UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Ingegneria dell'Informazione

SCUOLA DI DOTTORATO DI RICERCA IN: Ingegneria dell'Informazione

INDIRIZZO: Scienza e Tecnologia dell'Informazione

CICLO: XXXI

SUBJECT-INDEPENDENT FRAMEWORKS FOR ROBOTIC DEVICES: APPLYING ROBOT LEARNING TO EMG SIGNALS

Coordinatore d'indirizzo: *Ch.mo Prof. Andrea Neviani*
Supervisore: *Ch.mo Prof. Emanuele Menegatti*
Co Supervisore: *Ch.mo Prof. Enrico Pagello*

Dottorando:
Francesca Stival

to Stefano

Acknowledgements

At the end of this journey I have to thank my supervisors, Prof. Emanuele Menegatti and Prof. Enrico Pagello, for their guide and valuable advices, thanks for giving me the opportunity to work at the Intelligent Autonomous System Lab. I also want to thanks to my colleagues, and each one who spent some time at the lab, for the interesting discussions and for the pleasant moments. A special thank to my deskmate Morris, who constantly tried to stole my space and my pens, but gave valuable Bayesian advices too.

I want also to thank Prof. Manfredo Atzori and Prof. Henning Müller for having welcomed me at the HESSO, and for making me feel like home. Furthermore, I would like to thank Prof. Giuseppina Gini and Prof. Tetsuro Funato for their time as external reviewers, thanks for the precious advices.

A great thanks goes to my friends and my family for their constant support. A special thanks to my parents, to Laura and-of course-to my uncle, who was my mentor for scientific disciplines since when I was a child. Thanks for the support, the advices and the help in these years. I want to include my aunt in this list, she is not physically here anymore, but I am sure she is proud of this result, and thrilled for each new trip.

Last but not least, a huge thank you goes to Stefano, for his patience, his support, his kindness, and his generous help. Thank you for helping me to face calmly and rationally all the problems, for the constant encouraging, and for always believing in me.

Contents

List of Figures

List of Tables

Acronyms

ML Machine Learning

AIC Akaike Information Criterion

BIC Bayesian Information Criterion

CWT Continuous Wavelet Transform

DoF Degree of Freedom

DoFs Degrees of Freedom

DWT Discrete Wavelet Transform

EEG Electroencephalogram

EM Expectation-Maximization

EMG Electromyography

sEMG Surface Electromyography

FFT Fast Fourier Transform

FT Fourier Transform

GMM Gaussian Mixture Model

IGMM Incremental Gaussian Mixture Model

GMR Gaussian Mixture Regression

GMC Gaussian Mixture Classification

HMM Hidden Markov Model

HMM Hill Muscle Model

HPR Hierarchical Projected Regression

IAV Integral Absolute Value

MAV Mean Average Value

MoG Mixture of Gaussians

LOO Leave One Out

MDL Minimum Description Length

MML Minimum Message Length

MU Motor Unit

NMSE Normalized Mean Square Error

NN Neural Networks

RLfD Robot Learning from Demonstration

RMS Rooted Mean Square

ROS Robot Operating System

sEMG Surface Electromyography

YAML YAML Aint Markup Language

DTW Dynamic Time Warping

GoF Goodness of Fit

COW Correlation Optimized Warping

WT Wavelet Transform

EuRoC European Robotics Challenges

ROS Robot Operating System

UR Universal Robots

VS Visual Servoing

IBVS Image-Based Visual Servoing

PBVS Position-Based Visual Servoing

LfKD Learning from Kinesthetic Demonstration

GMM Gaussian Mixture Model

GMR Gaussian Mixture Regression

EM Expectation Maximization

MSE Mean Square Error

CCA Canonical Correlation Analysis

FSM Finite State Machine

SVM Support Vector Machine

PCA Principal Components Analysis

MoCap Motion Capture

AAL Active and Assisted Living

NMF Non-negative Matrix Factorization

VAF Variance-Accounted-For

ACC accelerometer

ADL Activities of Daily Living

PDF Probability Density Function

IMU Inertial Measurement Unit

TD Time Domain

WL Wavelet

FAI Finger Aperture Index

WFL Waveform Length

TSC Time Stamp Counter

MANOVA Multivariate Analysis of Variance

MAF Maximum Agreement Forest

SPR Subtree Prune-and-Regraft

MUAP Motor Unit Action Potential

Abstract

The capability of having human and robots cooperating together has increased the interest in the control of robotic devices by means of physiological human signals. In order to achieve this goal it is crucial to be able to catch the human intention of movement and to translate it in a coherent robot action. Up to now, the classical approach when considering physiological signals, and in particular EMG signals, is to focus on the specific subject performing the task since the great complexity of these signals.

This thesis aims to expand the state of the art by proposing a general subject-independent framework, able to extract the common constraints of human movement by looking at several demonstration by many different subjects. The variability introduced in the system by multiple demonstrations from many different subjects allows the construction of a robust model of human movement, able to face small variations and signal deterioration. Furthermore, the obtained framework could be used by any subject with no need for long training sessions.

The signals undergo to an accurate preprocessing phase, in order to remove noise and artefacts. Following this procedure, we are able to extract significant information to be used in online processes. The human movement can be estimated by using well-established statistical methods in Robot Programming by Demonstration applications, in particular the input can be modelled by using a Gaussian Mixture Model (GMM). The performed movement can be continuously estimated with a Gaussian Mixture Regression (GMR) technique, or it can be identified among a set of possible movements with a Gaussian Mixture Classification (GMC) approach. We improved the results by incorporating some previous information in the model, in order to enriching the knowledge of the system. In particular we considered the hierarchical information provided by a quantitative taxonomy of hand grasps. Thus, we developed the first quantitative taxonomy of hand grasps considering both muscular and kinematic information from 40 subjects. The results proved the feasibility of a subject-independent framework, even by considering physiological signals, like EMG, from a wide number of participants.

The proposed solution has been used in two different kinds of applications: (I) for the control of prosthesis devices, and (II) in an Industry 4.0 facility, in order to allow human and robot to work alongside or to cooperate. Indeed, a crucial aspect for making human and robots working together is their mutual knowledge and anticipation of other's task, and physiological signals are capable to provide a signal even before the movement is started. In this thesis we proposed also an application of Robot Programming by Demonstration in a real industrial facility, in order to optimize the production of electric motor coils. The task was part of the European Robotic Challenge (EuRoC), and the goal was divided in phases of increasing complexity. This solution exploits Machine Learning algorithms, like GMM, and the robustness was assured by considering demonstration of the task from many subjects. We have been able to apply an advanced research topic to a real factory, achieving promising results.

Keywords: Subject-Independence, Physiological Signals, EMG Signals, Quantitative Taxonomy of Hand Grasps, Human-Robot Interaction (HRI), Robot Programming by Demonstration (RPbD)

Sommario

La possibilità di collaborazione tra robot ed esseri umani ha fatto crescere l'interesse nello sviluppo di tecniche per il controllo di dispositivi robotici attraverso segnali fisiologici provenienti dal corpo umano. Per poter ottenere questo obiettivo è essenziale essere in grado di cogliere l'intenzione di movimento da parte dell'essere umano e di tradurla in un relativo movimento del robot. Fin'ora, quando si consideravano segnali fisiologici, ed in particolare segnali EMG, il classico approccio era quello di concentrarsi sul singolo soggetto che svolgeva il task, a causa della notevole complessità di questo tipo di dati.

Lo scopo di questa tesi è quello di espandere lo stato dell'arte proponendo un framework generico ed indipendente dal soggetto, in grado di estrarre le caratteristiche del movimento umano osservando diverse dimostrazioni svolte da un gran numero di soggetti differenti. La variabilità introdotta nel sistema dai diversi soggetti e dalle diverse ripetizioni del task permette la costruzione di un modello del movimento umano, robusto a piccole variazioni e a un possibile deterioramento del segnale. Inoltre, il framework ottenuto può essere utilizzato da ogni soggetto senza che debba sottoporsi a lunghe sessioni di allenamento.

I segnali verranno sottoposti ad un'accurata fase di preprocessing per rimuovere rumore ed artefatti, seguendo questo procedimento sarà possibile estrarre dell'informazione significativa che verrà utilizzata per elaborare il segnale online. Il movimento umano può essere stimato utilizzando tecniche statistiche molto diffuse in applicazioni di Robot Programming by Demonstration, in particolare l'informazione in input può essere rappresentata utilizzando il Gaussian Mixture Model (GMM). Il movimento svolto dal soggetto può venire stimato in maniera continua con delle tecniche di regressione, come il Gaussian Mixture Regression (GMR), oppure può venire scelto tra un insieme di possibili movimenti con delle tecniche di classificazione, come il Gaussian Mixture Classification (GMC). I risultati sono stati migliorati incorporando nel modello dell'informazione a priori, in modo da arricchirlo. In particolare, è stata considerata l'informazione gerarchica fornita da una tassonomia quantitativa dei mo-

vimenti di presa della mano. E' stata anche realizzata la prima tassonomia quantitativa delle prese della mano considerando l'informazione sia muscolare che cinematica proveniente da 40 soggetti. I risultati ottenuti hanno dimostrato la possibilità di realizzare un framework indipendente dal soggetto anche utilizzando segnali fisiologici come gli EMG provenienti da un grande numero di partecipanti.

La soluzione proposta è stata utilizzata in due tipi diversi di applicazioni: (I) per il controllo di dispositivi prostetici, e (II) in una soluzione per l'Industria 4.0, con l'obiettivo di consentire a uomini e robot di lavorare assieme o di collaborare . Infatti, un aspetto cruciale perché uomini e robot possano lavorare assieme è che siano in grado di anticipare uno il task dell'altro e i segnali fisiologici riescono a fornire un segnale prima che avvenga l'effettivo movimento. In questa tesi è stata proposta anche un'applicazione di Robot Programming by Demonstration in una vera fabbrica che si occupa di realizzare motori elettrici, con lo scopo di ottimizzarne la produzione. Il task faceva parte della European Robotic Challenge (EuRoC) in cui l'obiettivo finale era diviso in fasi di complessità crescente. La soluzione proposta impiega tecniche di Machine Learning, come il GMM, mentre la robustezza dell'approccio è assicurata dalla considerazione di dimostrazioni da parte di molti soggetti diversi. Il sistema è stato testato in un contesto industriale ottenendo risultati promettenti.

Parole chiave: Indipendenza dal soggetto, Segnali fisiologici, Segnali EMG, Tassonomia quantitativa dei movimenti di presa, Interazione uomo-robot, Robot Programming by Demonstration

Chapter 1

Introduction

During the last years, the dissemination of robots has exploded in many aspects of everyone lives. Up to now, we can meet robotic devices not only in the most advanced factories, but also in our houses. Nowadays, it is common to find a robot autonomously cleaning up a house, or assisting a surgeon during a medical operation. The main reasons are the price decrease and the boosted investments to develop technology in the field of robotics. All these new technologies have the goal of improving humans' quality of life, for example by reducing their workload, or by substituting the operator in dangerous and strenuous tasks. Furthermore, recent advancements are meant to help injured persons regain their lost functionalities, for example by providing robotized prosthesis, exoskeletons or advanced rehabilitation frameworks [1] [2]. These last solutions exploit physiological signals from the subject's body, in order to emulate and replicate the human behaviour. Physiological signals are generated from the nervous system and they cause muscle contractions, thus a movement. Considering physiological signals for controlling a device like a prosthesis has a twofold vantage. Firstly, they are generated before the actual movement, thus it is possible to detect the movement the subject will do before it is actually performed. The second vantage comes from the fact that we do not record the actual movement of a limb, but the muscular activation. This means that it is possible to collect this kind of signals even from amputated subjects, by recording the activation of their residual muscles. Despite the great progresses in the field, and the constant development of new, innovative solutions, we are still far from the realization of a prosthesis able to exactly reproduce a human limb. Physiological signals used to control prosthesis, like EMG signals, are sensitive to several factors [3]. Machine Learning can be used as a tool to generalize muscular behaviors in different conditions with the objective of inferring information useful to control a robotic device, e.g. the joints bending angles of a prosthesis [4]. In particular, Robot

Programming by Demonstration paradigm aims to train robotic devices through human demonstrations in order to teach them how to perform a task [5]. Usually, robots need a large number of demonstrations to learn how to perform a simple task. Selecting a specific Machine Learning algorithm could help to reduce the number of demonstrations in certain contexts, but otherwise improvements are limited. On the other hand, overcomplicated Machine Learning algorithms could end in overfitting. Consequently, the model could fail in predicting reliably future observations. The same problem may occur when data are limited and the model focuses on specific situations without the possibility to fit additional data. With overfitting, the framework lacks of abstraction and generalization capability and it will not be able to face even limited variations. Generalization is a key concept if the goal is to obtain a relation between movements. Nevertheless, very simple algorithms coupled with an excessively wide dataset could lead to underfitting. The resulting model will not capture the common characteristics among the data providing poor predictive performance. A good way to incorporate variability is to consider actions developed by many different subjects [6]. The same gesture can slightly change depending on who is doing it: the gesture remains correct, but the ways to perform it are almost infinite. However, the risk of underfitting is forestalled since there are common characteristics among the numerous ways different subjects perform the same movement.

From a kinematic point of view it is easy to prove the similarity among different subjects gestures. The joint angles are easily recordable, the information is smooth and regular, thus it is not hard to find a common path, even among different subjects. The situation definitely changes when moving to physiological signals. In particular, EMG signals are deeply affected from the source of the signal. Since EMGs records the muscular activation, different muscular conditions are translated in different signals. Each person is different from the other, this means that the muscular activation could vary among subjects, even if they are performing the same task. Furthermore, human body is a redundant system, thus each person interpret the movement in a peculiar way depending on the differences and characteristics that diversify every human being. Actually, the same movement could be generated by the activation of different EMG signals even for the same subject. Recordings of EMG signals can be affected from muscular fatigue, psychological and physiological stress and tiredness, presence of body fat, small changes in the sensors position, etc [7]. For these reasons, the classical approach when working with EMG signals is to focus on a specific subject [8]. Nevertheless, the usage of a subject-independent approach would provide many vantages. The first is the saving of a great amount of time and resources, since there is

no need for long and draining training phases to adapt the model to a new subject. This aspect is particularly important for injured subjects, for whom it is very painful and debilitating to repeat a task many times. Furthermore, the variability included in subject-independent frameworks ensures a greater ability of facing unexpected situations and reacting to limited variations in input.

A subject-independent approach is useful not only for controlling prosthesis or exoskeletons, but it can be applied also to industrial settings [9]. Besides physiological signals, other techniques can be used for robot learning. In kinesthetic demonstration [10], the robot is physically guided through the task by the humans. This technique benefits from its intrinsic subject-independence owed to robots capability of being a filter among different subjects. The subject-independence is harder to achieve when learning from visual information [11]. In fact, the constraints characterizing the movement should be extracted from a sequence of images, in a similar fashion to what happens when considering physiological information. Furthermore, industrial applications requires particular conditions on safety and efficiency. These aspects have to be taken into account when building the model to control a robotic device. The advent of Industry 4.0 brought new and innovative challenges [12] for robotics. The new concept of industry aims at reducing the waste, while maximizing the customization of the product, therefore a flexible and dynamic production line is essential. An efficient way to produce is necessary in modern factories, and the manufacturing system should be able to switch production in a very short time. The presence of intelligent and collaborative robots is a key factor for the fulfillment of these targets. Old-fashion robots are expensive devices, closed in a cage, repeating continuously the same task. Reprogram one of these robots takes time, money and requires the intervention of specialized programmers. New robots are more lightweight and no longer closed in cages, since they are equipped with force sensors, aware of possible contacts with the surrounding world. The characteristics of collaborative robots (Cobots) make them ideal to be Programmed by Demonstration. This programming paradigm reduces the time needed to program the robot, since there is no requirement of specialized personnel. On the contrary, the machine will learn the task by learning the demonstrations performed by skilled worker which knows well the tasks the robot should do. Furthermore, Cobots offer the possibility of having humans and machines working on the same workplace, an also to operate together to fulfill the same task. The closeness of the machine to humans arises several safety problems. The main issue is to avoid accidental contacts among humans and robots. Giving robot the capabilities perceive, understand and react to what happens in the environment will become essential in the factories of the future.

In other words, the robot should be intelligent, capable of interpreting feedbacks from outside, and it will need the ability to understand and probably also predict human movements.

In this thesis, we transferred an advanced research solution, like Robot Programming by Demonstration to a real industrial case. In particular, our goal was to boost the production of electric motor coils, by automatizing the copper winding procedure. The project was part of a European challenge aiming to encourage collaboration between academic and industrial counterparts [13].

1.1 Thesis outline

The topics introduced in the previous section will be deeply investigated in this thesis. The main idea connecting the different aspects among the whole thesis is the development of subject-independent solutions by exploiting Robot Programming by Demonstration frameworks. Being independent from the specific subject execution of a task ensures improved robustness and generalization capabilities. The obtained results will allow the usage of the developed technology by different subjects with no need of long training phases. Subject-independent frameworks can be applied to different situations. The wider part of this thesis regards the implementation of subject-independent models in the context of physiological signals to control prosthesis, exoskeletons or rehabilitation devices. Chap. 2 describes the existing solutions exploiting physiological signals for the control of robotics devices, with particular attention on subject-independent approaches. In the same Chapter the characteristics of EMG signals are illustrated, together with an overview of the datasets used for testing purposes. The intrinsic characteristics of EMG signals require a dedicated preprocessing phase. In Chap. 3, we discuss the preprocessing techniques ensuring a proper online signal elaboration. The processed physiological signals are then used to feed a probabilistic model, namely a GMM that is able to represent a connection between the physiological signals and the limbs movement. A classification technique, namely Gaussian Mixture Classification (GMC) allows the detection of the performed movement among a set of actions, while a regression approach, namely GMR, is able to continuously estimate the joint angles. GMM, GMC and GMR, together with the incremental version Incremental Gaussian Mixture Model (IGMM) are also expounded in Chap. 3. The experimental scenarios and the results obtained during the tests have been listed in Chap. 4. After an initial test involving a limited number of subjects to assure the feasibility of a subject-independent framework, we expand the approach

to a wider dataset. Initially, we focus on information coming exclusively from EMG signals, to investigate afterwards how the results change by including IMU information. We also propose a complete, low cost, classification framework, including signals recording, modelization and prediction applied to a low cost 3D printed prosthesis.

Including some a priori information in the framework would intuitively improve the predicted results. In this thesis, we speculate that the information from a hierarchic taxonomy of hand grasps could provide a guideline able to speed-up and improve classification results. Nonetheless, the many taxonomies proposed up to now are based on qualitative parameters, and a quantitative taxonomy is missing. Therefore, in Chap. 5, we introduce the first quantitative taxonomy of hand grasps. This taxonomy include quantitative information considering data by many different subjects. By doing so, we are able to integrate our taxonomy with a complete and robust description of hand grasps from a physiological point of view while ensuring repeatability and generalization. The quantitative taxonomy have been exploited in Chap. 6 providing a subject-independent binary classification framework.

Traditionally, physiological signals are used to control prosthesis or rehabilitation devices. However, the detection of human behaviour is useful also in other applications. In modern Industry 4.0 factories, intelligent robots work alongside humans in order to optimize the production flow. The concept of intelligence for robots comes from their capability to understand what happens in the world. Robots working alongside humans should always be aware of their human counterparts, in order to avoid dangerous situations. Knowing the position of the operator is not sufficient, the robot should be able to predict what the human will do. A possible way to predict the human movement exploits the use of physiological signals, as described in the first part of the thesis. Chap. 7 propose a framework based on physiological signals to predict human intention of movement in an industrial context in order to allow the robot to cooperate and work alongside human operators.

A general subject-independent approach has been applied also in other industrial settings, in particular to the EuRoC challenge. This challenge aims to boost the collaboration between research and industrial partners, in order to achieve innovative results. In Sec. 8.1, we illustrate in brief the EuRoC challenge and the importance of innovative solutions in industry. The challenge was divided in three phases of increasing difficulty. A key factor associating the solution proposed in the various phases is the usage of a general Robot Programming by Demonstration paradigm. Such paradigm exploits Machine Learning techniques to make the robot assimilate the tasks by observing human demonstrations. The three phases are accurately described in Sec. 8.3,

Sec. 8.4, Sec. 8.5.

Finally, in Chap. 9, we summarize the work and the results achieved during the three year gathered in this thesis.

1.2 List of Publications

Parts of the content presented in this thesis have been previously published or submitted for publication in national and international conferences and journals.

Physiological signals for the control of prosthesis

- ***F. Stival**, S. Michieletto, E. Pagello*, "Subject-Independent Modeling of sEMG signals for the Motion of a Single Robot Joint", in Workshop of Robotics: Science and Systems (RSS2015) on Combining AI Reasoning and Cognitive Science with Robotics. Rome, Italy, 2015.

- ***F. Stival**, S. Michieletto, E. Pagello*, "Online subject-independent modeling of sEMG signals for the motion of a single robot joint", in Proceedings of the 6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BIOROB), Singapore, 2016.

- ***F. Stival**, S. Michieletto, A. De Agnoi, E. Pagello*, "Toward a better robotic hand prosthesis control: using EMG and IMU features for a subject independent multi-joint regression model", in Proceedings of the 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (BIOROB), Enschede, Holland, 2018.

- ***F. Stival**, S. Michieletto, E. Pagello*, "Subject Independent EMG Analysis by using Low-Cost hardware", in Proceedings of the IEEE International Conference on Systems, Man, Cybernetics 2018 (SMC 2018), Miyazaki, Japan, 2018.

Quantitative Taxonomy of Hand Grasps

- ***F. Stival**, S. Michieletto, E. Pagello, H. Muller, M. Atzori*, "Quantitative hierarchical representation and comparison of hand grasps from electromyography and kinematic data", in Workshop Proceedings of the 15th International Conference on Autonomous Systems IAS-15, Workshop on Learning Applications for Intelligent Autonomous Robots (LAIAR-2018), Baden-Baden, Germany, 2018.

- ***F. Stival**, M. Moro, E. Pagello*, "A first approach to a taxonomy-based classification framework for hand grasps", in Workshop Proceedings of the 15th International Conference on Autonomous Systems IAS-15, Workshop on Learning Applications for Intelligent Autonomous Robots (LAIAR-2018), Baden-Baden, Germany, 2018.

- ***F. Stival**, S. Michieletto, M. Cognolato, E. Pagello, H. Mller, M. Atzori*, "A Quantitative Taxonomy of Human Hand Grasps", submitted to Journal of NeuroEngineering and Rehabilitation, 2018.

Physiological signals in Industrial Settings

- *S. Michieletto, S. Tortora, **F. Stival**, E. Menegatti*, "Low-cost Human-Robot Interface for Online Motion Intention Prediction", submitted to International Conference on Robotics and Automation (ICRA 2019).

EuRoC challenge

- *S. Michieletto, **F. Stival**, F. Castelli, E. Pagello*, "Teaching Door Assembly Tasks in Uncertain Environment", in Proceedings of the 47st International Symposium on Robotics (ISR 2016), Munich, Germany, 2016.

- *S. Michieletto, **F. Stival**, F. Castelli, M. Khosravi, A. Landini, S. Ellero,R. Land, N. Boscolo, S. Tonello, B. Varaticeanu, C. Nicolescu, E. Pagello*, "Flexicoil: Flexible robotized coils winding for electric machines manufacturing industry", in ICRA 2017 Workshop on Industry of the future: Collaborative, Connected, Cognitive, Singapore, 2017.

- ***F. Stival**, S. Michieletto, E. Pagello*, "How to deploy a wire with a robotic platform: Learning from human visual demonstrations", in Proceedings of 27th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM2017), ELSEVIER, Modena, Italy, 2017.

- *S. Michieletto, **F. Stival**, F. Castelli, E. Pagello* "Automated and Flexible Coil Winding Robotic Framework", in Proceedings of the 50st International Symposium on Robotics (ISR 2018), Munich, Germany, 2018.

Chapter 2

Physiological Signals: Analysis and Processing

2.1 Subject Independence for physiological signals

The interaction with robotic devices by means of physiological human signals has become of great interest in the last years. In particular it is significant their capability of catching human intention of movement and translating it in a coherent movement performed by a robotic platform. Moreover, the interest in wearable devices, the study of new materials, the improvements in mechanical design and the advancement of sensors are boosting the development of robotic prostheses as never before. Physiological characteristics of the human movement are more and more at the center of the technological improvements. For example, they can be exploited by injured subjects to replace lost limbs. In fact, physiological signals are usually applied to help amputees in gaining part of their lost functionality.

The information used to estimate and predict a human movement comes directly from the subject performing the task, regardless of the specific signal or the combination of them exploited in the process. The movement can be estimated with a regression approach, by continuously predicting the joints bending angles, or it is possible to classify and choose the movement performed by the subject among several sets of motion selected in advance. The majority of the studies in literature focus on classification problems which are able to determine the type of movement, but not the actual trajectories. On the contrary, the use of regression techniques allows a continuous and proportional control of robotic platforms.

In order to deeply understand and study the human behaviour, many physical and physiological information from subjects can be recorded both in invasive and non-

invasive ways. Different elements could provide different information, like the strength employed to fulfill a task, the joints angles along time, the muscular activation, and also the body position in the space. In order to exploit the human information to control a robotic device, the signals acquired from the human user should fulfill some requirements and overcome possible limitations:

- **Easy collection:** the signals should be collected with small devices, and in any condition. This means not only that the human user should not feel pain during the acquisition, but also that she/he should behave naturally, avoiding forced movements and positions. Invasive acquisition or dedicated surgery should not be excluded, unless the subject quality of life worsen.

- **Continuous acquisition:** in order to control a robotic device, also for long time periods, it should be possible to collect data continuously during the day. The subject should not feels the stress of the registration, or change dramatically his/hers behaviours. Furthermore, the user should be able to use the device everywhere, thus the signals recording setup should not be placed in an ad-hoc location, unless the robotic device that has to be controlled is in an industrial facility.

- **Effective signals:** collected signals should characterize profitably the information from the human. After a specific preprocessing the recorded signals should be informative enough to allow the recognition of the executed movement.

The majority of studies involved healthy subjects. Someone could foreseen a limitation in such approaches, since technologies like prostheses are developed mainly to help impaired subjects. Nevertheless, tests on injured people can be difficult to perform, since they are often affected by fatigue, both physical or psychological. Physical issues appear when people push too much on the impaired part, e.g. the subject cannot reach a specific pose and forces the muscles to unnatural positions. Psychological disturbs occur when they are under the effect of a major stress, e.g. the subject cannot accept the inability to perform a task, even simple. Working with healthy people can reduce the influence of fatigue, and provide a more robust way to collect data. These data are very important also to conduct comparative measurements between different techniques and to propose effective solutions to guide the research for injured patients. Anyway, studies on healthy subjects are less prone to physical or psychological problems, but not immune. This is particularly true when the system uses EMG or sEMG signals.

Many studies proved that EMG is an effective method to track and identify the muscular activity [7], thus they can be profitably used in order to predict the human movement. EMG signals provide good results even with low cost sensors [14], and sEMG have the supplemental vantage of allowing a non invasive data recording, since they are collected through superficial sensors applied on the skin. Nevertheless, this family of signals has the drawback of being conditioned by many common physical aspects [7]. They are influenced by muscular fatigue, amount of body fat, physical stress and tiredness. Furthermore, sEMG signals suffer from low voltage amplitude, broad bandwidth, and sensitivity to sensor placement. The result is a noisy signal which vary significantly during a relatively short period of time. Moreover, the information provided by EMG data is strongly dependent from the specific subject involved in the acquisition. Indeed, the muscular activation depends on the physical characteristics of the subject as well as on his mental conditions. As highlighted by Taylor *et al.* [15], repeatable and valid experiments can be obtained in strict conditions. Environment, observer, measurements and instruments should be the same. Furthermore, the same subjects must be involved in more that one experiment, so that a subject should not concentrate on a single action. The same actions should be repeated on a short period of time, to avoid major differences between iterations. This is the reason why the majority of studies concerning motion estimation by means of physiological signals are subject-specific, i.e. they focus on a determined subject in order to find a way to characterize and describe her/his motion. The subject-specific approach has the advantage of giving high accuracy, with the drawback of lacking in generality, since a new model has to be computed for every new subject. In addition, the accuracy of subject-specific models could decrease due to the deterioration of the signal in time, for example in case of continuous use of a prosthesis or for emotional instability. In order to overcome these limitations, we proposed a subject-independent approach, with the idea of selecting only the signal characteristics peculiar to several subjects. In literature, many different works studied subject-specific approaches involving EMG signals. A large number of them adopted classification techniques to establish a robust interaction with prosthesis devices. In [16] Ju *et al.* present the classification of 10 different grasps or in-hand manipulations using Fuzzy Gaussian Mixture Model, achieving an accuracy of 96.7%. Bu *et al*. [17] propose a framework based on Bayesian Networks for motion classification of a cooking task. The manually designed Bayesian Networks extract the statistical dependency between two continuous motions and it is combined with the output of a probabilistic Neural Network classifier, to improve stability and accuracy. In classification frameworks only a limited subset of movements are considered. In

real daily applications movements are complex and a single movement is composed by many simpler sub-movements and classification approaches could not track all of them. Instead, regression techniques are the ideal solution for estimating continuously a given movement. Krasoulis *et al* [18] proposed a regression technique for the continuous estimation of finger movements in a subject-specific framework. They proved that regression methods could generalize to novel movements, not included in the training dataset. LLoyd [19] and Gerus [20] were able to estimate forces by tuning the model parameters to fit the motion of a particular person. Among regression techniques, [21] should have a particular mention. In their work, Valentini *et al* processed signals online using Wavelet Transform focusing on subject-specific framework. EMG data have been elaborated in order to obtain comparable information through different subjects. The computed information has been used to train a GMM, resulting in a lightweight model with a reduced number of parameters to be kept, while GMR provided fast regression that perfectly matched the needs of online applications. In this thesis, we took inspiration from [21] to develop our original subject-independent frameworks able to process online signals through Wavelet Transform.

Subject-independent models are a quite new argument in the field of rehabilitation robotics. Nevertheless, some significant works have proved the effectiveness of these methods in finding a common pattern between distinct individuals. Furthermore, executing a particular task intuitively leads to some constraints that could be extracted by looking to different interpretations of the motion to obtain a subject-independent model. Creating a subject-independent model enables to generalize the control procedure by extracting specific features of EMG signals coming from multiple individuals. Studies in this field are few and relatively recent. First attempts focuses mainly on classification, they presented some inter-subject analysis alongside subject-specific approaches. Orabona *et al* [2] proposed a way to provide patients with a pre-trained model, which will be subsequently refined and adapted to the specific subject to shorten the training phase. Castellini *et al* [22] performed a cross-subject analysis as additional study by comparing the performances of models built on single subjects when fed with data from different users. More recent studies have highlighted that an underling common behavior can be identified between different subjects in order to obtain a subject-independent solution. Matsubara *et al* [6] developed a multi-user interface which can classify different movements using a bilinear model, achieving an accuracy of 73%. Khushaba [23] described a method based on Canonical Correlation Analysis (CCA) capable of adapting to new users while maintaining good performances. Antuvan *et al* [24] present a subject-independent classifier, for eliminating the calibration phase.

A hierarchical decision tree model classified 8 different movements involving two degrees of freedom, namely wrist flexion/extension and hand open/close. The model has been built on 4 subjects and tested on other 6 subjects. Hartwell *et al* [8] overturn this concept by proposing a way to optimally select a set of movements for a specific subject, in order to develop a personalized prosthetic control. Yang *et al.* [25] proposed a technique for the classification of real Activities of Daily Living, while Khushaba *et al.* [1] focused on individual and combined fingers control rather than on fixed, rough movements. Their solution reaches an accuracy up to 90% by using a Bayesian data fusion post-processing approach to maximize the probability of correct classification. The idea proposed by Gibson *et al*. [26] reached an accuracy of 79% by using an optimized decision tree able to generalize among users with no need of an additional training phase.

According to our knowledge, the only previous attempts of mixing together subject-independent and regression techniques have been proposed by Tommasi *et al* [27]. Their aim was to shorten the training procedure starting from a similar known model by minimizing the Mean Square Error (MSE) of the features measured with respect to the ones already processed for a single individual or a combination of subjects. While the method proposed by Tommasi *et al* mainly aims at improving an already existing model, our goal is building a "ready-to-use" model able to guarantee good performances since the first trials of a new user.

In this thesis, we exploited EMG signals from many different healthy subjects in order to build a probabilistic model that characterize the movement. We focused on different techniques for the preprocessing phase, and we estimated the performed movement with both regression and classification techniques. In particular, we developed different approaches to boost the classification of many hand movements by considering prior information. It is worth to notice that our study was almost completely oriented on healthy subjects. We decided not to cope with the limitations and difficulties connected to injured subjects, in order to avoid the long procedures needed to obtain data from patients. Furthermore, promising solutions for healthy subjects suggest possible applications in industrial environments.

2.2 Electromyography

The human body is composed by three major muscle types: skeletal, cardiac and smooth muscle. As their name suggests, skeletal muscles are attached to bones by tendons, and they are used to effect skeletal movement and to maintain posture. The

activation of skeletal muscles is voluntary, contrariwise to the other muscle types, whose activation is not under conscious control. In skeletal muscle, contraction is stimulated by electrical impulses transmitted by the motoneurons nerves, and the contraction translates into movement. Nerves move muscles in response to both voluntary and autonomic (involuntary) signals from the brain, different muscles correspond with dedicated regions in the primary motor cortex of the brain. In particular situations, muscles can react to reflexive nerve stimuli where the signal from the afferent fiber does not reach the brain, but produces the reflexive movement by direct connections with the efferent nerves in the spine. An essential contribution in the movement is given by the proprioperception, i.e. the unconscious awareness of where the various locations of the body are located in the space at any time. Several areas in the brain coordinates the movement thanks to the feedback given by proprioperception.

Signals sent from the brain are transmitted through neurons as electrical signals. This activity can be recorded through a technique called Electromyography. The EMG measures muscle response or electrical activity in response to a nerve's stimulation of the muscle. EMG signals derive from potential generated through muscular unit activation.

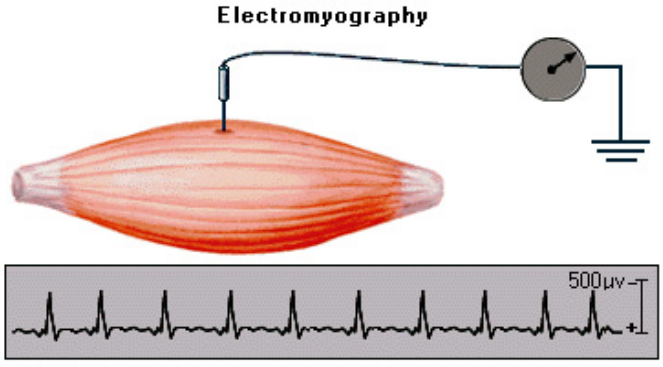


Figure 2.1: EMG acquisition

The analysis of EMG signals have a large variety of applications [7]. They are widely used in medicine for detecting neuromuscular diseases, and they are also an useful instrument for controlling external devices, such as prosthesis, exoskeletons, but also collaborative robots. The idea behind the analysis of such signals is to imitate the human body behaviour. The movement in limbs is due to muscular contraction caused by electric signals emitted from the brain. In order to emulate this behaviour on artificial devices, the goal is to create a mathematical and statistical model of EMG signals, in order to create a connection between electrical signals and movements.

EMG signals can be recorded both with invasive (intramuscular EMG) and non-

invasive (surface EMG) techniques Fig. 2.2. Invasive approaches have the vantage of providing more robust results, since many potentials noisy factors are bypassed, like the presence of body fat or skin damages. Furthermore, this technique enables the identification of the single muscles involved in the movement through the direct access to muscular fibers. The intuitive drawback of this solution is the necessity of a delicate and invasive procedure, which can be performed only by expert staff, and which can lead to possible damages for the subject, as for all invasive procedures. Furthermore, the control of robotic devices collecting data from invasive sensors could be difficult in everyday life.

The drawbacks and innate difficulties of this technique makes it relatively not widespread in research laboratories, where the preferred technique is the surface EMG. sEMG assesses muscle function by recording the activation of muscular fibres close to sensors. Thus, it is not possible to record exactly the single muscle involved in the movement, but a combination of them. Nevertheless, a proper sensors placement will ensure a good muscular isolation capability, at least for major muscles. Usually a large number of sensors are placed on the skin, and the redundancy is often required, in order to obtain a robust output. Unfortunately the sensors positioning on the skin arises other issues connected with the interference of the skin, fat and body hair positioned between the sensors and the muscles.

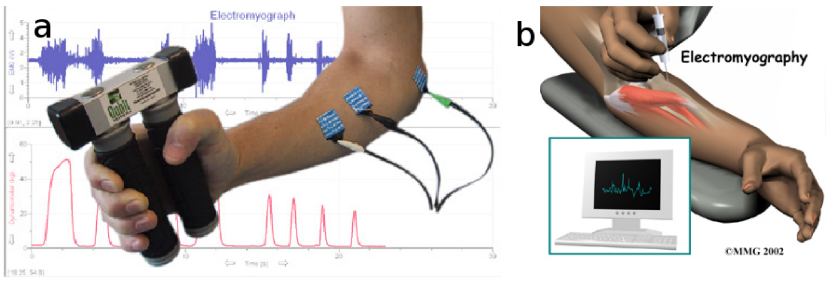


Figure 2.2: a) Surface recording of EMG signals, b) Invasive recording of EMG signals.

In general, EMG signals are deeply affected by physical and physiological fatigue, tiredness, stress, involuntary spasms and small shift in the sensors position. Therefore, using EMG data to model human motion is a though task. The problems affecting EMG signals show the key role of preprocessing. In fact, this phase aims at removing noise and outliars while highlighting the common constraints in several trials. If prepossesing is important when considering a specific subject, it becomes fundamental when dealing with a number of different individuals. It is worth to notice that physiological and structural differences among humans make subject-independent modelling

an ambitious goal. In particular, we are looking for a transformation of the raw EMG information able to represent the movement as a set of features shared among several people. The original signal has to be cleaned from noise and artifacts, since these components could spoil the outcome. Our final objective is to extract the peculiar characteristic of a certain motion by observing many different examples performed by different subjects.

2.3 Experimental data

In this section, we will present the general characteristics of the two datasets we will use for testing purposes. Depending on the objectives of the study we will use one or the other, mainly working with upper limbs movements. It is worth to notice that the differences between the two datasets will lead to distinct preprocessing phases described in the following of this work.

