one of the sensors is not giving good signals, then the prediction considered for stacking will be bad, and the ensemble learning will be also not good. Maybe it will be better than the EMG and FSR approaches, but not good enough compared with the stack approach.

Even the ensemble learning could have a better performance (basing on the off-line analysis with the first experiment data), the long experiment performance threw some interesting limitations of this method. Since it was tested in the off-line analysis, ensemble learning apparently was inheriting some EMG prediction features, like the oscillating output signal behavior. The fatigue sensibility could not be noticed, as there were just 10 repetitions in the previous experiment.

Another point to consider is the way the learning is done in this approach. As was previously defined, first it is trained in the same way as the other approaches, but for the input space construction for the third machine learning model the entire data set is considered. Then, what is happening in the rise, hold and fall states on the training stage, is affecting the ensemble learning. So the training should be really clean, without any mistakes, following the stimulus. What happened in the experiments, is that experienced users were performing a really good training, but the others training were varying. This fact is actually good as it is giving robustness to the experiment. The main feature that affects the training are the EMG impulses. When the subject performs a movement, in the rise part there was a high impulse on the EMG signal that after some seconds decreases almost to rest (Figure 11). Some users could keep the high amplitude of the EMG by tensing the movement, but some others could not, then the capture stage was performed in a really different amplitude compared with the previous high impulse.

Furthermore, the repetitions performed were so low compared with the first experiment ones (3 against 9), this could be also affecting the ensemble learning performance. However one of the implicit goals of the online experiment was to try in some way the robustness of the methods for real applications, following this line it is not practical for a user to perform 10 repetitions of the five actions just for training.

Even though there was the phenomena showed on the *Performance evolution across the tasks set* section, where the stack method was not having a good enough performance on the first section of the experiment, must be considered, and not just for this but for all the analyses, the length of the task set used for the analysis. For instance, the total number of tasks performed by all the users per method was 360, the complete overall approach is considering the entire set, but the complete action approaches are considering just 72 tasks per method, an overall level analysis is taking into account 24 ,the section evolution overall analysis was done over 90 tasks per method and finally the section evolution action analysis was just considering 18 tasks.

In spite of this fact and basing on the overall performance section evolution, can be stated that the stack method needs some adaptation in order to have a considerable good performance, based on the bad performance showed on the first section of the task set.

Without leaving aside the EMG and FSR methods, there is still pending the explanation about the behavior of these method's predictions on the different levels. First of all EMG sensors were presenting really defined patterns. Even they were going weaker, when the muscle activation happened the patterns were completely different from each movement. On the other hand, FSRs also presented differentiable patterns but with an interesting behavior, each movement was activating in different intensities the FSR sensors, but most of them were activating almost the same sensors, in other words the shape on the radial plot was almost the same for each action, the only thing that was changing was the size(amplitude) of that shape. This behavior can be appreciated even in Figure 11 from the first experiment. On the FSR signal, the highest amplitude sensor was always the same for each action, of course with different amplitudes, while in the EMG the highest amplitude sensor was different for almost all the actions. That is why FSR had better performance in high levels, as their patterns were differentiable. The problem appeared on intermediate levels, where, for instance, an intermediate level of one movement could have similar shape with a less intense movement, so the prediction was not done properly. This characteristic could be even more evident due to position of the bracelet, closer than usual to the elbow.

Summarizing, both multi-sensor learning approaches were better than the single-sensor ones. Ensemble learning could have a good performance, but it depends on the way the training is done, what is related with the user expertise and also could be related with the number of repetitions for training. Additionally, ensemble approach presents the drawback of fatigue sensibility, and unstable output (that could be fixed with by filtering it). Stack approach seems to be more robust and evidently with better performance, even there was the fact of bad performance when it was done first.

5.5.2 *User experience surveys*

Starting with the online survey, the results are supported by the ideas established in the first part of this discussion section. Considering that this survey was applied at the end of the experiment, after the 120 tasks, it is understandable that the EMG and ensemble methods were the worst evaluated, as they are influenced by the fatigue. Also the unstable output feature of both methods was reflected not just in the stability statement but also in the smoothness one. Additionally as the ensemble method is more sensible in intermediate levels compared with the EMG, some users interpreted this fact as an unreal behavior. On the other hand the results obtained on the FSR and stack methods just complement the conclusions obtained since the first experiment analysis.

