Repeatability analysis of hand movement recognition for control of robotic prosthesis based on sEMG data

Candidata
Francesca Palermo

Relatrice
Prof.ssa Barbara Caputo

Correlatore
Manfredo Atzori

Matricola 1401074
Anno Accademico 2016/2017
Contents

1 Introduction 7
  1.1 Outline ............................................. 10

2 Problem Statement 11
  2.1 Prosthesis and Control ............................... 11
  2.2 Experimental framework ............................... 13

3 Acquisition Setup and Protocol 15
  3.1 Acquisition Setup ...................................... 15
     3.1.1 Hardware ......................................... 15
     3.1.2 Software .......................................... 18
  3.2 Acquisition Protocol .................................... 19
  3.3 Database ............................................. 20

4 Data representation and Modeling 23
  4.1 Features ............................................. 23
  4.2 Classifiers ........................................... 25
     4.2.1 Decision Tree ..................................... 26
     4.2.2 Kernel Regularized Least Squares .................. 28
  4.3 Dataset ............................................. 32
     4.3.1 Setup 1 ........................................... 32
     4.3.2 Setup 2 ........................................... 32
     4.3.3 Setup 3 ........................................... 33
     4.3.4 Setup 4 ........................................... 34

5 Data Analysis ............................................. 35

6 Results .................................................. 39
  6.1 Setup 1 ............................................. 39
  6.2 Setup 2 ............................................. 42
  6.3 Setup 3 ............................................. 45
CONTENTS

6.4 Setup 4 ........................................... 48
7 Conclusions and Future works .......................... 53
Bibliography ........................................... 57
Abstract

Control methods based on surface electromyography (sEMG) obtained promising results for robotic hand prosthetics. However, control system robustness is still inadequate and does not allow the amputees to perform a vast number of movements useful for everyday life. Moreover, few studies analyzed the repeatability of sEMG classification of hand grasps. The main goals of this paper are, first, to explore repeatability in sEMG data and, second, to release a repeatability database with the recordings of the experiments. The data are recorded from 10 intact subjects repeating 7 grasps 12 times in 5 different days. The data were recorded twice a day (morning and afternoon), thus leading to 10 acquisitions for each subject. The data are publicly available on the Ninapro web page and follow protocols comparable with other publicly available data sets. The analysis for the repeatability is based on the comparison of movement classification accuracy in different data acquisitions and for different subjects. The analysis is performed using mean absolute value, waveform length and discrete wavelet transform features and Random Forest and Kernel Regularized Least Squares classifiers. Four distinct setups are created. Each one consists of unique training and testing dataset and it gives different understanding in how much the accuracy classification of hand grasps changes with data acquired from different acquisitions. The inter-subject variability is remarkable, suggesting that specific characteristics of the subjects (e.g. size of arm, body hair, etc.) can affect repeatability as well as sEMG classification accuracy. In conclusion, the results described in this thesis can contribute to develop more robust control systems for hand prosthetics by exploiting previous data acquisitions, while the presented data can allow researchers worldwide to test repeatability effects in further analyses.
Chapter 1

Introduction

In the United States, there are approximately 1.7 million people living with limb loss. It is estimated that one out of every 200 people in the U.S. has had an amputation. The statistics of COPC (Center for Orthotic & Prosthetic Care) show that the level of amputation concerns the upper limb in 14% of cases and lower limb in 86%. Among upper limb amputees, the trans-radial ones make up the 60% of total wrist and hand amputations. Each year the majority of new amputations occur due to different reasons. The main causes are vascular disease (54%) including diabetes and peripheral arterial disease, trauma (45%), and cancer (less than 2%) (in according to statistics of Advanced Amputee Solutions, LLC).

![Pie chart showing causes of amputations]

Figure 1.1: Advanced Amputee Solutions statistics referred to lower limb and upper limb amputations
A wide variety of mechanically advanced myoelectric prosthetic hands are now available on the market. Despite the mechanical advancements made over the years, the built-in sEMG control is often limited to opening and closing grasp. Recently, several improvements on sEMG control have been made applying modern machine learning techniques. However, pattern recognition techniques are often not robust enough for a scenario in daily life [27, 26]. The position of the sensors is one of the main factors influencing the sEMG signals and, as a consequence, control robustness. Thus, the analysis of repeatability in sEMG hand grasp classification can help to improve the robustness of robotic prosthetic hands when external factors (such as electrode re-positioning) can affect the sEMG signal.

Both the market and science are complex and changing quickly. The first commercial products exploiting pattern recognition to recognize the hand grasps have been released (e.g. CoAptEngineering[1] and TouchBionics[2]). However, the most common control systems still require long training times [27].

One of the main goals of the research community working on surface electromyography controlled hand prostheses is to improve the everyday life of the amputees. Non-invasive methods have been developed, that use sEMG electrodes to record muscular activity and pattern recognition algorithms to classify hand movements. The algorithms usually show average classification accuracies of up to 80-90% [8], while results over 90% can be reached in some cases on very few movements (e.g. [6, 7]).

Healthy subjects are often chosen to acquire data in scientific experiments since performing and repeating complex movements can be strenuous for the amputees. Scientific literature showed that intact subjects can be used as a proxy measure for amputees [29]. However, parts of the muscles can be missing in amputated subjects, thus the results of amputees are often lower.

Despite the results described by the literature in sEMG hand prosthetics, there are still several obstacles to overcome. The movement accuracy is never high enough to avoid misclassification on large sets of movements. Castellini et al. [13] show that nine postures can be classified with remarkable accuracy. In the studies for the NinaPro database the number of gestures was extended to 52 movements (including grasping), showing that using machine learning it is possible to classify a large number of tasks with an accuracy of over 80% [24, 21].

As highlighted by several papers (e.g. [27, 26]), achieving a robust control is one of the main obstacles to bring sEMG pattern recognition to real life use.

---

1http://www.coaptengineering.com/
2http://www.touchbionics.com/
sEMG signals can be influenced by several external factors that can affect control robustness, such as muscle fatigue or movements of the electrodes on the skin [14]. Thus, intersubject variability, muscle fatigue and electrode displacement should always be considered when working with sEMG [13].

Recent papers have stated that the classification accuracy of the proposed classifiers is high enough to effectively perform EMG pattern recognition with accuracy of around 90% [8]. As Shin et al. [10] suggest, the accuracy of a classifier is not the only factor to fully estimate the performance of a classifier for prosthesis control applications. Other parameters also exist. This paper deals with repeatability of grasp recognition for robotic hand prosthesis. Repeatability is defined as the variation in repeated measurements made of the same subject, under identical conditions and in a short period of time. As reported by Taylor and Kuyatt [15], the following conditions must be fulfilled to successfully complete repeatability experiments: same experimental tools, same observer, same measuring instrument (used under the same conditions), same location, repetition over a short period of time and same objectives. Studies on repeatability of sEMG classification of hand grasps could improve the knowledge on the effect of external factors on robustness. Radmand et al. [11] suggest that when the arm is moved to a position different from the one in which the classifier is trained the repeatability of the data decreases. However, training in multiple positions is stressful for the amputees. Repeatability studies may help the producers of prosthesis to settle sets of gestures that can be controlled robustly, while also being helpful in activities of daily life. For instance, Xiang et al. [9] suggest that hand gesture tasks with low repeatability should be avoided in myoelectric control systems. He et al. [17] investigated the variation in EMG classification over 11 consecutive days. They observed that, when they trained the classifier on data from one day and they used the following day as testing set, the classification error decreased exponentially but it stabilized after four days for healthy subjects. These results show that, when the set of days during which the subjects perform the defined motions is enlarged, changes in EMG signal features over time become gradually smaller. Amsüss et al. [18] were able to obtain an accuracy within days per subject of 97.9% ± 0.8 through five days and five subjects. They found that the classification accuracy decreased monotonically. It dropped by 4.17% per day between training- and test days. Unlike this thesis, in both articles the exact locations of the electrodes were marked through a pen and renewed every day for accurate repositioning of the electrodes.