2.3.1 NinaPro dataset

The NinaPro [28] [29] (Non Invasive Adaptive Prosthetics) database is a robust and wide dataset, made with data collected from many different subjects, which perform several hand and wrist movements. The database enables the comparison of classification and regression performances obtained using various techniques by providing sEMG, hand/arm kinematics, dynamics and clinical parameters. The database contains data obtained from 40 intact subjects (28 males, 12 females; 36 right handed, 4 left handed; age 29.9 $\pm$ 3.9). Each subject performed 6 repetitions of 50 different movements. Hand kinematics has been measured using a 22-sensor CyberGlove II (CyberGlove Systems LLC, www.cyberglovesystems.com) to provide joint-angle information at slightly less than 25 Hz. A 2-axis IS40 inclinometer with a range of 120° and a resolution of less than 0.15° (Fritz Kbler GmbH, www.kuebler.com) has been added to measure the wrist orientation with a frequency of 100 Hz. Muscular activity has been measured using Delsys double-differential sEMG electrodes sampling signals at a rate of 2 kHz with a baseline noise of less than 750 nV RMS. These electrodes integrate also a 3-axes accelerometer sampled at 148 Hz. Eight electrodes were equally spaced around the forearm at the height of the radio-humeral joint; two electrodes were placed on the main activity spots of the *flexor digitorum superficialis* and of the *extensor digitorum superficialis*, two electrodes were also placed on the main activity spots of the *biceps brachii* and of the *triceps brachii*. An accurate timestamp has been

associated to each data sample to properly synchronize the information collected. An example of the recording procedure is shown in Fig. 2.3.

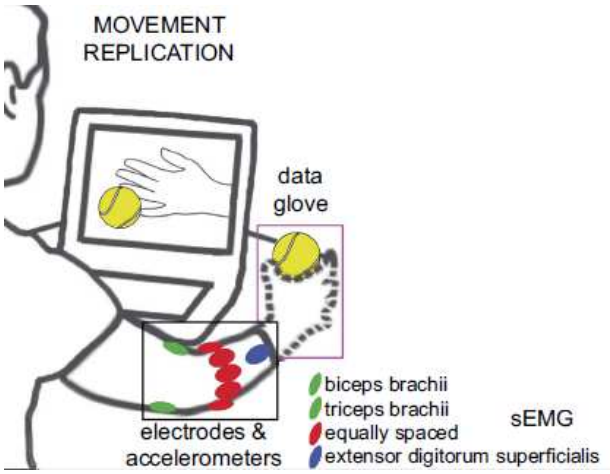


Figure 2.3: NinaPro recording procedure.

2.3.2 Myo dataset

The Myo dataset is based on data provided by a pair of Myo armbands[1]. It has been collected entirely during the period of this thesis in our laboratory, the Intelligent Autonomous Systems Laboratory (IAS-Lab) at the Department of Information Engineering (DEI) of the University of Padova. The Myo Armband is a quite diffuse low cost all-in-one sensor developed by Thalmic Labs, and it is shown in Fig. 2.4. Originally, its scope was to let users control technological devices by means of a set of hand motions. It should be considered as a black box for non expert users, which can use the sensor for high level tasks. Nevertheless, it has a great potential for developers and scientists interested in physiological signals, since it has 8 EMG sensors, combined with gyroscope, accelerometer and magnetometer to recognize gestures. The main reasons for selecting this hardware are:

- the high number of sensors with respect to other competitors;
- the easy access to ROS compatible drivers;
- the simple mechanism for adjusting the sensor on different subjects.

In fact, the sensors are uniformly distributed and arranged in an adjustable band. Ten sizing clips make the Myo expandable between 7.5 - 13 inches (19 - 34 cm)

[1] https://www.myo.com/

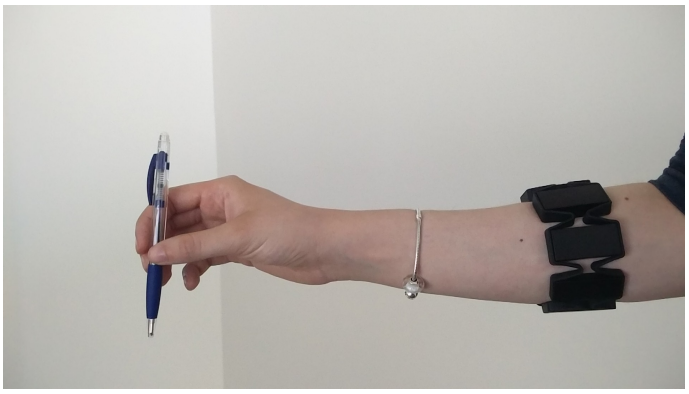


Figure 2.4: Myo armband.

forearm circumference, allowing its usage for a large amount users. Furthermore, its weight of only 93 grams and 0.45 inches of thickness make it easy to use daily without annoying the user. The EMG signals has the drawback of being very position-related, this means that small changes in the position of the sensors would cause a large variation in the recorded signal. For avoiding this problem, we introduced a simple setup procedure based on the constraints given by the arrangement of the sensors in a fixed position to reduce the possibility of signals misplacing, accelerating the setup time and improving the signal quality. The sensor communicates with many compatible operating systems (Windows, Mac, IOS and Android) by using the Bluetooth Smart Wireless Technology and there is a wide community of developers and a large support network constantly working on bug fixing. Starting from the packages compatible with the ROS middleware available in Myo community, we developed a specific driver for enabling a multi-device communication with a single PC.

Chapter 3

Signals Analysis Methods

The generic structure of the frameworks developed in this thesis is illustrated in Fig. 3.1. The overall procedure aims at estimating the control of a robotic device (output data) starting from human information (input data). Two main phases can be recognized, i.e. an offline and an online elaboration. During the offline phase, data are collected from many subjects while performing a certain task. In this phase, the information available for each trial should contain both input (e.g. EMG or accelerometer (ACC) signals) and output data (e.g. robot joint angles or kind of movement). Data undergo into a preprocessing phase in order to remove artifacts and noise from the signals. Finally, a probabilistic model is trained in order to represent the processed information with a limited set of parameters. The online phase considers data directly recorded from subjects, exploiting the model previously computed to estimate the corresponding robot motion. Again, data undergo into a preprocessing phase before entering regression or classification phases. All the concepts introduced so far will be accurately described in this chapter.

The complete procedure has been developed in C++ language. The communication with the different devices (i.e. sensors and robots) exploits Robot Operating System (ROS), a standard de facto in robotics.

3.1 Signal processing

EMG signals need to be elaborated in order to highlight the muscular activation during a task, and in particular to extract the common behaviours among different subjects. Selecting an extremely sophisticated machine learning technique does not guarantee high accuracy in data estimation. The aim of the preprocessing procedure is to obtain a set of significant features to estimate or in some cases predict robustly

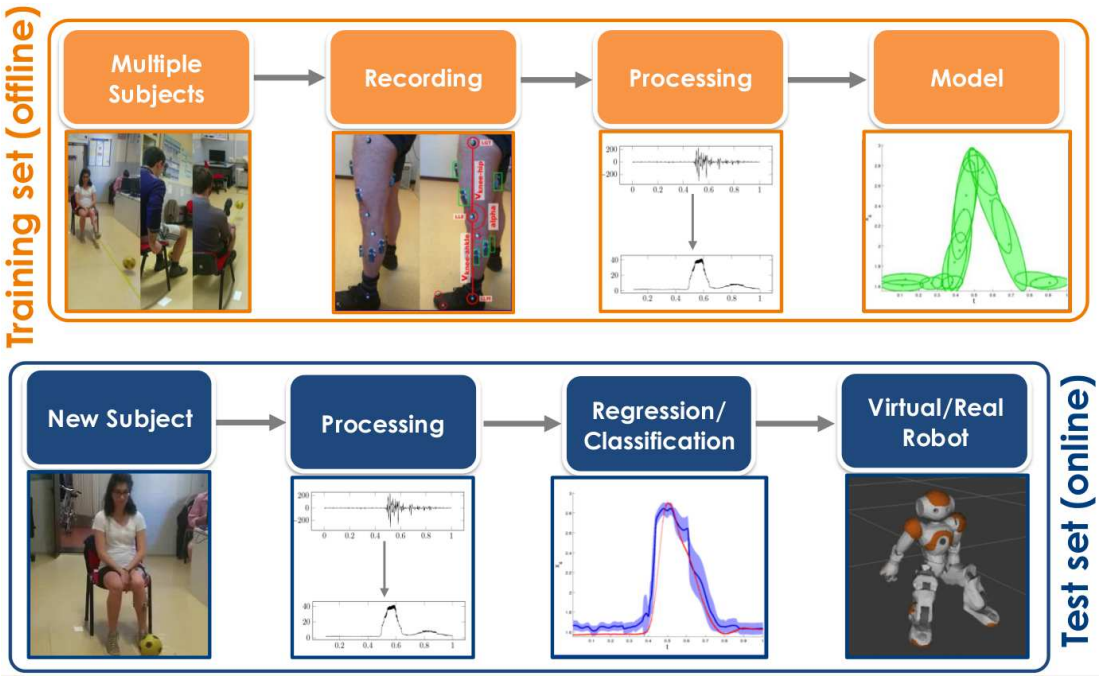


Figure 3.1: Sequence of activities necessary in order to achieve the robot motion starting from physiological signals.

and online the movement performed by a subject. In fact, the process of filtering the significant information from input signals can affect the actual success of the entire framework. The preprocessing phase is essential to obtain a well-balanced combination of similarity and variability within the signal. On one hand, if the considered signals have nothing in common, the final model would not work properly. On the other hand, a certain amount of variability should be integrated in the system, in order to build a general model which can work with new, unseen subjects. Moreover, we want the robotic devices to react almost immediately to human commands. Offline preprocessing methods are not suitable to direct control, consequently it would be almost impossible to use the resulting framework in real life applications. The signal analysis should be an online process, extracting useful information from the raw signal in real-time and with no dependency from the specific portion of the data in time. The online constraint limits the running time and the signal portion at the disposal along all the framework phases, and reduces the number of approaches for preprocessing. Indeed, some techniques could be useful even when applied offline only to training data to achieve a better regularization. For these reasons, the preprocessing phase is the combination of two main techniques:

- Wavelet Transform.

- Smoothing and normalization.

An additional and powerful offline tool for improving data regularization is Dynamic Time Warping, which works on the signal extension, thus obtaining an adaptation to changes in the velocity of the performing task.

As a further consideration, the peculiar characteristics of the two datasets considered for testing lead to slightly differences in the preprocessing operations. Nevertheless, the major steps are the same independently from the dataset.

All the techniques introduced so far are deeply described in the following subsections.

3.1.1 Wavelet Transform

Variations along time are extremely common in EMG signals. Extracting a smooth behaviour from the raw signal is a first important step to control a robotic device. The usual approach is to compute some interesting features from EMG signals, in particular a common solution regards the application of transformation methods. Some widespread techniques applied to EMG signals are Fourier Transform (FT) [30], Integral Absolute Value (IAV), variance and zero crossing [31], Mean Average Value (MAV) [32], Rooted Mean Square (RMS), Mean Power Frequency (MPF) [33], or as proposed in [34] full wave rectification, filtering and normalization.

The major drawback of these transformation methods, especially for the fast and short-term Fourier Transform, is that they assume signal to be stationary [35]. Since EMG signals are nonstationary, we investigated other solutions, and in particular a promising tool, namely the Wavelet Transform. Daubechies adopted WT to analyze time series that contain non-stationary power at many different frequencies [36]. Laterza [37] showed that WT is a valuable alternative to represent time frequency signals, since WT allows a linear multiresolution representation of the original signal without crossterms affections. Furthermore, Guglielminotti [38] proved the existence of good matching properties between an EMG signal and its WT shapes, and these results have been confirmed by more recent works, like [39] and [40]. Among the different mother wavelet, Chowdhury [41] emphasized the good results obtained when adopting Daubechies functions by analyzing various studies on Wavelet Transform. Wavelet information can be synthesized by using statistical features as stated by Subasi [42], [43], thus obtaining linearity, multiresolution representation and cross terms resolution.

Definition

A wavelet is a term that indicates a wave-like oscillation with a zero average amplitude. Generally, wavelets are arranged in order to make them have specific properties that make them useful for specific signal processing. In particular, starting from a well known mother wavelet, the signal is transformed and convolutionary combined with portions of a known signal in order to extract information from the unknown signal. WT [44], [45], together with Fourier Transform (FT), are commonly used for signal analysis and processing, particularly when we are interested in the analysis of frequency components. Nevertheless, WT preserves the temporal aspects of the signal without resolution limits in frequency, thus allowing an analysis in both time and frequency, in different amplitude windows. WT can be thought as an extension of FT able to work on a multi-scale basis (i.e. time and frequency), thus allowing the decomposition of a signal into several scales. WT and FT are similar also from a mathematical point of view, but instead of using a basis composed by sine and cosine, WT uses particular functions that satisfy certain mathematical rules. Wavelet Transform [44] decomposes the signal into several kernel functions called wavelets. A base wavelet, called mother wavelet ($\psi(t)$), is scaled and translated by a scaling function to generate the set of M wavelets composing the original signal while providing multi-resolution analysis. Each wavelet is represented by a coefficient (γ_m).

The input signals are represented as a linear combination of a particular set of functions (Wavelet Transform) as illustrated in Eq. 3.1, obtained by shifting and dilating one single function called a *mother wavelet* ($\psi(t)$) by means of a scaling function ($\phi(t)$).

$$f[n] = \frac{1}{\sqrt{M}} \sum_{k} W_{\phi}[j_0,k]\phi_{j_0,k}[n] + \frac{1}{\sqrt{M}} \sum_{j=j_0}^{\infty} \sum_{k} W_{\psi}[j,k]\psi_{j,k}[n] \tag{3.1}$$

The decomposition of the signal leads to a set of coefficients called wavelet coefficients. Therefore the signal can be reconstructed as a linear combination of the wavelet functions weighted by an adequate number of wavelet coefficients. WT are a very powerful tool, yet flexible and general, there is a wide variety of wavelet functions that can be suitable for different applications.

There are two types of Wavelet transform methods:

- Discrete Wavelet Transform.
- Continuous Wavelet Transform.

In this thesis, we focused in Discrete Wavelet Transform, the most commonly used

in real-time engineering applications.

Mother Wavelet selection

A previous study [21] showed that choosing a good mother wavelet is particularly important, since every mother wavelet yields to different results even when applied to the same signal. Chowdhury [41] successfully used Daubechies family (db) function to analyze sEMG signals. His work focused on the processing of sEMG and its use in different applications. The signal has been processed by means of a set of specific functions (db2, db4, db6, db44 and db45) at decomposition level 4 in order to maintain the maximum amount of information. In a similar study, Phinyomark [40] was able to find good results by using db7 as mother wavelet.

A key contribution of our work is the online processing of the EMG data. In order to achieve this result, at each instant t, only a small window of the whole signal has to be considered. We used Wavelet Transform to extract significant information from the raw signal.

$$\begin{aligned} \phi_{j,k} = f[n] = \frac{1}{\sqrt{M}} \sum_{k} W_{\phi}[j_0,k]\phi_{j_0,k}[n] + \\ \frac{1}{\sqrt{M}} \sum_{j=j_0}^{\infty} \sum_{k} W_{\psi}[j,k]\psi_{j,k}[n] \end{aligned} \tag{3.2}$$

By looking at the good performances obtained for subject-specific cases in both accuracy and time [21], we selected the db2 mother wavelet from the Daubechies family for representing the EMGs in input. Synthesizing the coefficients provided by Wavelet Transform to a single value representing the wavelet decomposition allows us to compare different signals. The synthesis function should guarantee a certain level of smoothness in order to avoid sudden changes from one instant to another and being fast enough to be computed online. Mean Average Value (MAV) (Eq. 3.3) represents a good candidate given the results achieved in [21]. Fig. 3.2 shows the comparison between the raw EMG signals and after the application of WT.

$$\text{MAV} = \frac{1}{M} \sum_{m=1}^{M} |\gamma_m| \tag{3.3}$$

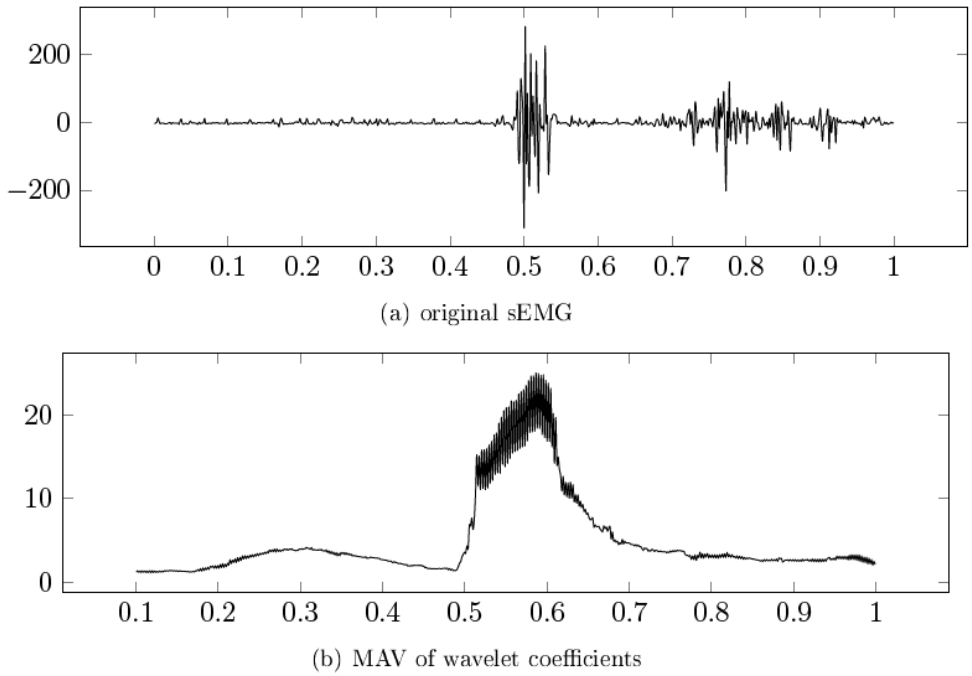


Figure 3.2: a) Original sEMG signal, b) sEMG signal after the application of WT and MAV.

3.1.2 Smoothing and Normalization

Nevertheless, data coming from Wavelet Analysis are still very jagged and they are not good enough to be used for a subject-independent modelization, since the great variability of the signal results in poor model performances. The WT of the EMG channels have been smoothed and normalized in order to obtain better and more robust models.

The smoothing function is based on a moving average filter. At the instant t, the average of S data points available within the windows is computed in order to smooth the data. This process is equivalent to lowpass filtering, with the response of the smoothing given by Eq. 3.4

$$\gamma_S(t) = \frac{1}{S+1} \sum_{s=1}^{S} \gamma(t-s) \tag{3.4}$$

The smoothing function used smooths the data using the loess method. It performs a local regression using weighted linear least squares and a second degree polynomial model. Fig. 3.3 shows the result of the smoothing on the EMG signal to whom has already been applied the WT.

Differences in the amplitude and in the mean of the signal requires the application

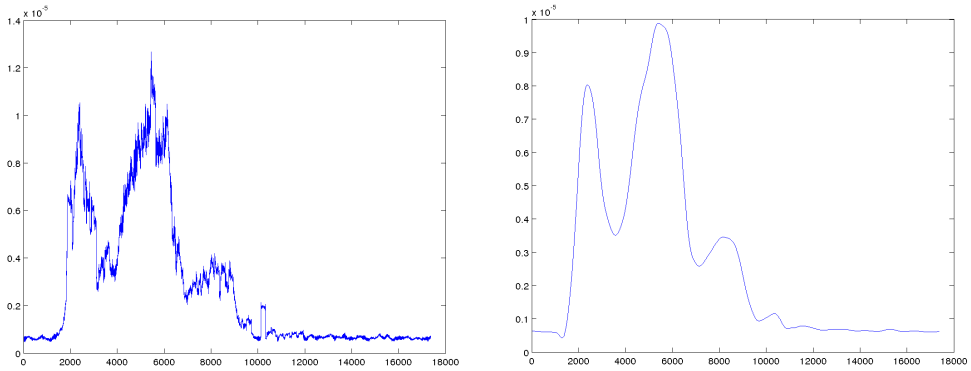


Figure 3.3: a) EMG signal before and after the application of smoothing.

of a normalization technique. The normalization process ensures the regularization of the signal in order to obtain a more robust model. The normalization has been implemented in two different manners for training and testing phases. Since the training phase is executed offline, the normalization for the data involved has been accomplished by using the relative maximum within the specific trial involved in the process. Instead, during the online testing procedure the information about the relative maximum is not available, and we needed to use a different method to be able to compute the normalization online. For obtaining this result, the mean of the relative maximums collected during the entire training set has been used as normalization factor.

3.1.3 Time Warping Techniques

Shift in time domain is a common occurrence in data analysis. It is almost impossible that two repetitions of the same task have the same duration. If limited variations could be ignored since the model would be able to abstract easily, this is not true if the changes are relevant. As a matter of fact, there must be a common behaviour among all the repetitions in order to model properly the input signal. If the differences among repetitions are relevant, data needs to be brought to a form where the observed variables of the samples under analysis express similar attributes. Warping is one of the numerous preprocessing techniques that have been proposed to correct shifts.

Two different warping algorithms have received much attention in recent years:

- DTW, was initially applied for aligning words pronounced by different speakers for speech to text recognition purposes.

- Correlation Optimized Warping (COW), was proposed more recently in order to correct chromatograms for shifts in the time axis prior to multivariate model-

ing [46].

Dynamic Time Warping

Dynamic Time Warping nonlinearly warps two trajectories in order to align similar movements by minimizing the distance between them. During the last years this algorithm has found application in many fields, like batch process monitoring, gene expression studies, temporal sequences of video, audio, and graphics data, automatic speech recognition and signature recognition. In general, any data which can be turned into a linear sequence can be analyzed with DTW. In time series analysis, DTW is an algorithm that compute the similarity among two sequences which may vary in time or speed. For instance, it could be used to detect similarities in walking patterns, even if one person walks faster than the other, or if there are accelerations and decelerations during the observation (Fig. 3.4, Fig. 3.5).

In general, DTW is a method that calculates an optimal match between two given sequences. In order to reach the best pairing. the sequences are ”warped” non-linearly in the time dimension. The signals distortion is computed in order to determine a measure of their similarity independently from non-linear variations in the time dimension.

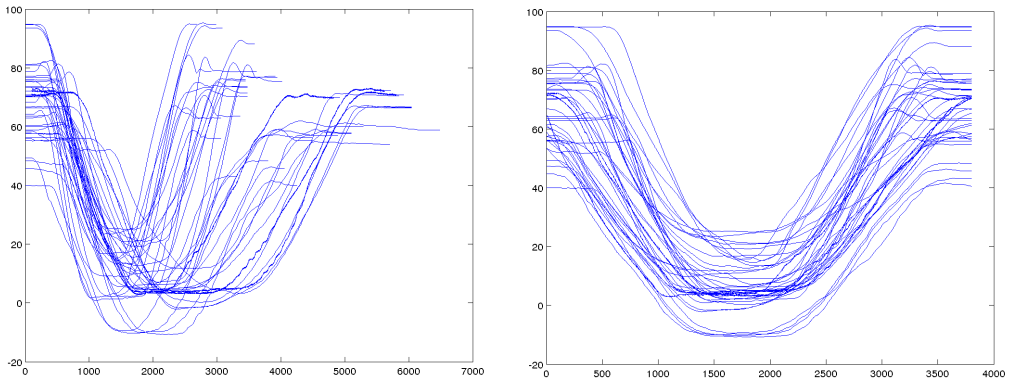


Figure 3.4: Signals before and after the application of the DTW algorithm.

Computing the DTW takes $O(N^2)$ and this particular algorithm cannot be used in real time, because it requires to have all the data at the beginning of the computation. This limitation makes it impossible to be used in online applications, but we have applied it successfully in preliminary studies and in the training phase.

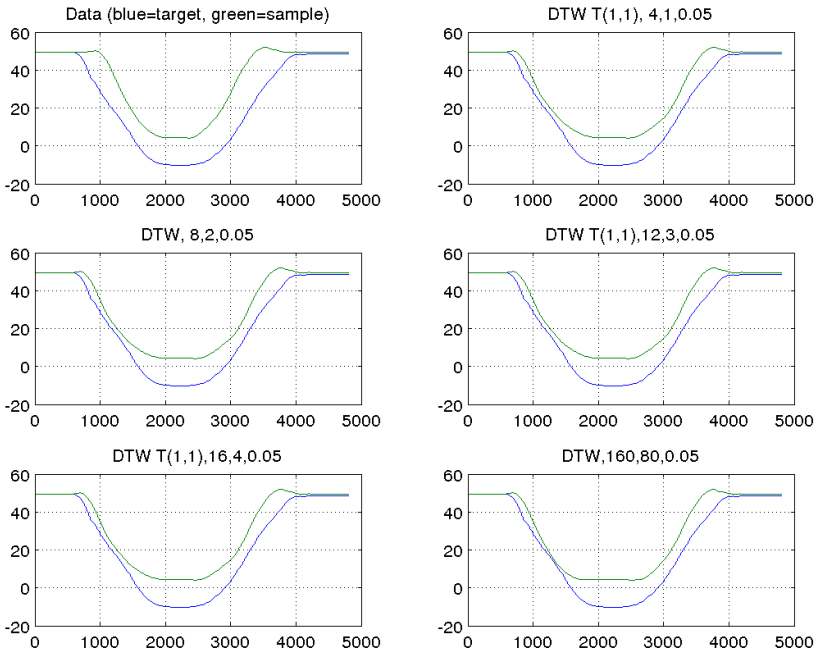


Figure 3.5: DTW on the angle of the knee with different parameters.

3.2 Gaussian Mixture Model

Gaussian Mixture Model is a parametric probabilistic model that assumes all data points are generated from a mixture of a finite number of Gaussian distributions. These distributions completely characterize the model, therefore it is composed by a weighted sum of Gaussian components. In particular, three parameters for each Gaussian component are sufficient to represent the whole information: *mean*, *covariance* and *weight*.

These parameters are estimated from training data using the iterative EM [47] algorithm. EM is a statistical algorithm that iteratively finds locally maximum likelihood parameters of a probabilistic model when equations can not be solved directly. The locally maximum likelihood is obtained repeating cyclically two phases:

- Expectation (E) step creates a function for the expectation of the log-likelihood evaluated using the following estimate of the components parameters:

$$p_{k,j}(t+1) = \frac{\pi_k(t)\mathcal{N}(\zeta_j;\mu_k(t),\Sigma_k(t))}{\sum_{i=1}^{K}\pi_i(t)\mathcal{N}(\zeta_j;\mu_i(t),\Sigma_i(t))} \tag{3.5}$$

- Maximization (M) computes parameters maximizing the expected log-likelihood

found during last E step:

$$\pi_k(t+1) = \frac{1}{N}\sum_{i=1}^{N} p_{k,j}(t+1) \tag{3.6}$$

$$\mu_k(t+1) = \frac{\sum_{i=1}^{N} p_{k,j}(t+1)\xi_j}{\sum_{i=1}^{N} p_{k,j}(t+1)}$$

$$\Sigma_k(t+1) = \frac{\sum_{i=1}^{N} p_{k,j}(t+1)(\zeta_j - \mu_k(t+1))(\zeta_j - \mu_k(t+1))^\top}{\sum_{i=1}^{N} p_{k,j}(t+1)}$$

The result is a continuously improving adaptation to the best representation of the input data as it is shown in Fig. 3.6.

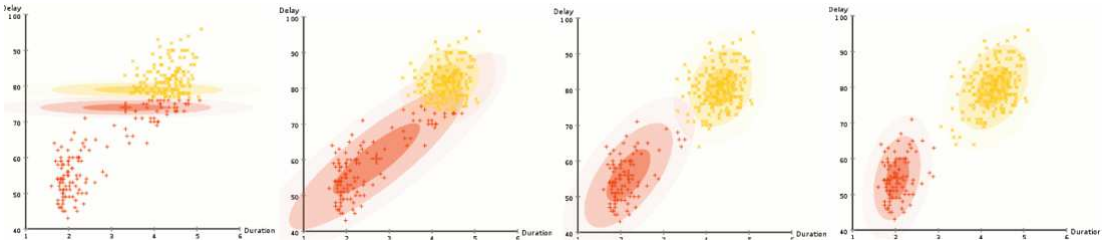


Figure 3.6: Example of the EM algorithm. The red and yellow ovals show how the algorithm adapt the parameters to fit the data (the red and yellow crosses)

The EM loop stops when the increment of the log-likelihood $\mathscr{L} = \sum_{j=1}^{N} \log\left(p\left(\zeta_j|\theta\right)\right)$ at each iteration becomes smaller than a defined threshold ε, i.e. $\frac{\mathscr{L}(t+1)}{\mathscr{L}(t)} < \varepsilon$.

A possible limitation in the learning process is the fact that EM requires a priori specification of the number of Gaussian components K. Selecting the correct K is a crucial task. On one hand, an overestimation of this parameter might lead to over-fitting and, consequently, to a poor generalization. On the other hand, an underestimation will result to poor predicting performances.

Several entropy based model selection techniques has been proposed in literature to estimate this parameter (e.g. Bayesian Information Criterion (BIC) [48], Akaike Information Criterion (AIC) [49], Minimum Description Length (MDL) [50], and Minimum Message Length (MML) [51]). In this thesis, we choose a standard approach based on BIC (Eq. 3.7).

$$S_{BIC} = -2\mathscr{L} + n_p \log N \tag{3.7}$$

with:

- $\mathscr{L} = \sum_{j=1}^{N} \log\left(p\left(\zeta_j|\theta\right)\right)$, the log-likelihood for the considered model θ.

- $n_p = (K-1) + K(D + \frac{1}{2}D(D+1))$, the number of free parameters required for a mixture of K components with full covariance matrix.

In our work, the physiological information from the subjects has been used to train a GMM. The data used for the learning process are composed by both input (e.g. the physiological signals) and output (e.g. joints angles). Instead, during the online testing phase, only the physiological signals are known, and the output is estimated. Following the example, the probabilistic algorithm is able to estimate the bending angle α of different joints during the movement using a regression technique, or alternatively to choose between a set of possible movements performed by the subject by following the classification approach.

Considering:

- H, number of subjects involved in the study.
- n, number of trials per subject used to train the system.
- T, number of repetitions of each trial.
- $N = nT(H-1)$, total number of data samples. The number of subjects is decreased by one, since the model is trained on $H-1$ subjects and then tested on the excluded subject h.

Then a single data $\zeta_j, 1 \leq j \leq N$ in input at the framework can be written like in Eq. 3.8.

$$\begin{aligned} \zeta_j &= \{\xi(t), \alpha(t)\} \in \mathbb{R}^D \\ \xi(t) &= \{\xi_c(t)\}_{c=1}^{C}, \\ \alpha(t) &= \{\alpha_g(t)\}_{g=1}^{G}. \end{aligned} \tag{3.8}$$

where:

- $C = |\xi|$, number of physiological signals considered.
- $\xi(t) \in \mathbb{R}^C$, the set of values assumed from all the considered physiological signals at the time instant t, with $\xi_c(t) \in \mathbb{R}$, the value assumed from the c^{th} physiological signal at the time instant t.
- $G = |\alpha|$, number of joint bending angles for regression, $G = 1$ for classification.
- $\alpha(t) \in \mathbb{R}^G$, for regression the set of values assumed from the considered joint bending angles at the time instant t, with $\alpha_g(t) \in \mathbb{R}$, the value assumed from g^{th}

joint bending angle at the time instant t. For classification $\alpha(t) \in \mathbb{N}$ the index of the performed movement.

- $D = C + G$, the dimensionality of the problem.

It is worth to notice that there is no direct use of the time instant t in the data provided to the model.

The final resulting probability density function is computed as:

$$p\left(\zeta_j\right) = \sum_{k=1}^{K} \pi_k \mathcal{N}\left(\zeta_j; \mu_k, \Sigma_k\right) \tag{3.9}$$

where

- π_k priors probabilities.
- $\mathcal{N}\left(\zeta_j; \mu_k, \Sigma_k\right)$ Gaussian distribution.
- μ_k mean vector of the k-th distribution.
- Σ_k covariance matrix of the k-th distribution.
- K number of Gaussian components.

Fig. 3.7 shows an example of the input signals modelization through GMM.

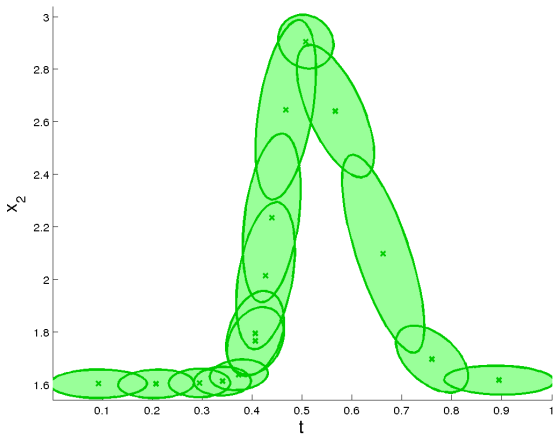


Figure 3.7: Example GMM, the green ovals represents the Gaussian components.

3.3 Incremental Gaussian Mixture Model

The construction of the probabilistic model is a time consuming task. Furthermore, often we are not interested in the building of a new model from scratch. On the contrary, in some cases we want to update the model, adding the information from new subjects, in order to make the model fit on the subject performing the task, without losing the knowledge, the robustness and the generality acquired from previous subjects. For these reasons, we implemented an incremental version of GMM, namely IGMM, able to update the model as new demonstrations are received from the subject. We implemented and tested the *Generative method* described in [52].

The first step consists of building a GMM with the classic EM algorithm as described in Sec. 3.2. When new data are available ξ_i, they undergo the following passages:

1. Synthetic data are stochastically generated with by performing a regression on the current GMM. The generated data are a compact representation of the previous data distribution.

2. A new GMM is computed on the whole set composed by new data ξ_i and the stochastically generated ones.

3. A learning rate $\alpha \in [0,1]$ is introduced to modulate the contribution from the new data and the stochastically generated ones. $\alpha = \frac{\tilde{N}}{\tilde{N}+N}$, with $\tilde{N}$ number of new datapoints available, and N number of datapoints from previous demonstrations.

4. Given $n = n_1 + n_2$ number of samples for the iterative learning procedure, with $n_1 \in \mathbb{N}$ number of trials from the new observations, and $n_2 \in \mathbb{N}$ number of trials generated from the previous model. The new training set is then defined by:
   $\xi_{i,j} = \tilde{\xi}_j$, if $1 < i \leqslant n_i$
   $\xi_{i,j} = N(\hat{\mu}_j, \hat{\Sigma}_j)$, if $n_i < i \leqslant n$
   $\forall j \in \{1,...,T\}$, with T number of timestamps, with $n_1 = [n\alpha]$ ([.] nearest integer function).

5. The training set of n trials is used to refine the model by updating the current set of parameters (π_k, μ_k, Σ_k) by using the EM algorithm.

3.4 Gaussian Mixture Classification

This section is dedicated to an accurate analysis of the online classification phase. A classification technique allows us to predict which kind of movement the subject is performing starting from a previously trained model and the EMG signals collected from the subject's muscles. We developed three different approaches, all based on Mixtures of Gaussian Components. One of the proposed techniques is a simple classification providing instantaneously the estimated class for each sample in input at the process. The other two methods follow an accumulation approach, where the classification of a certain sample ξ_k depends also on the classifications of previous samples ξ_1^{k-1}, with $1 \leq k \leq S$, being S the number of samples for a certain movement of a specific subject. Taking into account the previous classification outcomes leads to a more robust estimation less prone to misclassification problems.

3.4.1 Instantaneous Classifier

Considering the data of a certain trial performed by a specific subject, the samples belonging to a trial can be denoted as $\xi_0^t = \{\xi_0, \xi_1, ..., \xi_t\}$. Following this approach, at each instant, the classification depends only from the last sample ξ_t received in input to the classifier. In order to estimate the correct class, we compute the Probability Density Function (PDF) of each new sample as

$$\rho_{i,k} = PDF(\xi_k|\gamma_i) \tag{3.10}$$

where $1 \leq k \leq S$ is the index of the considered sample. γ_i with $1 \leq i \leq M$ indicating the index of all the possible classes of movements, being M the number of considered movements. The selected class is the one with the highest PDF, the formal equation is reported in Eq. 3.11.

$$\phi k = i : max\left\{\rho_{i,k}, 1 \leq i \leq M\right\} \tag{3.11}$$

3.4.2 Normalized Accumulation Classifier

The first accumulation approach extends the instantaneous classifier. The result of Eq. 3.10 is normalized $\forall i$ with $1 \leq k \leq S$, obtaining $\widetilde{\rho}_i$. Then, we compute the mean

between all the already considered normalized samples:

$$\mu_{i,k} = \frac{\sum_{j=1}^{k} \widetilde{\rho}_{i,j}}{k}, \forall i, 1 \leq i \leq M \tag{3.12}$$

The mean all the PDF computed from the collected samples represents the contribution of previous classifications until reaching the last sample.

The most probable class is chosen like in the previous technique:

$$\phi_k = i : max\left\{\mu_{i,k}, \forall i, 1 \leq i \leq M\right\} \tag{3.13}$$

3.4.3 Bayesian Accumulation Classifier

Up to now, we tried to compute the probability of each sample ξ_k of belonging to a certain class α. An alternative approach can be obtained by calculating the probability of the current class ϕ_k of being the class α, given that we received in input the sequence of data ξ_0^k. Consequently, we are interested in computing $p(\phi_k|\xi_0^k)$, we can apply Bayes' rule, obtaining:

$$p(\phi_k = \alpha|\xi_0^k) = \frac{p(\xi_k|\xi_0^{k-1}, \phi_k = \alpha)(\phi_k = \alpha|\xi_0^{k-1})}{p(\xi_k|\xi_0^{k-1})} \tag{3.14}$$

We can make two assumptions:

- *First order Markov assumption:* $p(\xi_k|\xi_0^{k-1}, \phi_k = \alpha) = p(\xi_k|\phi_k = \alpha)$, since the measurement of ξ_k is conditionally independent from previous measurements. In other words, next samples are not strictly determined from previous data.

- *Smoothness of the posterior assumption:* $p(\phi_k = \alpha|\xi_0^{k-1}) \approx p(\phi_{k-1} = \alpha|\xi_0^{k-1})$.

Thanks to the previous assumptions, we can write Eq. 3.14 as:

$$p(\phi_k = \alpha|\xi_0^k) = \frac{p(\xi_k|\phi_k = \alpha)(\phi_{k-1} = \alpha|\xi_0^{k-1})}{p(\xi_k|\xi_0^{k-1})} \tag{3.15}$$

Then, we have:

$$p(\phi_{k-1} = \alpha|\xi_0^{k-1}) = \frac{p(\xi_{k-1}|\phi_{k-1} = \alpha)(\phi_{k-2} = \alpha|\xi_0^{k-2})}{p(\xi_{k-1}|\xi_0^{k-2})} \tag{3.16}$$

Substituting the equations and putting everything together we have:

$$p(\phi_k = \alpha|\xi_0^k) = \frac{p(\xi_k|\phi_k = \alpha)p(\xi_{k-1}|\phi_{k-1} = \alpha)...p(\xi_1|\phi_1 = \alpha)}{p(\xi_k|\xi_0^{k-1})p(\xi_{k-1}|\xi_0^{k-2})...p(\xi_1|\xi_0^1)} p(\phi_0 = \alpha|\xi_0) \tag{3.17}$$

We can expound $p(\phi_0 = \alpha|\xi_0)$ by applying Bayes' equation.

$$p(\phi_0 = \alpha|\xi_0) = \frac{p(\xi_0|\phi_0 = \alpha)p(\phi_0 = \alpha)}{\sum_{m=1}^{M} p(\xi_0|\phi = m)p(\phi = m)} = \lambda_0 \tag{3.18}$$

where λ_0 is the initial probability that the class being α given the first sample.