About the workload survey, the main remark to do is the high effort demand, which is close to the fatigue on the EMG and ensemble methods, this survey is not showing anything new, but just support the statements already defined.

6 SOCIAL, ENVIRONMENTAL AND ECONOMICAL COMMITMENTS

The scope of this and all the projects done in robotic prosthetics field, goes beyond a scientific research. They must have social, environmental and economical commitments so then they can be considered, along with the technical achievements, real life applications. In this chapter are presented the ways this project impacts the three fields exposed.

6.1 SOCIAL COMMITMENT

As mentioned on the thesis introduction, one of the motivations of the robotic prosthetics researches is the life quality improvement of the amputee people, specially when the hand is a limb used for precise and daily life interaction tasks. The statistics show how the patients are affected in an emotional way. If the researches results are recovering at least some of the amputees usual tasks and activities, then these works are going on the right direction, and they are supported by a humanitarian cause.

6.2 ENVIRONMENTAL COMMITMENT

The work done on this project was not presenting any menace to the environment, moreover was not considered as a potential danger project according the Directive 2011/92/EU of the European Parliament and of the Council on the assessment of the effects of certain public and private projects on the environment [25]. Although the project is not dangerous with the environment, it is taking care of it. As an example, the battery used for the wireless device is a rechargeable LiPo one, which can work, according to the provider, for a year. A considerable lifespan difference compared with the disposal alkaline batteries. Even if the LiPo life cycle finishes, it is possible to send it to a recycle batteries collector.

6.3 ECONOMICAL COMMITMENT

Even though the work done in the DLR is focused on a high-technology device, which normally is not cheap, also focuses on a potential reduction of the device price. The main example of this commitment is the interest to research on new sensing methods different from the sEMG, not just for the performance improvement, but for the cost reduction (as mentioned on Chapter 2, each sEMG sensor costs around 150 €). In this way the possible future devices for real life application could be affordable to the people whom need it.

As part of this section, an estimated cost analysis was done. It is important to say that this analysis is oriented just for this project research development, were not included device design, neither previous or parallel works costs.

6.3.1 *Costs*

In order to have a total cost approach (Table 23), the cost calculation was divided in four different categories: hardware (Table 20), software (Table 21), human resources (Table 22) and indirect costs. However, it is important to make some remarks for each of the categories for a better understanding.

All the costs calculations are subject to the Spanish and the European Union legislation, as the master's degree university is located on these country and political-economical union. For the hardware and software costs, was considered the amortized amount corresponding to the weeks of use for each element, there was applied the constant amortization approach, as the depreciation across time is considered as continuous and homogeneous. As lifespan of some elements is undefined or unknown, were considered for the lifetime the maximum allowed for tax deductions in Spain according to the 27/2014 Law about the corporate tax. This law is allowing for instance a maximum lifetime of 10 years for electronic devices, 8 years for IT devices and 6 years for software.

On the hardware costs were considered an approximate cost of the wireless device for a usage approach and not for a construction one, as this last task was not part of this project. Furthermore as the battery has a lifespan of 1 year [2], was considered as a separated element. All the elements of the wireless device were considered as electronic devices. The PC used was considered as IT device for the amortization calculation.

For the human resources was considered an average engineer salary of 35 000 €per year, also there was just considered one person working on the project (the thesis author), even there were actually supervisors and collaborators contributing to the project accomplishment. Considering a full time work (1800 billable hours per year), then the cost for each worked hour is 19.44 €. This amount was considered for the human resources costs, as the project tasks duration considered were defined by hours as the time unit.

Finally as this thesis is a research project, it is difficult to determine indirect costs involved on the development. In order to help to the definition of those costs, every country establishes indirect costs rate that could be applied on research projects. In the case of the European Union countries, the Horizon 2020 framework programme for research and innovation, proposes an indirect cost flat rate of 25% of the direct costs [7]. As a final consideration the Spanish VAT (21%) was applied to get the project final estimated cost, which amounts to 28536.03 €.