This project focuses on noninvasive prosthesis, more precisely on myoelectric prosthetic hands. It deals with the repeatability of data acquisitions through sEMG sensors and it has two main goals:
CHAPTER 1. INTRODUCTION

1. the release of a publicly available database to study repeatability in hand movement sEMG;

2. the analysis of repeatability in sEMG, which is based on the comparison of movement classification accuracy, through machine learning techniques, in several data acquisitions and in different subjects.

1.1 Outline

The content of this thesis is organized as follows:

In chapter 2 some landmarks are given. These are useful to have an introduction of the problems undertaken in the following and a general vision of the entire project. It starts with an overview of different kind of prosthesis on the market and the general problems and challenges related to their usage and control.

Chapter 3 describes the MeganePro project which gathered data from intact and amputated subjects. First are explained the acquisition setup and the acquisition protocol exploited during the work to acquire the emg data from the 10 subjects of the project. In the last part of the chapter, the used database (available online http://ninapro.hevs.ch/) is illustrated in detail.

Chapter 4 gives a theoretical point of view of all the machine learning features implemented in the thesis. The first part focuses on the features and classifiers in a general way. Later on, the features and classifiers, exploited in the project, are deeply explained. Since more than one setup was created following different parameters for building training and testing set, the setups are described.

In chapter 5 the experiment is presented. The work is done on the database illustrated in chapter 3 exploiting the features and classifiers described in chapter 4. Four different setups are analyzed.

In chapter 6 the results of the experiment for the four setups are individually presented and analyzed.

In chapter 7 the final review on the work is provided along with ideas for possible future works and researches.
Chapter 2

Problem Statement

This chapter gives a general vision of the entire works: it aims to focus the problems linked to prosthesis and our approach to tackle them.

The first part of the chapter focuses on different non-invasive prosthesis currently available on the market. After an introduction, the problems, that an amputee tackles when he/she learns how to use a new prosthesis, are listed.
The second part is a summary of the work developed.

2.1 Prosthesis and Control

As explained in the introduction, this thesis focuses on non-invasive prosthesis of upper limbs. The most common non-invasive prosthesis available on the market are:

- Cosmetics. They are only aesthetic hands, they are comfortable but, as the name suggests, they don’t offer any type of grasping or movement.
CHAPTER 2. PROBLEM STATEMENT

• Body powered. They are mechanical prosthesis which work through cables to link the movement of the body to the prosthesis and to control it. When an amputee moves the body in a certain way he/she drives the cables, consequently the prosthetic hand is opened, closed, or bended. Only few movements are available.

• Myoelectrics. They use several electrodes placed in contact with the stump. These electrodes are called electromyography surface and they detect electrical activity produced by muscles when a movement is performed. The recorded electromyographic signal (EMG) varies in: $\sim 10\mu V \div 10mV$. The electrodes record the muscular activity, thus the signal is analyzed in order to decide the intentional movement and, theoretically, infinite positions can be performed.

Myoelectrics prosthesis could potentially improve the quality of life of an amputee, but the control system is still too difficult. The open challenge in the scientific world is the creation of an accurate control system to make them easy to use by the patient and, especially, to reduce the training time, i.e. the time required from the subject to learn how to use the prosthesis. This is still a very long and tiring process, often with great mismatch between desired and performed movements. Moreover this is generally perceived as very tiring and sometimes painful by the users. These reasons make the use
of myoelectric prosthesis still limited in practice. Often the amputees replace the myoelectric prosthesis with cosmetic ones [21].

2.2 Experimental framework

In this section are showed all the steps used to improve the control of non-invasive myoelectric prosthesis. For this project only intact subjects were considered. The electromyographic data are collected using surface electrodes (sEMG) placed on the right forearm. These electrodes detect the electromyographic signal generated by muscular contractions, where each signal corresponds to a different movement. The goal of the robotic prosthesis control is to establish the best output movement for each given input signal. To this purpose, first, the input data are processed to remove a noise component. Second, is performed the features extraction, which consist in extracting valuable informations from the data, in order to determine the discriminant characteristics of signal that must be classified. After being appropriately processed, the data are used as input of different classifiers. These will solve the classification problem, which consists in finding the right output movement for a given input signal. The choice for the output class is predefined, in other words is not possible to select other classes than the initial ones. All the described steps are reported schematically in the figure below.

![Figure 2.1: Steps in a complete experiment.](image-url)
Chapter 3

Acquisition Setup and Protocol

In this chapter are presented the acquisition setup and protocol of this work with particular attention to the used devices to collect data from subjects. The last part of this section focuses on the organization and structure of the database. The database used in this work is part of the MeganePro project which is the next phase of the NinaPro project. It aims to aid research on advanced hand myoelectric prosthetics with public databases. The databases are obtained by recording multi-modal data, including e.g. surface electromyography (sEMG) signals and hand kinematics while the subjects perform a predefined set of hand movements. The data exploited in this work have been acquired at the Technopole of Sierre between November and December 2016.

3.1 Acquisition Setup

The acquisition setup is based on the setup used for previous Ninapro datasets \[21\]. It can be split into hardware and software.

3.1.1 Hardware

The hardware acquisition setup consists of:

- DELL Latitude E5520: the laptop used to perform the data acquisitions and to record the data;
- Tobii Pro Glasses II: a wearable eye tracking system used to record the eye movements and field of view;
- 14 Delsys Trigno double differential sEMG Wireless electrodes: used to record the muscular activity of the forearm.
CHAPTER 3. ACQUISITION SETUP AND PROTOCOL

The Tobii Pro Glasses II (Tobii AB, http://www.tobii.com/), Figure 3.1, are composed of the head unit (having an eyeglass design) and a recording unit (that is used to record and store the videos on an SD card). Using four infrared cameras embedded in the frame of the glasses, the device can estimate where the subject is looking within his field of view, which is recorded via a full HD camera. During data acquisition, the device is connected to the laptop through a wireless network.

![Tobii Pro Glasses II](image)

Figure 3.1: Tobii Pro Glasses II.

The muscular activity is measured with 14 Delsys Trigno sEMG Wireless electrodes (Delsys Inc. www.delsys.com). The electrodes are connected through a wireless protocol to their base station (Figure 3.2). The base station is connected to the laptop via a USB cable. The sEMG signals are sampled at 2 KHz while the 3-axes accelerometer in the device is sampled at 148.148 Hz.

![Trigno Wireless EMG](image)

Figure 3.2: Trigno Wireless EMG.

The electrodes are equally spaced in two rows around the forearm. Figure 3.3 shows how electrodes are placed. The first row is composed of eight electrodes that are arranged in correspondence to the radio-humeral joint as described in [21]. The second row is composed of six electrodes that are placed just below the first row, in correspondence to the empty spaces of the
first row and positioned in order to avoid positioning over the ulna. Finally, an elastic latex-free band is placed around the electrodes to avoid falls and reduce their movement.