Therefore, the probabilities for the following samples can be computed as a function of the previous one:

$$\begin{aligned} p(\phi_1 = \alpha|\xi_0^1) &= \frac{p(\xi_1|\phi_1 = \alpha)}{p(\xi_1)}\lambda_0 = \lambda_1, \\ p(\phi_2 = \alpha|\xi_0^2) &= \frac{p(\xi_2|\phi_2 = \alpha)}{p(\xi_2)}\lambda_1 = \lambda_2, \\ &\dots \\ p(\phi_n = \alpha|\xi_0^n) &= \frac{p(\xi_n|\phi_n = \alpha)}{p(\xi_n)}\lambda_{n-1} = \lambda_n \end{aligned} \tag{3.19}$$

Eq. 3.19 can be easily computed by discretizing the input signal, i.e. by assigning each value to a class with a clustering solution. Finally, after the training phase we will be able to compute the probability of each class, thus the class belonging probability for each sample.

3.5 Gaussian Mixture Regression

Beside the discrete classification among a set of possible movements, it is also interesting to continuously estimating the joint bending angles by considering solely

physiological signals from the subjects. The angles can be estimated continuously by using a regression technique. The Gaussian Mixture Regression (GMR) provides a smooth generalized version of the signal starting from the GMM. GMR estimates the joints angles $\hat{\alpha}$ and their covariance from the EMG ξ (and eventually accelerometers φ) signals known a priori, respectively using Eq. 3.20 and Eq. 3.21.

$$\hat{\alpha} = E[\alpha|\xi,\varphi] = \sum_{k=1}^{K} \beta_k \hat{\alpha}_k \tag{3.20}$$

$$\hat{\Sigma}_s = Cov[\alpha|\xi,\varphi] = \sum_{k=1}^{K} {\beta_k}^2 \hat{\Sigma}_{\alpha,k} \tag{3.21}$$

where:

- $\beta_k = \frac{\pi_k \mathcal{N}(\xi,\varphi|\mu_{p,k},\Sigma_{p,k})}{\sum_{j=1}^{K} \mathcal{N}(\xi,\varphi|\mu_{p,j},\Sigma_{p,j})}$, the weight of the k^{th} Gaussian component through the mixture.

- $\hat{\alpha}_k = E[\alpha_k|\xi,\varphi] = \mu_{\alpha,k} + \Sigma_{\alpha p,k}\Sigma_{p,k}^{-1}\{\{\xi,\varphi\} - \mu_{p,k}\}$, the conditional expectation of α_k given $\{\xi,\varphi\}$.

- $\hat{\Sigma}_{\alpha,k} = Cov[\alpha_k|\xi,\varphi] = \Sigma_{\alpha,k} + \Sigma_{\alpha p,k}\left(\Sigma_{p,k}\right)^{-1}\Sigma_{p\alpha,k}$, the conditional covariance of α_k given $\{\xi,\varphi\}$.

Assuming that the parameters (π_k, μ_k, Σ_k) defining the k^{th} Gaussian component are decomposed as follows:

$$\mu_k = \left\{\mu_{p,k}\,\mu_{\alpha,k}\right\} \qquad \Sigma_k = \begin{bmatrix} \Sigma_{p,k} & \Sigma_{p\alpha,k} \\ \Sigma_{\alpha p,k} & \Sigma_{\alpha,k} \end{bmatrix} \tag{3.22}$$

in which the mean and the covariance of the known a priori information $p = \{\xi,\varphi\}$ have been represented respectively with μ_p and Σ_p. Thus, the model is completely defined by the Gaussian components composed solely by weights, means and covariances obtained by means of the EM algorithm. Subsequently, the information composing the model allows us to calculate a generalized motion $\hat{\zeta} = \{\xi,\varphi,\hat{\alpha}\}$ starting from physiological data provided by the sensors (Fig. 3.8).

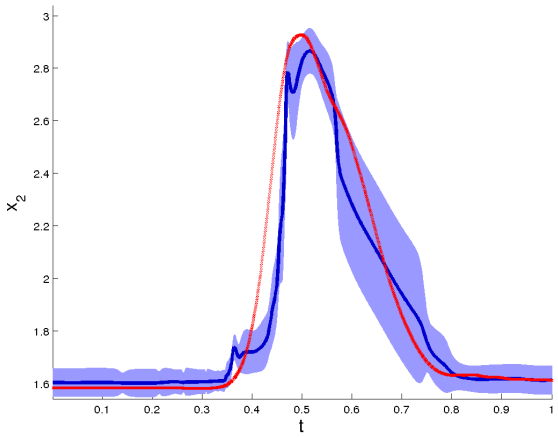


Figure 3.8: Signal estimation with GMR, the estimated movement is represented by the red line.

3.6 System effectiveness

3.6.1 Classification

For the classification, a measure for the quality of the prediction is obtained by computing the ratio between the number of correct predictions and the total number of examples. The obtained result is a number n, with $0 \leq n \leq 1$, where 0 indicates no correct prediction, while 1 means that all the predicted movements were correct. In general, the closer the accuracy to 1, the better the classification.

Often the accuracy is represented in percentage, with the results varying between 0% and 100%.

3.6.2 Regression

A common measure widely used for evaluating the goodness of the predicted measure [2] [27] is the correlation coefficient $\rho_{\alpha,\hat{\alpha}}$. This value is calculated between the predicted output $\hat{\alpha}$ and the real one α (Eq. 3.23), and it gives a measure of the model performances by means of the statistical relationships between different signals and different subjects. In particular, the correlation coefficient is a measure of the degree of linear dependence between two variables, and it is based on the covariance ($Cov(\alpha,\hat{\alpha})$) and the standard deviations (σ_α and $\sigma_{\hat{\alpha}}$) of the considered variables. The resulting formula is reported in Eq. 3.23.

$$\rho_{\alpha,\hat{\alpha}} = \frac{Cov(\alpha,\hat{\alpha})}{\sigma_\alpha \sigma_{\hat{\alpha}}} \tag{3.23}$$

The correlation coefficient can assume all the values between 1 and -1, where 1 is total positive correlation and indicates a perfect direct linear relationship (correlation), 0 is no correlation, and -1 is total negative correlation, or a perfect decreasing linear relationship (anticorrelation). The closer the coefficient is to either -1 or 1, the stronger is the correlation between the variables, while the closer it is to zero, the weaker is the correlation. When the correlation reaches zero the variables are independent.

Another common measure of the effectiveness of GMM-based systems is NMSE. This function measures the goodness of fit between test and reference data. NMSE (Eq. 3.24) costs vary between -∞ (bad fit) to 1 (perfect fit).

$$\text{NMSE}(t) = 1 - \left\| \frac{\hat{\alpha}(t) - \alpha(t)}{\hat{\alpha}(t) - \mu_t(\alpha)} \right\|^2 \tag{3.24}$$

where t is the temporal instant from the beginning of the trial; $\hat{\alpha}(t)$ is the estimated output at the instant t; $\alpha(t)$ is the groundtruth at instant t; $\mu_t(\alpha)$ is the mean along the time of considered quantity.

Chapter 4

Movements Prediction: Experiments and Results

4.1 Preliminary study

A preliminary study involved a limited number of subjects. We initially focused on a small dataset to test the feasibility of the subject-independent framework. In particular, we considered a very simple movement. Three healthy subjects (S1-S3; age 30±4; one female) were asked to naturally kick a ball from a sitting position Fig. 4.1. EMG signals were acquired with an active 8-channel wireless EMG system at 1000 Hz to cover the principal muscular groups active during the kick task, namely *Rectus femoris* (Ch1), *Vastus lateralis* (Ch2), *Vastus medialis* (Ch3), *Tibialis anterior* (Ch4), *Gastrocnemius lateralis* (Ch5), *Gastrocnemius medialis* (Ch6), *Biceps femoris caput longus* (Ch7), *Peroneus longus* (Ch8). Synchronously, six infrared digital cameras recorded at 60 Hz the kinematic of the knee-joint angle from the position of 6 markers on the subjects leg. Each person repeated the movement about 60 times. EMG data has been processed by means of signal rectification and smoothing in order to highlight the muscular activation during the kick tasks. The preprocessing has been deeply described in Chap. 3.

The information extracted from EMG has been used as input of a GMM to estimate its correlation with the knee bending angle α. GMM have been trained with data from couple of subjects (S1+S2, S1+S3, S2+S3). For every couple, different sizes of training set have been considered (10, 30, 60, 120), half from the first subject and half from the second one. For the testing phase, we used 10 trials coming from the remaining subject in order to verify the generality of the model. The described procedure has been applied to all the collected EMG channels. Fig. 4.2 represents the Goodness of

Figure 4.1: A person naturally kicking a ball from a sitting position.

Fit (GoF) computed for the couple S1+S2 varying the EMG channel and the number of data used for training. Results showed a good estimation even with few input data for almost all the channels. Increasing the cardinality of the training set, we generally obtained similar or better performance with some exceptions. In fact, it was worth to use a number of trials varying between 30 and 60 in order to obtain more stable results during the testing phase, while using a greater number of trials raised significantly the time for model training without any apparent benefit in efficiency.

Tab. 4.1 compare the results between subject-specific and subject-independent models regarding the most informative EMG channels.

Subject	Specific	Independent
S1	0.9238	0.8335
S2	0.9700	0.8543
S3	0.9570	0.9476

Table 4.1: NMSE values comparing results from subject-specific and subject-independent models.

Looking at the whole set of results, they were generally quite similar, except for specific channels in some rare cases. As regards the best EMG channels, generally

Rectus femoris, *Vastus lateralis*, *Biceps femoris caput longus* and *Vastus medialis* were the muscles bringing most information. In fact, the NMSE for the EMG channels related to these muscles resulted among the most informative even in subject-specific models for all the three subjects. Moreover, the cited muscles are the principal actors of the considered movement from a biological point of view. Coherently, tests showed that the best trade off in terms of both stability and efficiency has been obtained using three out four of the already cited EMG channels (NMSE = 0.9257 for Ch1+Ch2+Ch3 and NMSE = 0.9409 for for Ch1+Ch2+Ch7).

In this dataset, the high number of repetitions of the same movement (about 60) performed by three subjects gave us the possibility to study the performances of the model adaptation to a novel person. The selected channels (1, 2 and 8) recorded the activity of the muscles *Rectus femoris*, *Vastus lateralis* and *Peroneus longus*. A leave-one-out approach has been adopted by building the model on 30 trials coming from 2 subjects and tested on the remaining one. We obtained 3 models on the movement trained by using a total of 60 repetitions. For analyzing the progressive model adaptation, the data of every subject has been divided in block of 10 repetitions. In fact, 10 repetitions bring a valuable amount of information to the system, good enough to add a substantial contribution to the previous model. In the first part of the analysis, the different blocks of movements of a certain subject are used to test the model built on the other two subjects. This is represented in the Fig. 4.2 by the red lines. Instead, the blue line shows a model tested on the same data of the previous case, but using an updated model with the previous testing data added to the training set. In this way, the model has been updated with the data from the third subject.

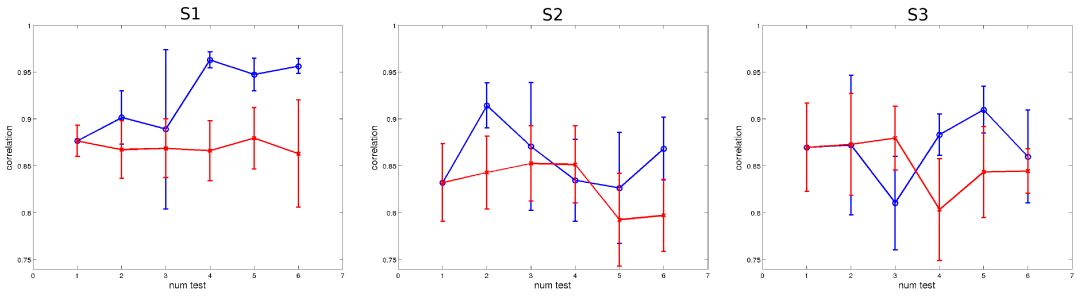


Figure 4.2: Correlation and Standard Deviation for the model of a kicking movement. The red line represent the results of the model built on two subjects and tested on different data from the third subject, without updating the model. The blue line represent the results of the model tested on the same data than the previous case, but updating the model with the data of the third person.

As the new data have been added to the model, the results improved, giving gen-

erally better results with respect to the first part of the analysis. The model generally decreases its performances when the testing data shows more variability. Anyway, in a long term perspective the adaptation characteristics of the proposed framework could be a great resource for rehabilitation purposes.

The time needed in each phase is presented in Tab. 4.2, together with the total amount of time needed for the signals elaboration and the joints angle prediction. The analysis of the computing time at each step has been conducted by using an Intel® 64-bit computer with *i3* quad core CPU of 2.13 GHz and 4 GB of RAM.

Step	Method	Time(μs)
WT	db2	581.3955
Feature extraction	MAV	0.8448
Regression	-	1774.6145
Angle remapping	-	7.4357
Total	db2 + MAV	2408.9559

Table 4.2: Analysis of the computational time (μs) needed from the framework at each phase. The times have been obtained with a

The generated motion has been successfully tested on a humanoid, namely a Aldebaran NAO Fig. 4.3. The EMG signals were sent via software and estimated angle was computed to actuate the robot through TCP/IP protocol. Our software is able to send pose messages to robot at 240 Hz, although in practice the rate has been reduced to satisfy NAO bound of 50 Hz.

The main purpose of testing the whole procedure on a humanoid robot (Aldebaran NAO) by remapping the human motion to the robotic platform was to verify the proper execution of the original movement. The robot execution properly mimic the movements of a person whose signals are not included in the model.

4.1.1 Conclusions

Tests have showed that our learning framework produced results comparable with subject-specific models. Besides the good results, the dataset was composed by only 3 people. Since the population is so poor, bad performances of a subject could ruin the whole model. Nevertheless, the good results we achieved with a first attempt of subject-independent framework proved us that we can expand the work by testing other dataset composed by more subjects and by applying the described method to a multiple joint motion. In fact, a bigger dataset could lead to even more general models. The

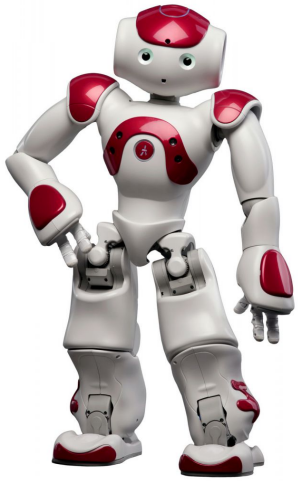


Figure 4.3: Humanoid robot on which has been tested the framework (Aldebaran NAO)

results showed good trends for both low and high variability in the task execution. This is a very important feature for a model to be used in rehabilitation contexts, since it can evolve with the patient without any external intervention.

4.2 Multi Joint Subject Independent Regression

The promising results of the preliminary work encourages us to expand our study to a greater number of subjects, in particular we considered data from the NinaPro dataset. In order to obtain comparable results between the considered datasets we applied a series of standardizing approaches. A similar number of samples ($\simeq 2000$) for trial has been considered by down-sampling the information available in the NinaPro database by a factor of ten. We looked at the most informative EMG channels by conducting a preparatory study. The study aimed to measure the engagement of each considered channel in the performed movement. By looking at the measure of engagement, we selected an equal number of channels for each motion. The db2 *mother wavelet* and MAV synthesis feature have been applied to the raw signal provided from every single channel. The resulting values have been associated to the corresponding bending angle along time. A model for each channel has been trained and GMR has been used to retrieve the estimated bending angle to be compared with a testing set. The three channels offering the best performances have been selected in order to obtain similar models. It is worth to notice that more channels could be considered for the NinaPro dataset, resulting in a more accurate estimation. Anyway, a subset of the significant

channels have been selected to simplify the comparison with the models produced with the Kicking Movement dataset, for which only three channels provided significant correlation when considering information from available subjects.

The high number of subjects involved in this dataset is ideal to analyze the robustness of the proposed framework. Two different movements of upper limbs have been selected between the different movements contained in the dataset. In particular, the wrist flexion (Movement 13) and the flexion of ring and little finger while extending others, namely "Three posture" (Movement 3), have been analyzed Fig. 4.4.

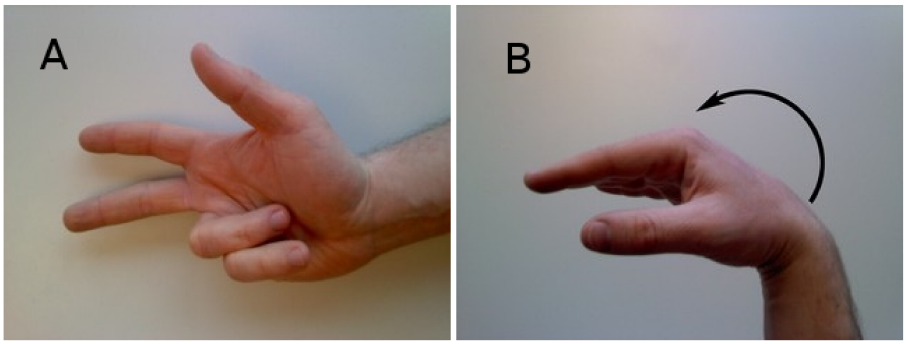


Figure 4.4: A) "Three posture" (Movement 3), B) Wrist flexion (Movement 13)

Both motions are quite simple and only some significant joints have been analyzed. The wrist has been considered for Movement 13, while for Movement 3, we selected the motion of Interphalangeal Joints as well as Metacarpophalangeal Joints of Thumb, Index, Middle, and Ring, using a total of 8 joints. The purpose of this choice is to focus on the model's independence from the subject more than on the complexity of the motion, while proving that our framework works on both single and multi-joints movements. The 3 most significant channels have been selected, according to the results coming from the preliminary study. The selected channels (3-5-7 for Movement 13, 2-3-7 for Movement 3) brought information from muscles around the forearm. A leave-one-out approach has been adopted by building the model on 35 right-handed subjects and testing it on the remaining one. We obtained 36 models for each movement to be tested on the 6 repetitions of the testing subject. For each repetition, we compared the estimated bending trajectory with the actual measured angles by computing the correlation coefficient. The mean and the standard deviation of the correlation coefficient have been estimated for each joint. Fig. 4.5 shows the correlation coefficient averaged between 6 trials of Movement 13 for all the subjects considered in the study. Fig. 4.6 regards Movement 3, the correlation coefficient has been averaged also on the joints, since many of them were involved in the movement.

The generated motion has been successfully tested on a simulated hand by Shadow

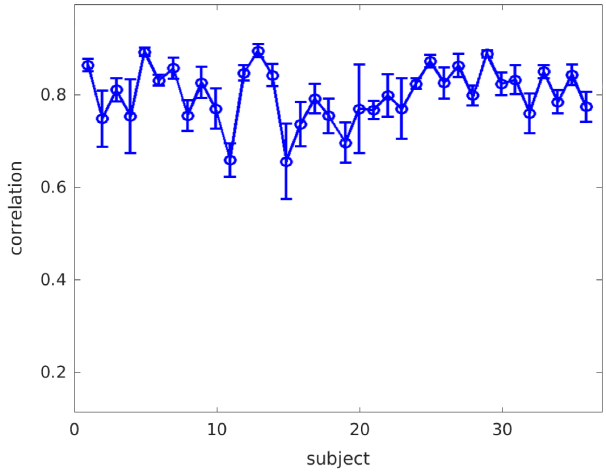


Figure 4.5: Correlation and Standard Deviation for the model of a wrist flexion movement. The model was built on $n-i$ subjects and tested on the ith. For every subject the correlation is the mean on 6 trials.

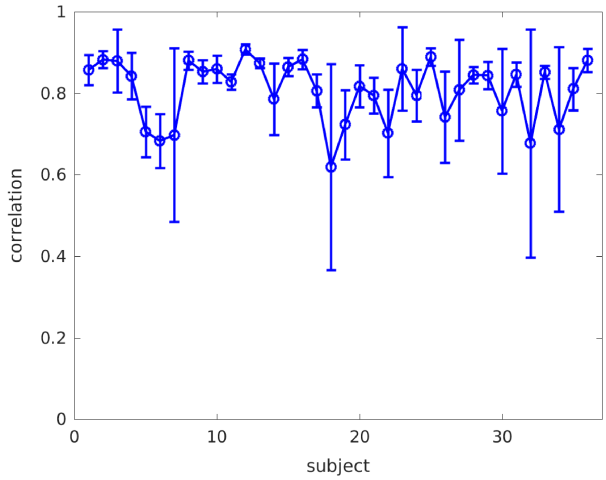


Figure 4.6: Correlation and Standard Deviation for the model of the "Three posture". The model was built on $n-i$ subjects and tested on the ith. For every subject the correlation is the mean on 6 trials and 8 joints.

Robots Fig. 4.7.

Figure 4.7: Simulated hand by Shadow Robots

The EMG signals were sent via software and estimated angle was computed to actuate the robot though TCP/IP protocol. The robot motion generated at 240 Hz was micro-interpolated from the simulated controller of the robot to match the actual rate of 1kHz. The results showed a good correlation resulting from the created GMM/GMR framework. Both the movements reached a statistically significant mean correlation coefficient ($\rho_{\alpha,\hat{\alpha}} \geq 0.8$), with good results for both single joint estimation ($\rho_{\alpha,\hat{\alpha}} = 0.8224$) for Movement 13 (Fig. 4.5), and multi-joints estimation ($\rho_{\alpha,\hat{\alpha}} = 0.8067$) for Movement 3 (Fig. 4.6). The performance reached for multi-joints motion was particularly good even if a bit lower than the single one, since the model has showed consistent results with similar correlation coefficients for all the considered joints.

4.2.1 Conclusion

The framework obtained significant results on new, unseen data, with great variability of subjects and few repetitions of the movements. It was able to estimate the motion of both single ($\rho_{\alpha,\hat{\alpha}} = 0.8224$) and multiple ($\rho_{\alpha,\hat{\alpha}} = 0.8067$) joints for different movements. The estimated joint angles have been remapped to control online a simulated hand for testing the effectiveness of the estimated motion on both the considered movements.

4.3 Integration of IMU data

In the previous chapters, we have already introduced how the comprehension of physiological characteristics of the human movement is gaining more and more interest in the scientific community. The advancements in understanding the mechanisms behind human motion can be exploited by injured subjects to replace lost limbs. In fact, physiological signals are usually applied to help amputees in gaining part of their lost functionality. Moreover, the interest in wearable devices, the study of new materials, the improvements in mechanical design and the advancement of sensors are boosting the development of robotic prostheses as never before. In Chap. 2, we listed the limitations of EMG signals, like the drawback of being conditioned by many common physical aspects [7], together with the strong dependency from the specific subject involved in the acquisition. These are the reasons why the majority of studies concerning motion estimation by means of physiological signals are subject specific. Nevertheless, the results showed in this thesis proved the effectiveness of the subject-independent framework, and the presence of a common pattern between distinct individuals.

In this section, we aim to improve previous results by enriching the model through the integration of accelerometer information from IMUs, alongside the already considered sEMG data. IMU data have been used in robotics in many different ways. In our case, we treated them as a resourceful way for investigating human movement. As a matter of fact, the more information related to motion can be extracted from the subject to be analyzed, the more likely the final result will be accurate. Many studies explored the IMU contribution in estimating people motion, and several of them paired accelerometers and sEMG to obtain more precise results. Keil *et al.* [53] proved that EMG and accelerometers capture different aspects of the movement, thus they can be considered complementary. Gijsberts *et al.* [54] tested data acquired from 20 subjects, building individual models on 40 different movements in order to highlight the contribution given by accelerometers to the classification of hand movements. Their tests compared modality based on accelerometers and sEMG, showing that the accelerometer modality outperformed the sEMG modality, but performances increased when integrating both the signals in an unified model. Liu *et al.* [55] proposed the combined use of sEMG and accelerometers in order to improve the upper limb rehabilitation. The rehabilitation process has been made more interesting and interactive by using video games. Khushaba *et al.* [56] compared five EMG features, one of them including data from accelerometers. They studied how the training set can be generalized in order to classify upper limb movements varying the forearm orientation and the muscular effort. Their tests proved that the inclusion of accelerometers improved the

movement classification accuracy. Anyway, all the presented works studied the contribution of accelerometers in subject-specific sEMG-based models, and as far as we know no attempt has been done in combining accelerometer and sEMG information to boost accuracy in subject-independent approaches.

The aim of this study is to improve the capabilities of a probabilistic model based on multiple subjects to estimate the human motion. The main contribution of this section is comparing, for the first time, the performances of a model built on physiological information with and without taking into account of the IMU as input data. Therefore, sEMG and accelerometer streams are collected to train a probabilistic model, namely a GMM, with the final scope to estimate online the joint bending angles involved in the motion by means of a regression-based approach, namely a GMR. The proposed solution implements a regression technique in order to obtain a continuous control of the movement with the main purpose of actuating a robotic hand or a prosthesis, so that the motion is not only a pre-defined qualitative classification, but an actual interpretation of the movement the patient would like to perform. Moreover, a regression-based approach has the advantage of providing more details regarding the response of the model in time, with the possibility of a more accurate analysis with respect to the results provided by classification methods. It is worth to notice that the proposed framework is not trained online on the data of the user, but it requires an offline phase for generating the GMM model. Anyway, the regression phase can be performed online to obtain the joint angles directly from the data collected on the subject. In order to better understand the contribution of accelerometer data, the results obtained from the novel model integrating both sEMG and accelerometers have been compared with the outcome of the model created in our previous work which involve sEMG data only. Both the models are subject-independent and the same movements have been considered, in order to allow us to compare the results effectively. In particular, we focused on two hand movements, involving a different number and type of joints, in order to test the robustness of the framework in different situations. Based on the results obtained on previous studies mixing together EMG and IMU signals, we expect to be able to improve the accuracy with respect to our previous work. This idea is supported from the fact that IMU sensors are not strongly influenced from subjects' fatigue. On the other hand, different problems could condition the recording of IMU data, such as drifting or sudden bumps. Furthermore, the contribution of accelerometers is larger for wide movements.

4.3.1 Data and methodologies

The dataset used in this study comes from the already described NinaPro dataset, which provide us not only the EMG signals, but also the IMU information.

Signal analysis

In order to be able to extract significant information from the raw signal, we need an appropriate elaboration offering the possibility to process online the input signals. We treated separately EMG data from the rest. Joint angles and IMU data have been aligned to zero and smoothed by means of Eq. 4.1, while EMGs have processed exactly like in the previous experiments.

$$\gamma_S(t) = \frac{1}{S+1} \sum_{s=1}^{S} \gamma(t-s) \tag{4.1}$$

The information from EMG, accelerometers and the joint angles have been used to train a probabilistic model, namely a Gaussian Mixture Model, accurately described in Chap. 3. The probabilistic model is able to estimate the bending angle α of different joints during the movement using a regression technique considering both EMG and ACC information. Once completed the offline modeling phase, the framework can estimate online the joint bending angles by considering solely EMG and accelerometers as input. The angles are estimated continuously by using a regression technique based on the GMM, i.e. Gaussian Mixture Regression (GMR). The Gaussian Mixture Regression (GMR) provided a smooth generalized version of the signal starting from the GMM.

The goodness of the prediction is estimated by computing the correlation coefficient $\rho_{\alpha,\hat{\alpha}}$ which has been accurately described in Chap. 3.

4.3.2 Results

A leave-one-out approach has been used to test the results coming out from the framework. The models has been built on data from $H-1$ subjects and tested on the 6 repetitions of the remaining one. In this way, we have been able to emulate and evaluate the usage of the model by a novel subject. For each repetition, we computed the correlation coefficient between the estimated bending angles and the actual joints angles recorded together with the EMG signals. For each joint, we computed the mean and the standard deviation of the correlation coefficient, averaging the results among the 6 trials. Furthermore, for Movement 3, the correlation coefficient has been

averaged also on the 8 joints involved in the test, in order to summarize the estimated information as a whole.

Corresponding vs *Best*

We performed the whole set of tests for both the *EMG + corresponding IMU* and the *EMG + best IMU* approaches obtaining good results for both the movements in terms of accuracy. In particular, for the *EMG + corresponding IMU* approach, Movement 13 reaches an accuracy of 0.8634 by averaging between all the considered models. 14 subjects out of 35 reach an accuracy above the 0.9, exceeding 0.95 in three cases. With this approach, Movement 13 showed a high regularity on the results among the different subjects, with the unique exception of Subject 26. In fact, the accuracy of 0.3968 reached by Subject 26 decreased the average value. This low accuracy value goes along with a high variance between the trials (0.5112 versus the average variance of 0.0789). This means that a subset of movements have been performed by the subject in a very different or wrong manner. With the same approach, Movement 3 reaches a lower accuracy with respect to Movement 13, obtaining a mean accuracy between the subjects of 0.7659 and a variance equals to 0.0673. The reduced accuracy is mainly due to the higher number of joints involved in Movement 3. Subject 13 only exceeds 0.9 in accuracy (0.9073), while the worst case is Subject 7, reaching an accuracy of 0.3207. Since the low variance between the different trials within the same subject, it is not possible to explain the low accuracy as an isolate error. More likely the low accuracy is due to differences in the way the subjects perform the movement during the entire set of trials. Using this approach, the variance is very similar and quite low for both the tested movements, showing a good generalization capability of the framework. For the *EMG + best IMU* approach, Movement 13 reaches an accuracy of 0.8200 by averaging between all the considered models. 12 subjects out of 35 reaches an accuracy above the 0.9, exceeding 0.95 in six cases. In this case, Movement 13 showed a lower regularity with respect to the previous approach. Again, Subject 26 reached the worst accuracy and confirmed the problems in the specific executions, with a mean of 0.3684 and a variance of 0.4628 (average variance of 0.1041). Regarding Movement 3, again the accuracy decreased with respect to Movement 13, reaching a mean of 0.7426 and variance of 0.0756. The evolution confirms that the higher number of joints involved in Movement 3 affects the accuracy. In general, the performances between the two approaches are not so different in average. On the other hand, it is clear that the use of the best pair of accelerometers improved the accuracy in terms of both mean and variance for the majority of the subjects, but they are not informative

enough to maintain the trend for the entire population. Fig. 4.5 and Fig. 4.9 summarize the tests performed on Movement 13 and Movement 3 respectively, by comparing the results obtained with the *EMG + corresponding IMU* approach, in blue, and the one computed with the *EMG + best IMU* approach, in green.

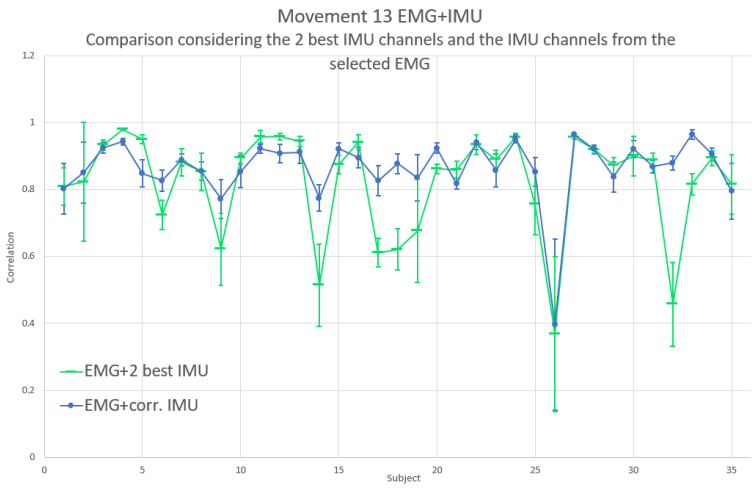


Figure 4.8: Correlation and Standard Deviation for the model of a wrist flexion movement. The model was built on $H-1$ subjects and tested on the remaining one. For every subject the correlation is the mean on 6 trials. Comparison between the *EMG + corresponding IMU* approach (blue) the *EMG + best IMU* approach (green).

Analysis of IMU contribution

A further analysis regards the comparison of the results achieved considering only EMG signals with the ones considering both EMG and IMU signals. In this case, we limited the analysis to the best performing approach: the *EMG + corresponding IMU*. Again, we considered Movement 3 and Movement 13 as testbeds. For both movements, the integration with the accelerometer signals augmented the model accuracy. For Movement 13 (Fig. 4.10), the average accuracy went from 0.8172 (EMG only), to the 0.8634 (EMG+IMU). For Movement 3 (Fig. 4.11), the increment is present, but limited, since the accuracy passed from 0.762 to 0.7659. In some cases, in Movement 13 accuracy increased remarkably when including accelerometers, while for Movement 3 accuracy decreases considerably in three subjects.

	Movement 3	Movement 13
sEMG	0.762	0.8172
sEMG and IMU	0.7659	0.8634

Table 4.3: Average accuracy for the considered movements. EMG only vs EMG +IMU

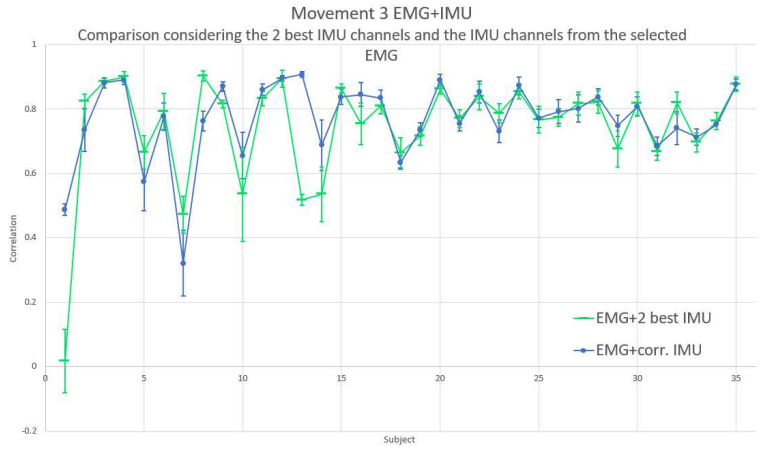


Figure 4.9: Correlation and Standard Deviation for the model of the Three posture. The model was built on $H-1$ subjects and tested on the remaining one. For every subject the correlation is the mean on 6 trials and 8 joints. Comparison between the *EMG + corresponding IMU* approach (blue) the *EMG + best IMU* approach (green).

This difference in the improvement can be explained by looking at the nature of the considered movements. Indeed, the contribution from the accelerometers is more valuable when the motion is wide, since they are placed nearby the forearm and they are not strongly sensible to finger motion. On the contrary, it is possible that a small movement corresponds to a considerable muscle contraction, resulting in a characteristic muscles activation visible in the analysis of EMG signals. Therefore, it is reasonable that IMUs data are more affected by movements involving the wrist rather than the fingers. This explains why Movement 3, which involves the movement of several fingers, does not receive a large benefit from the contribution related to accelerometers. Generally, the mean accuracy increased slightly with the introduction of the accelerometers. Moreover, the variance usually decreased showing a significant improvement corresponding to a better generalization capability of the model. This is probably, the most important result of this study, since this is one of the more desirable characteristics of a subject-independent model.

Tests on a robotic device

Finally, the predicted joint bending angles have been used to control a simulated hand by Shadow Robots. The objective of these tests was to check the correspondence between the actual movement and the motion performed by the robotic device. In this work, we wanted to focus on the contribution of IMU data to the model generalization capabilities rather than on an accurate motion estimation on robotic devices. Therefore, the consideration regarding this section are very limited. The estimated

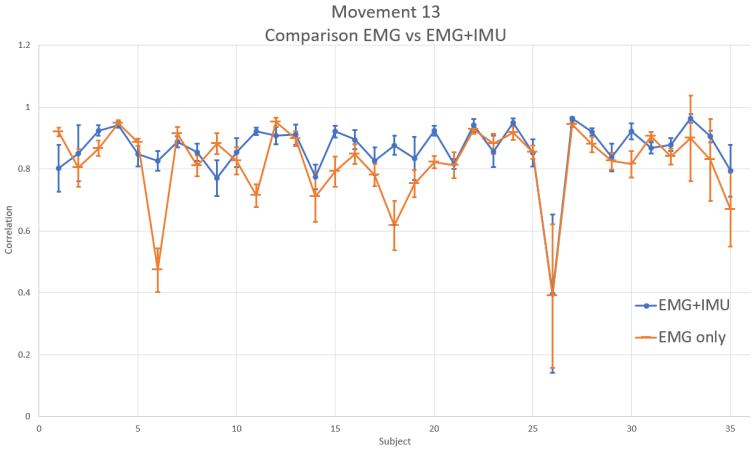


Figure 4.10: Correlation and Standard Deviation for the model of a wrist flexion movement. The model was built on $H-1$ subjects and tested on the remaining one. For every subject the correlation is the mean on 6 trials. The graph shows a comparison between the correlation of the model built using only the EMG signals (orange) and the model built considering also the IMUs (blue).

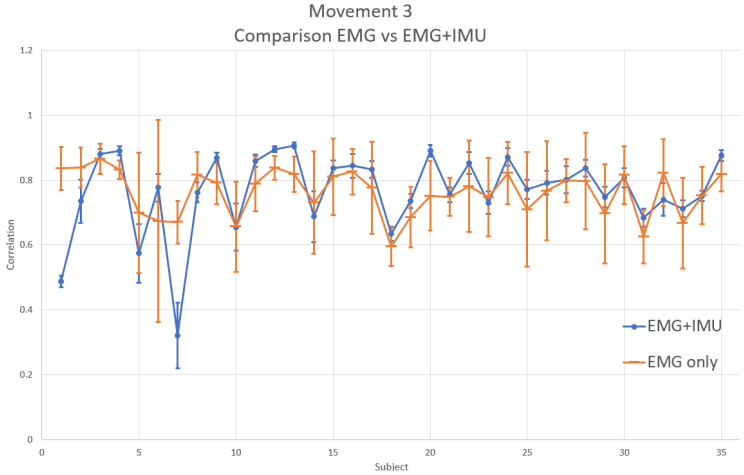


Figure 4.11: Correlation and Standard Deviation for the model of the Three posture. The model was built on $H-1$ subjects and tested on the remaining one. For every subject the correlation is the mean on 6 trials and 8 joints. The graph shows a comparison between the correlation of the model built using only the EMG signals (orange) and the model built considering also the IMUs (blue).

joint angles have been sent to the robot through the TCP/IP protocol, with a frequency around 240 Hz. Subsequently, the robot motion has been micro-interpolated from the simulated controller of the robot to match the final rate of 1kHz. At each trial, we observed the robot motion in order to compare it with with the actual movement and obtain a set of qualitative considerations on the feasibility of the task. In particular, we noticed that very few robot attempts resulted in an altered motion in both the considered movements. This is a first and significant result when dealing with hand gestures. On the other hand, it could not be sufficient if the objective of the motion is to grasp an object, and we are aware that more work could be done on this subject.

4.3.3 Conclusions

We extended our previous studies by integrating the EMG signal with information from IMUs. The final goal was to develop a subject-independent framework able to adapt quickly to new subjects not included in the model.

The signals have been preprocessed in order to be used online and to simulate correctly the usage of a prosthesis device by human users. We tested two different hand movements to prove the robustness and generality of our framework, by considering signals from 35 different subjects in a leave-one-out approach. The framework has been tested on the subject not included in the model.