Table 20: Hardware resources costs

Hardware resources costs				
	Unit cost (€)	**Units**	**Lifetime (years)**	**Amortization (€)**
Wireless Device				
sEMG sensor	150.00	10	10	51.92
FSR	5.00	10	10	1.73
Others (incl. electronics board, case, bracelets, housings)	110.00	1	10	3.80
Battery	19.00	1	1	6.57
PC	195.00	1	8	12.18
TOTAL	**479.00**	**-**	**-**	**76.20**

Table 21: Software resources costs

Software Costs				
	Unit Cost (€)	**Units**	**Lifetime**	**Amortization (€)**
MATLAB Standard	2000.00	1	6	128.20
Microsoft Visual Studio Express	0.00	1	-	0.00
Xamarin Studio	0.00	1	-	0.00
WinShell	0.00	1	-	0.00
ShareLaTeX	0.00	1	-	0.00
TOTAL	**2000.00**	**-**	**-**	**128.20**

Table 22: Human resources costs

Human resources costs		
Task	**Time (h)**	**Cost (€)**
Literature review	120	2332.80
Software and device familiarization and tests	136	2643.84
Experiment 1		
Development	64	1244.16
Test	32	622.08
Implementation	10	194.40
Results analysis	46	894.24
Experiment 2		
Development	274	5326.56
Test	40	777.60
Implementation	24	466.56
Results analysis	30	583.20
Documentation	184	3576.96
TOTAL	**960**	**18662.40**

Table 23: Total costs

Total costs	
Concept	**Cost (€)**
Hardware	76.20
Software	128.20
Human resources	18662.40
Total	18866.80
Indirect costs (25%)	4716.70
Subtotal	23583.50
VAT (21%)	4952.53
TOTAL	**28536.03**

7 CONCLUSIONS AND FUTURE WORK

The development of this project focused on the improvement of robotic hand prosthesis joints movement prediction, going further the single-sensor researches previously done in the area. Additionally, the project worked with wrist movements, that actually have not been enough evaluated in previous publications, focusing them on individual fingers movements. The fact of working with wrist actions opens a new path that can help patients to perform more complex tasks, which, along with the fingers movements could offer a more realistic approach.

The thesis followed a defined path to accomplish its objectives by performing two experiments. On the first one, there was explored the use of FMG as an alternative or complementary signal to the often used sEMG, trying to present a cheaper and simpler system compared not just with the ones used on research, but with the ones offered on market.

On this first experiment the FSR signals showed better and more significant properties than the sEMG ones. They had more stable signals that derived in a more constant experiment performance through the subject, a higher number of activated sensors and finally more separated movements. However FSR signals were not perfect at all, they were affected by some drawbacks, for instance their unstable rest position which is measuring the small movement variations that the hand could have on this position. Furthermore, FSRs detect the activity from muscles other than the sEMG ones. The FSR disadvantages along with the significant differences between both signals, supported the main thesis idea of fusing both input sensors.

Staying on the first experiment, the device used and tested led to interesting conclusions. Besides the possibility of using the device for future experiments, research and perhaps commercialization; the main goal of using a wireless device is to improve the portability and maneuverability of an acquisition device, then the user could carry it in his or her daily life activities without any problems.

On the second experiment, two multi-sensor learning models were developed and then tested with the single-sensor ones. The stack method, which actually was the simplest and the lowest computation time and resources consumption from both of the multi-sensors approaches, was the method with the best overall performance (even if it perhaps needs more adaptation time to reach its best performance). For this result were considering the completion time of the experiment, the successful tasks performed and the subjects opinion about the stability, the realism and the smoothness of the method prediction. The stack method success is based on two facts: first the use of a bigger input space, allowing the probability to perform more different patterns and in consequence more separated actions. The second fact that propitiated the stack method success was the lack of stability and good performance from other methods. For instance, the fatigue sensibility of both EMG and ensemble methods and the unstable output signal of the same methods.

Talking about the FSR method, even if it was having a good performance on high values, the bad prediction on the intermediate ones, produced by the already mentioned measurement features, derived on an overall bad result for this method.