Figure 3.3: Example of the final position of the electrodes. First row made up of eight electrodes arranged in correspondence to the radio-humeral joint. Second row composed of six electrodes placed just below the first row.

Figure 3.4 shows the complete setup for the acquisition, including the computer, the Trigno Base, the Tobii Pro Glasses II and some of the objects used for the experiments.
3.1.2 Software

The software acquisition is made up of two parts that work together:

1. the software used to simultaneously record the data from all the sensors;
2. the software that guides the subjects during the data acquisitions.

The first part of the software is a custom-made multithreaded application based on a producer-consumer pattern software written in C++ by Stefano Pizzolato [25] (Figure 3.5). When the data are recorded, a timestamp is assigned to them, in this way it is possible to synchronize the data acquired from the devices.
3.2. ACQUISITION PROTOCOL

The second part of the software was developed to guide the user during the data acquisition via visual and audio commands. The language of the audio commands can be chosen among four options (Italian, English, French and German) at the beginning of the acquisition.

![Figure 3.6: Acquisition Software : GUI](image)

3.2 Acquisition Protocol

The acquisition protocol is an improvement of the acquisition protocol used to record the previous Ninapro datasets [21]. The researcher explains the experiment to the subject, asks him to respond to a few questions (including age, gender, height, weight and laterality) and measures the length (wrist to elbow) and circumference of the subject’s forearm. The subject sits in front of a table with the forearm leaning on it. The experiment consists of 12 repetitions of 7 grasps performed on a set of 14 objects (Table 3.7), chosen as a subset of a more complete collection of 52 hands movements [24]. The set of hand grasps was chosen from the robotics and rehabilitation literature [24] [19] [20] with the goal of covering several hand movements exploited in activities of daily living (ADL). The grasp to be performed is shown to the subject with two videos (in first and third person perspective). Afterwards, a set of audio commands explains the subject the task to be performed (i.e. grasping the object, releasing the object and returning to the rest position).

While performing the experiment, a fixed image representing the grasp is shown on the screen of the laptop, to avoid distraction from the user. In this phase the signals provided by the electrodes are saved in a stimulus file with a time-stamp and the grasp label. The number of repetitions is equally...
distributed among two objects (6 for each object), as shown in Figure 3.7. Each repetition lasts for 4 seconds and is followed by 4 seconds of rest. The data recorded from each subject are uploaded to Ninaweb and publicly available as the 6th Ninapro dataset.

![Figure 3.7: Set of Hand Movements and Objects in the experiment](image)

### 3.3 Database

The database exploited in this work was acquired with the setup and procedure explained in previous section. The database contains data from 10 intact subjects: 3 females and 7 males, all right-handed with 27 ± 6 years. The total number of movements is 7 with 12 repetitions for each hand grasp. All the details about the participants are reported in Table 3.1.

![Table 3.1: Description of database 6](table)

The raw data of the acquisitions are not online but available on request. For each subject, it is possible to find online [http://ninapro.hevs.ch/](http://ninapro.hevs.ch/) a Matlab file with the following data:

20
3.3. DATABASE

- **subj**: the number of the subject;
- **daytesting**: the number of the day in which the data were acquired;
- **time**: the time (1=morning, 2=afternoon) in which the data were acquired;
- **emg** (16 columns): sEMG signal of the 14 electrodes (2 columns are equal to 0 since electrodes 9 and 10 are not used). Columns from 1 to 8 include the signals from the 8 electrodes positioned in the first row. Columns from 11 to 16 includes the signals from the 6 electrodes of the second row;
- **object**: gives the id of the object taken during the grasp phase;
- **reobject**: same as object but after the relabeling phase;
- **stimulus**: the original label of the movements repeated by the subject;
- **restimulus**: label of movements after the relabeling phase;
- **repetition**: repetition of the stimulus;
- **rerepetition**: repetition of restimulus;
- **acc** (48 columns): three-axes acceleration values of the 16 electrodes (to consider 6 columns equal to zero, which corresponds to electrodes 9 and 10).
Chapter 4

Data representation and Modeling

In this chapter are presented all the machine learning features implemented in the thesis from a theoretical point of view. The first part focuses on the features and classifiers in a general way. Later on, the features and classifiers exploited in the work are deeply explained. In the end, the four analyzed setups are described.

4.1 Features

In this section the set of available features will be presented and described.

Raw sEMG signal are mapped into smaller-dimension feature vectors since features define the information content of the signal more efficiently than complex raw signal. In consequence of the reduced size of the feature vectors, classification of the grasp recognition is performed faster than on raw signals. This improves the real-time properties of the system. EMG features can be grouped in four main categories depending on the domain where they are computed [2]:

1. **Time Domain** (TD) features. These are the most common and easiest features in sEMG signal classification. Their advantage is that they are fast to calculate since they require no mathematical transformation. On the other hand, they are sensitive to noise. They have been divided in four categories [1]:

   (a) energy and complexity information methods;
   (b) frequency information methods;
(c) prediction model methods;
(d) time-dependence methods.

The most common used TD features are Mean Absolute Value (MAV), Waveform Length (WL), Zero Crossing (ZC) and Signal Slope Changes (SSC). MAV estimates an average of absolute value of the EMG signal amplitude while WL the cumulative length of the waveform over the time segment.

2. **Frequency Domain (FD)** features. These features can be exploited to estimate muscle fatigue and force production. FD features are computed from power spectral density (PSD), which can be estimated through Periodogram or parametric methods. FD features are inferior in respect to TD features. They are more computationally more expensive than TD features, moreover their classification accuracy is lower. Nevertheless, combining FD and TD features is a good strategy to generate more robust classification.

3. **Time Frequency Domain (TFD)** features. These set of features includes Short Time Fourier Transform (STFT), Continuous Wavelet Transform (CWT), Discrete Wavelet Transform (DWT), Wavelet Packet Transform (WPT) and Stationary Wavelet Transform (SWT). Even if TFD features are more complex than TD features, they can be implemented with fast algorithms that have shown to meet in real-time requirements in sEMG classification when appropriate dimensional reduction and segmentation techniques are used. Thus, these features may be able to improve the robustness of the system in regard of TD and FD features. DWT, being the most computationally efficient, has become the most exploited TFD feature in sEMG interfaces.

4. **Spatial Domain (SD)** features. The features improve the differentiation between postures and force levels providing information about spatial distribution of the motor unit action potentials and of the load-sharing between muscles.

The features exploited in this work are Mean Absolute Value, Waveform Length and Marginal Discrete Wavelet Transform.
4.2. Classifiers

After the features extraction, the set of data obtained must be classified. In this section the exploited classifiers for grasp recognition will be illustrated in detail.