Tests compared three different models: the first built using only EMG signals, the second trained with EMG plus data from the 2 best IMUs available, and the third created with EMG and IMUs corresponding to the same 3 electrodes. Both the approaches using IMU information improved the correlation with respect to the one considering the EMG signals solely. The introduction of the IMU data helped in improving significantly the generalization capabilities of the framework and consequently in obtaining a better subject-independent model.

In the best case, the mean correlation was 0.8634 for the single joint motion and 0.7659 for the multi joints movement, with an associated variance of 0.0789 and 0.0673 respectively. The improvement in accuracy has been higher in the first case, due to the characteristics of the considered movements: accelerometers are located on the forearm so they are not strongly involved in movements using mainly fingers.

4.4 Low cost framework

The introduction of robotic prosthetic devices can improve the quality of life in amputees, helping to interact with the world around them in simple activities without

depending on other people [57]. Robotics prosthesis could be divided into upper and lower limbs. Hand devices are more complex than legs due to the large number of Degrees of Freedom (DoFs), furthermore they are crucial for interacting and performing daily activities. With hands, it is possible to grasp different objects, and to manipulate them, in order to achieve the most disparate tasks which can be extraordinarily complex and composite. Even if we choose to focus on a subset of movements, it requires a certain amount of dexterity to obtain a renewed upper limbs mobility for amputated subjects. Anyway, hand prosthesis do not suffer of stability problems and there are minor risks of endangering the patient. In fact, prosthetic legs or feet do no need to focus on grasping or manipulation problems, nevertheless a wrong weight distribution or a faulty movement during the walking can cause a damage to the musculoskeletal system. Due to these reasons, hand prosthesis have a major role in the research community [58].

Usually, sophisticated and expensive technologies are needed in order to make robotic hands interact with the environment. Furthermore, the classical approach is based on the prosthesis customization depending on the specific subject that will use it. The customization process is applied on the three fundamental aspects that characterize the rehabilitation system, i.e. the hardware, the sensors used to record the physiological signals that the subject will use to control the device, and the software connecting the information from the sensors to the physical device. The highly technological process and the ad hoc software make the whole prosthesis a very expensive device. A purely cosmetic arm can cost up to $5000, that become $20.000-$100.000 for an advanced myoelectric arm [59]. Many potential users can not afford such pricey products. The users penalized by the high costs are not only private, public clinics or rehabilitation centers, but also research groups that would like to work on this field, focusing mainly on the software rather than on the mechanical or medical side. Furthermore, a great number of injured subjects come from poor countries which cannot afford investments on these technologies [60]. For people with economical difficulties it is even more important being reintroduced in the labour market. The possibility of being able to carry out a physical job is essential for the subjects survival and nourishment, but this is not possible with the classical approach. Different considerations are necessary for children [61]. This kind of subjects have peculiar physiological characteristics that distinguish them from adults, therefore a separate study is needed. Despite the differences between adult and young subjects, it is clear that both the groups would benefit from a low cost prosthesis system. Children grow quickly, thus they need to change prosthesis often, causing a double drawback: high costs for new prosthesis,

while the old ones can not be used by other patients, since they are fitted on the original subject.

For these reasons we focused on the development of a low cost prosthesis framework as described in Fig. 4.12, due to the great benefits brought by a low cost technology. In particular, we consider three different aspects:

- Underactuated 3D printed hand prosthesis.
- Compact and low cost arm band for EMG signals recording.
- Subject-independent probabilistic framework.

Anyway, it is worth to notice that despite the great number of advantages related to a low cost approach based on these aspects, it has also some limitations:

- An underactuated hand does not allow an exact movement replication, since a certain number of joints are bonded, thus it has a lower degree of freedom with respect to a real hand.
- A low cost recording system gives usually worst results than more expensive technologies. The classical solution for the control of the prosthesis is by using EMG signals from the residual muscles. EMG signals can be recorded easily with non invasive technologies, but they have the drawback of being variable during time, dependent on the considered subject, and sensitive to human fatigue and stress. The limitations of the EMG signals are emphasized by using low cost sensors, and they should be filled by using a robust software framework.
- The software framework should handle the limitations of the low cost EMG sensors, while being able to differences of the human hand joints with respect to the underactuated hand. Furthermore, the proposed software solution is subject-independent, which means that each subject can use the generic and robust model with no need of long training phases. This approach has the vantage of zero or rapid training phase for new subjects while maintaining a generic and robust modelization, with the drawback of a slightly worst accuracy.

Up to now, a wide number of robotic hands have been proposed by the scientific community. Different approaches have been explored as regards the previously listed elements, i.e. prosthesis hardware, myoelectric sensors for the recording of physiological signals, and software frameworks for combining the previous two points. An increasing number of hardware solutions have been proposed by both research groups

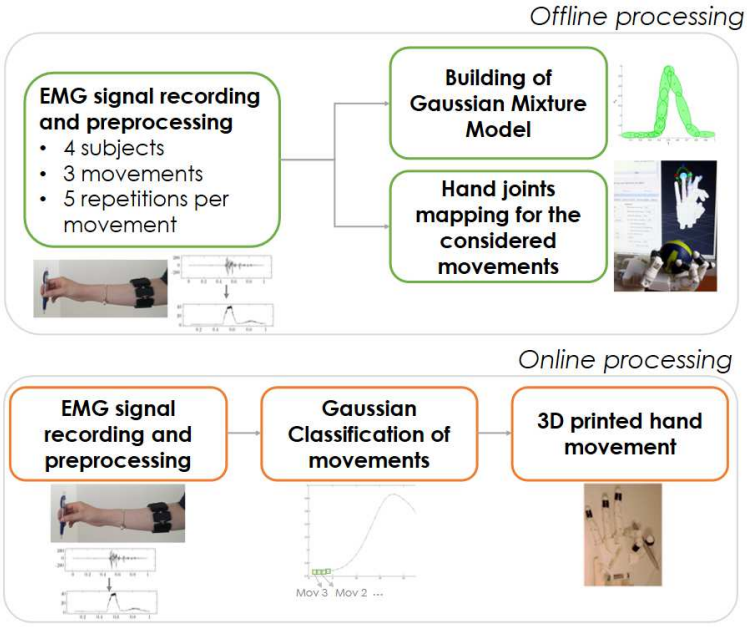


Figure 4.12: The figure presents the sequence of operations for data recording, analysis and modeling. A preliminary offline phase used for training puposes is followed by an online phase for testing the classification framework.

and people became fond on the argument, the cost can vary from very expensive and complex mechanisms to low-cost devices. This great variability depends mainly on the materials selected for building the prostheses, the technologies adopted for moving the fingers, and the presence of sensors. Exploiting these factors, the proposed solutions can be applied in industrial facilities, used as prostheses, or assembles as part of a whole humanoid robot. Looking at low-cost solutions, Yale University proposed an open source project, namely the Yale OpenHand Project [62], to produce a set versatile and customizable robotic hands at low cost. Several different configurations are available, with 2, 3 or 4 fingers, each finger is motor-powered and it has 2 joints with a single degree of freedom. The hands can grasp many different objects of varying size, but none of them has a layout similar to the human hand, since they are mainly oriented to industrial applications. Yang et al. [63] introduced the design of a multi-fingers hand, each finger having 2 joints, with the only exception of the thumb, which had 3 joints. The finger motion was based on the loading of a compression spring, thus the solution is relatively low cost, compared to other technologies. In their paper, Scarcia et al. [64] presented a novel 3D printed hand, where the finger motion is actuated by means of a Rotational Joint, produced as a single piece.

Regarding EMG sensors, the two alternatives for obtaining a low cost system are self-made infrastructures, like [65], or all-in-one solutions, as DTing Gesture Control

Wristband[1].

The main purpose of the software is to process the sensor data to consequently control the finger joints [66]. The majority of solutions focus on properly classify the movement the user wants to perform, in order to reproduce the same task on the robotic device. The traditional approach is a subject specific solution, where the model is trained and tested on the data from the same subject, since EMG signals can strongly differ from one subject to another. Nonetheless, recent studies have highlighted that an underling common behavior can be identified between different subjects in order to obtain a subject-independent solution. This work aims to develop a complete framework bringing together low cost hardware accompany by a suitable software able to work in practice with no need for long training sessions from a specific subject.

4.4.1 Experimental setup

Underactuated 3D printed hand

Between the large number of low cost 3D printed robotic hands proposed in the past few years, we selected the hand design from an open project, called InMoov[2]. InMoov is a wide project that propose a low cost, 3D printed humanoid robot by providing the CAD models for the whole body, as well as the assembly scheme. The goal of the project is to supply universities and research groups with a low cost device, while providing constant improvements. For our work, we chose to focus on the hand instead than on the whole robot. The main reasons for selecting this specific hand were (i) the open-source nature of the project providing 3D models ready for printing, (ii) the human-like hand design, robust and simple to print with almost every kind of 3D printer (iii) the wide community supporting the project.

The hand has 5 fingers, each of them has 3 joints with one degree of freedom, i.e. one less than the human hand. The joints cannot be controlled singularly, two wires connected to the extremity of each finger control their opening and closure. It is not possible to move a single joint without moving the others connected to the same wire. Joints are passive and depends on the movements of the electric motors that control the wires, and there is no tactile or force feedback on the fingers.

With respect to the original project, we created a ROS interface which let the hand being easily integrated with other robots or systems based on ROS. Furthermore, we created a virtual model of the hand which can be useful in order to test the movements

[1] http://www.dtingsmart.com/
[2] http://inmoov.fr/

without using the real device. We sent commands to the electric motors by means of an embedded sheet, namely an UDOO QUAD[3]. UDOO is compatible with Arduino and allows the execution of an operating system, i.e. Ubuntu, with ROS installed. The main purpose of using this device is to control the hand without need of a second connection to a computer.

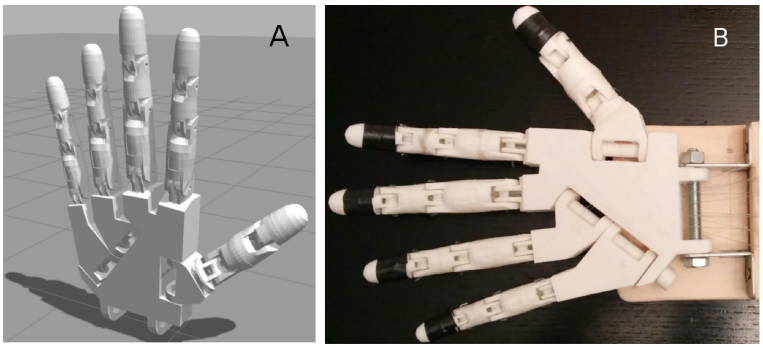


Figure 4.13: Hand simulated in MoveIt (A) and 3D printed hand (B)

Myo armband

The dataset was recorded by using a lowcost device, namely the Myo armband. The description of the dataset is accurately described in Chap. 2.

4.4.2 Methodology

EMG signals from Myo armband uderwent the usual preprocessing phase i.e.:

- Wavelet transform.
- Smoothing and normalization.

The processed signal has been used to train offline a GMM. Once the model is ready, it is possible to classify the movements online by using uniquely the EMG signals coming directly from the subject, after the preprocessing phase. The classification phase exploited a GMC described in Chap. 3.

Considered movements

We tested the framework on a limited set of significant movements. In particular, we focused on three significant hand grasps with peculiar characteristics and useful for

[3]https://www.udoo.org/

help amputated subjects interact with the outer world. The selected grasps (Fig. 4.14) are:

- Grasp a can (A). This grasp is known as Fixed Hook Grasp in the hand grasp taxonomy by Feix et al. [67].
- Grasp a small ball (B), known as Sphere Grasp.
- Grasp a pen (C), known as Writing Tripod Grasp.

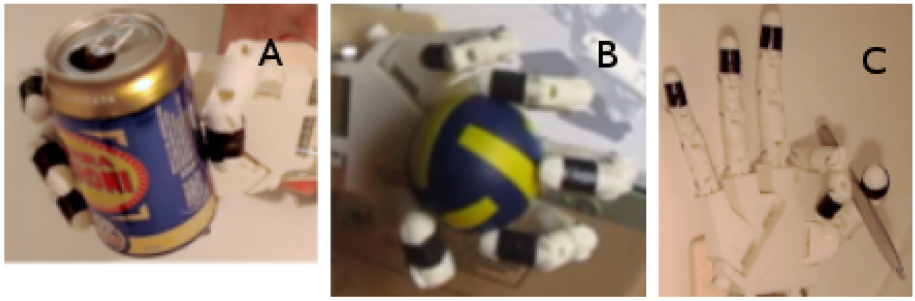


Figure 4.14: Significant grasps of different objects. (A) Grasp a can; (B) Grasp a small ball; (C) Grasp a pen

(A) and (C) involve a rigid and non-deformable object, while (B) involves a soft object. According to Feix's taxonomy, (A) and (B) are power grasps, in (A) the object is supported by the palm, while in (B) it is supported by the hand pad. On the contrary, (C) is a precision grasp and the object leans against the hand pad.

Hand poses and calibration

The set of movements selected during the setup of the experiments has being encoded by recording data from real objects to obtain the correct grasping poses. A preliminary calibration phase has been developed in order to avoid possible differences between two different 3D printed robotic hands. The calibration phase aims at obtaining a reliable grasp motion based on a prerecorded information without major problems. For calibrating the real hand, we developed a procedure able to register simulated and real hand motion by using the visual feedback from the user. We divided the motion of each finger in three steps: (i) flexion of *Distal Phalanx*, (ii) flexion of *Middle Phalanx*, (iii) flexion of *Proximal Phalanx*. The steps are subsequent, so it is not possible to go directly to (ii) or (iii), neither is feasible to pass from (i) to (iii) due to the mechanical behavior of the robotic device. During the calibration phase, the user is asked to provide the three angles corresponding to the end of each step, i.e. the

maximum flexion. These values are used to compute the nominal angles for each joint in the virtual model by means of Eq. 4.2 and Eq. 4.3 respectively for thumb and the other fingers.

$$d = \frac{\alpha}{\alpha_d}\frac{\pi}{2} \quad m = \frac{\alpha - \alpha_d}{\alpha_m - \alpha_d}\frac{\pi}{2} \quad p = \frac{\alpha - \alpha_m}{\alpha_p - \alpha_m}\frac{\pi}{4} \tag{4.2}$$

$$d = \frac{\alpha}{\alpha_d}\frac{\pi}{4} \quad m = \frac{\alpha - \alpha_d}{\alpha_m - \alpha_d}\frac{\pi}{2} \quad p = \frac{\alpha - \alpha_m}{\alpha_p - \alpha_m}\frac{\pi}{2} \tag{4.3}$$

where α is the generic rotation angle of the servo motor used to actuate the selected finger, d, m, and p are the joint angles assumed by *Distal Phalanx*, *Middle Phalanx*, and *Proximal Phalanx* respectively, while α_d, α_m, and α_p are the servo motor angles of maximum flexion for each of the phalanxes.

4.4.3 Results

We tested the framework on data recorded from 4 different subjects. We recorded 8 EMG channels by using a single Myo sensor placed in the proximity of the forearm. Each subject performed 5 times the 3 grasping movements. Four different GMMs have been trained starting from the data of 3 subjects and testing the classification accuracy on the information coming from the remaining subject, following a leave-one-out approach to validate the results. The classification accuracy has been computed by dividing the number of correct predictions over the total number of considered tests. We averaged the accuracy over the 5 trials to obtain a unique confusion matrix for each subject, as reported in Fig. 4.15.

Generally, the results are quite good, with a similar trend between the four subjects, with a mean accuracy of 78.5%, 75.2%, 78.6%, and 74.9% respectively for the model tested on Subject 1, Subject 2, Subject 3, and Subject 4, and an overall accuracy of 76.8% (mean inter-subject accuracy: 85.4%). It is easier to classify the Writing Tripod Grasp, in which the model tested on Subject 1 reached the maximum accuracy of 85.6% and all the models performed best. There are more errors when comparing Sphere Grasp and Fixed Hook Grasp, the minimum accuracy has been reached on the model tested on Subject 4 with 72.3% in recognizing the Sphere Grasp, while some of the other models obtained their personal worst results in classifying the Fixed Hook Grasp. We expected such behavior, since Sphere Grasp and Fixed Hook Grasp are more similar with respect to Writing Tripod Grasp. In fact, the two former movements are both labeled as Power Grasps, while the latter is part of the so called Precision Grasps.

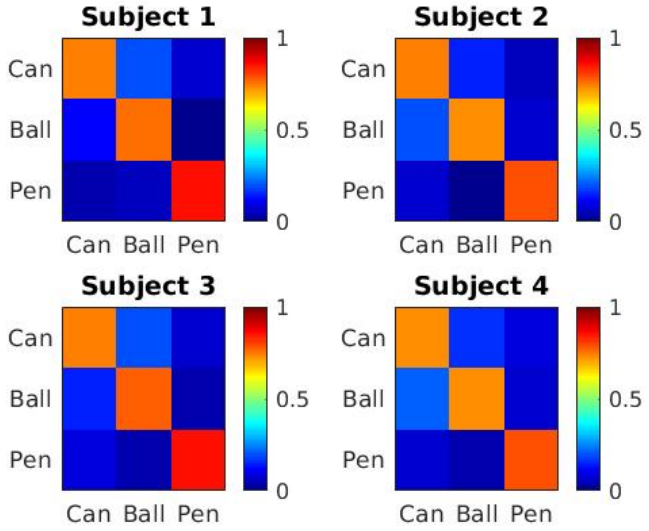


Figure 4.15: Confusion matrices related to the 4 subjects involved in this study. For each subject the classification model has been trained by using data coming only from the three remaining subjects. The results compared the predicted class within the 3 selected grasps with respect to the the actual one.

Once the movement has been classified by the framework, the information passed to the UDOO board to actuate the robotic device. Tests have been performed on real objects to prove the actual grasping capability of the 3D printed hand. We tried also a set of new objects, slightly different with respect to the original ones but with similar grasping characteristics, such as a bulb, a glass, and a cordless phone. The prerecorded motion were able to adapt to these small differences with no need of special interventions thanks to the flexibility of the 3D printed robotic hand when closing on the objects.

4.4.4 Conclusions

The system was able to reach a mean accuracy of 76.8% showing that it is not necessary an expensive setup nor long training sessions from the specific subject to control an hand prosthesis. In the near future, this technology has the potential to provide access to low cost and easy to use prostheses to everyone in need. Tests with a larger variety of grasping poses and with a wide number of subjects could provide useful data to enhance the results obtained up to now, but tests on different objects showed that a small set of basic grasps can actually improve the daily life of injured people without discriminating on their social status.

Chapter 5

Quantitative Taxonomy of Hand Grasps

The techniques proposed to classify or continuously estimate the human movements take into account only the physiological signals from the human, but we did not include any additional information. Intuitively, taking into account some a priori information would boost the estimated results accuracy, since the additional data would bring a greater knowledge of the observed parameters.

Some useful information could be provided by a taxonomy of movements, which can give a structured dependence relationship among movements. This information can be exploited as an useful guideline in the prediction phase. Furthermore, a proper modeling of human grasping and hand movements is fundamental for robotics, prosthetics, physiology and rehabilitation. The taxonomies of hand grasps that have been proposed in scientific literature so far are based on qualitative analyses of the movements, thus they are usually not quantitatively justified. In order to overcome this limitation, we developed the first quantitative taxonomy of hand grasps based on biomedical data measurements. The taxonomy is based on electromyography and kinematic data recorded from 40 healthy subjects performing 20 unique hand grasps. For each subject, a set of hierarchical trees are computed for several signal features. Afterwards, the trees are combined, first into modality-specific (i.e. muscular and kinematic) taxonomies of hand grasps and then into a general quantitative taxonomy of hand movements. The modality-specific taxonomies provide similar results despite describing different parameters of hand movements, one being muscular and the other kinematic.

The general taxonomy merges the kinematic and muscular description into a comprehensive hierarchical structure. The obtained results clarify what has been proposed in the literature so far and they partially confirm the qualitative parameters used to cre-

ate previous taxonomies of hand grasps. According to the results, hand movements can be divided into five movement categories based on the overall grasp shape, finger positioning and muscular activation. Part of the results appears qualitatively in accordance with previous results describing kinematic hand grasping synergies. The taxonomy of hand grasps clarifies with quantitative measurements what has been proposed in the field on a qualitative basis, thus having a potential impact on several scientific fields.

5.1 The quantitative taxonomy of hand grasps

In 1989 Cutkosky [68] said that the main goal in the field of rehabilitation robotics was to build a robot capable of deciding autonomously how to pick up and manipulate objects to perform everyday tasks just like humans do. However, the human hand can perform an almost infinite number of movements. Structuring and organizing the hand grasps into a hierarchical taxonomy can be useful to better understand how the hands interact with different objects as well as to advance and evaluate devices that try to imitate them.

A taxonomy of hand movements is important for several scientific fields, including robotics, prosthetics, physiology and rehabilitation. In robotics, it can be useful to compare the functionality of robotic hands with real human hands. In prosthetics, very advanced myoelectric hands have been developed from a mechanical point of view but they are usually not well accepted by amputees [69–71]. A taxonomy of hand grasps can foster the development of prosthetic hands that perform movements corresponding to the taxonomic groups that are mostly useful in real life situations. In physiology, a comprehensive quantitative comparison of hand grasps may create a link between hand synergies [72] and real life needs. In rehabilitation, a proper taxonomy of human grasps allows priorizing the hand functionalities that need to be restored with the highest priority. Santello et al. [72] proposed an early approach into this direction by applying Principal Components Analysis (PCA) to digit joint angles during a significant set of hand postures. Their work, as many others that followed it [73–78], took inspiration from grasp taxonomies in order to properly select the set of hand movements.

Several attempts to build a complete taxonomy of hand grasps were published in the scientific literature during the last 30 years. However, all of the presented taxonomies were based on qualitative approaches and qualitative justifications. Most of the taxonomies of hand movements include a division between power and precision tasks. This idea was originally proposed by Napier et al. in 1956 [79] and influenced most

authors afterwards (e.g. [80–83]). Cutkosky [83] organized 16 hand grasps into a hierarchical tree according to the adaptability required by small-batch tasks. The grasps were characterized using several qualitative measures (such as compliance, connectivity, grasp isotropy, resistance and other parameters) and they were split into power and precision tasks. Feix et al. [67] compared several previous taxonomies of hand grasps and created a taxonomy of hand grasps that they called the *GRASP* taxonomy. This taxonomy is organized in a matrix, with the grasps divided into several columns and in two rows according to four main parameters including power type, opposition type, position of the thumb and virtual finger assignments. Starting from Feix's work, Wolf et al. [84] considered composed tasks in order to evaluate the micro-gestures that can be performed alongside the main grasp. More recently, Bullock et al. [85] decomposed manipulation tasks into simpler movements with an object-centric, environment-centric and hand-centric perspective. This taxonomy provides a structured way to classify 15 simple movements, where basic movements can be composed in order to build more complex movements.

Qualitative methods can provide useful perspectives of nature. However, quantitative measurements are strongly related to the scientific method and to the concept of science itself. Quantitative methods provide practical control over the subject studied, they make possible a formulation of principles that are capable of unambiguous confirmation or refutation (depending on experiments and measurements) possible and therefore very few investigations can be carried out without them [86].

A quantitative taxonomy of hand movements can therefore reduce ambiguity in the field, but it requires the measurement of specific biomedical data. Several parameters can be used to quantitatively characterize hand grasps, such as posture, muscular activity and force. Kinematic data are usually measured with two main techniques: visual or wearable systems. Visual systems can be affected by visual occlusion in the recording of hand grasps and the procedure to place the visual markers can be time consuming. Data gloves are a common alternative that is sufficiently precise [87] and extremely easy to record. Thus, they are suitable for studies involving many subjects. The joint angles were previously used as features in order to compare model estimations with real position measurements [88], [89]. In the comparison of movements, synthesis functions are often applied to represent the entire motion with fewer data. For instance, Finger Aperture Indexs (FAIs) were used to represent long finger opening starting from joint angles collected by a Motion Capture (MoCap) system composed of nine infrared cameras and 17 retro-reflective hemispheric markers [90]. Normalized geometric distances were used as features for representing hand gestures [91].

Muscular data can be measured with Surface Electromyography (sEMG). The sEMG signal can be modeled as a superimposition of the Motor Unit Action Potentials (MUAPs) of the active Motor Units (MUs) [92, 93]. The MU recruitment and firing frequency are the major factors for both EMG amplitude and force exerted by the muscle [93, 94]. Thus, a qualitative relation between the sEMG signal amplitude and the force exerted by the muscle can be noticed [94]. Signal features based on sEMG signal amplitude can reveal the hand movement patterns based on the sEMG amplitude-force relation in both intact subjects [29] and hand amputees [95]. Muscle activation patterns can differ strongly between intact and transradial amputees, particularly in relation to clinical parameters such as phantom limb sensation intensity, remaining forearm percentage and time since the amputation [95]. This result leads to the fact that signal acquisition controls trained on intact subjects may not be valid for amputees [96]. Other parameters may be interesting but in this work we focus on kinematics and muscular activity because we mainly target the posture of the hands, but also due to practical data availability. In this thesis it is presented the first quantitative taxonomy of hand movements.The relative variations between joint bending angles (measured with a data glove) allow a quantitative characterization of the hand movement kinematics. The sEMG signals allow a functional analysis of the muscles involved in each grasp. The taxonomy is organized in a hierarchical structure and it is based on a signal feature extraction procedure that is common in sEMG literature. The analysis of the movements performed by 40 intact subjects allows to extract the common underlying patterns that characterize each grasp. The quantitative approach ensures the repeatability non-subjective perspective of this taxonomy, thus making it a reference for several scientific fields.

5.2 Methods

This section describes how kinematic and sEMG data were recorded and analyzed to create a quantitative taxonomy of hand movements. The data analysis procedure can be summarized as data acquisition (subsection 5.2.1), signal feature extraction (subsection 5.2.2, subsection 5.2.2), creation of the hierarchical trees (subsection 5.2.3) and fusion of the trees into super-trees (subsection 5.2.4), a procedure coming from genetics studies and leading to the general quantitative taxonomy of hand movements.

5.2.1 Data Acquisition

The used dataset is the second Ninapro dataset, including 40 intact subjects (28 males, 12 females; 36 right handed, 4 left handed; age 29.9 $\pm$ 3.9 years). The Ninapro database[1] [28, 29], is a publicly available resource aiming at improving the control of robotic hand prostheses. The data comprise 50 hand and wrist movements, including basic motions (e.g. flexion, extension) as well as 20 grasps.

Acquisition setup

The acquisition setup includes a data glove and a set of surface electromyographic electrodes with built-in accelerometer. Hand kinematics were measured using a 22-sensor CyberGlove II (CyberGlove Systems LLC [2]), providing data proportional to joint angles, sampled at slightly less than 25 Hz. Muscular activity was measured using a Delsys Trigno Wireless system. The sEMG electrodes are double-differential and measure the myoelectric signals at 2 kHz with a baseline noise of less than 750 nV RMS. The sEMG electrodes were placed using the hypo-allergenic Trigno Adhesive Skin Interfaces. Prior to electrode placement the skin was cleaned with isopropyl alcohol.

A hand movement is the result of an activation pattern potentially involving several muscles controlling hand and wrist. Therefore, in order to identify the hand movement from the sEMG signal by means of pattern recognition methods, the electrodes were placed around the subject's forearm combining a precise anatomical positioning strategy [94] with a dense sampling approach [97, 98]. An array of eight sEMG electrodes was applied at the height of the radio-humeral joint. The electrodes were equally spaced, creating an array covering the whole circumference of the forearm. Four electrodes were placed on the main activity spots of four specific muscles: the *flexor digitorum superficialis*, the *extensor digitorum superficialis*, the *biceps brachii* and the *triceps brachii*. The aforementioned strategy is widely used in the prosthetic field. It was shown that, in terms of pattern recognition accuracy for hand movement identification, the electrode position is not a crucial aspect as long as a sufficient number of channels is provided [99, 100].

[1] http://ninapro.hevs.ch/

[2] url: http://www.cyberglovesystems.com/

Acquisition protocol

During the data acquisitions, the subjects were sitting with the arms positioned in a relaxed way on a desktop. A laptop computer was used to show them the videos representing the movements to be performed and to record the data from the sensors. The subjects were asked to synchronously mimic the movements with their right hand. Each subject performed 6 repetitions of 49 movements plus rest. Each movement repetition lasted 5 seconds, alternated with 3 seconds of rest. Several precautions were taken in order to encourage a natural and spontaneous execution of the grasp. First, the subjects were instructed to perform the grasp movement as naturally as possible, without lifting the objects or exerting unnatural grip force. The movements were not randomized and the objects to be grasped were positioned as closely as possible to the hand of the subject. The latter also helped in minimizing the time of the reaching and releasing phases. The hand movements were selected from the hand taxonomy, robotics, and rehabilitation literature (e.g., [67, 68, 81, 101]) according to Activities of Daily Living (ADL) requirements. Everyday objects that can easily be found in daily life tasks were used in the experiment.

5.2.2 sEMG and Data Glove Signal Processing

In order to allow the creation of the sEMG based quantitative taxonomy, preprocessing and feature extraction were performed. First, data preprocessing was performed to assure good data quality. This phase included filtering and synchronization. Second, the information of the sEMG signals was made usable by extracting a set of signal features using a moving window technique [102, 103]. Finally, the signal features were used as input data to compute the hand movement taxonomies.

The CyberGlove data were analyzed with a procedure that takes inspiration from window based time series analysis and in particular from the literature in EMG data analysis [29, 102]. The procedure includes synchronization and feature extraction.

Filtering

The Delsys electrodes are not shielded against power line interference, so the power line interference was removed using an Hampel filter at 50 Hz [104].

Synchronization

A high-resolution timestamp based on the Time Stamp Counter (TSC) of the CPU was assigned to each sample recorded for both the sEMG and joint angle data. The

timestamp was used in the post-processing phase to synchronize the devices. To do so, all the modalities were up-sampled at the sampling frequency of the fastest device (2 kHz) using linear-interpolation. This is a well-known technique to increase resolution, avoid aliasing, and reduce noise [105]. Interpolation is particularly useful when the data collected with low frequency (kinematics) is considerably smoother than the data at high frequency (sEMG).

Feature Extraction

Signal feature extraction was performed applying the method described by Englehart et al. [102]. Each movement repetition was windowed using a 200 ms window, with 100 ms of overlap. As described in scientific literature, diverse signal features highlight different signal properties, leading for instance to varying classification performance (e.g. [29]). In order to make the taxonomy robust to differences between features, a selection of five time domain signal features was extracted on each time window.

The features were chosen according to use in the previous literature and include Rooted Mean Square (RMS), Mean Average Value (MAV), Integral Absolute Value (IAV), Time Domain (TD) [103] and Wavelet (WL) [102,106–111]. The Time Domain (TD) are composed of: Mean Average Value (MAV), Mean Absolute Value Slope (MAVS), Zero Crossings (ZC), Slope Sign Changes (SSC) and Wavelet (WL) [103]. Each feature was extracted from each signal x on each time window w of T samples in length.

Rooted Mean Square (RMS) is arguably one of the most common features to represent sEMG signals. RMS provides a useful measurement of signal amplitude and, under ideal conditions, it has a quasi-linear relationship with the force exerted by a muscle [94]. The RMS feature for a given time window w was obtained as:

$$RMS_w(x) = \sqrt{\frac{1}{T}\sum_{t=1}^{T} x_t^2};$$

where x_t is the t^{th} sample in the window w.

The Mean Average Value (MAV) and the Integral Absolute Value (IAV) are also popular features in sEMG signal analysis and for a given time window w they are defined as [103, 112]:

$$MAV_w(x) = \frac{1}{T}\sum_{t=1}^{T} |x_t| \qquad IAV_w(x) = \sum_{t=1}^{T} |x_t|$$

The Mean Absolute Value Slope (MAVS) is defined as the difference between the MAV value of two adjacent time windows, w and $w+1$ [103]:

$$MAVS_w(x) = MAV_{w+1}(x) - MAV_w(x)$$

The Zero Crossings (ZC) [103] feature gives an indication about the frequency of the signal by counting how many times the signal crosses zero. The ZC of a signal x in a given window w, $ZC_w(x)$, is increased by one if, given two consecutive samples x_t and x_{t+1}, $\{x_t > 0$ and $x_{t+1} < 0\}$ or $\{x_t < 0$ and $x_{t+1} > 0\}$ and $|x_t - x_{t+1}| \geq$ threshold.

Another feature related to the frequency of the signal is the Slope Sign Changes (SSC) [103] which is defined as the number of times the sign of the slope changes. The SSC of a signal x in a given window w, $SSC_w(x)$, is incremented if, given three consecutive samples x_{t-1}, x_t and x_{t+1}, $\{x_t > x_{t-1}$ and $x_t > x_{t+1}\}$ or $\{x_t < x_{t-1}$ and $x_t < x_{t+1}\}$ and $\{|x_t - x_{t+1}| \geq$ threshold or $|x_t - x_{t-1}| \geq$ threshold$\}$.

Wavelet (WL) returns a single parameter providing a measure of the waveform complexity and given a time window w it is defined as [103]:

$$WL_w(x) = \sum_{t=2}^{T} |x_t - x_{t-1}|$$

5.2.3 Hierarchical trees

The quantitative taxonomy of hand movements is based on a hierarchical structure in order to highlight dependencies and relationships between the different motions. For each subject, one hierarchical tree was computed for each modality-feature combination, thus leading to five hierarchical trees for the EMG data and five trees for the kinematic data. We adopted this approach, instead of building only one large hierarchical tree containing all the subjects in order to achieve a higher control of intermediate results and to be able to check the similarity across subjects. For each subject modality-feature combination, the hierarchical trees were computed by performing one-way Multivariate Analysis of Variance (MANOVA) [113] between the hand movements. This procedure allows us to test our hypothesis for all the movements at once to measure how much a grasp is correlated to the others. Therefore, MANOVA can provide a measure of similarity between the grasps that were considered in the study. Moreover, MANOVA is a standard, well accepted means of performing multivariate analysis. The signal features were grouped by movement and the means of the collected measures were compared by computing the Mahalanobis distance [114]. It is multi-dimensional, unitless, and scale-invariant. The Mahalanobis distance takes into

account the correlations coming from the MANOVA procedure to measure how distant a specific movement is from the distribution (the whole set of grasps) in terms of standard deviations. The distances between the movements were then used as a basis to build the dendrograms. The dendrograms were initially represented as binary trees composed of clusters of two movements combined depending on the distance. We followed a hierarchical agglomerative clustering or bottom-up approach. By doing so, we treated each movement as a singleton cluster and then agglomerate pairs of clusters until all clusters are merged into a unique tree containing all grasps. The initial set of grasps was a previous knowledge so it almost naturally implied the use of hierarchical agglomerative clustering. On the contrary using a divisive (or top-down) approach could have lead us to a different final number of grasps not corresponding to the initial set that was available. Subsequently, the dendrograms were converted into phylogenetic trees that are unordered rooted trees with unweighted edges, with the characteristic of having all the leaves at the same distance from the root. Part of the information contained in the dendrograms is lost when using unweighted edges. This is due to a limitations of the merging algorithm (subsection 5.2.4) that is not currently able to manage such information. Using a weighted structure may provide more accurate results than the current work, thus we are working on an improved version of the merging procedure. Nevertheless, an approach based on unweighted trees is important to have a proper understanding of the general methodology since this is the first attempt to obtain a quantitative taxonomy of hand grasps.

5.2.4 Computation of the muscular, kinematic and general quantitative taxonomies: hierarchical super-trees

The capability to merge several highly specific trees is the key idea in obtaining a unique hierarchical structure. This part of the data analysis is fundamental, since it allows us to create a general and global quantitative taxonomy that takes into account inter- and intra-subject variability. Inter-subject variability is due to the highly specific way in which each person performs hand movements. Intra-subject variability is due to the small differences between repetitions of the same movement by the same subject. Despite the inter- and intra- subject variability, each hand movement has common underlying kinematic and muscular patterns that can be extracted by analyzing several repetitions of the same movement performed by different subjects. The variability between subjects can be measured as *edit distance*, that is the minimal-cost sequence of node edit operations that transforms one tree into another [115, 116] (more details in subsection 5.2.5). The common characteristics emerged in a preliminary study [117],

where we focused on the taxonomies built on specific subjects (average edit distance of 4.312) and for specific features (average edit distance of 3.548), excluding the generalization phase. More detailed information about these results is reported in Table 5.1 and Table 5.2, it is worth to notice that we considered the weighted dendrograms for computing edit distances displayed in both tables. Starting from the initial results, in this paper we aim at expanding and enriching the knowledge in the field by merging several features in order to develop a unique and general taxonomy. Considering 40 subjects and 6 repetitions for each subject results in 240 repetitions of each movement, which is a sample large enough to create the taxonomy with the procedure described in this section. The procedure to compute the taxonomies of hand movements starts from the subject-specific hierarchical modality-feature trees and includes several phases. First, subject-independent hierarchical modality-feature trees are computed. Second, the general kinematic and muscular taxonomies of hand movements are computed. Third, a general taxonomy of hand movements is computed.

Table 5.1: Inter-subject variability in grasps across the different quantitative metrics expressed as edit distance. Rows represent the different modalities while columns represent modality features.

	IAV EMG	MAV EMG	RMS EMG	TD EMG	WL EMG
Muscular	4.60±1.68	4.60±1.68	4.05±1.24	4.36±1.48	4.51±1.60
	IAV glove	MAV glove	RMS glove	TD glove	WL glove
Kinematic	4.04±1.45	4.04±1.45	4.04±1.45	3.95±1.59	3.84±1.28

As said in the previous section, one hierarchical tree is computed for each subject and for each combination of modality-feature, thus leading to five hierarchical trees for the EMG data and five trees for the kinematic data. For each modality and for each feature, a supertree is computed by combining the data of all the 40 subjects, leading to a subject-independent hierarchical modality-feature tree. The procedure used to merge the hierarchical trees is based on the Subtree Prune-and-Regraft (SPR) distance [118]. The calculation of the SPR distance is computationally expensive. Thus, the algorithm combines the Maximum Agreement Forests (MAFs) approach and clustering to make the construction of the SPR-based supertree feasible. Clustering reduces the complexity of the input trees into sub-problems that can be solved recursively. The algorithm solves the MAF problem between a pair of rooted trees by recursively exploring all edge-cutting possibilities. The supertree is built in two phases: the construction of an initial SPR supertree and the SPR rearrangement. The final supertree is a binary rooted tree constructed iteratively by minimizing the SPR distance. This approach was demonstrated to be better than other common distance criteria on biological data

sets [118]. Merging the subject-independent hierarchical modality-feature trees of the same modality leads to two modality supertrees: the first one representing the quantitative kinematic taxonomy of hand movements (obtained by merging all the Cyberglove feature trees); the second one representing the quantitative muscular taxonomy of hand movements (obtained by merging all EMG feature trees).

Finally, the quantitative muscular and the kinematic taxonomies of hand movements were merged into the general quantitative taxonomy of hand movements. While the EMG tree gives a vision of muscular involvement in the movement and the kinematic tree shows the actual physiological movement performed by the subject, mixing the two allows a general analysis of the movement from both the muscular and the kinematic perspective.