Ensemble learning presented a high influence from the learning performance as the number of samples considered for training were higher than the other methods. This feature, as explained, could derive on a bad performance of the algorithm as it is depending on each of the sensor model prediction, but this fact is also producing an interesting behavior, the multi-DOF control training single DOF actions. Until now if a composed DOF action, for instance a pointing position (all fingers flexed except for the index one), was desired to train, then the way to do it was by defining the composed DOF action as a single action. This means, following the same pointing position example, that if each finger would be trained in a separate way, afterwards would be really difficult to perform the action, then the user must train the models by doing the pointing position. The fact that ensemble learning is considering two different model outputs produces that if an action is not significant for one of them and other action is performed at the same time, then the prediction could react with both them, creating a composed movement. This phenomena could be appreciated for instance on the wrist rotation actions, the hand could be also extended or flexed while the rotation was performed. This behavior could be also supported on the fact that each of the actions are performed on different axis.

Anyway, at the end the hypothesis was proved. It is possible to improve the prediction performance by combining two different input techniques (FMG and sEMG on this case). On this project the two multi-sensors methods were having better results than the single-sensor ones. Each of the multi-sensor approaches presented special properties to be used for certain application and also future developments.

The thesis contributions were: a comparison analysis of the sEMG and FSR signals properties. A user opinion about the wireless device previously developed and built, but tested on this project. Additionally were modified the single-sensor models to create a multi-sensor one (stack method) and was developed a firstly used on prosthetics field method also based on the RRRFF algorithm (ensemble method). The PC GUI was modified so now it is able to train with the four methods and test prediction on a experiment set or on a free mode. Finally was demonstrated that multi-sensors approaches were having better performance although there is still a big path for improvement.

FUTURE WORK There is still a long way to cover in the robotic prosthetics field. The gap defined on the thesis introduction between the research work and commercial one, is still big. However nowadays, as was demonstrated across this work, the research started to focus not just on the performance prediction, but also on the real life usability and applicability, going further through new input methods, acquisition devices and prediction algorithms.

Following this line there is an interesting hypothesis created from the experiments design and results. The performance of the robotic hand does not depend only of the sensors, the actuators or the learning algorithm (of course these elements are the basis of the prosthetics field and their importance is evident), but also the user is having a considerable impact on this performance. Considering for instance the different prediction behaviors on the experiments depending the subject, or how some of them were adapting their movements in order to accomplish the task, moreover the surveys applied gave an extra analysis source that could support or refute the qualitative results. Would be interesting to change in some way the approach by focusing on the subjects feedback to improve the prediction performance, considering them an important element just as the ones mentioned in the last paragraph.

Talking specifically about the technical part of this project development, of course the first task to do after this project, should be the implementation of the thesis results on a real prosthetic hand. As was mentioned previously on the work, it is not difficult to do as the stimulus vectors are analog to the outputs needed to move a robotic hand. Apart of that, future research must follow the goals defined on this project. For instance the ensemble learning approach could be improved by using a different machine learning algorithm on its third model (the one that is using the stacked outputs as the input space), also as was defined, the incremental feature of the RRRFF algorithm could be implemented for ensemble learning, and all the methods could present a better performance by conditioning the output signals (filtering them for example). What is even better, new sensors mixing approaches could be researched and tested in order to find the one that offers the best performance. This combination is not just limited to machine learning algorithms, but also to topics related with sensor fusion for instance.

Furthermore, the future work can be focused on exploring more hand actions, trying different kind of grasps and combining wrist movements with individual fingers one. Finally a last possibility is the exploration of the already mentioned multi-DOF control by training single DOF movements.

In this way, making improvements from different perspectives for sure will help to develop a robotic hand system that is not used just as a tool for certain tasks, but as a complete replacement of the lost limb.

Part I

APPENDIX. USABILITY AND USER EXPERIENCE SURVEYS

SUS

Please cross the cell that better fits with your device experience. (Leave the total column empty)

	Strongly disagree				Strongly Agree	**Total**
	1	2	3	4	5	
I felt comfortable with the device						
I found the device unnecessarily complex						
I thought the device was easy to use						
I think that I would need the support of a technical person to be able to use this device						
I found the various functions in this system were well integrated						
I thought there was too much inconsistency in this device						
I would imagine that most users would learn to use this device very quickly.						
I found the device very cumbersome to use						
I felt very confident using the device.						
I needed to learn a lot of things before I could get going with this device.						