Classification is the task of learning a target function $f$ that maps each attribute set $x$ to one of the predefined class label $y$. A classification technique (or classifier) is a systematic approach to building classification models from an input data set. Examples include decision tree classifiers, rule-based classifiers, neural networks, support vector machines. Each technique employs a learning algorithm to identify a model that best fits the relationship between the attribute set and class label of the input data. The model generated by a learning algorithm should both fit the input data well and correctly predict the class labels of records it has never seen before. Therefore, a key objective of the learning algorithm is to build models with good generalization capability. First, a training set consisting of records whose class labels are known must be provided. The training set is used to build a classification model, which is subsequently applied to the test set, which consists of records with unknown class labels. Evaluation of the performance of a classification model is based on the counts of test records correctly and incorrectly predicted by the model. These counts are tabulated in a table known as a confusion matrix. Each entry $f_{ij}$ in this table denotes the number of records from class $i$ predicted to be of class $j$. Although a confusion matrix provides the information needed to determine how well a classification model performs, summarizing this information with

<table>
<thead>
<tr>
<th>Feature</th>
<th>Definition (per channel)</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Absolute Value (MAV)</td>
<td>$\hat{x} = \frac{1}{T} \sum_{t=1}^{T}</td>
<td>x_t</td>
</tr>
<tr>
<td>Waveform Length (WL)</td>
<td>$\hat{x} = \sum_{t=1}^{T}</td>
<td>x_t - x_{t+1}</td>
</tr>
<tr>
<td>marginal Discrete Wavelet Transform (mDWT)</td>
<td>$\hat{x}<em>t = \sum</em>{\tau=0}^{T/2-1} \left</td>
<td>\sum_{t=1}^{T} x_t \psi_{l,\tau}(t) \right</td>
</tr>
</tbody>
</table>

Table 4.1: Algorithm for features extraction. C is the number of channels (10 or 12 depending on database). For mDWT, one uses $\psi_{l,t}$ to denote the mother wavelet with translation $l$. Source: [3].
a single number would make it more convenient to compare the performance of different models. This can be done using a performance metric such as accuracy, which is defined as follows:

\[
\text{Accuracy} = \frac{\# \text{ correct predictions}}{\# \text{ total predictions}}
\]

### 4.2.1 Decision Tree

Decision tree learning is a method for approximating discrete-valued target function, in which the learned function is represented by a decision tree. Decision trees classify instances by sorting them down the tree from the root to a leaf node. Each node of the tree characterizes a test of some attribute of the instance and each branch descending from that node corresponds to one of the possible value for that attribute. An instance is classified by starting at the root node of the tree, testing the attribute specified by this node, then moving down the tree branch corresponding to the value of the attribute in the given example. This algorithm is repeated for the subtree rooted at the new node. The algorithm is summarized in Figure 4.1: example of Decision Tree.

Decision trees represent disjunction of conjunctions of constraints on the attribute value of instances. Each path from tree root to a leaf corresponds to a conjunction of attribute tests and the tree itself to a disjunction of these conjunctions.

![Figure 4.1: example of Decision Tree](image-url)
4.2. CLASSIFIERS

Start at the root node;

\textbf{foreach} \( X \) \textbf{do}

find the set \( S \) that minimizes the sum of the node impurities in the two child nodes;

choose the split \( X^* \in S^* \) that gives the minimum overall \( X \) and \( S \);

\textbf{if} a stopping criterion is reached \textbf{then}

\hspace{1cm} exit;

\textbf{else}

\hspace{1cm} apply step 2 to each child node in turn;

\textbf{end}

\textbf{Algorithm 1:} Pseudocode for tree construction by exhaustive search

Random Forest

A Random Forest classifier uses a number of classification trees in order to improve the classification rate. As Breiman [4] suggests, a random forest is a classifier consisting of a collection of tree structured classifiers \( \{ h(x, \Theta_k), k = 1, \ldots \} \) where the \( \{ \Theta_k \} \) are independent identically distributed random vectors and each tree casts a unit vote for the most popular class at input \( x \). In other words, to classify a new object from an input vector, the latter is positioned down each of the trees in the forest. Each tree computes a classification and the forest chooses the most common one or the one with the higher classification accuracy. Each tree is grown in the following way:

1. If the number of cases in the training set is \( N \), \( N \) cases are sampled at random from the original data. This sample will be the training set for growing the tree.

2. If there are \( M \) input variables, a number \( m \ll M \) is specified in such a way that \( m \) variables are selected at random out of the \( M \) and the best split on these \( m \) is used to split the node, at each node. The value of \( m \) is kept constant during the growing.

3. Each tree is grown to the largest extent possible, without any pruning.

Given a set of classifiers \( h_1(x), h_2(x), \ldots, h_k(x) \) and with the training set draws at random from the distribution of the random vector \( Y, X \), the margin function for the random forest is defined as

\[
    mr(X, Y) = P_{\Theta}(h(X, \Theta) = Y) - \max_{j \neq Y} P_{\Theta}(h(X, \Theta) = j) \tag{4.1}
\]
The margin represents the extent to which the average number of vote at \( \mathbf{X}, Y \) for the right class surpasses the average vote for all the other class. The greater the margin, the larger the confidence in the classification is. The strength of the classifiers \( h(x, \Theta) \) is

\[
s = E_{\mathbf{X},Y} mr(\mathbf{X}, Y)
\]

The generalization error is given by

\[
P_{E}^{*} = P_{\mathbf{X},Y}(mg(\mathbf{X}, Y) < 0)
\]

In random forests \( h_k(\mathbf{X}) = h(\mathbf{X}, \Theta_k) \). If we are considering a vast number of trees, it follows from the Strong Law of Large Numbers that as the number of trees increases, almost surely all sequences \( \Theta_1 \ldots P_{E}^{*} \) will converge to

\[
P_{\mathbf{X},Y}(P_{\Theta}(h(\mathbf{X}, \Theta) = Y) - \max_{j \neq Y} P_{\Theta}(h(\mathbf{X}, \Theta) = j) < 0)
\]

From this result, is possible to infer that random forest don’t overfit as more trees are added, in contrast they produce a limiting value of the generalization error. The forest error rate depends on two things:

1. The correlation between any two trees of the forest. Rising the correlation, increase the error;
2. The strength of each individual tree in the forest. Increasing the strength of the individual trees, decreases the error.

### 4.2.2 Kernel Regularized Least Squares

The Kernel Regularized Least Squares (KRLS) classifier originates from the Regularized Least Squares (RLS). The "K" has been added to emphasize that, on the contrary of RLS, it exploits kernels. As explained in [5], this method belongs to a class of models for which marginal effects are well-behaved and easily obtainable due to the existence of a continuously differentiable solution surface, estimated in closed form. Any kernel method solution is made up of two parts:

- a module that performs the mapping into a feature space
- a learning algorithm designed to discover linear patterns in that space

In KRLS the initial mapping component is defined implicitly by a *kernel function* which will depend on the specific data type and domain knowledge concerning the patterns that are to be expected in the specific data source.
The pattern analysis algorithm component is robust, comes with a statistical analysis of its stability and is efficient, since it requires an amount of computational resources which is polynomial in the size and number of data items.

**Primal linear regression**  Considering the problem of finding a homogeneous real-valued linear function

\[ g(x) = \langle w, x \rangle = w^T x = \sum_{i=1}^{n} \omega_i x_i \]

that best interpolates a given training set \( S = \{(x_1, y_1), \ldots, (x_l, y_l)\} \) of points \( x_i \) from \( X \subseteq \mathbb{R} \). This task is called linear interpolation which geometrically corresponds to fitting a hyperplane through the given \( n \)-dimensional points. Figure 4.2.2 shows an example for \( n = 1 \). There are four different ways to solve this problem:

1. In the case of having exact \( l = n \) linearly independent points, it is possible to obtain the parameters \( w \) by solving the system of linear equations

\[ Xw = y \]

2. If there are less points than dimensions, there exist different \( w \) that describe the data exactly. In this case, the vector \( w \) with minimum norm will be chosen.

3. If there are more points than dimensions and there is noise in the generation process, the pattern with the smallest error will be chosen.

4. If the datasets are both noisy and small, the two previous strategies will be used. A vector \( w \) that has both small norm and small error will be preferred.