5.2.5 Supertree similarity measurements

Evaluating the similarity between the quantitative muscular and the kinematic taxonomies of hand movements can yield fruitful insight, particularly to measure whether the two agree or not. While a reasonable agreement between the two taxonomies may enforce their representativeness, a strong disagreement may lead to a limited acceptability. Intermediate situations can be interesting to highlight differences in the data acquisition modalities or highlight differences between the muscular activation and the actual performed movement. The tree edit distance is frequently used in the comparison of hierarchical trees [115, 116]. The measure is computed as the minimal-cost sequence of node edit operations that transforms one tree into another. The algorithm used to compute the tree edit distance was originally proposed by Pawlik and Augsten [115, 116]. It includes three possible edit operations: delete a node, insert a node and rename the label of a node. A cost was assigned to each edit operation. The cost of an edit sequence is the sum of the costs of its edit operations. The tree edit distance is the sequence with the minimal cost.

5.3 Results

This work presents a quantitative taxonomy of hand grasps based on muscular and kinematic data, described in detail in subsection 5.3.3. The general taxonomy is computed by merging the sEMG and kinematic taxonomies of hand grasps (that are described in detail in subsection 5.3.2) and it is compared in subsection 5.3.4 with a qualitative taxonomy of hand grasps that merges most of previous results described in literature.

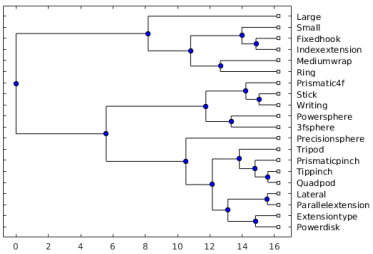
5.3.1 Preliminary study

Robots that aim at grasping and manipulating objects, such as industrial robotic hands and prostheses, all lack reliability and robustness in real life settings. In industry, automated warehouses are successful at removing processes such as walking and searching for items. However, automated handling of goods in unstructured environments still remains a difficult challenge.

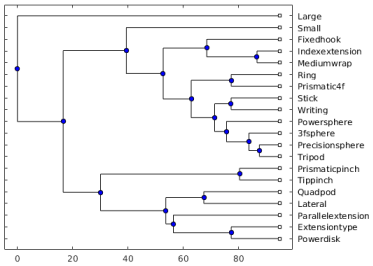
The human hand is highly dexterous and can be used for many diverse tasks. It includes 15 joints (not including carpus and metacarpus), leading to more than 20 degrees of freedom. The development and evaluation of quantitative representations of hand movements requires the measurement of specific biomedical data. Kinematic and muscular data can provide a complete view of hand functions and they have been widely studied for rehabilitative robotic applications.

Before focusing on the quantitative taxonomy of hand grasps, we studied the feasibility of our goal by quantitatively analyzing the similarity between hand movements in 40 subjects. Surface electromyography (sEMG) and the relative variations between joint bending angles of a data glove allow a hierarchical quantitative characterization of the hand movement dynamics and kinematics for each subject, in order to identify functional similarities and variability. The proposed approach suggests a way to perform a systematic, quantitative analysis of hand movements in order to develop a quantitative taxonomy.

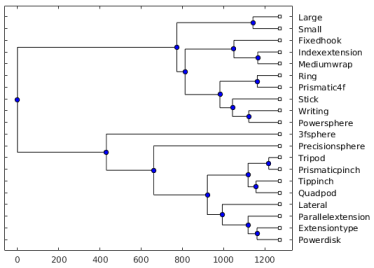
The obtained taxonomies have been compared both among subjects and among features. The comparison of the signal feature taxonomies show that IAV and MAV features produced the same tree for almost all the considered subjects for both EMG and glove signals. Anyway, the trees related to the other features have some common characteristics: for example *Extension Grasp* and *Power Disk* are closely related for Subject 2 (Fig. 5.1a and Fig. 5.1b). Furthermore, IAV and MAV are very similar to RMS results, while TD tree is the most different from the others (Fig. 5.1c). The previous assertion is valid for both EMG and GLOVE signals, confirming that the represented information is coherent among the two kind of signals. The comparison of the subject taxonomies shows that a common behaviour could be highlighted among the different individuals. Some movements are strictly related among almost every subject, this is the case of *Extension Grasp* and *Power Disk Grasp*. While the majority of subjects have a similar behaviour, some of them vary from the average, as it happens for Subject 9 (Fig. 5.1d). In particular, this subject's subtree is populated by a wide set of movements for both IAV and MAV features for EMG signals In fact, in this case movements like *Stick Grasp*, *Writing Grasp*, *Power Sphere Grasp* are in the same



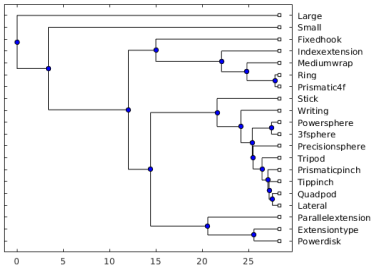
(a) Subj:2, Feat:MAV, Data:EMG



(b) Subj:2, Feat:MAV, Data:GLOVE



(c) Subj:2, Feat:TD, Data:EMG



(d) Subj:9, Feat:MAV, Data:EMG

Figure 5.1: Hierarchical trees created by using (a) MAV features on EMG data from Subject 2, (b) MAV features on Glove data from Subject 2, (c) TD features on EMG data from Subject 2, (d) MAV features on EMG data from Subject 9

subtree, while trey are usually separated in different branches.

The results suggest that quantitative hierarchical representations of hand movements can be performed with the proposed approach and the results from different subjects and features can be compared.

5.3.2 Muscular and kinematic taxonomies of hand grasps

The quantitative hand movement taxonomies based on EMG (Figure 5.2(a)) and kinematic data (Figure 5.2(b)) are in agreement and provide a similar representation of the hierarchical organization of hand movements. This result is confirmed by the edit distance between the two supertrees, which is 33 (a value within the range of the distances computed for hierarchical trees obtained in a specific modality by using different features). This fact enforces the validity of both taxonomies, that were computed using sensors measuring different parameters related to hand movements. The groups of movements defined in the two modality-specific taxonomies are often similar. For instance, the *large diameter* and *medium wrap* grasps are linked at the first level in both the EMG and the kinematic taxonomy. The same happens for several other groups of

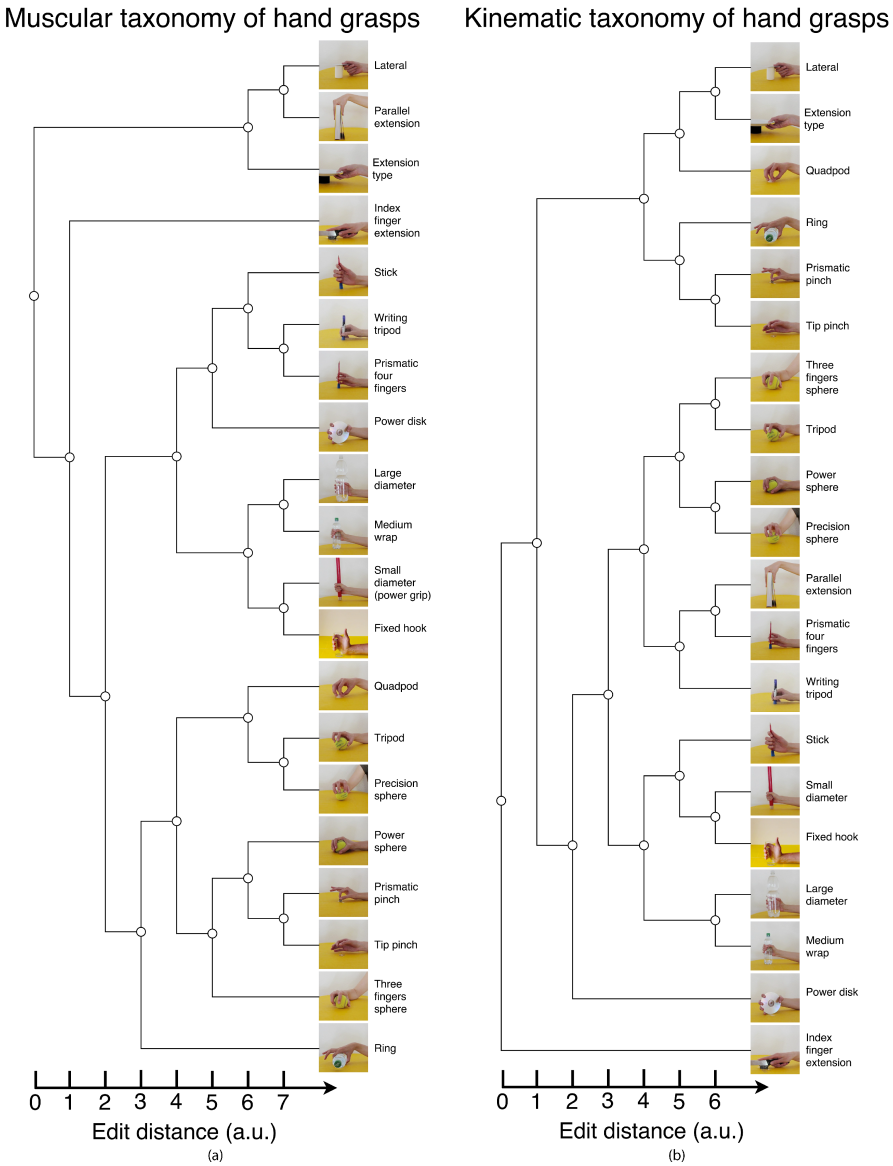


Figure 5.2: Modality specific quantitative taxonomies of hand grasps. (a) Muscular taxonomy computed from EMG data and (b) kinematic taxonomy computed from kinematic data recorded using the cyberglove. The taxonomies are similar (edit distance = 33). The differences highlight the differences between the two data acquisition modalities.

movements, such as the *small diameter* and *fixed hook* grasps, the *prismatic pinch* and *tip pinch* grasps. Other movements change from first level connections in one tree to second level connections in the other. This is the case, for instance, of *parallel extension* and the *lateral grasp*, *prismatic four fingers* and *writing tripod*, *precision sphere* and *tripod*. An interesting change happens considering the *power sphere*, *precision sphere*, *tripod* and *three finger* grasps. These grasps are strongly linked (i.e. they are very similar) considering the kinematic taxonomy. In the EMG taxonomy the *precision sphere* and the *tripod* grasps are closer to the *quadpod* movement, while *power sphere* and *three finger* grasps are closer to the *prismatic pinch* and *tip pinch* grasps. Few movements change the grasp group depending on the considered taxonomy. This is the case of the *power disk*, the *index finger extension* and the *parallel extension* grasp. The *power disk* is grouped with the *prismatic four fingers*, the *writing tripod* and the *stick* grasp in the EMG based taxonomy. On the other hand, only in the penultimate level in the kinematic taxonomy is linked to them. The *index finger extension* grasp is isolated in both trees. In the EMG tree, this grasp is in a single branch, close to the majority of graspings but linked at the higher level to the grasp group including the *parallel extension*, *lateral* and the *extension type* grasps. In the CyberGlove tree on the other hand, it is completely isolated from the others movements. A strong difference between the two taxonomies occurs for the *parallel extension* grasp. In the EMG tree, the grasp is isolated from the other movements, but grouped with the *lateral* and the *extension type* grasps. In the kinematic taxonomy it is close to the *prismatc four fingers* and the *writing tripod* grasp. There are two possible reasons to explain this difference. First, the EMG taxonomy considers the activation of wrist flexors/extensors, while the taxonomy based on the data glove does not consider them. The EMG signals measure all the muscular activity in the forearm, including the activity related to wrist movements while the Cyber Glove on the other hand is sensitive only to finger movements. Second, the difference can be due to variations in the force used to accomplish the movements, since the EMG signals are sensitive to it. In any case, except these few situations, the differences between the EMG and the kinematic taxonomy of hand movements are limited, confirming the validity of the proposed approaches and thus the validity of both taxonomies.

As previously said, the final EMG and glove taxonomies are assembled from supertrees built from single features. We measured how much the grasps are similar one to each other by using MANOVA starting from those features. The comparison involved the same features for different movements, so we computed a metric able to cope with the entire distribution in a multi-dimensional space as described in subsec-

tion 5.2.3. Mahalanobis distances were computed for all features, and for each modality, we averaged them to obtain a unique value representing how close a movement is to another. In order to provide an intuitive way to show similarity between EMG and glove information, we built two distance matrices, one for the muscular (Figure 5.3) and one for the kinematic (Figure 5.4) data. The two matrices show several similarities and, in general, they confirm the considerations derived from the respective supertrees. For example, the *index finger extension* is clearly distant from all the other movements, while *small diameter*, *fixed hook*, *large diameter*, and *medium wrap* are very similar grasps. As a further prove to sustain the idea of merging trees built from different sensors, Table 5.3 represents the edit distance between the modality-specific taxonomies and the supertrees built on each modality for each specific feature. Considering the EMG data, the IAV and the MAV based taxonomy are the most similar to the muscular taxonomy (edit distance = 22). The edit distance between the IAV and the MAV taxonomies is 0. The most different tree is the one based on WL (edit distance = 39). Considering the kinematic data, the RMS based taxonomy is identical to the kinematic taxonomy (edit distance = 0), while the TD tree is the most different.

5.3.3 General quantitative taxonomy of hand grasps based on muscular and kinematic data

The general quantitative taxonomy of hand grasps (Figure 5.5) is computed by merging the muscular and the kinematic taxonomies and offers a general and comprehensive description of hand movement similarities, thus overcoming the subjectivity of previous qualitative taxonomies as well as the limitations of the muscular and the kinematic taxonomies presented in Subsection 5.3.2. The general taxonomy of hand grasps is slightly closer to the EMG taxonomy (edit distance = 29) than to the kinematic taxonomy (edit distance = 42). Coherently, the supersupertree has more connections in common with the EMG one.

The general quantitative taxonomy of hand grasps presents a division into five categories that correspond to real finger positioning and muscular activation, reflecting the shape of the grasped object and balanced combinations of parameters rather than the force used or other specific single parameters. The categories were named as follows according to specific properties of each group: ***1) flat grasps; 2) distal grasps; 3) cylindrical grasps; 4) spherical grasps; 5) ring grasps***. ***Flat grasps*** are well separated from all the others and are characterized by an elongated (or "cupped") positioning of the palm with an abducted or adducted thumb. Parallel extension can be added to this group considering its similarity with the extension grasp and that the grasp is included

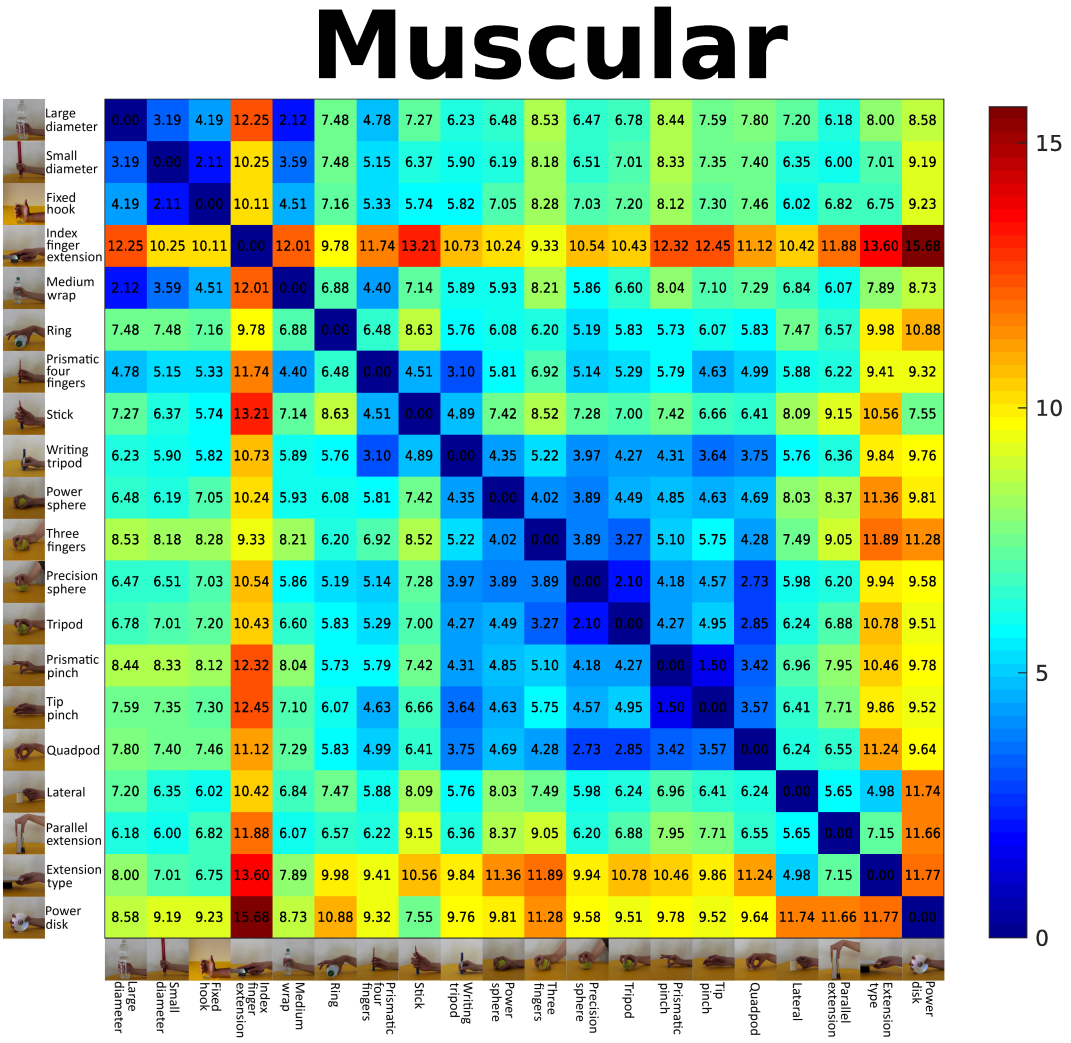


Figure 5.3: Distance matrix computed as the mean Mahalanobis distance between the considered grasps, using the available features computed on the EMG data.

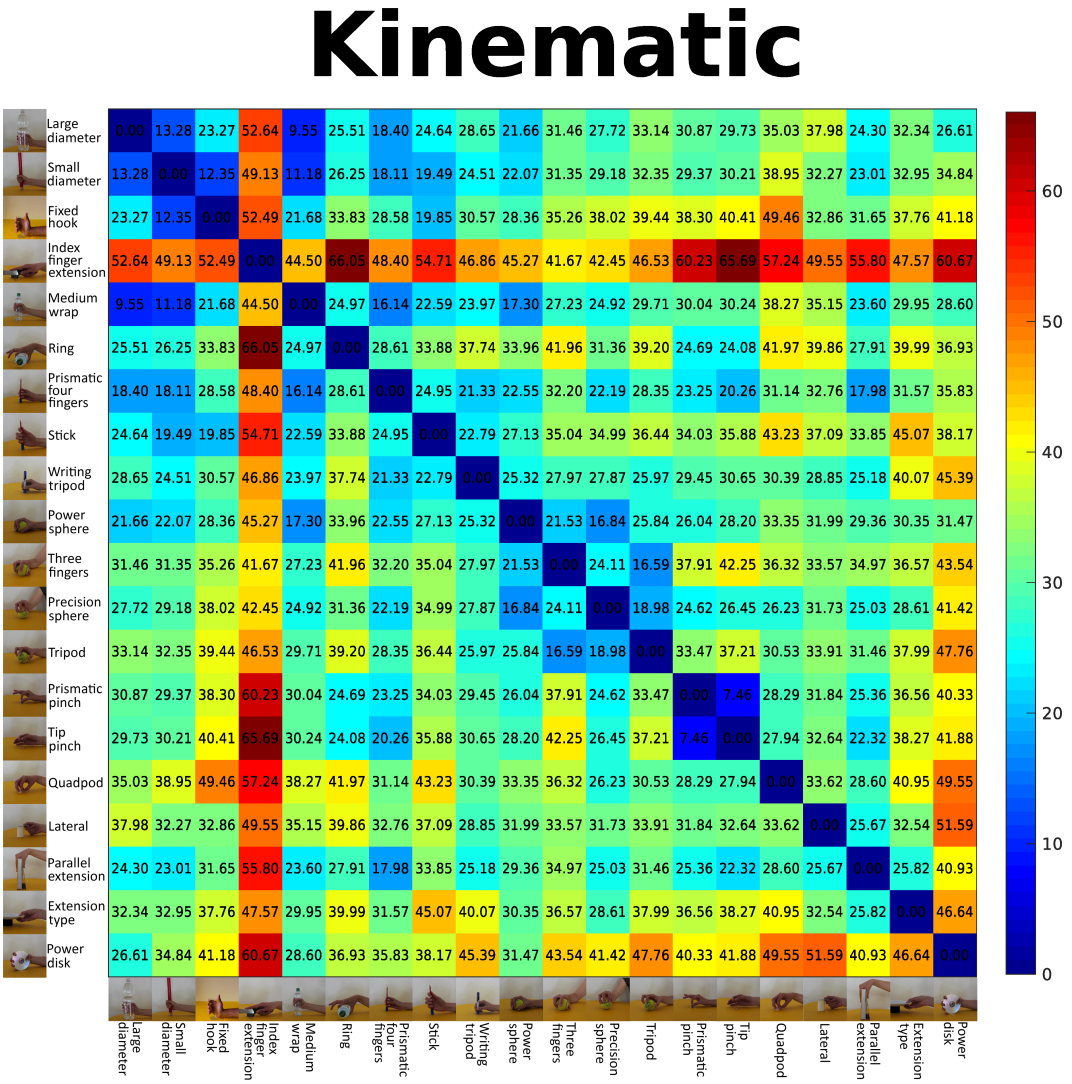


Figure 5.4: Distance matrix computed as the mean Mahalanobis distance between the considered grasps, using the available features computed on the CyberGlove data.

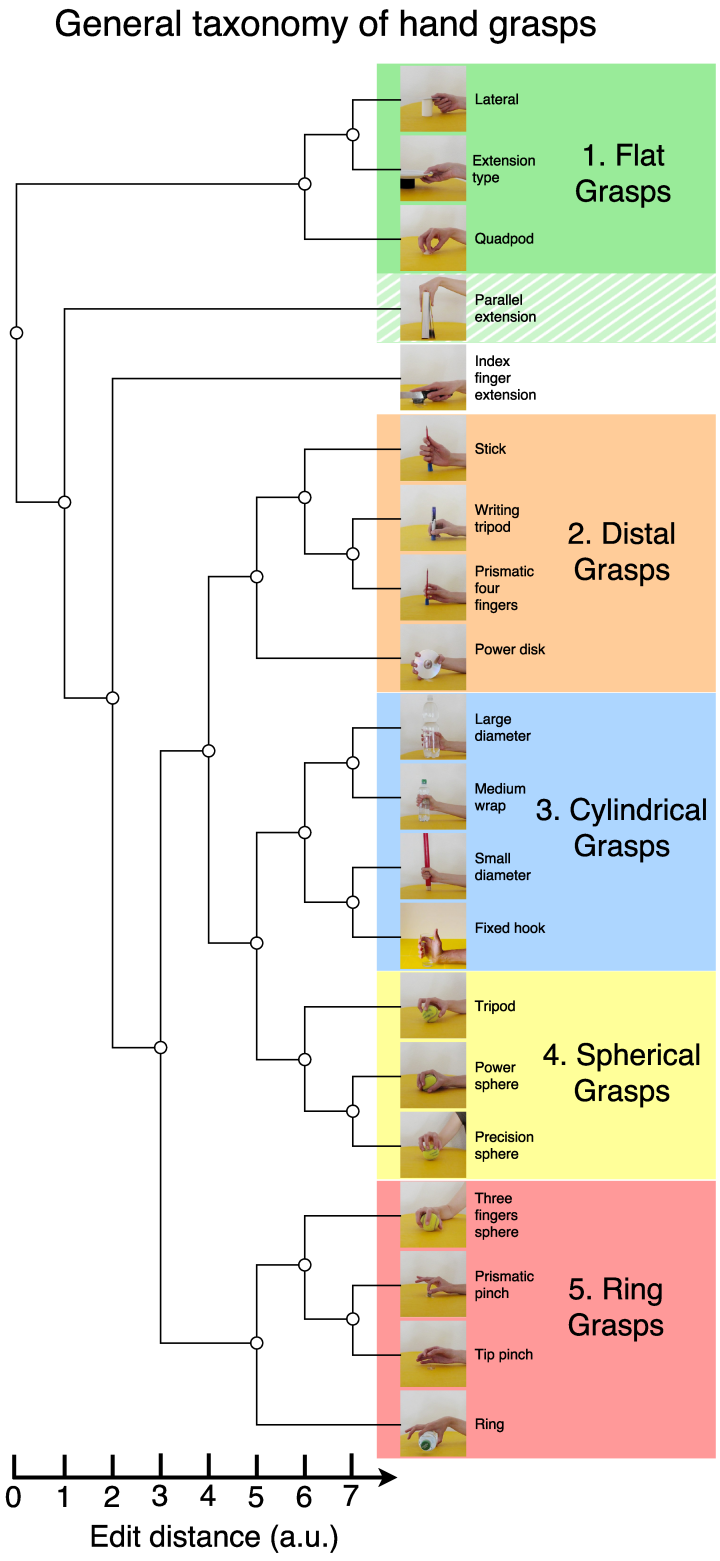


Figure 5.5: General quantitative taxonomy of hand grasps, obtained by muscular and kinematic data. The parallel extension grasp is not part of group 1 but it may be considered as related to it due to the similarity with the extension grasp and its positioning in the muscular taxonomy.

in the same group in the quantitative muscular taxonomy of hand grasps. ***Distal grasps*** are usually characterized by the strong involvement of distal phalanxes, thus of the flexor digitalis profundus. ***Cylindrical grasps*** are strongly linked to the shape of the object. They usually involve palm opposition with both adducted or abducted thumb and virtual fingers 2-5. ***Spherical grasps*** are strongly linked to the shape of the object as well. They involve both pad and palm opposition with virtual fingers 2-3, 2-4 and 2-5. ***Ring grasps*** are almost entirely in accordance with the *GRASP*'s taxonomy grasps with virtual fingers 2. This category includes as well the *three finger sphere* grasp, which is the only power, pad opposition grasp with virtual fingers 2-3 in the *GRASP* taxonomy. The *three finger sphere* grasp is grouped differently within the muscular and the kinematic taxonomy. These facts suggest that in static conditions the middle finger may have an accessory function in the grasp.

Cylindrical and spherical grasps can also be grouped into a macro-sub group at the third level. Qualitative comparison of the results with the kinematic hand grasping synergies [72] highlights an overall similarity between the cylindrical grasps and the first synergy obtained by Santello et al.(closure of finger aperture achieved by flexion at the pip joints of the fingers and thumb adduction and internal rotation) and between the spherical grasps and the second synergy (flexion at the mcp joint and adduction of the fingers).

The main differences between the general and the muscular taxonomies concern the grasps targeting spherical objects. In the general taxonomy, the *power sphere*, the *precision sphere* and *tripod* grasps are grouped together, similarly to what happens in the kinematic one. The *three finger sphere* is closer to *prismatic pinch*, *tip pinch* and *ring* grasp. Similarly to the kinematic taxonomy, the *extension type*, *lateral* and *quadpod* grasps are grouped separately from all the other movements.

Two more important differences between the general taxonomy and the modality-specific ones are related to two movements that have different connections in EMG and glove trees: *parallel extension* and *index finger extension*. The *parallel extension* grasp is grouped with the *lateral* and *extension* grasp in the EMG taxonomy, while it is grouped with the *prismatic four fingers* and the *writing tripod* grasp in the kinematic taxonomy. In the general taxonomy, the *parallel extension* grasp is located alone in its own branch, separated from almost any other grasp and separated from the *extension*, *lateral* and *quadpod* grasp. The *index finger extension* grasp is represented as completely separated from the others in the kinematic taxonomy. In the EMG and in the general taxonomy of hand movements on the other hand, the grasp is still quite isolated but it is grouped with classical grasps (such as the *stick*, *medium wrap* grasp) at a very

high level.

As previously mentioned, the differences between the muscular and the kinematic taxonomies are due to the properties of the movements that the different sensors can highlight. Thus, the general taxonomy of hand grasps provides a unified and general description of all of them.

5.3.4 Comparison with the *GRASP* taxonomy

Comparing the general quantitative taxonomy of hand grasps with previous taxonomies allows evaluating the considerations used to create the previous taxonomies and to better interpret the results achieved. Among the taxonomies presented so far, the *GRASP* taxonomy [67] is a a well accepted taxonomy that represents most of the previous studies and includes all the movements considered in this work. The authors divided the grasps into groups according to four main parameters: 1) power type; 2) opposition type (i.e. the direction in which the hand applies force on the object); 3) thumb position and 4) virtual finger assignments.

The quantitative taxonomy of hand grasps is partially similar to the *GRASP* taxonomy considering the sub-groups determined by the intersection of the *GRASP* parameters. However, it differs in the fact that the parameters considered in the *GRASP* taxonomy are differently (and only partially) represented by the quantitative taxonomy (Figure 5.6).

GRASP Taxonomy Parameters

The ***subdivision according to power type*** (power, intermediate or precision grasps) that strongly influenced the scientific literature in the past is not well supported by the general and modality-specific quantitative taxonomies of hand movements (power, intermediate and precision grasps are usually divided between the five groups presented in this work). This result is also confirmed considering only the quantitative muscular and kinematic taxonomies, that are more similar to the the general quantitative taxonomy. The ***subdivision according to opposition type*** (pad opposition, palm opposition or side opposition) is partially supported by the general quantitative taxonomy of hand grasps (palm and pad grasps are in most cases well divided). The ***subdivision according to thumb position*** (thumb abducted or adducted) is not well supported by the general taxonomy of hand grasps, even if the *index finger extension grasp* is constantly well separated from all the other grasps and the *lateral* and *parallel extension grasp* are grouped. The ***subdivision according to virtual finger assignments*** is well supported

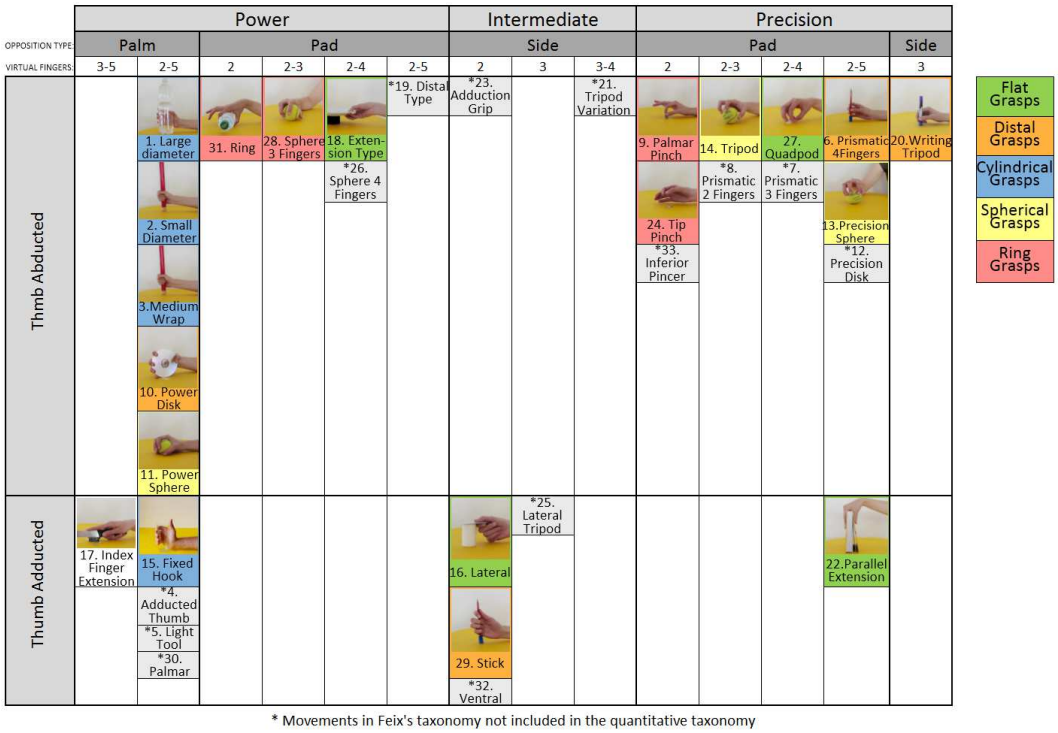


Figure 5.6: Comparison of the general quantitative taxonomy of hand grasps with the GRASP taxonomy.

by the general quantitative taxonomy of hand grasps. The *index finger extension grasp* is constantly well separated from the other grasps, coherently with the fact that it is the only grasp with virtual fingers 3-5 in the *GRASP* taxonomy. The general quantitative taxonomy category "ring grasps" is almost entirely in accordance with the *GRASP* taxonomy groups having virtual finger 2. This category includes as well the *three finger sphere* grasp, which is the only power, pad opposition grasp with virtual fingers 2-3 in the *GRASP* taxonomy. Two grasps of the virtual finger 2-4 category in the *GRASP* taxonomy (*extension type* and *quadpod*) are grouped in the category "flat grasps" of the general quantitative taxonomy. However, grasps that are grouped with virtual fingers 3, 2-3 and 2-5 are often mixed within the general quantitative taxonomy of hand grasps.

Sub-groups determined by the intersection of the GRASP Taxonomy Parameters

The ***thumb-abducted, palm opposition, power grasps*** of the *GRASP* taxonomy are grouped together in the general quantitative taxonomy (the only differences are for the *power disk* and the *power sphere*, that are separated in the in the general quantitative taxonomy). Coherently with the *GRASP* taxonomy, the *fixed hook* is grouped at the

second level with the thumb abducted palm opposition power grasps. The *index finger extension* grasp is separated from almost all the others, somehow corresponding to the *GRASP* taxonomy, in which this movement is alone in the power palm thumb adducted group having virtual fingers 3-5.

Coherently with the *GRASP* taxonomy, the *ring* and the *sphere three finger* grasps are grouped in the quantitative taxonomy. However, these grasps are grouped with the *prismatic (palmar) pinch* and the *tip pinch*, which previously were represented as part of the precision grasps. In the *GRASP* taxonomy, the ***thumb-adducted, side opposition, intermediate grasps*** include the *lateral* and *stick* grasps. These two grasps are not grouped in the general quantitative taxonomy of hand grasps. In fact, the *lateral* grasp is grouped with the *extension type* and the *quadpod* grasp. The *stick* grasp is grouped at the second level with the *writing tripod* and the *prismatic four fingers* and, at the third level, with the *power disk*. In the quantitative taxonomy, the ***thumb abduction, precision group*** of the *GRASP* taxonomy is divided into several sub-groups. The sub-groups are often grouped (also with other movements) according to the shape of the object rather than on the properties previously identified. The *prismatic four finger* grasp is grouped with the *writing tripod* at the first level, the *tripod* and the *precision sphere* are grouped at the second level, the *prismatic (palmar) pinch* and the *tip pinch* are grouped at the first level. This result shows that, on average, the shape of the object influences the positioning of the fingers and the muscular activity more than the usefulness for power or precision tasks and opposition type.

Finally, accordance between the general quantitative and the *GRASP* taxonomy is obtained for the *parallel extension* grasp, that is separated from most of the other movements in both of them.

5.4 Discussion

This work presents to the best of our knowledge the first quantitative taxonomy of hand grasps based on muscular and kinematic measurements of the hand (Figure 5.5). Several taxonomies of hand grasps were presented in scientific literature. All of them are based on rigorous qualitative descriptions of hand movements and valuable intuitions performed by scientists. They are capable to highlight intrinsically important characteristics of hand movements, nevertheless, a qualitative analysis is prone to subjectivity and it does not allow a demonstrable confirmation or refutation offered by quantitative methods [86]. The quantitative taxonomy is based on solid experimental measurements and statistical data analyses. The analysis is limited to the considered

data and time domain features, that determine the organization of the taxonomy. Further analysis (including other data and signal features, such as for instance frequency based features) is considered in future work. Such analyses may be able to highlight deeper or different relationships between muscular and kinematic properties of hand movements. This work sets the basis for such work by providing a quantitative description of the hand movements that are divided into five main groups, as presented in Figure 5.5. The general quantitative taxonomy of hand grasps is based on two modality specific taxonomies (based on EMG and kinematic data Figure 5.2). The results are interesting both considering the modality-specific and the general quantitative taxonomies.

The modality-specific taxonomies provide very similar representations of the hierarchical organization of hand movements (edit distance = 33), thus validating each other. The similarity between the muscular and the kinematic taxonomy confirms the existence of strong relationships between the muscular activity and the actual motion of the hand, as expected by anatomy. Small differences between the muscular and the kinematic taxonomies exist. Such differences can be due to the differences in the techniques used to record the data. The EMG based taxonomy considers the force exerted and the motion of the wrist, while the taxonomy based on the data glove does not consider these parameters. Nevertheless, these differences are in general small compared to the similarities and they can possibly be reduced by considering weights when merging hierarchical trees, as explained in subsection 5.2.3. Another possible source of difference can be related to non-linearities existing between some joint angles in the CyberGlove sensor output (e.g. abduction/adduction at the metacarpophalangeal joints) [119, 120]. Although this aspect can affect the kinematic data, the kinematic taxonomies are based on grasp similarities in the kinematic feature space (that take into account the distribution of the data) and not directly the joint angles, probably contributing to the similarity with the muscular taxonomy.

Several parameters can be used to quantitatively characterize hand grasps. This work considers kinematics and muscular activity in order to target the posture of the hands and due to practical data availability. The consistency between the muscular and kinematic taxonomy enforces the usefulness and reliability of the results. The analysis of other parameters can definitely be interesting and should be considered in follow-up work. The edit distance boundaries depend on the number of nodes (thus on the number of considered grasps). Intuitively, the larger the number of classes, the higher the possible number of variations that can occur between different trees. This fact can be one of the reasons behind the discrepancy obtained between the modality trees.

Future work should address this fact in detail, by applying additional or alternative operations, measures or approaches.

Depending on the domain, one specific taxonomy may be more useful than the other. While the kinematic taxonomy may be more useful for robotics, the muscular taxonomy may be more suitable for applications in prosthetics. Both taxonomies can have applications in rehabilitation, physiology and neuroscience. The general taxonomy aims at providing a solution that is intermediate to the different fields, allowing (and hopefully fostering) the collaboration among them on the basis of the first set of quantitative results in this challenging domain. The general quantitative taxonomy provides a comprehensive quantitative representation of hand grasps, overcoming the subjectivity of the taxonomies previously presented in literature and the limitations of the muscular and the kinematic taxonomies presented in this thesis. The general quantitative taxonomy suggests a division into five groups of grasps that were named after specific properties of each group: 1) flat grasps; 2) distal grasps; 3) cylindrical grasps; 4) spherical grasps; 5) ring grasps. Cylindric grasps and spherical grasps can also be grouped into a macro-sub group at the third level.

The division in categories is arbitrary, made in order to facilitate the comparison with previous taxonomies and to provide a further synthesis of the taxonomy. Future work could benefit from including quantitative approaches to perform the division in categories.