NASA TLX

How mentally demanding was the task?

Very Low Neutral Very High

How physically demanding was the task?

Very Low Neutral Very High

How hurried or rushed was the pace of the task?

Very Low Neutral Very High

How successful were you in accomplishing what you were asked to do?

Perfect Neutral Failure

How hard did you have to work to accomplish your level of performance?

Very Low Neutral Very High

How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low Neutral Very High

Travis Survey

Step 1: Read over the following list of words. Considering the device you have just used, tick those words that best describe your experience with it. You can choose as many words as you wish.

- o Unpredictable
- o Inadequate
- o Simple
- o Intuitive
- o Poor quality
- o Expected
- o Distracting
- o Ordinary
- o Secure
- o New
- o Accessible
- o Effective
- o Stimulating
- o Time-consuming
- o Busy
- o Credible
- o Slow
- o Sophisticated
- o Difficult
- o Rigid
- o Ineffective
- o Boring
- o Contradictory
- o Reliable
- o Familiar
- o Trustworthy
- o Comfortable
- o Efficient
- o Too technical
- o Satisfying
- o Meaningful
- o Usable
- o Creative
- o Non-standard
- o Comprehensive
- o Insecure
- o Clear
- o System-oriented
- o Professional
- o Useful
- o Engaging
- o High quality
- o Ambiguous
- o Confusing
- o Inconsistent
- o Frustrating
- o Unattractive
- o Stressful
- o Irrelevant
- o Organised
- o Easy to use
- o Convenient
- o Innovative
- o Effortless
- o Complex
- o Awkward
- o Incomprehensible
- o Fast
- o Attractive
- o Hard to Use
- o Impressive
- o Predictable
- o Motivating
- o Friendly
- o Stable
- o Fun
- o Unconventional
- o Understandable
- o Powerful
- o Intimidating
- o Advanced
- o Consistent
- o Illogical
- o Time-saving
- o Painful

Step 2: Now look at the words you have ticked. Circle five of these words that you think are most descriptive of the product and write a simple reason about your decisions.

Word	Reason
1.	
2.	
3.	
4.	
5.	

Learning Approach User Experience

APPROACH 1						
How real is the hand model movement?						
Unreal			Moderate			Real
How stable is the hand model movement?						
Completely unstable			Moderate			Completely stable
How smooth is the hand model movement?						
Low			Moderate			High

APPROACH 2						
How real is the hand model movement?						
Unreal			Moderate			Real
How stable is the hand model movement?						
Completely unstable			Moderate			Completely stable
How smooth is the hand model movement?						
Low			Moderate			High

APPROACH 3						
How real is the hand model movement?						
Unreal			Moderate			Real
How stable is the hand model movement?						
Completely unstable			Moderate			Completely stable
How smooth is the hand model movement?						
Low			Moderate			High

APPROACH 4						
How real is the hand model movement?						
Unreal			Moderate			Real
How stable is the hand model movement?						
Completely unstable			Moderate			Completely stable
How smooth is the hand model movement?						
Low			Moderate			High

BIBLIOGRAPHY

[1] Ethem Alpaydin. *Introduction to Learning Machine*. Second. The MIT Press, 2010, p. 487.

[2] Touch Bionics. *i-limb ultra ™ultra Clinician Manual*. 2013. URL: http://www.touchbionics.com/sites/default/files/i-limb_ultra_clinician_manual_issue._2.pdf.

[3] John Brooke. "Usability Evaluation in Industry." In: ed. by Bernard A. Weerdmeester Patrick W. Jordan Bruce Thomas and Ian L. McClelland. Taylor & Francis, 1996. Chap. SUS: A "quick and dirty" usability scale, pp. 189–194.

[4] M. C. Carrozza, B. Massa, S. Micera, M. Zecca, and P. Dario. "A "Wearable" Artificial Hand for Prosthetics and Humanoid Robotics Applications." In: *Proceedings of the 2001 IEEE-RAS International Conference on Humanoid Robots*. 2001.

[5] Cheng-Hung Chen. "Hybrid Control Strategies for Smart Prosthetic Hand." PhD thesis. Idaho State University, 2009.