Figure 4.2: A one-dimensional linear regression problem.
The distance $\xi$ is the output of the putative pattern function

$$f((x, y)) = |y - g(x)| = |\xi|$$

The goal is to find a function for which all of these training errors are small. The chosen measure is the sum of the squares of these error

$$\mathcal{L}(g, S) = \mathcal{L}(w, S) = \sum_{i=1}^{l} (y_i - g(x_i))^2 = \sum_{i=1}^{l} \xi_i^2 = \sum_{i=1}^{l} \mathcal{L}((x_i, y_i), g),$$

where $\mathcal{L}((x_i, y_i), g) = \xi_i^2$ denotes the squared error or loss of $g$ on example $(x_i, y_i)$ and $\mathcal{L}(f, S)$ denotes the collective loss of a function $f$ on the training set $S$. The learning problem is now to choose a vector $w \in \mathcal{W}$ which minimizes the collective loss. This problem is called least squares approximation and was firstly introduced by Gauss. The vector of output discrepancies may be written as

$$\xi = y - Xw.$$

The loss function can be written as

$$\mathcal{L}(w, S) = \|\xi\|^2 = (y - Xw)'(y - Xw) \quad (4.5)$$

The optimal $w$ is obtained by taking the derivatives of the loss with respect to the parameters $w$ and setting them equal to the zero vector

$$\frac{\partial \mathcal{L}(w, S)}{\partial w} = -2X'y + 2X'Xw = 0$$

obtaining the normal equations

$$X'Xw = X'y \quad (4.6)$$

If the inverse of $X'X$ exists, the solution of the least squares problem can be expressed as

$$w = (X'X)^{-1}X'y$$

The predicted output on a new data point can be computed through the prediction function

$$g(x) = \langle w, x \rangle$$

In some situations, there is no possibility to fit the data exactly. These problems are known as ill-conditioned, because the information available are not sufficient to precisely find the solution. In these circumstances the adopted
approach is to restrict the set of function through some methods. Such a restriction or bias is referred to as regularization. In the case of least squares regression, from this it is obtained the optimization criterion of ridge regression. The ridge regression consists on resolving the optimization

$$
\min_w \mathcal{L}_\lambda(w, S) = \min_w \lambda \|w\|^2 + \sum_{i=1}^l (y_i - g(x_i))^2
$$

(4.7)

**Ridge regression: primal and dual** The following equation are computed taking the derivative of the cost function with respect to the parameters

$$
X'Xw + \lambda w = (X'X + \lambda I_n)w = X'y
$$

(4.8)

in which $I_n$ is the $n \times n$ identity matrix. If $\lambda > 0$, the matrix $(X'X + \lambda I_n)$ is always invertible. Thus the solution is given by

$$
w = (X'X + \lambda I_n)^{-1}X'y
$$

(4.9)

The resulting prediction function is the following

$$
g(x) = \langle w, x \rangle = y'X(X'X + \lambda I_n)^{-1}x
$$

(4.10)

The equation (4.8) can be rewrited as

$$
w = \lambda^{-1}X'(y - Xw) = X'\alpha
$$

(4.11)

showing that $w$ can be reproduced as a linear combination of the training points, $w = \sum_{i=1}^l \alpha_i x_i$ with $\alpha = \lambda^{-1}(y - Xw)$. Hence it is obtained

$$
\alpha = \lambda^{-1}(y - Xw)
\Rightarrow \lambda\alpha = (y - XX'\alpha)
\Rightarrow (XX' + \lambda I_l)\alpha = y
\Rightarrow \alpha = (G + \lambda I_l)^{-1}y
$$

(4.12)

in which $G = XX'$. The resulting prediction function is given by

$$
g(x) = \langle w, x \rangle = \sum_{i=1}^l \alpha_i x_i, x \rangle = \sum_{i=1}^l \alpha_i \langle x_i, x \rangle = y'(G + \lambda I_l)^{-1}k
$$

(4.13)

where $k_i = \langle x_i, x \rangle$. Thus there are two methods to solve the ridge regression optimization of equation 4.7.
CHAPTER 4. DATA REPRESENTATION AND MODELING

1. The first equation \([4.9]\) is known as the \textit{primal solution} and calculates the weight vector explicitly.

2. The second \([4.12]\) is a linear combination of the training examples and is known as \textit{dual solution}.

4.3 Dataset

The exploited dataset is part of the Ninapro project and is publicly available as the 6\textsuperscript{th} Ninapro dataset on \url{http://ninapro.hevs.ch/}. To better comprehend the repeatability of the experiment, more than one setup were created following different parameters for building the training and testing set.

4.3.1 Setup 1

The first setup is created to understand the possibility to use only one session for the training data of the robotic prosthesis. The training set is produced using the mornings of the days, while the testing of the same day is created exploiting the afternoons.

<table>
<thead>
<tr>
<th>DATASET</th>
<th>TRAINING</th>
<th>TESTING</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Day 1 Morning</td>
<td>Day 1 Morning / Day 1 Afternoon</td>
</tr>
<tr>
<td>2</td>
<td>Day 2 Morning</td>
<td>Day 2 Morning / Day 2 Afternoon</td>
</tr>
<tr>
<td>3</td>
<td>Day 3 Morning</td>
<td>Day 3 Morning / Day 3 Afternoon</td>
</tr>
<tr>
<td>4</td>
<td>Day 4 Morning</td>
<td>Day 4 Morning / Day 4 Afternoon</td>
</tr>
<tr>
<td>5</td>
<td>Day 5 Morning</td>
<td>Day 5 Morning / Day 5 Afternoon</td>
</tr>
</tbody>
</table>

4.3.2 Setup 2

The second setup is similar to the first one but instead of training on the mornings and testing on afternoons, the training data is based on the morning (or afternoon) of the first day and the testing is implemented on the mornings (or afternoons) of the remaining days.
Table 4.3: Second setup. Training morning (afternoon) of the first day. Testing mornings (afternoons) of the remaining days.

<table>
<thead>
<tr>
<th>DATASET</th>
<th>TRAINING</th>
<th>TESTING</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Day 1 Morning</td>
<td>Day 1 Morning&lt;br&gt;Day 2 Morning&lt;br&gt;Day 3 Morning&lt;br&gt;Day 4 Morning&lt;br&gt;Day 5 Morning</td>
</tr>
<tr>
<td>2</td>
<td>Day 1 Afternoon</td>
<td>Day 1 Afternoon&lt;br&gt;Day 2 Afternoon&lt;br&gt;Day 3 Afternoon&lt;br&gt;Day 4 Afternoon&lt;br&gt;Day 5 Afternoon</td>
</tr>
</tbody>
</table>

**4.3.3 Setup 3**

The third setup is made to comprehend how much the classification accuracies improve using more than one session for the training of the robotic hand prosthesis. Thus, the training set is created using an increasing group of dataset. First it is considered only the morning (or afternoon) of Day 1, then the mornings (or afternoons) of Day 1 + Day 2, and so on. The testing is fixed and is made on the morning (or afternoon) of the last day.

Table 4.4: Third Setup. Training on an increasing set of mornings (afternoons). Testing morning (afternoon) of last day.