The comparison of the general quantitative taxonomy of hand movements with previous taxonomies is important because it allows to validate the parameters on which the previous taxonomies were based and to better interpret the achieved results. The *GRASP* taxonomy [67] represents a proper reference for the comparison because it is one of the most recent qualitative taxonomies of hand grasps and because it is based on the comparison of several previous taxonomies. The quantitative approach only partially confirms the parameters used to create the previous taxonomies (and thus the *GRASP* taxonomy), while it enforces movement groups defined on the basis of real finger positioning and muscular activation, reflecting often the shape of the grasped object and balanced combinations of parameters rather than specific single qualitative parameters. The intersections of different parameters in the *GRASP* taxonomy are partially similar to the general quantitative taxonomy of hand grasps. However, there are still important differences (Figure 5.6). Considering each parameter separately, some of the qualitative *GRASP* parameters are not well represented in the quantitative taxonomy and some others are predominant in a few categories. In particular, the subdivision of hand grasps according to power (which strongly influenced the scientific

literature in the past), is not well supported by the general quantitative taxonomy of hand movements, while the subdivision into opposition and virtual finger assignments are usually better represented in the general quantitative taxonomy (in particular for specific groups, such as ring grasps). We offer two possible interpretations of these results. First, human intuition and perception enrich previous taxonomies with alternative perceptions of the grasps, such as their usual aim, that is separated from a strictly kinematic or muscular representation of the grasps. Second, it can be important that the parameters of previous taxonomies are considered but it is not easy to balance and weigh the parameters in each movement and category properly only on a qualitative basis.

The hierarchical model of human manipulation and grasping improves several fields (including robotics, prosthetics, rehabilitation and physiology) with the quantitative analysis of relationships that were previously widely described in literature on the basis of qualitative parameters. In robotics, the five categories of movements defined in Figure 5.5 can help to describe and plan robotic hands according to a clear, solid and simple modular definition of movements. Moreover, the taxonomy can provide a priori information to improve classification algorithms, as proposed in [121]. In prosthetics, the general quantitative taxonomy can foster the development of prosthetic hands that are more suitable for real life situations in terms of both control and mechanical design. For instance, the five categories of movements can be compared with the mechanical properties of the prosthesis, as well as with the ADLs and the movements mostly needed by hand amputees in order to develop modular control systems based on the movement categories. In rehabilitation, the presented taxonomy of hand grasps can improve planning with a better scheduling that prioritizes the categories of movements that are more useful (or more realistically achievable) and thus need to be restored earlier. In recent years, hand synergies gathered importance in physiology, bioengineering, rehabilitation and robotics [72]. Comparing the quantitative taxonomy of hand movements with the hand synergies can highlight relationships between the two. The cylindrical grasps look qualitatively similar to the first kinematic hand grasping synergy obtained by Santello et al. [72], characterized by the closure of the finger aperture achieved by flexion at the pip joints of the fingers and thumb adduction and internal rotation. The spherical grasps look qualitatively similar to the second kinematic hand grasping synergy (flexion at the mcp joint and adduction of the fingers). These considerations provide a coherent relationship between the hand synergies and the quantitative taxonomy approaches.

In conclusion, this work presents the first quantitative taxonomy of hand grasps

based on muscular and kinematic data. The taxonomy clarifies with a solid quantitative approach what was proposed in the field so far based mainly on qualitative assumptions, thus unifying the diverse perspectives presented and offering a scientific reference for the taxonomies of hand grasps. The results were compared with previously presented taxonomies of hand grasps, improving them and clarifying the parameters used to define them. They appear at a first qualitative inspection in accordance with hand synergy studies.

5.5 Quantitative Taxonomy of Hand Grasps for Amputated Subjects

The previous work have been extended to amputated subjects. All the signal processing is the same of the previous section, but it has been applied to a different dataset. In particular, we considered a different dataset from the NinaPro projects. The data comes from 11 amputated subjects (11 males; 10 right handed, 1 left handed; age 42.36 $\pm$ 11.96 years), performing the same exercises of the healthy ones. Detailed information about the subjects, the movements and the acquisition setup can be found in [29]. Amputated subjects wear the sEMG electrodes on the stump, while the dataglove and the inclinometer were placed on the contralateral hand.

It is worth to notice that for amputated the generalization among subjects is even more challenging. As a matter of fact, many factors contributes to the quality of EMG signals, like:

- Years since the amputation.
- Cause of the amputation.
- Phantom limb sensation.
- Amputation level.
- Usage of other prosthesis.

Nevertheless, the needs of a subject-independent solution is even more important, since the major stress for the training procedure with respect to healthy subjects.

The study of a quantitative taxonomy of hand grasps, which considers both muscular and kinematic aspects, and its comparison with the one built on healthy subjects would highlight the differences and the common behaviour among these two kind of subjects. A coherent taxonomy for amputated subjects suggests a similarity with

healthy subjects, despite the physical and physiological differences, thus validate the subject-independent framework for amputated.

Figure 5.7, Figure 5.8, Figure 5.9 represents the quantitative taxonomy of hand grasps for amputated subjects. Figure 5.7 and Figure 5.8 shows the muscular and the kinematic taxonomy respectively, while Figure 5.9 represents the general taxonomy, that exploits both EMGs and glove data.

5.5.1 Analysis of Amputated Taxonomy

The obtained trees are slightly less balanced than the ones built on data from healthy subjects. It is interesting to study the analogies between the muscular and the kinematic taxonomy. For healthy subjects the similarity was high, confirming the connection among EMG signals and muscular activation. The results for amputated subjects are less robust, thus the damaged muscles are not able to control the muscular activation as profitably as for healthy persons. Further studies are needed, but patients who lost their forearm many years ago, and without an efficient rehabilitation are more inclined to lose their ability to control a limb. Nevertheless results are promising, since many movements are organized in the tree in the same way.

Comparison between Muscular and Kinematic Taxonomies for Amputated Subjects

For amputated subject the connection between muscular and kinematic taxonomies is not trivial. As a matter of fact, the kinematic is recorded from the only forearm still present, while EMG signals are recorded from the residual limb. For these persons is particularly complex try to perform a movement from a limb that does not exists, and this can also lead to frustration and stress, but mostly the movements and the muscular activation does not come from the same limb. For these reasons, the presence of common subtrees among these two taxonomies is very interesting. In particular *Tripod* and *Precision sphere* movements are very close in both the taxonomies. Also *Prismatic pinch*, *Quadpod*, and *Tip pinch* are strictly connected in both the taxonomies, but the subtrees containing them changes:they are considered similar to *Lateral grasp*, *Power disk* in the muscular taxonomy, while to *Tripod*, *Precision sphere*, *Power sphere*, and *Three fingers grasp* in kinematic taxonomy. Other similar movements are *Lateral grasp*, *Extension type*, and *Power disk*, but in this case too they are organized differently in the two taxonomies: in the kinematic taxonomy they are in the same subtree as *Parallel extension*, but that is not true for muscular taxonomy. On the contrary, as

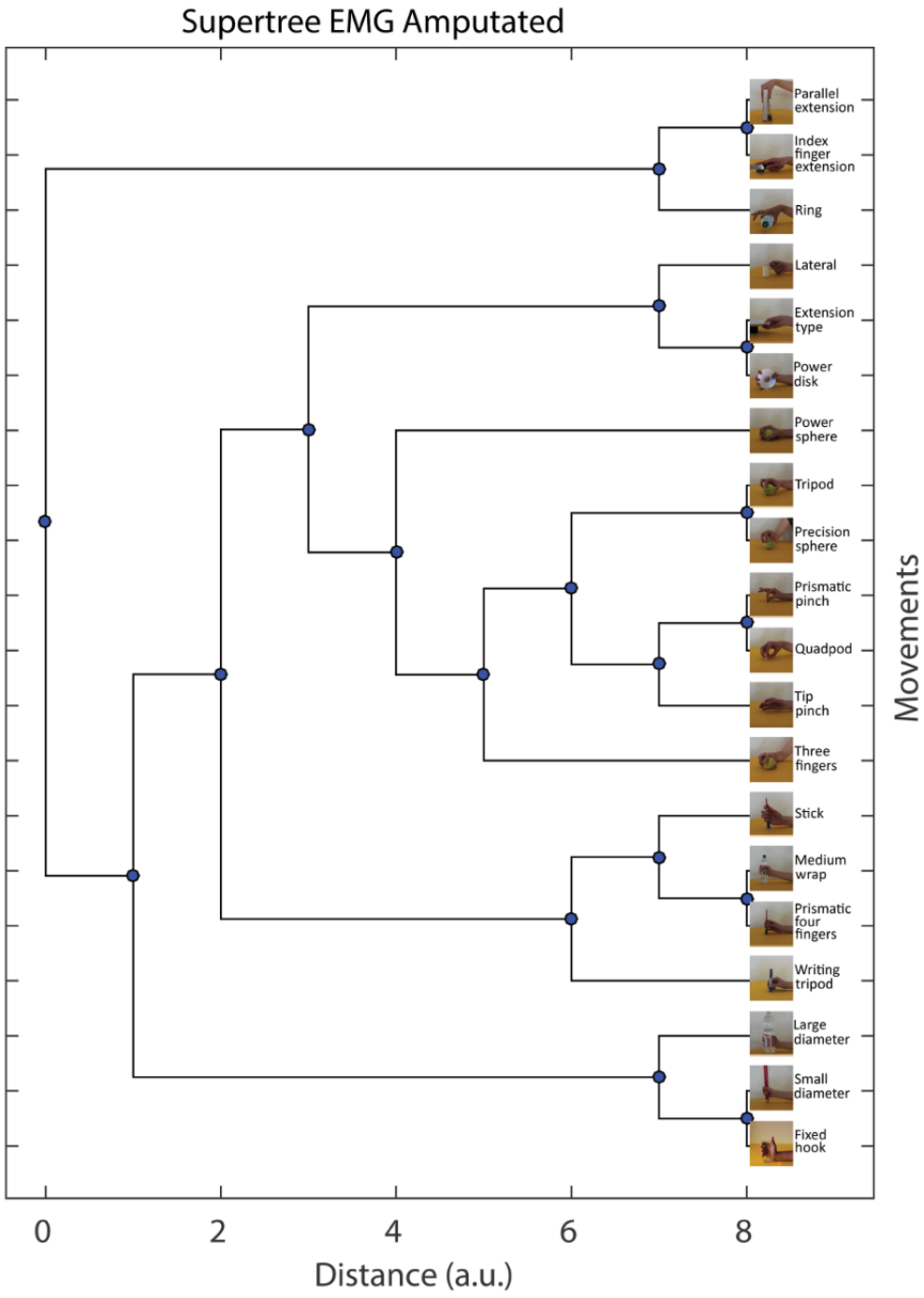


Figure 5.7: Quantitative taxonomy of hand grasps for amputated subjects built on EMG signals, thus representing the muscular activation.

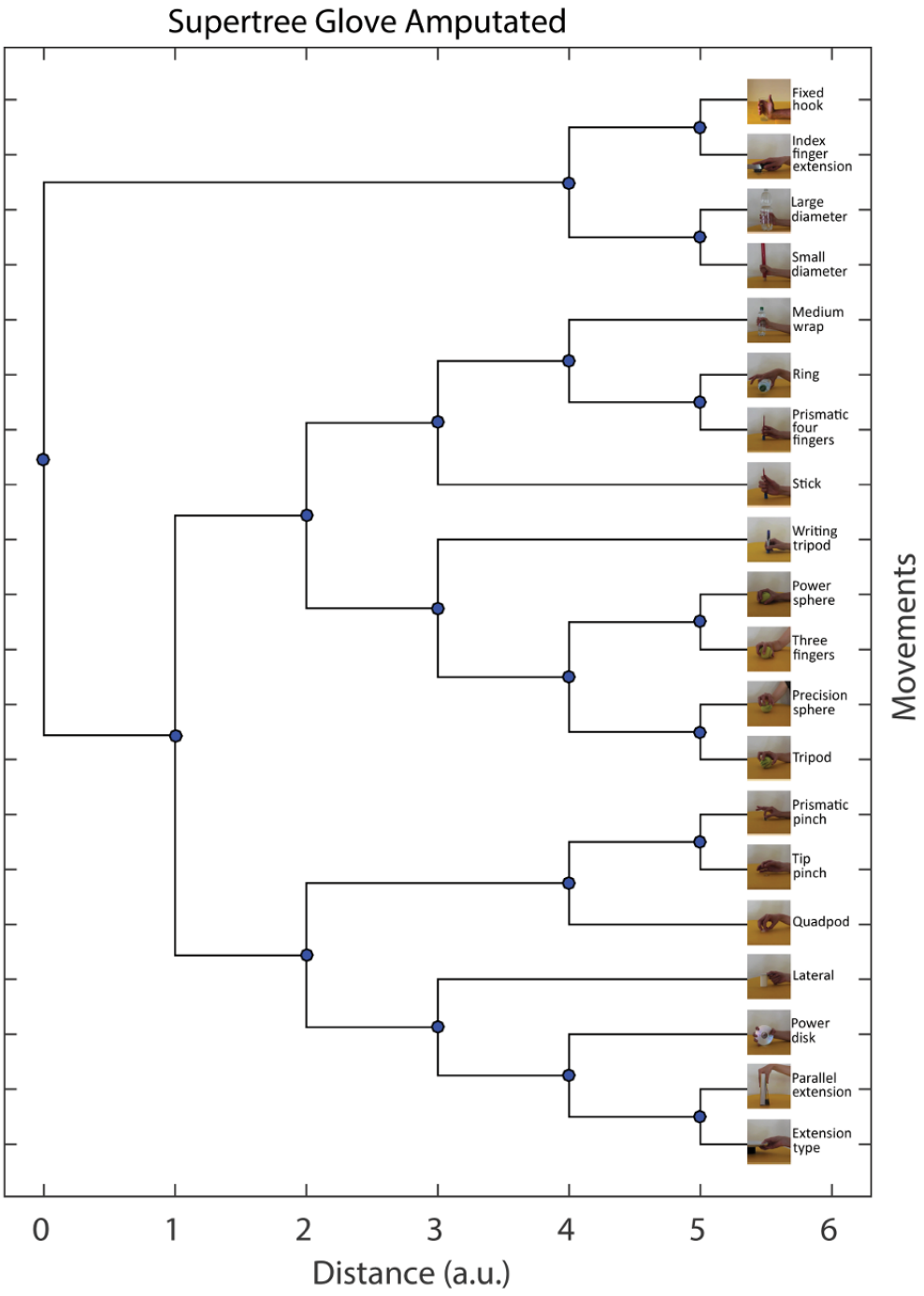


Figure 5.8: Quantitative taxonomy of hand grasps for amputated subjects built on glove information, thus representing the kinematic of grasps.

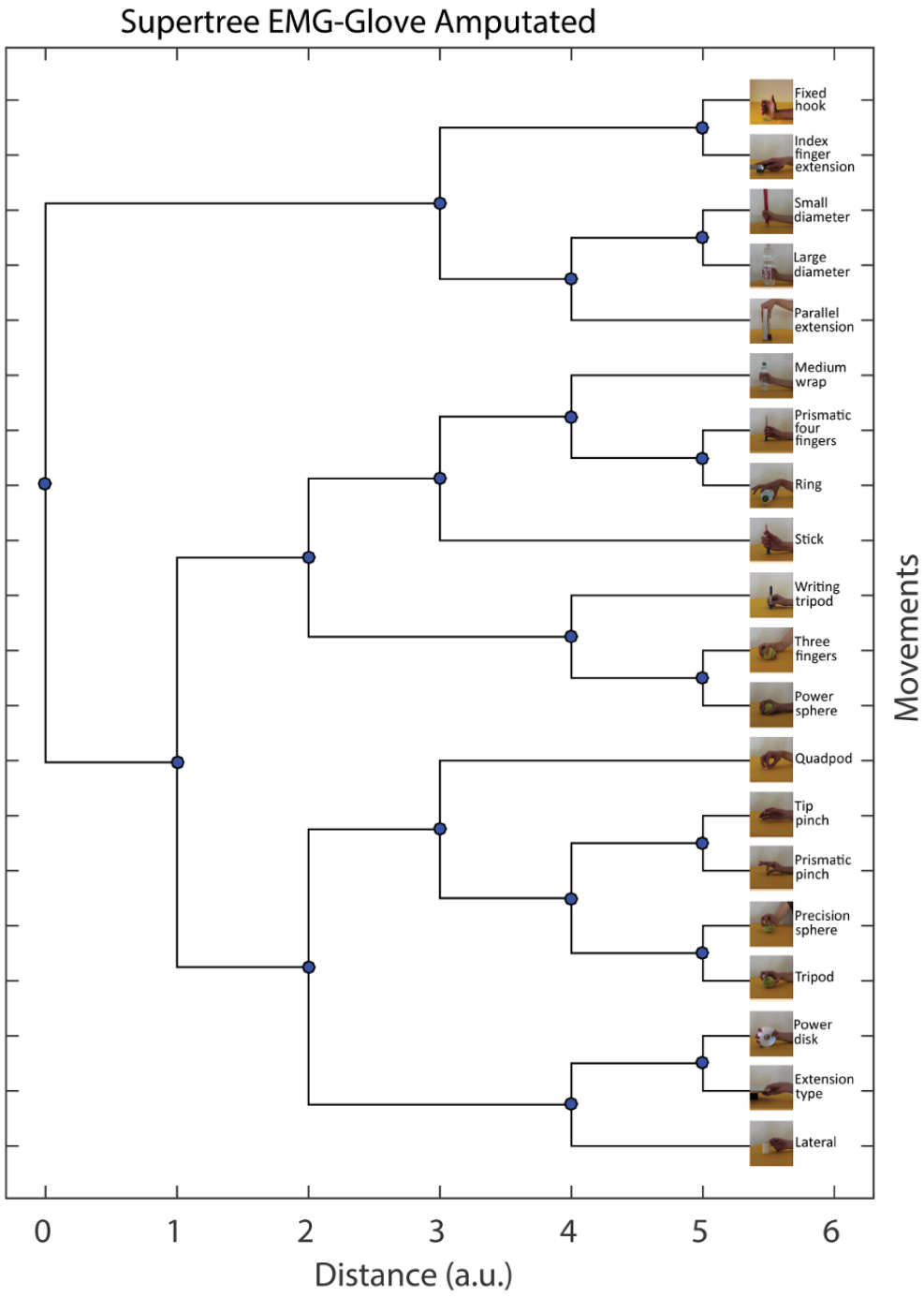


Figure 5.9: Quantitative taxonomy of hand grasps for amputated subjects built considering both muscular and kinematic information.

previously said, from a kinematic perspective these movements are in the same subtree as *Lateral grasp*, divided from "spherical" grasps, like *Tripod grasp*. Furthermore, *Medium wrap*, and *Prismatic four fingers* are close in both the taxonomies, but their belonging subtrees differs from the presence of *Ring grasp* exclusively in the kinematic taxonomy.

Nevertheless, both the trees have the same general structure: they have two subtrees, one rich of movements, while the other less populated, with four movements maximum (*Ring grasp*, *Index finger extension*, *Parallel extension* in the muscular taxonomy, *Fixed hook*, *Index finger extension*, *Large diameter grasp*, *Small diameter grasp* in the kinematic taxonomy). The remaining subtree is very balanced in the kinematic taxonomy, but is much less regular in the muscular taxonomy, probably a limitation given by the limb amputation.

Analysis of General Taxonomy for Amputated Subjects

The fusion of the two taxonomies overcome their limitations, by creating an unique, general tree, that incorporates the information of both the muscular and the kinematic perspective. The structure of the general taxonomy is similar to the two specific ones, thus it is composed by two subtrees, the smaller one composed of only five movements: *Fixed hook*, *Index finger extension*, *Large diameter grasp*, *Small diameter grasp*, and *Parallel extension*. *Large diameter grasp* and *Medium wrap* are separated, like happened in the kinematic taxonomy, nevertheless the result of the muscular taxonomy that groups them together seems more reasonable from a qualitative analysis. An other aspect that reflect what happens in the kinematic taxonomy is the fact that the larger subtree is well balanced. Furthermore, from a qualitative analysis emerges a correct separation of "palm grasps" (*Medium wrap*, *Ring grasp*) from "finger grasps" (*Quadpod*, *Tip pinch*), and a proper grouping of *Lateral grasp*, *Extension type*, and *Power disk*.

5.5.2 Comparison between Amputated and Healthy Subjects Taxonomy

The comparison of the results obtained by healthy and amputated subject is extremely useful in order to understand the characteristics of these groups of users, and most of all to find similarities which can suggest the feasibility of a subject-independent approach for injured persons.

Muscular taxonomies

The comparison of these taxonomies would intuitively give the most different results since EMG signals of non intact muscles from amputated subject could be extremely different from the ones obtained from healthy subjects. Despite the great differences among these groups of subjects, there are some comparable results in the two taxonomies. It is worth to notice that the healthy tree is more balanced, nevertheless also in this case there is a largely populated subtree and an other one with only three movements, but only *Parallel extension* is shared between the two trees. The larger subtree is more balanced for the healthy people, excluding *Index finger extension*. It is interesting to notice that this movement is organized in the smaller subtrees for amputated. The group of "spherical" grasps is organized in the same way for healthy and for amputated subject, with the only exclusion of *Ring grasp*, which is present only in the healthy subtree. However, a limitation of amputated taxonomy emerges for the movements *Large diameter*, *Medium wrap*, *Small diameter* and *Fixed hook*: as a matter of fact they are grouped together in the healthy taxonomy, and this seems correct from a qualitative analysis, but they are separated in the amputated taxonomy.

Kinematic taxonomies

It is interesting to notice differences among these taxonomies too, even if the amputated subjects performed the task with their intact limb. What emerges from an accurate analysis is that the healthy supertree is less balanced, mainly for the presence of *Index finger extension* movement completely isolated from the rest of movements, while for the amputated taxonomy this movement is grouped with *Fixed hook*. A similar behaviour involves *Power disk*, since for healthy subjects this movement is isolated, on the contrary then for amputated. The movements *Stick grasp*, *Small diameter*, *Fixed hook*, *Large diameter*, *Medium wrap* are grouped in the same subtree for healthy subjects, but there is not a coherent grouping in amputated taxonomy. Despite the differences, there are some common aspects among the taxonomies, for example the group composed by *Three fingers sphere*, *Tripod*, *Power sphere*, and *Precision sphere* is organized in the same way in the two taxonomies. Also *Prismatic pinch*, *Tip pinch*, *Quadpod*, *Extension type*, are grouped together in both the taxonomies, but for healthy subjects they are strongly related with *Ring grasp*, while a strong connection is highlighted with *Power disk* and *Parallel extension* for amputated subjects.

General taxonomies

In the comparison of the general taxonomies, it is possible to notice in the amputated one the absence of the class of movements highlighted for healthy subjects.

1. **Flat grasps** is present in both the taxonomies, with the exclusion of *Quadpod grasp*.
2. **Distal grasps**: the majority of movements are relatively close, with the exclusion of *Power disk*, which is in a completely different position.
3. **Cylindrical grasps**: in amputated taxonomy *Medium wrap* is substituted by *Index finger extension*.
4. **Spherical grasps**: in amputated taxonomy *Power sphere* is missing.
5. **Ring grasps**: only *Tip pinch* and *Prismatic pinch* are close in the general taxonomy for amputated.

5.6 Conclusions

In conclusion, this work presents the first quantitative taxonomy of hand grasps based on muscular and kinematic data. The taxonomy clarifies with a solid quantitative approach what was proposed in the field so far based mainly on qualitative assumptions, thus unifying the diverse perspectives presented and offering a scientific reference for the taxonomies of hand grasps. The results were compared with previously presented taxonomies of hand grasps, improving them and clarifying the parameters used to define them. They appear at a first qualitative inspection in accordance with hand synergy studies.

The analysis of taxonomy for amputated subjects highlight some interesting results. Generally the taxonomy is quite different from the one obtained on data from healthy subjects, and also from Feix's one. The results seems less coherent even from a qualitative analysis. Particularly interesting are the different results achieved for the kinematic taxonomy. This was unexpected, since the amputated subjects were using their intact limb for performing the tasks. These differences suggest a modification in the way subjects perform daily grasps due to the fact that they are forced to use only one hand.

Table 5.2: Intra-subject variability in grasps across the different quantitative metrics expressed as edit distance. Rows represent subjects while columns represent the different modalities.

	Muscular	Kinematic
Subject 1	4.00 ± 2.65	1.44 ± 1.33
Subject 2	2.48 ± 1.75	1.28 ± 1.87
Subject 3	0.96 ± 1.00	2.40 ± 2.59
Subject 4	2.08 ± 1.44	0.96 ± 1.40
Subject 5	3.04 ± 2.05	1.60 ± 2.33
Subject 6	3.84 ± 2.49	1.44 ± 1.65
Subject 7	2.96 ± 2.14	1.20 ± 1.30
Subject 8	2.24 ± 2.03	1.76 ± 2.10
Subject 9	4.32 ± 2.78	0.64 ± 0.93
Subject 10	3.12 ± 2.14	1.28 ± 1.43
Subject 11	2.80 ± 2.10	0.96 ± 1.40
Subject 12	0.32 ± 0.47	2.24 ± 1.99
Subject 13	2.48 ± 2.08	1.60 ± 1.26
Subject 14	3.12 ± 2.07	0.64 ± 0.93
Subject 15	2.72 ± 1.87	1.92 ± 1.74
Subject 16	3.36 ± 2.31	1.60 ± 1.88
Subject 17	2.16 ± 1.67	1.76 ± 1.73
Subject 18	2.96 ± 1.91	1.28 ± 1.87
Subject 19	1.44 ± 1.33	0.96 ± 1.40
Subject 20	2.32 ± 1.64	1.28 ± 1.87
Subject 21	2.72 ± 1.95	2.00 ± 1.81
Subject 22	2.24 ± 1.63	2.56 ± 2.30
Subject 23	2.32 ± 1.59	2.56 ± 2.30
Subject 24	1.92 ± 1.32	1.68 ± 1.49
Subject 25	2.96 ± 1.87	0.96 ± 1.40
Subject 26	3.52 ± 2.28	1.28 ± 1.87
Subject 27	2.80 ± 1.90	1.04 ± 0.96
Subject 28	3.12 ± 1.99	0.64 ± 0.93
Subject 29	3.28 ± 2.41	2.24 ± 1.99
Subject 30	3.68 ± 2.78	1.60 ± 2.33
Subject 31	2.72 ± 2.43	2.40 ± 2.30
Subject 32	1.84 ± 1.51	0.64 ± 0.62
Subject 33	3.28 ± 2.20	1.04 ± 0.96
Subject 34	2.00 ± 1.36	1.28 ± 1.87
Subject 35	2.48 ± 1.75	1.92 ± 2.80
Subject 36	2.56 ± 2.30	0.96 ± 1.40
Subject 37	0.32 ± 0.47	1.28 ± 1.87
Subject 38	2.56 ± 1.70	2.08 ± 2.15
Subject 39	2.24 ± 1.58	1.28 ± 1.87
Subject 40	3.36 ± 2.57	2.16 ± 2.27

Table 5.3: Edit distance between the modality-specific taxonomies of hand grasps and each modality feature supertree

	IAV EMG	MAV EMG	RMS EMG	TD EMG	WL EMG
Muscular Tax.	22	22	34	24	39
	IAV glove	MAV glove	RMS glove	TD glove	WL glove
Kinematic Tax.	19	31	0	34	26

Chapter 6

Taxonomy-Based Classification

New and innovative prosthesis devices can improve the life quality of amputated subjects, helping them to interact with the surrounding world. Hands are one of human principal ways to interact with the surrounding world, and in particular with objects. The three most important hand functions are to explore, to restrain and to manipulate objects. Indeed, the several ways of grasping an item are part of the ADL that allow us to interact naturally and easily with the world. These activities are crucial for each individual and a great effort has been put to gain back lost functionality for injured, amputated, or impaired subjects. Everyday hands grasp an high number of objects with different shapes, dimensions and weight. These characteristics lead to disparate ways of interacting with objects by using different grasping approaches. The scientific community have selected a number of essential grasps, organizing them in taxonomies depending on several qualitative factors. Many different taxonomies have been proposed during the years, considering different aspects of hand motion. One of the most complete and detailed taxonomy has been proposed by Feix et al. in [67] by arranging together many different proposals. The result, namely the GRASP taxonomy, organizes a matrix of 33 hand grasps, arranging the movements according to qualitative force parameters and finger positions. Nevertheless, a qualitative approach can reduce ambiguity, since it depends on measurable and repeatable experiments. For this reason we exploited the quantitative taxonomy of hand grasps (see Chap. 5) as a guideline. In particular, the previously described taxonomy has been reduced extracting 8 significant hand grasps. The movements are arranged in a binary tree, where movements belonging to the same subtree have a common underlying behaviour. The taxonomy have been built by considering data from many different subjects, in order to obtain a robust and general solution. The considered taxonomy is based on quantitative parameters, i.e. it includes the information from hand and fingers kinematics, and the muscular

activation during the motion. In particular, the muscular activation is represented by sEMG signals. These signals are also commonly used for the control of prosthesis, we considered sEMG signals from the subjects with the goal of exploiting the physiological behaviour in order to emulate what happens in the human body. sEMG signals have the drawback of being non stationary, and very sensitive to the physical and physiological state of the subject. In particular, studies shown a sensibility for muscular fatigue, as well as for stress and physical weariness [122]. These characteristics result in a great variability of the sEMG signals, even if they are collected from a single subject in a limited period of time. As a consequence, the traditional approach when using sEMG signals is to focus on the signals of a specific subject, without mixing signals from several subjects. Nevertheless, despite the prevalence of subject-specific approaches, in literature there are some examples of subject-independent approaches. An extended literature review can be found in Chap. 2, here we report only some examples of the techniques proposed during the years to establish a robust interaction with prosthesis devices [16] [17]. Castellini *et al.* [22] developed a subject-specific (train and test data from the same person) and a cross-subject (train data from one person and test from another) analysis in both controlled (baseline) and ADL conditions. Matsubara and Morimoto [6] implemented a bilinear model able to reach an accuracy of 73% by classifying different movements in a multi-user context.

In these studies, the fusion of signals from several different subjects produces generally slightly worst performances, but has the vantage of obtaining a more robust and general model. Thanks to this generalization, a new subject can use the framework without the needs of long training phases, or with no training phase at all. In fact, the model has larger variability, to let it embody a wide number of possible subjects behaviours. Previously, we described a set of additional tools capable of conforming EMG signals by means of an online preprocessing phase (Chap. 3). The necessity of such tools is more relevant for subject-independent approaches rather than subject-specific ones due to the greater signals variability. The crucial and fundamental role committed to the preprocessing phase is to even differences between signals due to noise, muscular and physical fatigue, while highlighting the intrinsic characteristic of the considered movement.

Besides the preprocessing phase, better results can be obtained by implementing new classification techniques. In classification frameworks, only a limited subset of movements are considered. In [25] Yang et al. studied how the composition of a movement affect the movement classification, with the aim of using this technology in real ADL. They proved that better results can be achieved by including dynamic arm

postures and varying muscular contractions in the training phase. Khushaba et al. [1] worked on a similar problem, focusing on individual and combined fingers control rather than on fixed, rough movements. They used a Bayesian data fusion postprocessing approach to maximize the probability of correct classification, obtaining an accuracy of about 90%. Abdullah et al. [123] asserted that the great variability of EMG signals affects the classification performances. The Wavelet Packet Decomposition (WPD) has been used for features extraction from surface EMG signals, with the vantage of avoiding EMG variations. Random Forest, Rotation Forest and MultiBoost have been used for classification, reaching a classification accuracy of 92.1% with Random Forest. Chan et al. [124] developed a fuzzy approach to classify single-site EMG signals. The results have been compared with a neural network-based approach, obtaining better results since the fuzzy system gave more consistent classification results being insensitive to over-training. Tang et al. [125] raised the problem of multiple hand motion identification, since error rate increases with the addition of hand motions, producing more overlapping areas for the projecting features. They proposed a classifier dividing the classification procedure into several levels, where at each level different features are located and projected in different spaces, obtaining an accuracy greater than 89% for identifying 11 gestures.

Our previous studies [126] [127], described in detail in Chap. 4, shown good results in the continuous online estimation of both upper and lower limb movements, considering up to 40 different subjects. Furthermore, we used a freely available online dataset (Chap. 2), the data was not registered ad-hoc for our intentions assuring general results, and the availability of the dataset makes the experiments comparable and reproducible. The studies focused on subject-independent regression solutions for the continuous estimation of several joints for a fixed movement. In this work, we aim to approach to hand grasp classification by using a subject-independent framework. In order to be able to control a prosthetic device we need a twofold approach:

- obtaining good prediction of the movement, in order to be able to reproduce correctly the requested movement;
- being able to process the information and the prediction in a short time, in order to be able to work online.

We propose to exploit prior information regarding the tasks to refine and fasten the classification process. The approach is based on a hierarchical cascade-based classification technique by following the idea from Tang et al. [125], to both reduce elaboration time and improve the prediction accuracy. The classification algorithm is developed to re-

duce the number of movement classes at which the input samples could belong while descending into the depth of a hierarchical tree. Thus, comparisons are made only within a subset of the initial movements representing the classes. This subset is continually refined and reduced until arriving at the classification between two movements, and eventually to a final decision. The binary classification follows the structure of the quantitative taxonomy of hand grasps previously introduced. This means that the classification will be repeated among the taxonomy subtrees. With this approach we divide the classification in levels of increasing complexity, by following the taxonomy branches structure. In particular, level 0 indicates a classification among a couple of leaves, while higher levels classify among groups of movements. The basic idea behind this approach is that close movements in the taxonomy are more hard to distinguish than distant ones. At the same time, a misclassification between two movements closed one to the other is less questionable than a misclassification between two distant motions. Moreover, this approach is easily to generalize and scale. New movements can be easily added to the classfication framework if they are in the taxonomy structure.

6.1 Signal analysis

The preprocessing phase is designed to enable the online elaboration of the input signal. Data used in this study come from the NinaPro dataset, a freely online available dataset [28]. Information has been collected from N= 40 healthy subjects performing 8 different hand grasps, each movement was repeated 6 times. The dataset contains both muscular activation and physical movements, but we focused only on the physiological signals from 8 sensors placed around the forearm.

The preprocessing has the role to detect a clear underlying behaviour between the considered signals. The process provides conditions for good classification performances. This goal is even more important to reach since we are considering signals from a large number of different subjects. As described in Chap. 3, we started by applying the WT [44] to the signal in order to obtain an analysis in time and frequency, thus removing the dependence from time and granting online processing. We used the db2 mother wavelet from the Daubechies family and MAV as synthesis function, applying the WT to consecutive windows of 200 samples. Finally, data has been smoothed by applying a moving average lowpass filter and normalized, in order to regularize the output.

6.1.1 Gaussian mixture model and classification

The EMG signals from a selected number of channels plus the additional information of the movement type, have been used to train offline a probabilistic model, namely a GMM. The GMM approximates the input by using a weighted sum of K Gaussian components which better represent the input data, used to train the model. A complete overview of the GMM can be found in Chap. 3.

After training the model, we proceeded with the online classification phase. We will perform tests with three methods in order to compare the use of preliminary information from the taxonomy in different contexts. The classification techniques have been described also in Chap. 3. makes us predict the kind of movement the subject is performing starting from EMG signals from the subject's muscles.

6.1.2 Taxonomy

To simplify the work, we considered a subset of 8 movements extracted from the complete taxonomy. The movements have been chosen among classical every day grasps to interact with common objects. In particular, we considered the following movements: *Index Finger Extension*, *Medium Wrap*, *Prismatic Four Fingers*, *Stick*, *Writing Tripod*, *Power Sphere*, *Extension Type*, and *Power Disk*. The taxonomy for the considered subset of chosen movements is represented in Fig. 6.3.

6.2 Results

We tested the framework with a Leave-One-Out approach. The model has been built on N-1 subjects and tested on the remaining one. The whole process has been repeated for all the N subjects. The testing phase has been organized in levels of increasing complexity. *Level 0* includes the classification between couples of movement, i.e. *writing tripod* and *prismatic four fingers*, *medium wrap* and *power sphere*. *Level 1* rises of one level in the binary tree and compares a couple with a single movement. Particularly, the movements of the previous level are considered as the same "movement group". Therefore, in *level 1* we classified between *stick* vs. *writing tripod* plus *prismatic four fingers*, and so on. A graphical example of this procedure is depicted in Fig. 6.2.

Fig. 6.2 summarizes the achieved results, showing the accuracy obtained during the tests. In particular, we computed the accuracy for each trial and movement, in every classification level among all the 40 considered subjects. The final accuracy has

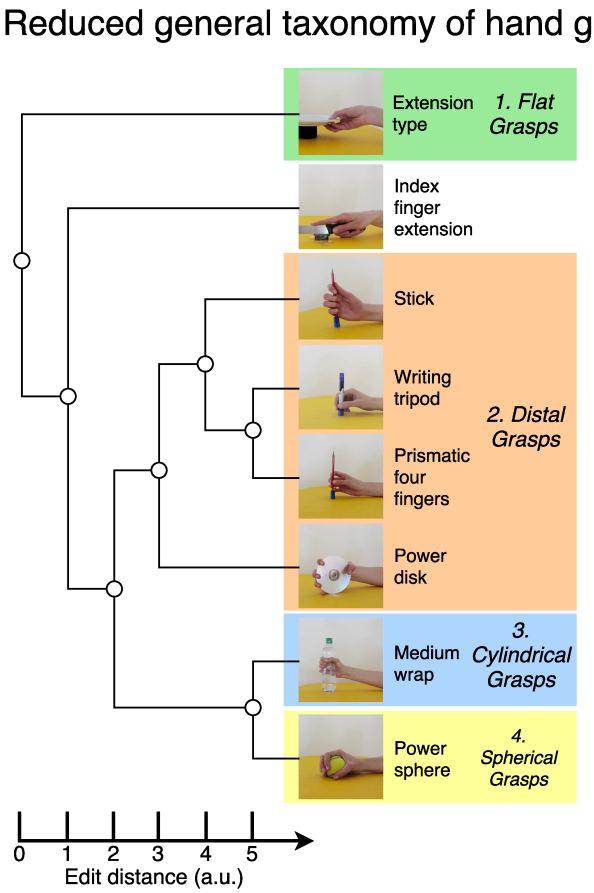


Figure 6.1: Subset of the General Taxonomy of Hand Grasps with the movements considered for this experiment.

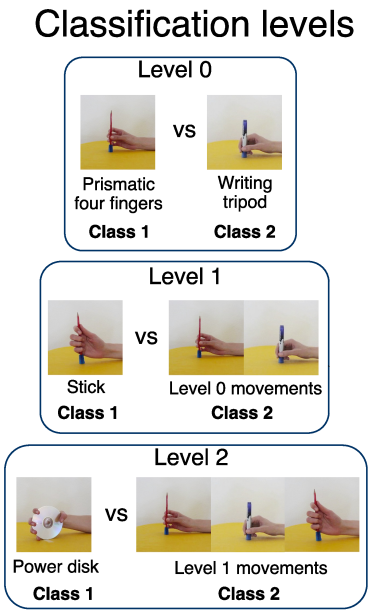


Figure 6.2: Example of the levels organization in the taxonomy-based classification.

been computed by fixing the level and averaging among all the remaining factors. The obtained results are generally good, and they confirmed our expectations. Lower levels achieves worst results, this is probably due to the similarity of the two movements considered at these levels. In fact, they showed very similar sEMG signals, thus it is easier to missclassify them. This hypothesis is supported by the fact that high level reaches higher accuracy. In particular, level 3 reaches an accuracy of 97.29%.