[6] Amputee Coalition. *Limb Loss in the U.S.A.* Web Page Infographic. 2016. URL: http://www.amputee-coalition.org/wp-content/uploads/2016/04/LLAM-Limb-Loss-in-the-USA-Web.pdf.

[7] European Comission. *Horizon 2020, The New EU Framework Programme for Research and Innovation*. Informative presentation. 2013. URL: http://ec.europa.eu/research/horizon2020/pdf/press/horizon2020-presentation.pdf.

[8] Mathilde Connan, Eduardo Ruiz Ramírez, Bernhard Vodermayer, and Claudio Castellini. "Assessment of a wearable force- and electromyography intent-detection system for prosthetic myocontrol." In: *Frontiers in Neurorobotics* (2016). under review.

[9] William Craelius. "The Bionic Man: Restoring Mobility." In: *Science* 295 (2002), pp. 1018–1021. DOI: 10.1126/science.295.5557.1018.

[10] Thomas G. Dietterich. "Ensemble Learning." In: *The Handbook of Brain Theory and Neural Networks*. Second. 2002.

[11] Richard L. Drake, A. Wayne Vogl, and Adam W. M. Mitchell. *Gray's anatomy for students*. Second. Churchill Livingstone, 2010, 650 to 755.

[12] Yinfeng Fang, Nalinda Hettiarachchi, Dalin Zhou, and Honghai Liu. "Multi-modal Sensing Techniques for Interfacing Hand Prostheses: A Review." In: *IEEE Sensors Journal* 15.11 (2015), pp. 6065–6076. DOI: 10.1109/JSEN.2015.2450211.

[13] Arjan Gijsberts and Giorgio Metta. "Real-time model learning using Incremental Sparse Spectrum Gaussian Process Regression." In: *Neural Networks* (2012). DOI: 10.1016/j.neunet.2012.08.011.

[14] Arjan Gijsberts, Rashida Bohra, David Sierra González, Alexander Werner, Markus Nowak, Barbara Caputo, Maximo A.Roa, and Claudio Castellini. "Stable myoelectric control of a hand prosthesis using non-linear incremental learning." In: *Frontiers in Neurorobotics* 8 (2014). DOI: 10.3389/fnbot.2014.00008.

[15] A S Jain and Douglas P H Robinson. *Synopsis of Causation, Upper Limb Amputation*. Ministry of Defence (UK). 2008.

[16] Ning Jiang, Strahinja Dosen, Klaus-Robert Müller, and Dario Farina. "Myoelectric Control of Artificial Limbs-Is there a Need to Change Focus?" In: *IEEE Signal Processing Magazine* (2012), pp. 148–152. DOI: 10.1109/MSP.2012.2203480.

[17] J C K Lai, M P Schoen, A Perez Garcia, D S Naidu, and S W Leung. "Prosthetic devices: challenges and implications of robotic implants and biological interfaces." In: *Proceedings of the Institution of Mechanical Engineers*. Vol. 221. 2007, pp. 173–183. DOI: 10.1243/09544119JEIM210.

[18] Maurice LeBlanc. *"Give Hope - Give a Hand" The LN-4 Prosthetic Hand*. Lecture slides for the Prespectives in Assistive Technology course. Stanford University. 2011. URL: https://web.stanford.edu/class/engr110/2011/LeBlanc-03a.pdf.

[19] Joan Lobo-Prat, Peter N Kooren, Arno HA Stienen, Just L Herder, Bart FJM Koopman, and Peter H Veltink. "Non-invasive control interfaces for intention detection in active movement-assistive devices." In: *Journal of NeuroEngineering and Rehabilitation* 11.1 (2014), p. 168. DOI: 10.1186/1743-0003-11-168. URL: http://dx.doi.org/10.1186/1743-0003-11-168.

[20] R. Merletti, A. Botter, C. Cescon, M.A. Minetto, and T.M.M. Vieira. "Advances in surface EMG: Recent progress in clinical research applications." In: *Critical reviews in biomedical engineering* 38.4 (2011), pp. 347–379.