<table>
<thead>
<tr>
<th>DATASET</th>
<th>TRAINING</th>
<th>TESTING</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Day 1&lt;br&gt;Day 1 + 2&lt;br&gt;Day 1 + 2 + 3&lt;br&gt;Day 1 + 2 + 3 + 4 Morning</td>
<td>Day 5 Morning</td>
</tr>
<tr>
<td>2</td>
<td>Day 1&lt;br&gt;Day 1 + 2&lt;br&gt;Day 1 + 2 + 3&lt;br&gt;Day 1 + 2 + 3 + 4 Afternoon</td>
<td>Day 5 Afternoon</td>
</tr>
</tbody>
</table>
4.3.4 Setup 4

The fourth and last setup is the opposite of the third one. It is implemented to understand how much the classification accuracy decrease through the use of the robotic limb. The training set is fixed and is created using the morning (or afternoon) of the first day. Whereas the training set is generated exploiting an increasing set of mornings (or afternoons) of the remaining days.

Table 4.5: Fourth Setup. Training on morning (afternoon) of the first day. Testing on an increasing set of mornings (afternoons) of the remaining days.

<table>
<thead>
<tr>
<th>DATASET</th>
<th>TRAINING</th>
<th>TESTING</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Day 1 Morning</td>
<td>Day 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Day 2 + 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Day 2 + 3 + 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Day 2 + 3 + 4 + 5 Morning</td>
</tr>
<tr>
<td>2</td>
<td>Day 1 Afternoon</td>
<td>Day 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Day 2 + 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Day 2 + 3 + 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Day 2 + 3 + 4 + 5 Afternoon</td>
</tr>
</tbody>
</table>
Chapter 5

Data Analysis

The analysis of repeatability is based on the comparison of movement classification accuracy in several data acquisitions and for several subjects. The data are recorded from 10 subjects (3 females, 7 males, average age 27 ± 6 years). Before the raw data could be used for classification, several steps are necessary. The movement classification follows the procedure suggested by Englehart et al. [28]. It includes preprocessing, relabeling, feature extraction, and classification.

**Preprocessing** First, the data are preprocessed. This step is composed of synchronization and filtering. The signals representing the movement stimuli, the accelerometers and the sEMG are synchronized to the highest frequency (2 kHz) by interpolating the timestamps with piecewise linear models. Then, the EMG signals are filtered from interferences with a Hampel filter at 50 Hz.

![Figure 5.1: Movements division process for a little part of total raw signal. Yellow squares contain rest postures, red squares contain a movements.](image)

35
CHAPTER 5. DATA ANALYSIS

Relabeling  Data relabeling is required because the subjects do not always react promptly to the voice commands. Often, the real duration of the movement is not the same as the video. The relabeling is performed following the procedure exploited in previous work [24].

Figure 5.2: Example of the effect of relabeling. The original window is highlighted in blue while the detected window (relabeling window) of movement is showed in green.

Figure 5.3: Each field of the .mat file is shifted in relation of the detected windows.
**Feature Extraction**  Feature extraction is performed on 200 ms time windows, with an increment of 10 ms. The features chosen are the Mean Absolute Value (MAV), Waveform Length (WL) and marginal Discrete Wavelet Transform (mDWT) which previously obtained good results on sEMG \[22\] \[23\]. The feature extraction algorithms are based on the work of Chan et al. \[16\].

![Figure 5.4: Example of feature MAV from channel 1](image1)

![Figure 5.5: Example of feature WL from channel 1](image2)
CHAPTER 5. DATA ANALYSIS

Classification  Random Forests with 100 trees and Kernel Regularized Least Squares (KRLS) are adopted as classifiers. For KRLS, the exploited feature is the mDWT with Exp Chi-Square kernel. Moreover, the script optimizes all hyperparameters including weights of the kernel combination. Regarding Random Forest, the used features are MAV and WL. The training and the testing sets differs in relation to the setup used. Thus, it will be better explained in the section [6].

Analysis tests  Afterwards, only for the first two setups, the Friedman test is used to test for differences between groups. It was performed to compare the classification results obtained in the training and test sets coming from the same acquisitions with training and test sets coming from different acquisitions. Then, the Kruskal-Wallis test, which is a non-parametric test for examining if the samples originate from the same distribution, was performed to compare the accuracies obtained on several subjects.
Chapter 6

Results

In this chapter are presented the results of the analysis of data acquired through the setup explained in Chapter 3. The analysis follows the guidelines of Chapter 5. The main goal is the analysis of repeatability in sEMG, which is based on the comparison of movement classification accuracy in several data acquisitions and in several subjects. In other words, we are interested in understanding how much the accuracy classification of hand grasps changes using data acquired from different acquisitions. In each section is presented an unique setup which exploits different training and testing dataset.

NOTE. During the acquisition of subject 2 day 2 afternoon, the Trigno base disconnected from the laptop, thus reducing the accuracy for the session and increasing the standard deviation of the overall accuracy for the subject.

6.1 Setup 1

The first setup is explained in Section 4.3.1. It is made on training on the morning of the days and testing on the afternoon of the same day (Table 4.2). Figures 6.1 and Figure 6.1 show the accuracy of the morning and afternoon acquisitions with classifier Random Forest and feature MAV and WL, respectively. Whereas 6.1 illustrates the accuracy of the morning and afternoon acquisitions with classifier KRLS and feature mDWT. Better accuracies are obtained with training and test sets coming from the same acquisitions (morning). The classification accuracies decrease by an average of 27.03% with training and test sets of different acquisitions (training from the morning acquisitions, testing on the afternoon data). The comparison of the classification accuracies obtained with training and test sets of the same ac-
acquisitions with those obtained from different acquisitions show a significant difference (Friedman test, \( p < 0.001 \)). This is likely due to the positioning of the electrodes, which changes between acquisitions. Nevertheless, the accuracies with training and testing from different acquisitions is higher than the chance level for the considered number of movements (12.5%), thus suggesting that different acquisitions can be useful to train the control systems of the prosthesis. Although, only one session is not enough to obtain results sufficiently robust to control the prosthesis.

![Figure 6.1: Classification accuracies for setup 1 exploiting Random Forest classifier and MAV feature.](image1)

![Figure 6.2: Classification accuracies for setup 1 exploiting Random Forest classifier and WL feature.](image2)
6.1. SETUP 1

(a) Mornings  
(b) Afternoons

Figure 6.3: Classification accuracies for setup 1 exploiting KRLS and mDWT feature.

A median test on the overall accuracies per day obtained with the two features MAV and WL shows that the results retrieved with the WL feature are slightly higher compared to MAV. On the other hand, the results obtained with classifier KRLS and feature mDWT are much more higher than those reached by random forest classifier with MAV and WL feature. The best accuracy is reached on the same dataset (Subject 5, Day 1, Morning) with all the three features: 93, 90% for mDWT, 81, 94% for WL and 81, 80% for MAV. Table 6.1 and 6.2 show the overall accuracies and standard deviations for each day and for each subject. Only the feature mDWT is exploited since it’s the one, among the other features, which reaches the best results. The Kruskal-Wallis test was performed on the morning acquisitions and the value obtained ($p < 0.001$) indicates that the null hypothesis of having all data samples from the same subject is rejected. Thus, there are significant differences between subjects. The variability within each subject is in general low, suggesting that external factors (e.g. size of the arm, muscle fatigue, ecc.) may contribute to determine the results.