LEV 0		LEV 1		LEV 2	
GROUP A	GROUP B	GROUP A	GROUP B	GROUP A	GROUP B
Medium wrap vs. *Power sphere*	*Prismatic four fingers* vs. *writing tripod*	*Index finger extension* vs. Gr. A-LEV 0	*Stick* vs. Gr. B-LEV 0	*Extension type* vs. Gr.A-LEV 1	*Power disk* vs. Gr. B-LEV 1
80,21%	**57,50%**	**60,56%**	**68,05%**	**77,39%**	**73,12%**

LEV 3	LEV 4	LEV 5
Gr. A-LEV 0 vs. Gr. B-LEV 2	LEV 3 vs. *Index finger extension*	LEV 4 vs. *Extension type*
97,29%	**84,17%**	**90,21%**

Figure 6.3: Accuracy obtained among all the considered levels for the subject-independent taxonomy-based classification.

6.2.1 Test of different classification techniques

The previous tests have been expanded to different classification approaches, and they have been compared with the results achieved without the taxonomy contribution. In particular, we considered:

- 8 movements.
- 40 subjects.
- 5 taxonomy levels.
- Three classification approaches.
- Comparison with the classification of 8 movements without exploiting the taxonomy.

The preprocessing phase is described in Chap. 2, while the three classification approaches have been described in Chap. 3. We will refer to the following techniques:

- Instantaneous Classifier.
- Normalized Accumulation Classifier, the approach used in the previous tests.
- Bayesian Accumulation Classifier.

The following tables report the accuracy achieved with the three different classification techniques. Furthermore, for each technique there is a comparison between the results obtained exploiting the prior information from the taxonomy and without it. As in the previous results, we tested the framework with a Leave-One-Out approach, building the model on N-1 subjects and testing it on the remaining one.

Instantaneous Classifier

	Movements 4, 5, 10, 19			**Movements 7, 8, 9, 20**						
	Level 0	*Level 1*	*Level 2*	*Level 0*	*Level 1*	*Level 2*	*Level 3*	*Level 4*	*Level 5*	***No Taxonomy***
Movement 4		0.0792	0.9667					0.025	0.975	0.1458
Movement 5	0.7875	0.85	0.925				0.9917	0.8792	0.8875	0.525
Movement 7				0.4583	0.825	0.95	0.9417	1	1	0.2833
Movement 8					0.325	0.925	0.9625	1	1	0.3833
Movement 9				0.6917	0.8917	0.975	0.9708	1	1	0.575
Movement 10	0.8167	0.9042	0.9875				0.9917	0.9875	1	0.7583
Movement 19			0.2167						0.3417	0.4042
Movement 20						0.075	0.9792	1	0.9958	0.233
Mean accuracy	0.8021	0.6111	0.7739	0.575	0.6805	0.7312	0.9729	0.8417	0.9	0.4135

Table 6.1: Average results obtained by applying Instantaneous Classifier technique

Normalized Accumulation Classifier

	Movements 4, 5, 10, 19			**Movements 7, 8, 9, 20**						
	Level 0	*Level 1*	*Level 2*	*Level 0*	*Level 1*	*Level 2*	*Level 3*	*Level 4*	*Level 5*	***No Taxonomy***
Movement 4		0.0917	0.9375					0.0417	0.9625	0.1625
Movement 5	0.7	0.8042	0.9083				0.9875	0.8625	0.8833	0.525
Movement 7				0.5042	0.8917	0.9792	0.9542	1	1	0.2542
Movement 8					0.4458	0.9292	0.975	1	1	0.4333
Movement 9				0.7083	0.9125	0.9708	0.9875	1	1	0.5917
Movement 10	0.8625	0.8917	0.9792				1	0.9833	0.9958	0.6958
Movement 19			0.2083						0.3458	0.4417
Movement 20						0.0875	0.9875	1	0.9958	0.275
Mean accuracy	0.7812	0.5958	0.7583	0.6062	0.75	0.7217	0.9819	0.8979	0.8979	0.4224

Table 6.2: Average results obtained by applying Normalized Accumulation Classifier technique

Classification techniques

By observing the results achieved with the three different classification techniques, it is possible to notice that there is no substantial differences among the different techniques, both considering the taxonomy or not. Normalized Accumulation Classifier reached slightly better results when considering the taxonomy, while without taxonomy, Normalized Accumulation Classifier and Bayesian Accumulation Classifier obtained substantially the same mean accuracy. In general, when considering the taxonomy, the maximum difference among the tested classification techniques is only

Bayesian Accumulation Classifier

	Movements 4, 5, 10, 19			**Movements 7, 8, 9, 20**						
	Level 0	*Level 1*	*Level 2*	*Level 0*	*Level 1*	*Level 2*	*Level 3*	*Level 4*	*Level 5*	***No Taxonomy***
Movement 4		0.0833	0.9333					0.0792	0.9667	0.075
Movement 5	0.7375	0.8583	0.9042				0.9958	0.8208	0.8417	0.3125
Movement 7				0.5583	0.7917	0.9667	0.9417	1	0.9958	0.125
Movement 8					0.5042	0.9083	0.9333	0.9958	0.9667	0.3625
Movement 9				0.6708	0.8208	0.9708	0.9667	1	0.9917	0.3708
Movement 10	0.8542	0.8875	0.975				1	0.975	0.9958	0.8708
Movement 19			0.3						0.4125	0.65
Movement 20						0.1	0.9708	0.9958	0.9833	0.6125
Mean accuracy	0.7958	0.6097	0.7781	0.6146	0.7056	0.7365	0.968	0.8381	0.8943	0.4224

Table 6.3: Average results obtained by applying Bayesian Accumulation Classifier technique

0.89%, while without introducing the taxonomy information the maximum discard is 0.73%. Thus, we can assume that the three different classification approaches are comparable, both for the achieved results and the computational time.

Table 6.4: Average results with the various techniques

	Instantaneous Classifier	Normalized Accumulation Classifier	Bayesian Accumulation Classifier
Tax	0.7654	0.7727	0.7712
No Tax	0.4135	0.4224	0.4224

Taxonomy vs. general classification

The discussion becomes more interesting when analyzing the differences between the classical classification and the one that exploits the prior information from the taxonomy. The introduction of taxonomy hierarchy boosts the accuracy achieved in all the considered classification methods of about 35% for a subject-independent solution.

General considerations

This wider analysis confirms the results obtained in the initial study. Again, higher levels in the taxonomy reached higher accuracy, while the prediction is less precise in the lower levels. Nevertheless, a misclassification is less problematic in the lower levels, since the possible movements are close.

6.3 Conclusions

In this work we exploit the first results of the hierarchical and dependency relationship between the movements in order to develop a robust classification framework, able to identify online the movement performed by a subject. Classifications are more difficult as the number of considered class increases. In fact, the higher number of classes lead to more misclassifications and lower accuracy. Our solution aims to limit this problem by exploiting a hierarchical quantitative taxonomy of hand grasps. The binary tree structure of the taxonomy comports a classification between two groups of movements, close to each other. The classification becomes more precise descending the tree and reaching the root, where the classification is restricted between a couple of movements. The proposed solution is subject independent, and it works also with new, unseen subjects. The developed GMM based classification allows an incremental and progressive classification of the samples, granting robust and online results. In particular, we considered data from 40 healthy subjects performing 8 common grasps. Human information has been modeled and the classification process determined online by using a Gaussian-based framework. Results are promising: we obtained a mean accuracy of 76.5%, reaching 97.29% in one of the higher levels. Our approach showed an improvement of performances when considering a-priori information from the quantitative taxonomy of hand grasps.

Chapter 7

Physiological Signals in Industrial Settings

The advent of Industry 4.0 brought a wide number of innovative solutions for manufacturing [12], and the interaction between man and machines is a fundamental part in this process. The availability of affordable and reliable collaborative robots opens new and interesting perspectives. Human and robot can work together in the same area, on the same product. Operators can easily maneuver powerful manipulators, while robots can observe several human demonstrations to learn how to perform a task [128] [129] [130]. Up to now, many solutions propose to observe and track the body by means of visual systems, for example by exploiting 3D camera networks, or markers attached to the body [131]. This solution has the drawback of being sensitive to camera occlusions, light variations, and motion blur [132]. IMUs are probably the main alternative to cameras, they are effectively used to learn new behaviors [133] and control robots in industrial setups [134]. Moreover, Venek *et al.* [135] evaluated the use of IMUs as body tracking system under 5 major requirement areas to classify Active and Assisted Living (AAL). They were able to distinguish among different activities, with a limited amount of errors. In many cases, a multi-modal approach can be used to enrich the information and overcome limitations of unimodal systems [136]. Many solutions propose the introduction of physiological signals, recorded directly from the human body. EMG signals have been rarely considered as a unique tool for the motion prediction [137] due to their non-stationarity and sensitivity to muscular fatigue and stress [7], but they are quite popular for controlling exoskeletons or prosthesis [126]. Nevertheless, these signals have proved to be a valuable source of information when used in conjunction with other measurement units. Peternel *et al.* [9] aimed to read human intentions during a cooperating task through force/torque sensors, operator muscle

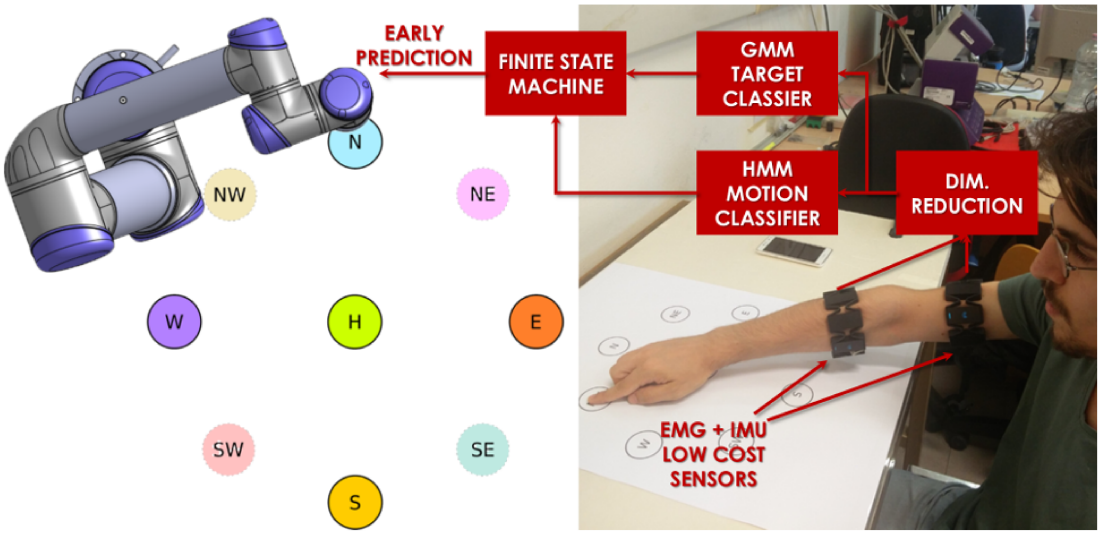


Figure 7.1: Human-robot cooperation framework for early prediction of human intention of movement by using solely EMG and IMU information. The dimensionality of the data is reduced to improve the leaning capability of the probabilistic models used to detect motion (HMM) and predict the correct motion (GMM). The robot reaction to the human action is triggered by a FSM receiving in input the classification results.

signals and dynamic of motion, obtained by a MoCap system. In [138], they extended their work to estimate the operator muscular fatigue in order to obtain an improved human-robot interaction in a leader-follower setup.

In this work, we aim to extend the state of the art by proposing and evaluating a technique to track and predict human upper-limb movements in an industrial facility, exploiting IMU and EMG data. The proposed system aim to (i) detect when the subject is moving, (ii) anticipate the identification of the movement direction, and (iii) provide the proper commands to the robot according to the application, while being robust to possible misclassifications. In particular, the system exploits the information registered through a pair of Myo armbands, by Thalmic Labs. As major advantage, the Myo sensor allows the simultaneous acquisition of kinematic and muscle activity information, at a very affordable cost. In a preliminary work [139], we developed a predictive system for healthcare based on GMM. An evidence accumulation framework selected the target starting by rich EMG signals. Now, we aim at applying such methods to human-robot collaboration by pairing IMU information to the reduced set of EMG data. Furthermore, we introduce two novel criteria to allow a dynamic stopping of the evidence accumulation, independently from movement speed or distance. IMU information have been also used to train a Hidden Markov Model (HMM) for motion intention detection. These two classifiers run in parallel to provide external inputs to a FSM that controls the robot according to operator's action, inaction, or motion

direction. Finally, the proposed system has been tested by teleoperating a UR10 collaborative robot with online hardware-in-the-loop simulations. However, the framework is general, thus it can be applied in many Industry 4.0 applications, where autonomous and collaborative robots are an essential part of the innovative facilities.

7.1 Experimental setup

The data for models training has been registered from four healthy right-handed subjects (age 26 $\pm$4, one female) performing reaching movements with their right upper-limb. At the beginning of the protocol, the subject is asked to comfortably sit on a fixed chair, with his right arm lying on a working bench. At each session, the subject is asked to move his hand from a home position ('H') in the middle of the workspace, to one of the targets placed at a distance of 15 cm in cardinal directions ('N','E','S','W'), as shown in Fig. 7.1. Each trial consists of a movement towards the target, followed by about 2 seconds of rest on the target, and a movement backwards the home position. These main directions have been primarily chosen to compare the results of this study with a previous work [139]. In addition to the cardinal directions, four secondary directions ('NE','SE','SW','NW') have been registered to test the robustness of the proposed system when increasing the number directions to detect. For each subject, six sessions have been performed, resulting in 30 trials for each target, thus 240 trials in total. The whole protocol lasted about 40 minutes per subject.

Angular velocity, linear acceleration and muscle activity information have been registered with two Myo armbands, from Thalmic Labs, worn on the upper arm and on the forearm. Both Myo devices have been connected and synchronized to the ROS middleware by means of a custom software library developed by our research group enabling the use of more than one device on the same PC[1]

7.2 Methods

7.2.1 Offline movement segmentation

The data collected from the subjects during the experimental protocol are composed of several trails acquired on after the other. Therefore, they present an alternation of states: motion (i.e. forward and backward movements) and rest (i.e. on the

[1]The library code is publicly available at `https://github.com/ste92uo/ROS_multipleMyo`

home position and on each target). We segmented the dataset offline in order to distinguish among the two states and divide each sample accordingly. The segmentation consisted of two cascading steps, both composed by a filtering and a thresholding. The procedure has been applied to the sum of the angular velocity magnitude of both arm and forearm. The filters are respectively a zero-phase eight-order Butterworth filter at 0.01 Hz between 0.01 Hz and 3 Hz, and a moving average filter with 1s sliding window. The first threshold is computed as the average value plus six times the standard deviation of the first and last seconds of each session, where subjects are known to be in rest. The second threshold is fixed to 0.1.

7.2.2 Dimensionality reduction

All the registered dataset consisted of inertial and muscular information. From each Myo device we saved 3-axis angular velocity, 3-axis linear accelerometer and 8 built-in surface EMG electrodes, resulting in a total of 28 features. Thus, we compared a number of dimensionality reduction techniques in order to limit the system complexity.

PCA

One of the most common unsupervised method is PCA [140]. PCA is based on an orthogonal linear transformation of the data in a different coordinate system, where the first components represent the most of the variance of the original dataset. This can be achieved by iteratively look for a linear mapping matrix W that maximizes the trace-norm of the multiplication between W and the sample covariance matrix of the M-dimensional original dataset. The lower embedding dimension m is chosen as the minimum number of principal components necessary to explain 90% of the dataset variance. In this thesis, PCA has been applied on the dataset including all the 28 features from IMU and EMG channels ('PCA').

Non-negative Matrix Factorization

Non-negative Matrix Factorization (NMF) [141] is a unsupervised dimensionality reduction algorithm giving particularly good results with dataset including only non-negative values, such as images and muscle activity envelopes. Given a dataset $\mathbf{X}$ of non-negative values, NMF extracts H and W by minimizing the divergence $D(\mathbf{X}||HW)$ between the original and the reconstructed datasets. Where, H is the subject-specific synergy matrix and contains m time-invariant and task-independent synergy modules.

And W is the m-dimensional matrix of activation coefficients over time. The m embedding dimension has been chosen by looking to the Variance-Accounted-For [142] in order to have a robust agreement between the original and the reconstructed dataset. In this work, NMF has been applied on the dataset including the 16 EMG features concatenated to the features extracted by PCA on IMU ('PCANMF').

Fisher's Discriminant Analysis (FDA)

Here, we refer to FDA not as the classification method, but as the dimensionality reduction method based on the Fischer's score. The aim of FDA is to project the data samples in a subspace with embedding dimension m where the within-class variance is minimized, while the between-class variance is maximized, in order to improve class separability. Given a multi-class problem of L classes, linear mapping W can be extracted from the first m eigenvectors $\mathbf{v}_i$ solving the system of linear equations:

$$S_W^{-1}\bar{S}\mathbf{v}_i = \lambda_i \mathbf{v}_i, \quad i = 1, ..., m \tag{7.1}$$

$$\bar{S} = S_W + S_B \tag{7.2}$$

where S_W and S_B are the *with-class scatter matrix* and the *between-class scatter matrix*, respectively and λ_i are descending eigenvvalues. This solution is limited to $m < L$ cases, since the rank of S_B is $L-1$, thus all the eigenvalues from L to M are the same and equal to 1 [143]. In this work, $m = L-1$ has been chosen and applied on two different datasets: the one including all the 28 features from IMU and EMG channels ('FDA'), and the one including only the 12 IMU features ('FDA-IMU').

7.2.3 GMM for direction identification

The feature extracted from the dimensionality reduction phase have been used to train a probabilistic model, namely a Gaussian Mixture Model, which is able to estimate the subject's chosen direction, by classifying among all the possible positions. This classification framework has been accurately described in (METHODS). For this particular work, a single data $\zeta_j, 1 \leq j \leq N$ in input at the framework can be written as Eq. 7.3.

$$\zeta_j = \{\xi(t), \gamma\} \in \mathbb{R}^D, \quad \xi(t) = \{\xi_c(t)\}_{c=1}^{C}. \tag{7.3}$$

where $C = |\xi|$ is the number of selected features, $\xi(t) \in \mathbb{R}^C$ is the set of values assumed from the considered features, and γ is the class of movement. $D = C+1$ is the

dimensionality of the problem.

7.2.4 HMM for motion prediction

A HMM has been implemented to predict if the subject is in a state of *REST* (i.e. the hand is stopped on the home position or on one of the available targets) or in a state of *MOTION* (i.e. towards a target or backwards to the home position). As observed variable $y(t)$ emitted by the state $\mathbf{x}(t)$ at time t, the sum of filtered angular velocity magnitude of arm and forearm has been exploited. The model has been trained by the *Baum-Welch* algorithm [144]. At each time step t, the posterior probability of the $i-th$ state (i.e. rest or motion) is computed as:

$$p(x_i(t)|y(t)) = p(y(t)|x_i(t)) \sum_{j=1}^{2} p(x_i(t)|x_j) p(x_j(t-1)) \quad (7.4)$$

$$i = 1,2$$

7.2.5 FSM for robot control

The output of HMM and GMM have been used to trigger a Finite State Machine with the following states:

1. *Home*: the subject is not moving from the home position, the GMM probabilities are not accumulated and the robot keeps its state;

2. *Evidence Accumulation*: the subject is moving towards a target and the output probabilities of the GMM are accumulated over time;

3. *Send Command*: the predicted movement direction is sent to the robot and the consequent action is activated;

4. *On Target*: the subject is on the target position;

5. *Back Movement*: the subject is moving back to the home position.

The transition map of the proposed FSM is shown in Fig. 7.2. At the beginning of each session, the subject is supposed to be in *Home* state and the HMM is initialized to REST. When MOTION is detected by the HMM, the machine immediately transits to *Evidence Accumulation* state. If the accumulated evidence fulfill the criteria for dynamic stopping, the machine goes to *Send Command* state. On the other hand, if the

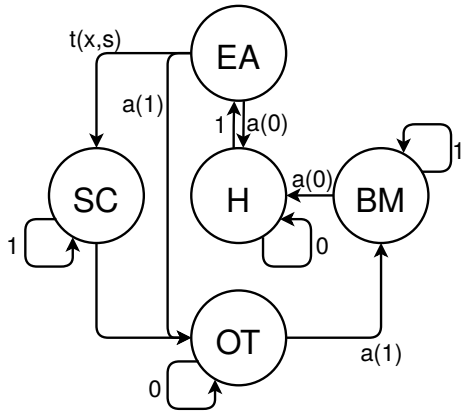


Figure 7.2: FSM Transion map. The transition valus 0 and 1 corresponds to REST and MOTION respectively. The transition function a enables changing the state after a predefined number of samples corresponding to the value in input. $T(s,r)$ enables a state transition only if the criteria for dynamic stopping is satisfied.

HMM identifies a REST and it is kept for at least X samples, the machine goes back to *Home* state. Finally, if the accumulated evidence does not satisfy the criteria for longer than Y samples, the FSM jumps directly to *On Target* state. The state goes to *On Target* also from *Send Command* state just after receiving a MOTION. If the machine is in *On Target* state and the HMM detects a MOTION for at least X samples, state transits to *Back Movement*. It stays in *Back Movement* until going back *Home* once HMM detects a REST for at least X samples. The values of X and Y have been calculated as the average value plus six times the standard deviation of the mean number of MOTION samples within a motion state and the number of REST samples in a rest state.

7.2.6 Criteria for dynamic stopping

The transition rule from *Evidence Accumulation* state to *Send Command* state determines the amount of time the classifier accumulates the GMM posterior probabilities before classifying the correct direction. However, it is difficult to determine *a priori* the minimum amount of time the classifier should accumulate to guarantee a certain accuracy, since the movement is performed at self-selected speed. For this reason, two criteria have been introduced to determine dynamically the time of accumulation at each trial. Given the vector of normalized accumulated posterior probabilities $\mathbf{nP}(t)$ for each of the L classes, the first criterion, namely the *ratio criterion*, is defined as:

$$C_r(t) = 1 - \frac{k_2(t)}{k_1(t)} \tag{7.5}$$

where

$$k_1(t) = nP_{i_1}(t) \quad with \quad i_1 = \arg\max_{i \in L} nP_i(t) \tag{7.6}$$

$$k_2(t) = nP_{i_2}(t) \quad with \quad i_2 = \arg\max_{i \neq i_1 \in L} nP_i(t) \tag{7.7}$$

and it represents the ratio between the probabilities of the two most probable directions. This criterion has been introduced so that the system sends a command to the robot only if the confidence on the selected command is sufficiently high. On the other hand, given the not normalized accumulated probabilities $\mathbf{P}(t)$, thus the cumulative sum of the raw posterior probabilities of the GMM for each of the L classes, the second criterion, namely *sum criterion*, is defined as:

$$C_s(t) = \sum_{i=1}^{L} P_i(t) \tag{7.8}$$

Combining the two criteria allows the construction of the transition rule $T(C_r, C_s, th_r, th_s)$ as

$$T(C_r, C_s, th_r, th_s) = \begin{cases} true, & \text{if } C_r > th_r \vee C_s > th_s \\ false, & \text{otherwise} \end{cases} \tag{7.9}$$

Thus, a command is sent as soon as one of the criteria is verified (i.e. $T = true$). A grid search has been conducted with a 5-fold cross-validation for each subject to determine the thresholds th_r and th_s.

7.3 Results

To assess the performances of the system, the classification accuracy over the reaching distance has been computed and averaged across subjects. The performance in discriminating between the four main directions for the different dimensionality reduction methods, all coupled with the GMM, are shown in Fig. 7.3 (top). As expected, the accuracy increases over time, reaching more than 90% of accuracy for all the methods. However, 'FDA' shows the highest performance in the first half of the reaching distance when compared to the other methods. The accuracy for 'FDA' reached more than 90% of accuracy already at 30% of reaching distance, and with a remarkably lower variability between subjects. To assess robustness and scalability of the classification methods, we tested their performance over the extended set of eight classes. The results, shown in Fig. 7.3 (bottom), reveal a higher robustness with the number of

classes of the supervised dimensionality reduction methods ('FDA' and 'FDA-IMU'). In fact, the accuracy is higher than 80% at 50% of reaching distance, and it is up to 90% at the end of the movement. Again, FDA applied on both EMG and IMU data showed a slightly better behaviour and lower variability between subjects. For this reasons, the couple FDA-GMM will be considered as the selected classifier for the following analysis.

The results of the grid search for the stopping criteria thresholds are shown in Fig. 7.4 for one of the tested subjects. Ideally, we would like to find the pair of thresholds that maximizes the accuracy while minimizing the time, thus maximizing the detection speed. It can be noticed that, at increasing of both thresholds, the accuracy increases, as well as the time to send the command. The selection of the thresholds allows a flexible design of the classifier's performance, adjusting the speed-accuracy trade-off according to the application. In this context, the pair of thresholds has been selected as the one that guarantees an average accuracy of at least 95% in the minimum amount of time. The selected thresholds for each subject and their corresponding accuracy-time performance are provided in Table 7.1. In the final testing, the two GMM and HMM classifiers have been used as external inputs to the finite state machine, implemented on ROS to control a physical manipulator robot, namely the UR10. The characteristics of this robot make it the ideal choice for safe collaboration between man and machines, since it is easy to control, even by non-expert users. Furthermore, the force sensors block the robot when an impact is perceived, in order to avoid damages to both humans and objects. In the tested application, the robot end-effector has been teleoperated on a horizontal plane, moving in four possible directions, forward ('N'), backward ('S'), right ('E') or left ('W'). The robot does not move while the machine is in *home* state. As soon as the FSM transits to *Send Command* state, the robot moves in the detected direction and keeps moving in the same direction as long as the machine is in *On Target* state. The robot stops as soon as the machine transits to *Back Movement* state, meaning that the user does not want to send the command anymore. The results of this testing are shown in Fig. 7.5. The figure shows four consecutive trials, that consists of four commands to control the robot. The blue line represents the offline movement segmentation, and can be either REST or MOTION. Two consecutive motions represent a forward movement towards one of the target, and a backward movement to the home position. The green line represents the output of the HMM, that can predict either a status of REST or a status of MOTION, resulting in an average accuracy of 82.5$\pm$4.8% across subjects. The red lines represents the transitions of the FSM through its states (*Home*, *Evidence Accumulation*, *Send Command*, *On Target*,

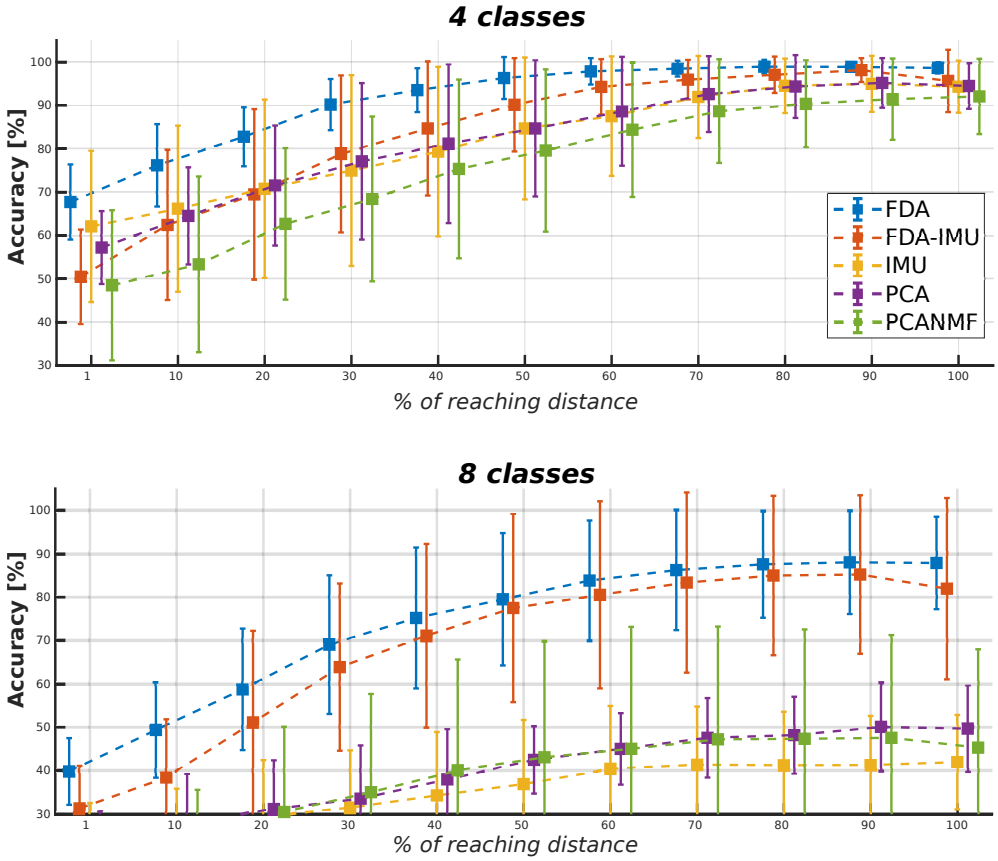


Figure 7.3: Classification accuracy over percentage of reaching distance of Gaussian Mixture Model coupled to 5 dimensionality reduction algorithms. The classifier has been tested with 4 classes (top) and 8 classes (bottom).

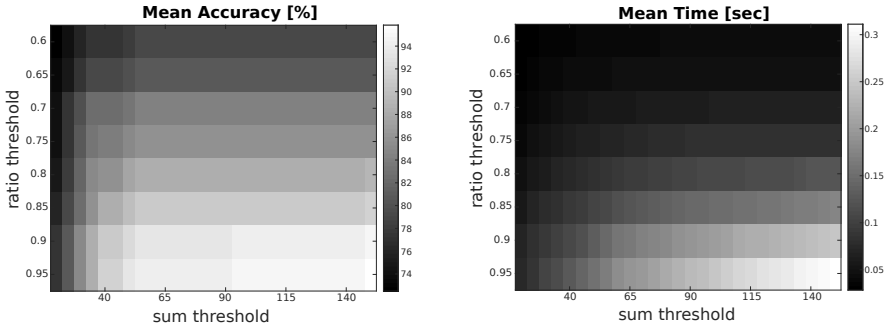


Figure 7.4: Visualization of the grid search results using the combined transition rule T. Each point of the picture reflects the performance that would have been achieved with the corresponding thresholds. A total of 120 trials per subject each with a different thresholds combinations are shown. Performance have been measured in terms of mean accuracy (left) and mean time to send a command (right).

Back Movement). It can be seen that the FSM driven by the GMM and the HMM correctly transits through the states, following the user's movement, with delays that are principally caused by the HMM transitions. Quantitatively, the performance of the whole system averaged across the tested subjects is shown in Table 7.2, in terms of percentage of correctly sent command, time to send a command and number of erroneous transitions of the FSM.

Table 7.1: Thresholds that have been selected for each subject to achieve at least 95% of accuracy in the minimum amount of time. Their corresponding performance are also shown

Subjects	th_r	th_s	**accuracy [%]**	**time [sec]**
s1	0.95	95	95.0 ± 6.8	0.24 ± 0.25
s2	0.95	35	95.0 ± 1.8	0.37 ± 0.24
s3	0.95	45	95.0 ± 3.5	0.25 ± 0.23
s4	0.95	25	95.0 ± 3.5	0.16 ± 0.17

Table 7.2: Performance of the whole human-machine interface during hardware-in-the-loop simulations on UR10 robot. The results are evaluated in terms of percentage of correctly sent commands (Accuracy), mean time to send a command and percentage of FSM erroneous transitions, averaged across subjects

Accuracy [%]	**Mean time [sec]**	**Transitions err. [%]**
94.3 ± 2.9	0.16 ± 0.08	1.6 ± 1.5

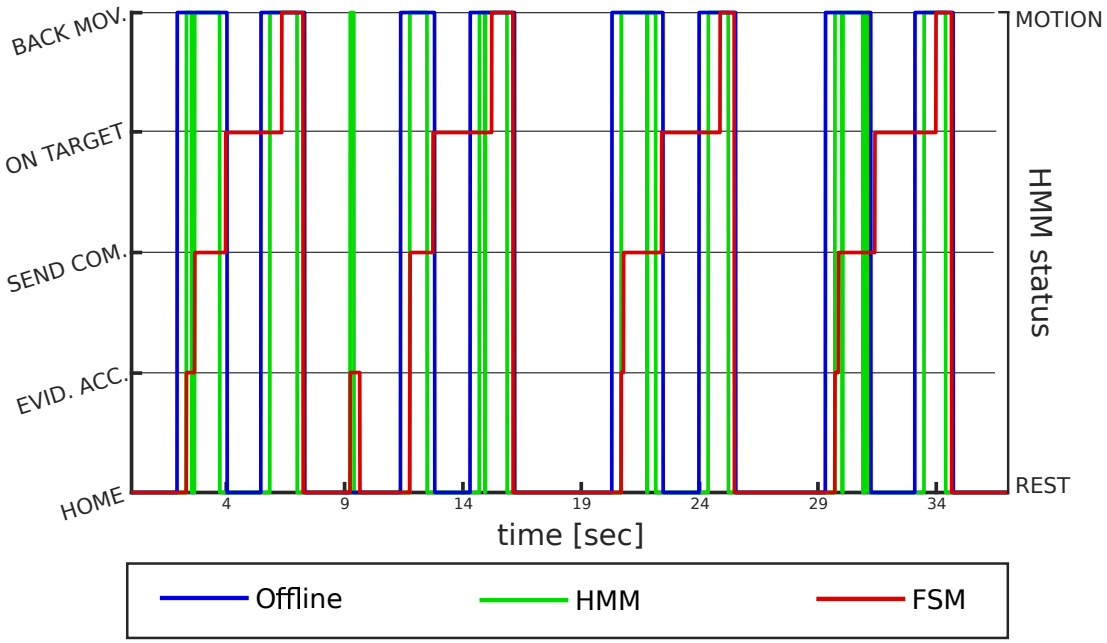


Figure 7.5: A sample of four trials taken from the final experiment with the FSM used to control the UR10 robot. The offline segmentation (blue) is shown as ground truth for the evaluation. Two consecutive state of MOTION represent a movement towards the target and a backward movement to the home position. The trained HMM (green) has been used to predict states of MOTION from state of REST. The FSM (red) correctly transits through its possible states according to the outputs of the HMM and the GMM.

7.4 Discussion

Different dimensionality reduction algorithms have been tested to evaluate the performance of unsupervised approaches respect to supervised approaches in motion direction prediction. In particular, the Non-negative matrix factorization has been tested since it has previously shown to be efficient in extracting motion primitives from EMG envelopes and to improve classification accuracy and speed in a similar context [139]. However, NMF on EMG data, coupled to PCA on IMU data, performed poorer in this application compared to other methods. On the other hand, FDA has been tested with and without the contribution of EMG channels, to see if inertial information alone would be enough for motion classification. Interestingly, in both the 4 classes and 8 classes cases, the FDA with additional information on muscle activity improved the classification accuracy of the FDA on IMU alone, of about 15% and with evidently lower variability across subjects. These findings strengthen the hypothesis that multi-model approaches, enriched with the introduction of physiological signals, can overcome the limitations of traditional uni-modal approaches. Of interest is the filtering effect of the FSM on the HMM misclassification, discarding fast and unstable transitions from REST to MOTION and viceversa, according to the machine state. This can be clearly seen between the first and second trials of Fig. 7.5, where a MOTION peak of the HMM activates the *Evidence Accumulation* state immediately, but it does not last enough to allow the GMM to verify the stopping criteria. As a consequence, the machine goes back to the *Home* state, without sending an undesired command to the robot. The evidence accumulation is a very well-known solution to improve the accuracy of a classification systems, particularly useful in applications where the driving signals are very noisy [145]. However, it could be challenging to determine *a priori* when to stop the accumulation. To this aim, the proposed dynamic stopping criteria allows a flexible design of the system performance, adjusting the speed-accuracy trade-off according to the application. In the tested teleoperation application, we were more interested in the accuracy of the system, thus a command should have been sent only if the direction has been predicted with 95% of confidence. On the other hand, if fast detection of the motion is more important than the confidence on the predicted direction (i.e. safety applications), the thresholds of the two criteria can be adjusted to met the desired performance. With the selected subject-independent thresholds, the system shows very high online accuracy in about 100 ms from movement onset. The major limitation concerns the way the FSM reacts to erroneous transitions. In fact, during the simulation, if a transition errors would have occurred, the FSM has been manually reset to the *Home* state and the simulation restarted from the following trial.

The response to wrong transitions, even if it should be improved in future works, does not really affect the system performance, since 1 error every 100 commands happens on average.

Chapter 8

Robot Programming by Demonstration in Industrial Settings

8.1 European Robotic Challenge

The European robotic industry needs new, innovative, ideas to be globally competitive. Creative and avant-garde solutions usually emerge in the academic world, which is commonly more inclined to abstract experimentation, while the industrial world is more fond to traditional and safer solutions. The EuRoC project aims to promote collaboration between industrial and research communities by solving industrial relevant problems with innovative proposals while increasing the state of the art in corresponding fields. The project is developed as a challenge. Even though there is no failure in losing the competition, the desire to win motivates the teams to improve their work, and the money prize encourages them to invest funds in the challenge. Challenges are very popular in the robotic field since 1977 and they evolved alongside robotics. From Christiansen [146] proposing the "Amazing Micro-Mouse Maze Contest", in which the task was accomplished by solving a path planning problem to "DARPA Robotics Challenge" involving multi-joints robots interacting at very high level with humans. The increasing interest in robotic challenges has several motivations. The main reason is the capability to attract and interest a great audience. Gaining the attention of a wider and non-specialized public is very important to enhance the knowledge of emerging fields and new research areas, as expressed in [147]. Presenting research and scientific topics as an involving and catchy competition is also a good way to collect private funding from firms having a vision in the sector or simply interested in a cool manner of promoting their products. Furthermore, challenges promote interactions and collaborations between different areas and aspects, improving the state of the art in the

robotic field. In fact, having a great number of specialists able to concentrate on the same issue, would lead to novel ideas and provide innovative solutions. In some cases, challenges focus on problems not showing a direct improvement in the real life. For example, at the very beginning, RoboCup [1] [147] aimed at developing a robot team able to play soccer. The direct advantages in industrial or service robotics were not immediately clear. Anyway, since from the first edition, RoboCup has leaded to several evolutions in real-time recognition, planning, reasoning and acting in dynamic environments. In the recent past, RoboCup has included different areas from rescue missions to home robot companions, passing through industrial settings. Therefore, research has been enhanced to consider robots working as team with both other robots or humans by learning behaviours for complex, cognitive modelling in the more different scenarios. Another very famous competition is the DARPA Robotics Challenge [2] [148]. It also evolved along time and right now it focuses on disaster or emergency-response scenarios, asking robots to autonomously drive cars, move into terrains with debris while interacting with the environment. Even if the tasks proposed in these challenges are very complex, teams have the vantage of using a completely free and customizable framework.