[21] R. Merletti, M. Aventaggiato, A. Botter, A. Holobar, H. Marateb, and T.M.M. Vieira. "Advances in surface EMG: Recent progress in detection and processing techniques." In: *Critical reviews in biomedical engineering* 38.4 (2011), pp. 305–345.

[22] NASA. *NASA TASK LOAD INDEX (TLX) v 1.0 Paper and Pencil Package*. Human Performance Research Group, NASA Ames Research Center. Moffett Field, California, 1986.

[23] D. Subbaram Naidu, Cheng-Hung Chen, Alba Perez, and Marco P. Schoen. "Control Strategies for Smart Prosthetic Hand Technology: An Overview." In: *30th Annual International IEEE EMBS Conference*. 2008.

[24] Christian Nissler, Nikoleta Mouriki, and Claudio Castellini. "Optical Myography: Detecting Finger Movements by Looking at the Forearm." In: *Frontiers in Neurorobotics* 10.3 (2016). ISSN: 1662-5218. DOI: 10.3389/fnbot.2016.00003. URL: http://www.frontiersin.org/neurorobotics/10.3389/fnbot.2016.00003/abstract.

[25] European Parlament and Council of the European Union. *Environmental Impact Assessment (EIA) Directive (2011/92/EU)*. Official Journal of the European Union. Amended on April 2014. 2011.

[26] Eric C. Polley and Matk J. van der Laan. "Super Learner in Prediction." In: *U.C. Berkeley Division of Biostatistics Working Paper Series*. University of California, Berkeley, 2010.

[27] Institute for Quality and Efficiency in Health Care. *How does the hand work?* https://www.informedhealth.org/how-does-the-hand-work.2223.en.html. [Online; accessed August 20th, 2016]. 2010.

[28] Ali Rahimi and Ben Recht. "Random Features for Large-Scale Kernel Machines." In: *Advances in neural information precessing systems*. 2007, pp. 1177–1184.

[29] Vikram Ravindra and Claudio Castellini. "A comparative analysis of three non-invasive human-machine interfaces for the disabled." In: *Frontiers in Neurorobotics* 8 (2014). DOI: 10.3389/fnbot.2014.00024.

[30] Martin Sewell. *Ensemble Learning*. Tech. rep. UCL Department of Computer Science, 2011.

[31] David Sierra González and Claudio Castellini. "A realistic implementation of ultrasound imaging as a human-machine interface for upper-limb amputees." In: *Frontiers in Neurorobotics* (2013). DOI: 10.3389/fnbot.2013.00017.

[32] Erin Strait. *Prosthetics in Developing Countries*. Prosthetic Resident Paper (American Academy of Orthotists & Prosthetists). 2006.

[33] Freek Stulp and Oliver Sigaud. "Many Regression Algorithms, One Unified Model-A Review." In: *Neural Networks* 69 (2014), pp. 60–79. DOI: 10.1016/j.neunet.2015.05.005.

[34] David Travis. *Measuring satisfaction: Beyond the usability questionnaire*. http://www.userfocus.co.uk/articles/satisfaction.html. [Online; accessed June 10th, 2016]. 2008.

[35] Richard F. ff. Weir. "Design of artificial arms and hands for prosthetic applications." In: *Standard Handbook of biomedical engineering and design*. McGraw-Hill, 2004.

[36] Michael Winiger, Nam-Hun Kim, and William Craelius. "Pressure signature of forearm as predictor of grip force." In: *Journal of Rehabilitation Research and Development* 45.6 (2008).

[37] David H. Wolpert. "Stacked Generalization." In: *Neural Networks* 5 (1992), pp. 241–259.

[38] Zehn G Xiao and Carlo Menon. "Towards the development of a wearable feedback system for monitoring the activities of the upper-extremities." In: *Journal of NeuroEngineering and Rehabilitation* 11.2 (2014). DOI: 10.1186/1743-0003-11-2.

[39] Loreda Zollo, Stefano Roccella, Eugenio Guglielmelli, M. Chiara Carozza, and Paolo Dario. "Biomechatronic Design and Control of an Anthropomorphic Artificial Hand for Prosthetic and Robotic Applications." In: *IEEE/ASME Transactions on Mechatronics* 12.4 (2007), pp. 418–429.