Table 6.1: Overall Accuracy Par Day. mDWT Feature. Setup 1

<table>
<thead>
<tr>
<th>Day</th>
<th>Average</th>
<th>S.D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>58,02%</td>
<td>23,35</td>
</tr>
<tr>
<td>2</td>
<td>59,72%</td>
<td>28,47</td>
</tr>
<tr>
<td>3</td>
<td>64,73%</td>
<td>25,46</td>
</tr>
<tr>
<td>4</td>
<td>65,11%</td>
<td>24,70</td>
</tr>
<tr>
<td>5</td>
<td>64,02%</td>
<td>22,86</td>
</tr>
</tbody>
</table>
Table 6.2: Overall Accuracy Per Subject. mDWT Feature. Setup 1

<table>
<thead>
<tr>
<th>Subject</th>
<th>Average</th>
<th>S. D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>65.99%</td>
<td>5.58</td>
</tr>
<tr>
<td>2</td>
<td>67.04%</td>
<td>16.49</td>
</tr>
<tr>
<td>3</td>
<td>62.67%</td>
<td>8.18</td>
</tr>
<tr>
<td>4</td>
<td>61.98%</td>
<td>5.18</td>
</tr>
<tr>
<td>5</td>
<td>73.47%</td>
<td>7.79</td>
</tr>
<tr>
<td>6</td>
<td>66.82%</td>
<td>5.58</td>
</tr>
<tr>
<td>7</td>
<td>60.10%</td>
<td>4.75</td>
</tr>
<tr>
<td>8</td>
<td>61.35%</td>
<td>9.04</td>
</tr>
<tr>
<td>9</td>
<td>46.36%</td>
<td>10.43</td>
</tr>
<tr>
<td>10</td>
<td>57.42%</td>
<td>2.56</td>
</tr>
</tbody>
</table>

6.2 Setup 2

The second setup is explained in Section 4.3.2. It is made on training on the first day and testing on the rest of the days one by one, separating the mornings from the afternoons (Table 4.3).

What we would expect from this setup is to obtain best results with testing acquisition that immediately follows the training sessions (e.g. Day 2 should have better classification accuracy than Day 5). Figures 6.2 and Figure 6.2 show the accuracy of the morning and afternoon acquisitions with classifier Random Forest and feature MAV and WL, respectively. Whereas 6.2 illustrates the accuracy of the morning and afternoon acquisitions with classifier KRLS and feature mDWT.
6.2. SETUP 2

Figure 6.4: Classification accuracies for setup 2 exploiting Random Forest classifier and MAV feature.

Figure 6.5: Classification accuracies for setup 2 exploiting Random Forest classifier and WL feature.
CHAPTER 6. RESULTS

(a) Mornings
(b) Afternoons

Figure 6.6: Classification accuracies for setup 2 exploiting KRLS and mDWT feature.

Table 6.3 and 6.4 show the overall accuracies and standard deviations for each day and for each subject. Only the feature mDWT is represented, for brevity. With Table 6.3 is possible to prove our previous thesis, namely that better results are obtained with testing acquisitions that immediately follow the training sessions. As a matter of fact, the average of the overall accuracy per day gradually decreases as we distance ourselves from the training dataset. The best classification accuracy is reached on the first day (81.46%) and it decreases of a maximum of 45.04% on the last day of acquisition (36.42%).

Table 6.3: Overall Accuracy Par Day, mDWT Feature, Setup 2

<table>
<thead>
<tr>
<th>Day</th>
<th>Average</th>
<th>S.D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>81.46 %</td>
<td>1.72</td>
</tr>
<tr>
<td>2</td>
<td>42.38 %</td>
<td>1.50</td>
</tr>
<tr>
<td>3</td>
<td>40.53 %</td>
<td>2.35</td>
</tr>
<tr>
<td>4</td>
<td>40.32 %</td>
<td>4.22</td>
</tr>
<tr>
<td>5</td>
<td>36.42 %</td>
<td>3.67</td>
</tr>
</tbody>
</table>
Table 6.4: Overall Accuracy Per Subject. mDWT Feature. Setting 2

<table>
<thead>
<tr>
<th>Subject</th>
<th>Average</th>
<th>S. D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>54.99%</td>
<td>15.30</td>
</tr>
<tr>
<td>2</td>
<td>53.98%</td>
<td>20.09</td>
</tr>
<tr>
<td>3</td>
<td>41.19%</td>
<td>21.91</td>
</tr>
<tr>
<td>4</td>
<td>43.25%</td>
<td>19.60</td>
</tr>
<tr>
<td>5</td>
<td>51.73%</td>
<td>23.59</td>
</tr>
<tr>
<td>6</td>
<td>56.46%</td>
<td>17.24</td>
</tr>
<tr>
<td>7</td>
<td>46.74%</td>
<td>17.13</td>
</tr>
<tr>
<td>8</td>
<td>46.35%</td>
<td>17.53</td>
</tr>
<tr>
<td>9</td>
<td>43.35%</td>
<td>20.77</td>
</tr>
<tr>
<td>10</td>
<td>43.73%</td>
<td>21.44</td>
</tr>
</tbody>
</table>

6.3 Setup 3

The third setup is explained in Section 4.3.3. It is made on training first only on the morning (afternoon) of the first day. Then the training set is expanded to the set of the rest of the mornings (afternoons). The testing is fixed to the morning (afternoon) of the fifth day (Table 4.4). Figures 6.3 and Figure 6.3 show the accuracy of the morning and afternoon acquisitions with classifier Random Forest and feature MAV and WL, respectively. Whereas 6.3 illustrates the accuracy of the morning and afternoon acquisitions with classifier KRLS and feature mDWT. From the graphs it is already possible to state that adding days to the training set improves the classification accuracies, although there are some outliers, especially exploiting KRLS classifier and mDWT feature. Thus, this is not a fixed rule.
CHAPTER 6. RESULTS

Figure 6.7: Classification accuracies for setup 3 exploiting Random Forest classifier and MAV feature.

Figure 6.8: Classification accuracies for setup 3 exploiting Random Forest classifier and WL feature.
6.3. SETUP 3

(a) Mornings

(b) Afternoons

Figure 6.9: Classification accuracies for setup 3 exploiting KRLS and mDWT feature.

Table 6.5 and 6.6 show the overall accuracies and standard deviations for each day and for each subject. Only the feature mDWT is represented, for brevity. Analyzing Table 6.5, it is possible to state that enlarging the training set with dataset coming from other days, beside the first one, allows to reach better results. In fact, the classification accuracies gradually increase from 36.32% to 49.10%, growing of a total of 12.78%.

Table 6.5: Overall Accuracy Par Day. mDWT Feature. Setup 3

<table>
<thead>
<tr>
<th>N. Training Days</th>
<th>Average</th>
<th>S.D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>36.32%</td>
<td>3.46</td>
</tr>
<tr>
<td>2</td>
<td>42.18%</td>
<td>2.43</td>
</tr>
<tr>
<td>3</td>
<td>45.85%</td>
<td>1.33</td>
</tr>
<tr>
<td>4</td>
<td>49.10%</td>
<td>2.33</td>
</tr>
</tbody>
</table>
CHAPTER 6. RESULTS

Table 6.6: Overall Accuracy Per Subject. mDWT Feature. Setting 3

<table>
<thead>
<tr>
<th>Subject</th>
<th>Average</th>
<th>S. D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>46.07 %</td>
<td>3.92</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>36.10 %</td>
<td>10.97</td>
</tr>
<tr>
<td>4</td>
<td>43.86 %</td>
<td>11.18</td>
</tr>
<tr>
<td>5</td>
<td>59.63 %</td>
<td>5.44</td>
</tr>
<tr>
<td>6</td>
<td>42.70 %</td>
<td>0.95</td>
</tr>
<tr>
<td>7</td>
<td>46.70 %</td>
<td>6.21</td>
</tr>
<tr>
<td>8</td>
<td>41.06 %</td>
<td>5.83</td>
</tr>
<tr>
<td>9</td>
<td>31.23 %</td>
<td>7.06</td>
</tr>
<tr>
<td>10</td>
<td>34.15 %</td>
<td>4.10</td>
</tr>
</tbody>
</table>

NOTE. Subject 2 is missing from the afternoons of classifier KRLS since the data of Day 2 are corrupted and the classifier is not able to work with this condition.

6.4 Setup 4

The fourth setup is explained in Section 4.3.4. The training set is fixed and is created using the morning (or afternoon) of the first day. Whereas the training set is generated using an increasing set of mornings (or afternoons) of the remaining days (Table 4.5).

Figures 6.4 and Figure 6.4 show the accuracy of the morning and afternoon acquisitions with classifier Random Forest and feature MAV and WL, respectively. Whereas 6.4 illustrates the accuracy of the morning and afternoon acquisitions with classifier KRLS and feature mDWT. From the graphs it is possible to state that adding days to the testing set decreases the classification accuracies, although there are some outliers, e.g. subject 10. Thus, this is not a fixed rule.
6.4. SETUP 4

Figure 6.10: Classification accuracies for setup 4 exploiting Random Forest classifier and MAV feature.

Figure 6.11: Classification accuracies for setup 4 exploiting Random Forest classifier and WL feature.
CHAPTER 6. RESULTS

Figure 6.12: Classification accuracies for setup 4 exploiting KRLS and mDWT feature.

Table 6.7 and 6.8 show the overall accuracies and standard deviations for each day and for each subject. Only the feature mDWT is represented, for brevity. Analyzing Table 6.7, it is possible to state that enlarging the testing set with dataset coming from more than one day brings worst results. In fact, the classification accuracies gradually decrease from 41.61% to 39.82%, dropping of a total of 1.79%, which is not a big fall.

Table 6.7: Overall Accuracy Per Day. mDWT Feature. Setup 4

<table>
<thead>
<tr>
<th>N. Testing Days</th>
<th>Average</th>
<th>S.D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>41.61 %</td>
<td>2.59</td>
</tr>
<tr>
<td>2</td>
<td>41.38 %</td>
<td>0.31</td>
</tr>
<tr>
<td>3</td>
<td>41.01 %</td>
<td>1.59</td>
</tr>
<tr>
<td>4</td>
<td>39.82 %</td>
<td>2.00</td>
</tr>
</tbody>
</table>
Table 6.8: Overall Accuracy Per Subject. mDWT Feature. Setting 4

<table>
<thead>
<tr>
<th>Subject</th>
<th>Average</th>
<th>S. D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>48.78 %</td>
<td>2.56</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>37.02 %</td>
<td>4.13</td>
</tr>
<tr>
<td>4</td>
<td>37.78 %</td>
<td>3.08</td>
</tr>
<tr>
<td>5</td>
<td>40.72 %</td>
<td>2.09</td>
</tr>
<tr>
<td>6</td>
<td>50.64 %</td>
<td>1.23</td>
</tr>
<tr>
<td>7</td>
<td>40.37 %</td>
<td>2.04</td>
</tr>
<tr>
<td>8</td>
<td>40.35 %</td>
<td>1.40</td>
</tr>
<tr>
<td>9</td>
<td>34.39 %</td>
<td>1.05</td>
</tr>
<tr>
<td>10</td>
<td>34.48 %</td>
<td>2.25</td>
</tr>
</tbody>
</table>

**NOTE.** Subject 2 is missing from the afternoons of classifier KRLS since the data of Day 2 are corrupted and the classifier is not able to work with this condition.
Chapter 7

Conclusions and Future works

In this thesis, we analyze the repeatability of grasp recognition for robotic hand prosthesis control based on sEMG data. The project has two main goals:

1. To release a repeatability database with the data recorded during the experiments, available online for everyone;
2. To explore repeatability in sEMG data through movement classification accuracy.

The data were recorded from 10 subjects (3 females, 7 males, average age 27 ± 6 years) in a period of 5 days, twice on each day (morning and afternoon). The repeatability database about sEMG hand movement recognition, obtained from these acquisitions, is publicly released on the NinaPro website.¹

Before the raw data can be used for classification, several steps are necessary. The movement classification includes preprocessing, relabeling, feature extraction, and classification. The classifiers chosen for the project are:

- Random Forest with Mean Absolute Value and Waveform Length;
- Kernel Regularized Least Squares and marginal Discrete Wavelet Transform feature.

Since we are interested in understanding how much the accuracy classification of hand grasps changes with data acquired from different acquisitions, it was decide to exploits four different setup which use unique training and testing dataset:

¹url: http://ninapro.hevs.ch/
1. The first setup is created to understand the possibility to use only one session for the training data of the robotic prosthesis; The training is produced using the mornings of the days, while the testing of the same day is created exploiting the afternoons. The movement classification accuracies obtained when training and test sets are from the same acquisitions are 27.03% higher than those obtained when training and test sets are from different acquisitions. The accuracies with training and testing from different acquisitions is higher than the chance level for the considered number of movements (12.5%), thus suggesting that more than one session of different acquisitions could be useful to train the control systems of the prosthesis.

2. The second setup is similar to the first one; the training data is based on the morning (or afternoon) of the first day and the testing is implemented on the mornings (or afternoons) of the remaining days. As a result it is obtained that the average of the overall accuracy per day gradually decreases as we distance ourselves from the training dataset. The best classification accuracy is reached on the first day and it decreases of a maximum of 45.04% on the last day of acquisition.

3. The third setup is made to comprehend how much the classification accuracies improve using more than one session for the training of the robotic hand prosthesis. The training set is created using an increasing group of dataset. The testing is fixed and is made on the morning (or afternoon) of the last day. Enlarging the training set with dataset of more than one day allows to reach better results. The classification accuracies gradually increase from 36.32% to 49.10%, growing of a total of 12.78%.

4. The fourth setup is implemented to understand how much the classification accuracy decrease through the use of the robotic limb. The training set is fixed and is created using the morning (or afternoon) of the first day. Whereas the training set is generated exploiting an increasing set of mornings (or afternoons) of the remaining days. Enlarging the testing set with dataset coming from more than one day brings worst rate of accuracy for hand grasps. The classification accuracies gradually decrease from 41.61% to 39.82%, dropping of a total of 1.79%.

Each setup allowed to obtain an important result for the development of the thesis and can provide additional informations to develop more robust control systems for robotic prosthesis. At the same time, the acquired data...
can support researchers to analyze repeatability in future work and to better comprehend effects of outside factors on the resulting data.

There are some common findings for many experiments:

- The Kruskal-Wallis test shows that there are significant differences between the subjects ($p < 0.001$);

- The variability within each subject is quite low, suggesting that outside factors (e.g. size of arm, muscle fatigue, etc.) may contribute to determine the results.

Future applications could make use of the total set of 52 hand movements suggested in [24] and use the data recorded with the Tobii glasses in order to improve the classification of the hand grasps through object recognition. Moreover, the time of acquisitions could be increased to thrice a day, this could help on tracking the various reasons that cause the large variability on the classification rate and therefore improve the results. Last, the subjects exploited in this work were all healthy people, thus it would be interesting to use disabled people with prosthesis hands as subjects to obtain result more fitting to the research field.
Bibliography


[23] Jie Liu, Adaptive myoelectric pattern recognition toward improved multifunctional prosthesis control, Medical Engineering and Physics, 2015