A new project called EuRoC has been proposed in 2014 to boost robotics and manufacturing in Europe by means of three challenges. Each challenge consists of a series of stages over the 4 years of the project life leading to a progressive selection among the participants. The EuRoC project is accurately described in [13]. The EuRoC project is organized in four levels of different complexity, and it is divided in three different challenges:

- Reconfigurable Interactive Manufacturing Cell (RIMC).
- Shop Floor Logistics and Manipulation (SFLM).
- Plant Servicing and Inspection (PSI).

As it is shown on figure 8.1, each challenge is divided in levels of increasing complexity and competitiveness, which ended with the selection of a winner. Participant teams belong to different backgrounds, there are technology developers, system integrators, and end users.

103 teams joined the Stage I, i.e. the Simulation Contests, but only 41 teams did advance to match making, in particular 11 teams for Challenge 1, 15 teams for Challenge 2 and 15 teams for Challenge 3. For Stage II, groups belonging to different areas

[1] http://www.robocup.org/
[2] http://www.theroboticschallenge.org/

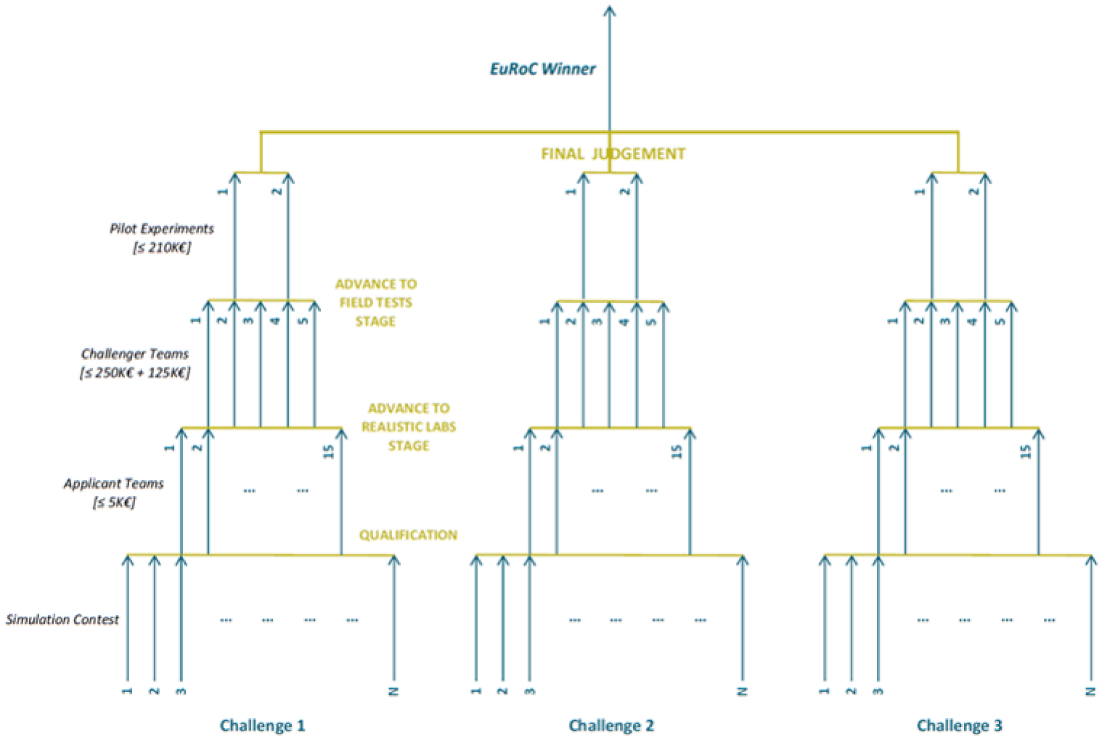


Figure 8.1: Chart of EuRoC Challenge

work together in the same team, with a unique project and common goal. Combining the different characteristics and capabilities within the teams helps in boosting the productivity, obtaining new and innovative solutions. The final goal of the challenge is the cooperation in the development of an innovative solution for the end user. 15 proposals were particularly innovative thus they have been funded from the European Union and admitted to Stage II. The funded proposals were equally divided among the three sections, which means that five teams have been funding to participate to Challenge 1 (RIMC), Challenge 2 (SFLM) and Challenge 3 (PSI).

IAS-Lab of the University of Padova joined the Reconfigurable Interactive Manufacturing Cell as R&D partner of ITRXCell. ITRXCell team is composed by four members:

- **IAS-Lab** from University of Padova as R&D partner. The group takes advantage of its expertise in machine learning, people detection and tracking system.
- **IT+Robotics** is a spin-off of the University of Padova and in the team is the industrial partner. They will exploit their knowledge on inverse kinematics engine and advanced quality inspections.
- **Stam** is a system integrator based in Genova, Italy. Their role involves industrial

design of advanced tools and cell layout.

- **ICPE** is the end user. This Romanian company is specialized in the production of electric motors coils.

The expertise of ICPE in electric motor bobbins oriented ITRXCell project in this field, by proposing and innovative technology for developing an automated coils winding framework. Major details are provided in the following sections.

As it was previously said, EuRoC challenge is structured in stages of increasing complexity:

- Stage II: Benchmarking.
- Stage III: Freestyle.
- Stage IV: Showcase.
- Stage V: Pilot (with access limited to the two best teams of each challenge).

The benchmarking phase is common among the five teams of a certain challenge. For the Reconfigurable Interactive Manufacturing Cell (RIMC), the benchmarking goal is the autonomous mounting of a plastic module on a car door. Each team tests its solution in the same facility in Stuttgart, and the proposed solution is evaluated by a team of experts which judges several parameters.

Higher other stages depends on the projects proposed by the teams, and they are evaluated by a group of experts which assess the obtained results by considering both the innovation of the solution, and the fulfilment of results. The tests are carried out in Fraunhofer IPA in Stuttgart for Stage II up to Stage IV, thus the working cell have to been transported and rebuilt in Stuttgart. For the finalist teams Stage V is tested at the end user facility. Unfortunately ITRCell has not been admitted to the final phase, but the scientific value of the proposal was assessed by the large number of scientific publication, thus we are still collaborating with our partners to develop alternative solutions.

8.2 FLEXICOIL Project

The Challenger team ITRXcell composed by UniPD and IT+R, the System Integrator STAM and the End User ICPE propose the FLEXICOIL project, which aims at developing a learning-based approach for robotised coils winding, to be used in the

electric machines manufacturing industry. In order to exploit the electro-magnetic effect, wound coils can be found in several products (motors, generators, sensors), for a wide range of applications. Smaller lot sizes and higher product flexibility can not be achieved with conventional winding systems. In fact, highly productive automated winding systems are not flexible enough, and manual labour causes high product costs.

FLEXICOIL aims at overcoming this drawback by developing a robotic cell for coils winding. Three subsequent sets of activities allows the developing of a reconfigurable interactive manufacturing cell with learning capability, suitable to wind the coils of several kind of motors already on the market, basing on a simple teaching interface that can be easily used by operators without specific skills in robotics. Adaptive perception will in fact recognize the type of motor, in order to select the most suitable winding procedure. The system will be able also to propose an effective solution in case of motors never seen before. The solution proposed by FLEXICOIL will be affordable from small-medium enterprises (SMEs), producing small batches of motors and frequently changing product designs, to big companies having a market request of several thousand standard units. This result will be possible by merging together ICPEs experience in the electric motor field, the research skills of UniPD in the field of human-robot interaction and robot learning, the technology transfer and industrial vision capabilities of IT+R, and the ability of STAM in integrating robotic and automation systems.

The benefits in terms of the partners relative position in the competitive environment are:

- ICPE: increased market share, introduction in new market segments (product diversification), increased quality and performances of products.
- STAM and IT+R: creation of two new products (the end effector and the vision system) which equip the FLEXICOIL cell.
- UniPD (and all): increase of scientific reputation. All the partners are be co-authors/contributors of any publication related to the FLEXICOIL technologies.

FLEXICOIL ultimately produces the following scientific and technical benefits:

- Reduction of setup times and cost of the winding machine.
- Increase of product performances and quality.
- Reduction of environmental impact of the production process.
- Winding operations can be easily parallelized.

- Reduced number of defected cores.

8.2.1 Scientific and technical quality

Proposal target

Objectives and relevance to the Reconfigurable Interactive Manufacturing Cell Challenge FLEXICOIL project perfectly matches the focus of EuRoC Challenge 1:

- Safety of the operator is addressed through the development of two sets of safety protocols (basic ones in the freestyle demo and advanced one in the pilot).
- Human/robot interaction is the cornerstone upon which FLEXICOIL is based, as the worker teaches the robot how to properly wind the coils of stator/rotors, by showing the process.
- The cell is based on the human operator teaching the winding process to the robot.
- The whole project is aimed at developing a cell to robotize the manufacturing process of electric machines, in particular winding of coils on stator or rotor cores.
- The cell is based on artificial vision systems to implement the learning system, the safety protocols and the quality inspection, thus coping with process uncertainties.

The desired goal is reached by increasing complexity phases: the freestyle and the showcase.

The Freestyle Demo The proposed freestyle activity is the cornerstone of FLEXICOIL and it paves the way for the development of the robot learning system. The challengers UniPD and IT+R exploit their experience in robotic learning and artificial vision systems. They develop the learning system of the cell, thanks to whom the robot is able to learn from a human operator, how to correctly wind coils of electric motors (stators and/or rotors, depending on the motor technology). The learning system is complemented by the following activities:

- A people detection and tracking system, which puts the basis for the implementation of a safe and fruitful human/robot interaction.

- A wire deployment system, to be installed on the robot end-effector and capable of winding the wire at the prescribed tension level.
- An inverse kinematic engine, to reconstruct the configuration requested to the robot in order to suitably position the end-effector.

The freestyle activity starts with an analysis from ICPE aimed at identifying basic requirements for the motor winding operations. The cell layout is designed by the system integrator STAM, it allows demonstration and proof of concept activities, to be performed after hardware prototyping and system integration. In a nutshell, the learning system is used to prove that, on the base of the visual input provided by the human operator (Fig. 8.2, part **a**), the robot can wind wires on complex paths (Fig. 8.2, part **c**).

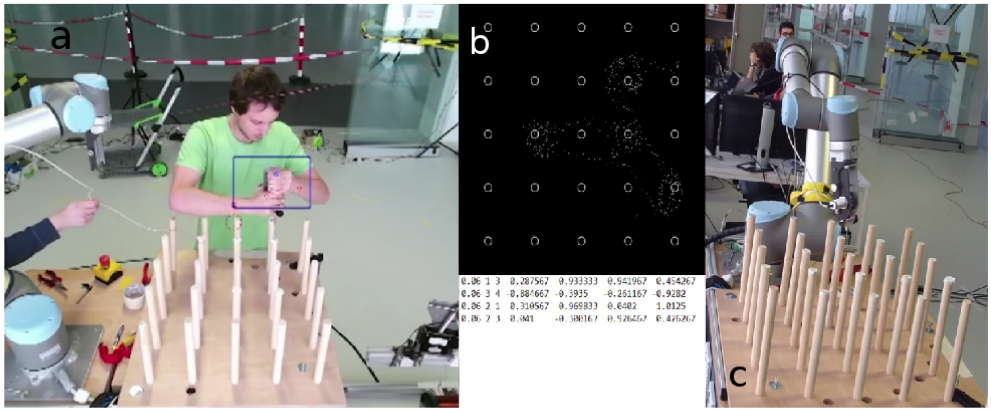


Figure 8.2: Sequence of operations in the Freestyle phase: a) the subjects perform the task while cameras track their body; b) the system infer the human movement; c) the robot replicates the human demonstration.

The Showcase Demo The showcase round aims at transferring the know-how generated in the freestyle activity to a manufacturing scenario. The winding operations learned by the robot are performed on real coils, to be used as sub-assemblies of real electric motor stator. STAM redesigns the layout of the cell, in order to include the additional complexities of the real manufacturing case, such as the support for the motors stator/rotor, and the clamp to grip the end of the wire to be wound. The end-effector of the robot is equipped with a sensor system capable of monitoring and providing real-time feedback on the wire tension during the winding operation and also with a pick & place gripper, which enables the placement of the coil on the stator/rotor. The development of paths control for pick & place operations complements the hardware design. UniPD and IT+R works on the winding learning system, in order to take into

account several motor parameters including the wire tension feedback provided by the sensor. Furthermore, a basic quality inspection protocol is implemented: the quality of the generated coil is estimated by monitoring turns count and the wire tension. Finally, the demonstration activity aims at assembling the poles of a stator, by winding each coil and assembling them in the stator through the pick & place 8.3.

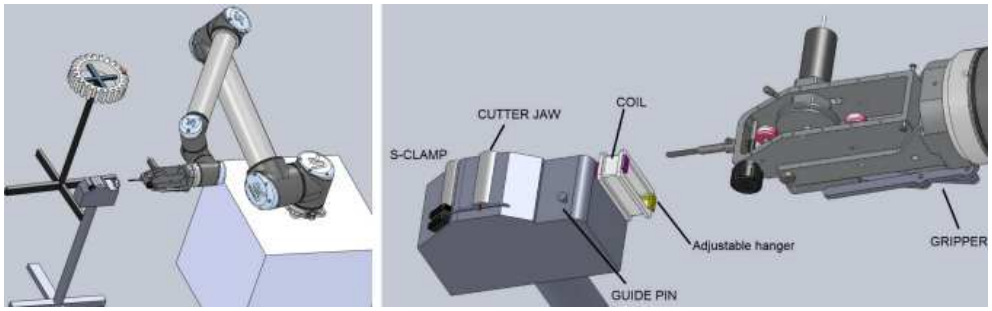


Figure 8.3: Simulation of the robotized coil winding and stator assembling for Showcase phase.

Progress beyond the state-of-the-art

The low flexibility of automated winding machines, i.e. the time and costs required to switch from one design to another, coupled to high cost of machinery (up to 100 kEuros), force small manufacturers (especially SMEs) to employ human operators in this task. The handcrafted job is obviously much more flexible, but more expensive (because of labour cost and equipment), and for the worker it is distressing, frustrating and repetitive. Few attempts of robotic cell for coil winding have been made. Although the approach proved to be competitive, none of these projects has eventually turned into a real commercial product, mainly because of the following reasons:

- Low increase of flexibility, because reprogramming the cell to produce a new coil design requires an operator skilled in robotics. This prevented small manufacturers, that often change their production batch, from preferring the robotic approach to the manual one.

- Low production rate compared to automated winding machinery. This prevented large manufacturers, that produce several thousands of products with the same design, from preferring the robotic approach to the traditional one.

FLEXICOIL aims to overcome all bottlenecks shown by previous solutions for coil winding, whose breakthrough is represented by the strong human/robot interaction,

that allows the robot to learn from the operator how to wind the coils. FLEXICOIL increases the automation of manufacturers exploiting wound coils, without limiting the flexibility to address fast production changes. Furthermore, the solution is complemented by implementation of safety procedures and quality inspection protocols,

Previous works FLEXICOIL builds upon the previous work and relevant experience developed by all partners:

- The scientific know-how of UniPD in robotic learning systems.
- The synergic industrial background of the SME IT+R in artificial vision systems.
- The technical experience of the SME STAM in integration of robotic manufacturing cells.
- The knowledge about electric machines market and technical solutions of the end user ICPE.

The FLEXICOIL approach to robotize coil winding was suggested by the end-user ICPE, that, thanks to his profile of research centre and motors manufacturer, understood that a robot could be the solution for the problem of coil winding. Two main needs were identified:

- A fast procedure to develop the robot control at any change of production reference, not requiring operators skilled in robotics and ensuring safety of human workers.
- An accurate set of automated quality inspections to verify that the coil meets specifications.

FLEXICOIL consortium selected an approach based on the following works:

- UniPDs learning system for industrial robots and people detection [11], [149] and tracking systems [150].
- IT+Rs quality inspection and motion planning systems based on visual feedback.
- STAMs quality sensor for optical fibres and grasping/tilting gripper for small toys.
- ICPEs high performance outer rotor motor.

The chosen approaches The approach proposed by FLEXICOIL aims to dramatically simplify the robot control by deploying a learning system based on artificial vision. In the last years, Robot Learning from Demonstration (RLfD) has become a major topic in robotics research. In fact, direct programming a robot motion can be a very difficult and time-consuming task. Acquiring examples from humans provides a powerful mechanism to simplify the process of programming complex robot motions. In this work, we will focus on allowing non-expert users to naturally interact with robots to teach them new behaviours. The learning system is based on a Gaussian Mixture Model /Gaussian Mixture Regression framework. This framework has been proven to produce good results with a relative low number of demonstrations in repeatable industrial tasks [11]. We considers several variables connected to the winding task in order to obtain a greater flexibility in terms of industrial application basing our work on related experiences in the field. Such amount of task constraints and the natural human-robot interface is the major contributions of this work to the state of the art. An on-line trajectory correction systems based on vision feedback is coupled with the learning module to avoid possible collisions. Using this technique allows to reach a significant increase of the coil density and consequently higher standards in coil quality and in engine efficiency. Reliably detecting and tracking movements of nearby workers on the factory floor is crucial to the safety of advanced manufacturing automation in which humans and robots share the same workspace. We address the problem of multiple people detection and tracking in industrial environments in real-time by using a network of cameras covering the workspace area. The entire project is developed on top of the ROS framework [151], a robotics middleware which can be considered as a standard de facto in the research community. It implements a good infrastructure for network communication and provides all the tools necessary to a modern distributed system. In this way, the created system would be easily scaled to different type of robots with the only constraint of a proper wrapper for ROS.

Impact

Expected results FLEXICOIL project aims to produce a number of concrete results in different research areas such as: manufacturing processes, learning systems, artificial vision devices and testing in a complex machinery. The exploitable results expected are:

- The FLEXICOIL prototype is the main concrete result of the Project.
- The Pick & Place end-effector equipped with the wire deployment system and

tension sensor.

- The learning system and teaching human-robot interface based on artificial vision system.
- The advanced quality inspections based on artificial vision systems.

The benefits in terms of the partners relative position in the competitive environment are:

- ICPE: increased market share, introduction in new market segments (product diversification), increased quality and performances of products.
- STAM and IT+R: creation of new products (the end effector and the vision system) which equip the FLEXICOIL cell; the two products could also be applied for other applications.

As stated by the European Motor & Motion Association (EMMA) in its annual report on industrial automation: *to remain competitive in the global arena, future manufacturing scenarios will have to combine highest productivity and flexibility with minimal lifecycle-cost of manufacturing equipment*. European electrical machines manufacturers need to increase the flexibility of production process, due to the high cost of equipment setup at the beginning of each new production batch. Overall, most of these European manufacturers are striving to reduce costs while preserving the quality of products, in order to face the competition by Far East companies. The FLEXICOIL concept is expected to produce the following scientific and technical benefits, each of them contributing to address the need for cost reduction and increase of flexibility:

- Reduction of setup times of the winding machine by 50%, thanks to the removal of auxiliary special wire guides, which are necessary with conventional fly winding machines.
- Reduction of setup cost, mainly in terms of effort, by 70%, since the robotic-based system does not require any machining of new fittings for every production batch.
- Thanks to the improved stability of the process, the slot fill ratio can be increased (increasing the performances) and less copper wire is wasted (greener process).
- Bobbins can be wound directly on a workpiece pallet, reducing the handling steps. Cycle times can be reduced by about 30% and system stability is increased.

- Winding operations can be easily parallelized, dramatically increasing production rate.
- The robot can also be used for assembly tasks (e.g. winding an armature and assembling the motor), reducing equipment costs or labour costs and allow better exploitation of robots.
- The function integration reduces the number of handling and loading operations and hence the winding process chain can become more robust.
- The number of defected cores is expected to decrease from current 0.5% to 0.05%.

The FLEXICOIL project will affect the European market by increasing the competitiveness in the electric vehicles and motors manufacturing, where automation at different levels is already applied. This would have a major impact on economy at the European level, because these markets, where automation is already implemented, are facing strong competition from Countries with low labour cost. Increasing the process efficiency would therefore strengthen a sector that is very critical for Europe as it represent almost 2% of the EU GDP. The introduction of this innovative technology could then considerably increase the competitiveness of European enterprises which operate in the electrical motors manufacturing field.

The role of this thesis into the EuRoC project regards the development of Robot Programming by Demonstration solutions, in order to handle the most complex and delicate part of the motion by exploiting the human knowledge. For this reason, part of the work involves motion modelization and prediction, while an other part deals with robot movement.

8.3 EuRoC Benchmarking Phase

In the EuRoC benchmarking phase the teams have to perform the same task, i.e. automatically assembly a car door with its module. All the teams have to accomplish the same tasks in a period of few months and they are evaluated in the same facility in Stuttgart through some fixed parameters. In this manner, it is possible to guarantee success and repeatability of each task in a quantifiable way.

Constraints imposed by the challenge are similar to the constraints present in a real factory. All the teams have to use a predetermined system, without the possibility of adding new sensors or modifying the already existent ones. Similarly, any integration in a factory supply chain has to cope with a previous installed system, probably obsolete or outdated. Moreover, complex and challenging workcell layouts, changing illumination, and tight workplaces are very common in both benchmarking and industrial settings. On the other hand, a dynamic environment with people working alongside robots in a collaborative manner, uncertain position of parts to be assembled and teaching of novel configurations are still open problems in research.

As suggested by Chen *et al* [152], flexible assembly applications are actually uncommon and only a small portion of industrial robot are used to perform tasks with variability in the parts. In fact, conventional industrial robots are not able to adapt to changes in the assembly processes. On the other hand, Goya *et al* [153] indicated flexibility in the automotive manufacturing as one of the more competitive weapons in the economical analysis of North American automotive industry. They proposed the possibility of switch easily and with a lower risk from a production line to an other as main advantage in future achievements with respect to foreign competitors. The reduced risk allows the industries to invest in low volume-high risk products, since the money and time loss would be minor and the production line would remain the same.

In the EuRoC challenge both hardware and firmware composing the robotic system are fixed. Anyway, they have to be programmed with ROS, the standard de facto in research robotic frameworks and only recently introduced in industry. The importance of ROS has been expressed in [154] by Tavares *et al*. They analyzed a pick and place task by combining several layers of control. Using ROS in developing industrial applications gives the possibility of efficiently divide layers in standardized and compact blocks able to interact one with each other to autonomously correct errors during the accomplishment of the task. In our solution, we take advantage of the modularity and standardization ROS characteristics to fuse together visual and robot learning techniques in order to face the variability in the system configuration.

Stability and reliability of each method have to be enhanced to meet the bench-

marking requirements in terms of elapsed time and hardware compatibility. We also develop a human-robot interface able to easily teach the system with novel door assembly combinations. Vakanski *et al* [155] suggest to take advantage of robots ability of learning from what surrounds them, transferring skills to a robot thanks to multiple demonstration of the same skill under similar conditions. Vakanski's idea is to learn the robot trajectory by observing a subject moving a tool. Instead, in our solution the robot motion is acquired by looking at a person actually moving the robot. The robot has to infer a generalized trajectory obtained extracting relevant unknown and hidden features from the demonstrations.

8.3.1 Task and System Description

In the EuRoC benchmarking phase the teams have to automatically assembly a car door with its module. The module and the door are represented on Fig. 8.4. The positions of door and module could vary of few centimeters in translation and few degrees in rotation in each direction. The testbed is composed by three tasks:

1. Pick and insert door module: in the Pick phase challengers have to locate the door module by using visual and force information, pick it up and reach a reference position. Then the robot has to place the module into the door, and come back to reference position without detaching it.

2. Screw door module: consists of detecting, picking up, inserting three screws into three relevant holes to fix the module on the door.

3. Teach and assemble unknown door: teams have to teach to the system how to perform the whole assembly for a novel pair of module and door.

Available points and maximum execution time for each task are listed in the benchmarking rules alongside with constraints about hardware and software.

The available hardware in the Fraunhofer IPA facility consists of a lightweight collaborative robot (Universal Robots UR10) equipped with three sensors: a 3D sensor camera (PMD CamBoard Nano), a stereo camera (VRMagic D3), and a 6D force-torque sensor (Robotiq FT150). A vacuum gripper, composed by 6 suction cups, and a screwing tool (Weber Pluto 6D) are available and they could be automatically attached or detached from the robot flange by using a tool changing rack (Schunk SWS011).

We replicated the setting in our laboratory, using as a basis the same lightweight collaborative robot, the Universal Robots UR10, but equipping it with different sensors. A PMD CamBoard Pico has been used as 3D camera, a pair of Philips SPZ5000

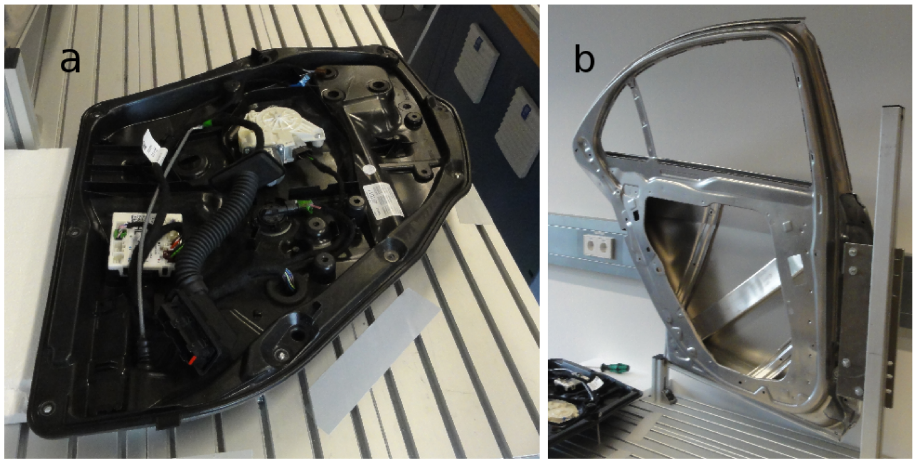


Figure 8.4: a) Plastic module, b) Car door where the module has to be inserted.

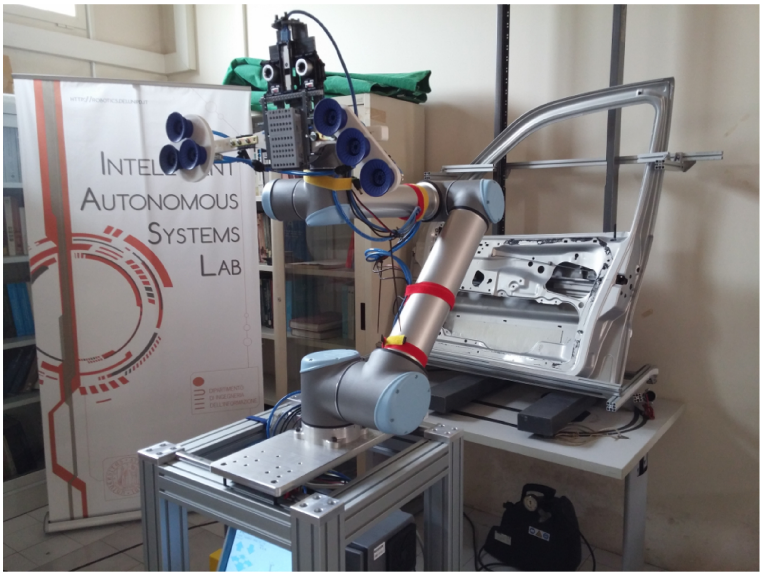


Figure 8.5: The replica of EuRoC benchmarking platform built in our laboratory.

webcams have been calibrated to work as stereo camera, while no force sensor has been mounted on the robot. A vacuum gripper and a screwing tool have been built by means of 3D printed materials to mimic the functionality available in the original system. The tools could be manually changed from one to another.

A central working station connected through Gigabit Ethernet collects data coming from sensors and going to controllers. ROS is used as framework to enable interaction with the system. Topics, services, and actions coming from the system, as well as device drivers, have been established in advance and everyone has to use the same setting. Positions and orientations of every relevant frame are published and updated by using the ROS TF package. A rough virtual model of the environment is also available and objects are coherently placed in the scene depending on the position published. 2D

and 3D cameras are provided with a default intrinsic and extrinsic calibration.

8.3.2 Methodology

The proposed tasks are connected to 3 main constraints:

- Limited time available for developing the solution.
- Flexibility needed to deal with position tolerances and unknown modules and doors.
- Usability and reliability of the teaching procedure.

These characteristics lead us to propose a solution able to face both known and unknown door assembly in a very similar manner. In fact, different modules and doors rely on similar structures, and these features can be used as input for the framework. We want to extract these common characteristics to simplify and speed-up the new module and door identification. In order to successfully solve the previously described tasks we used the following pipeline:

- Learn the relative positions of each screw hole in both module and door.
- Learn the gripping and inserting trajectories through human demonstrations.
- Identify both module and door real positions by visual inspection.
- Pick and place of the module by transposing the learned motion to real position.
- Identify screw positions.
- Screwing.

In the learning phase, we collect template images for each part to identify. The idea is to extract relevant visual information in order to recognize them during the part inspection in order to build a coarse virtual model of the environment. A combination of Robot Learning from Kinesthetic Demonstration and Inverse Kinematics is used to learn how to pick and place the module. The variability and robustness of the system are granted by collecting several repetitions of the same action, performed by different subjects. A visual system is used also for finding the screws and pick them up and fixing the module. A Template Matching approach is used for the screwing operation. Again, a door inspection is performed looking for screw holes positions. Once a matching has been identified, the system will align the screw with the hole by using the previously acquired template in order to perform the insertion.

Learning Phase

The learning phase mainly involves how to correctly pick the module and insert it in the door. These operations could be very easy or critical depending on the context and on the complexity of parts to be assembled. The module is composed by several elements including some flexible cables possibly assuming different configurations while the task evolves. The gripper provided for performing the picking action has a fixed base, while suction cups attached on its extremities can slightly change in position. Positioning the gripper in a consistent manner obtaining a robust layout is essential to assure a safe pick. This operation has to be performed in advance, since the surface of the module vary a lot and it is not always smooth. Moreover, the module is quite heavy while the suction cups do not have great gripping power in case of unbalanced loads. Selecting a wrong gripping position can determinate a loss in vacuum system and the consequential module falling due a displacement in weight or position of the suction cups. Once the module is properly picked, it has to be placed into the door. A series of coupling pins should be inserted into slots in the door to hold up the module, while avoiding cables to obstruct the movement. The motion should be executed precisely by placing the cables, and proceeding diagonally to insert the first coupling pin. Finally, the module should be straighten, and the rest of the pins could be set into place. Pinching the cable or failing to place a coupling pin lead to incorrect insertion and consequential falling of the module. Obviously, the placing trajectory is really dependent from the initial gripping position. Therefore, a successful picking does not correspond necessarily to a good performance in placing the module.

The described constraints could be easily met when the task is performed by a human being. Indeed, people have the capabilities to understand the task, test the selected strategy in few trials, and move the robot accordingly to fulfill the objectives. Nevertheless, recording a single execution is not enough for achieving a smooth generalized trajectory able to take into account to the intrinsic variability in the tasks. In order to obtain such motion, we used a Robot Learning by Demonstration paradigm able to build a robust model of the movement starting from a series of demonstrations. In particular, we recorded data directly from the robot while a subject is free to physically guide the manipulator following the desired path (Fig. 8.6).

It is worth to notice that the use of this technique, namely Kinesthetic Demonstration, is strictly connected with the use of a compliant robot with the capability of being externally guided by releasing motor brakes. Moreover, lightweight collaborative robots such us UR10 can be easily used in a real factory with no need of safety guarding, since they are intrinsically safe. Therefore, also unskilled workers can coop-

Figure 8.6: The robot is physically guided to follow the desired motion through Kinesthetic Demonstration.

erate with the robot by showing it what to do without any risk. In this way, it is possible to let the robot learn a novel task in little time and without the need of additional staff for robot programming.

Only few demonstrations are necessary to build from scratch a robust model of an unknown movement. In order to avoid unnecessary variability in the motion and reduce the complexity of the system we decided to keep human demonstrations as short as possible. Short trajectories are computed quicker resulting in a more standardized movement, while allowing a simpler and consequently safer robot activity. We mixed together Probabilistic Robot Learning with Inverse Kinematics to take advantage from both of them. The robot reaches fixed and safe positions close to the targets by using an Inverse Kinematics engine obtaining better performances in both reliability and time. The last part of the movement, namely the most complex one, is performed by using inferred trajectories computed through Robot Learning. Ten repetitions of the movement performed by different subjects has been recorded from an arbitrary selected initial position. The angles assumed by each joint while the robot is manually controlled have been considered for building the probabilistic model. Since door and module positions are not known a priori, a visual feedback is used to compute the actual configuration of the system before proceeding with the real picking and placing motions.

The raw data recorded from robot encoders have to be preprocessed in order to be able to generate a good probabilistic model. As a first step, they have been filtered to remove artifacts, such as periods in which all the joints were still. Doing so all the data not correlated with the movement have been eliminated, maintaining exclusively motion information. This process leaded to more robust and smooth trajectories while speeding up the creation of the model. Once the model is built, the final motion is estimated thanks to a regression technique.

GMM has been used as probabilistic model to predict angles of $G = 6$ robot joints. GMM is a parametric probability density function represented as a weighted sum of K Gaussian components which best approximate the input dataset. As described in Chap. 3, an advantage of using GMM is the few parameters needed to represent the whole model (i.e. the mean, the covariance and the prior probability of every Gaussian component), resulting in a lightweight representation of the movement.

Naming n the number of repetitions of the task, and T the number of observations acquired during each trial, the total number of data samples is $N = nT$. A single data in input at the framework $\zeta_j, 1 \leq j \leq N$ is described in Eq. 8.1.

$$\begin{aligned} \zeta_j &= \{t, \alpha(t)\} \in \mathbb{R}^D \\ \alpha_x(t) &= \{\alpha_g(t)\}_{g=1}^{G}. \end{aligned} \tag{8.1}$$

with:

- G, number of joint bending angles.
- $\alpha_g(t) \in \mathbb{R}$, the value assumed from g^{th} joint bending angle at the time instant t.
- $\alpha(t) \in \mathbb{R}$, the set of values assumed from the considered joint bending angles at the time instant t.
- $D = G + 1$, the dimensionality of the problem.

Finally, the resulting probability density function is computed:

$$p\left(\zeta_j\right) = \sum_{k=1}^{K} \pi_k \mathcal{N}\left(\zeta_j; \mu_k, \Sigma_k\right) \tag{8.2}$$

with π_k priors probabilities, and $\mathcal{N}\left(\zeta_j; \mu_k, \Sigma_k\right)$ Gaussian distribution defined by μ_k and Σ_k, respectively mean vector and covariance matrix of the k-th distribution.

Empirical experiments shows that, in this case, good results can be achieved with $k = 10$ Gaussian components. Using few Gaussian components cause the generation of a too general model, unable to handle the variability of the signal. Contrariwise, if k is big the final model will be too complex.

The Gaussian Mixture Regression (GMR) provided a smooth generalized version of every joint angle starting from the GMM (Fig. 8.7). Every joint angle $\hat{\alpha}$ and its covariance are estimated from the known a priori time instant t respectively using Equation 8.3 and 8.4.

$$\hat{\alpha} = E\left[\alpha \,|t\,\right] = \sum_{k=1}^{K} \beta_k \hat{\alpha}_k \tag{8.3}$$

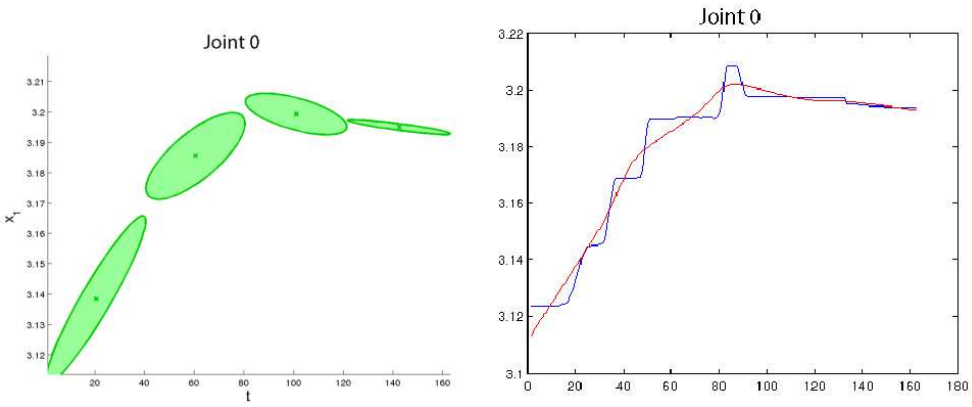


Figure 8.7: Modelization of Joint1 with GMM and continuous estimation of Joint1 angle retrieved with GMR.

$$\hat{\Sigma}_s = Cov[\alpha|t] = \sum_{k=1}^{K} \beta_k{}^2 \hat{\Sigma}_{\alpha,k} \tag{8.4}$$

with:

- β_k, the weight of the k^{th} Gaussian component through the mixture.
- $\hat{\alpha}_k$, the conditional expectation of α_k given t.
- $\hat{\Sigma}_{\alpha,k}$, the conditional covariance of α_k given t.

assuming that the parameters (π_k, μ_k, Σ_k) defining the k^{th} Gaussian component are decomposed as follows:

$$\mu_k = \{\mu_{p,k}\,\mu_{\alpha,k}\} \qquad \Sigma_k = \begin{bmatrix} \Sigma_{p,k} & \Sigma_{p\alpha,k} \\ \Sigma_{\alpha p,k} & \Sigma_{\alpha,k} \end{bmatrix} \tag{8.5}$$

with μ_p and Σ_p respectively the mean and the covariance of the known a priori information. Thus, the generalized form of the motions $\hat{\zeta} = \{t, \hat{\alpha}\}$ required only weights, means and covariances of the Gaussian components calculated through the EM algorithm.

The described framework could be used with known setting as well as with novel unknown door-module pairs. It gives good results both in time needed to teach the tasks and in robustness in reaching the goals. Nevertheless, it is hugely dependent from the information provided by the visual counterpart system.

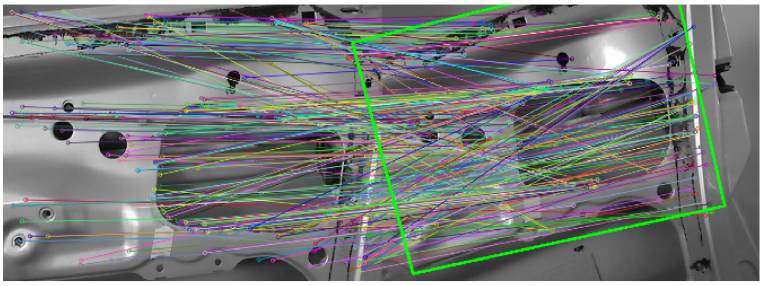


Figure 8.8: Keypoints extraction for the current frame with respect to the desired template.

Module and Door Identification

As before, we tried to adopt the same base procedure for both known and unknown objects. We decided to implement a reliable and adaptive identification system working for both door and module. A stereo camera has been used for obtaining a visual servoing procedure able to align the robot with respect to an object.

A simple Image-Based Visual Servoing (IBVS) is really hard to implement due to lack of common robust set of visual features to use. On other hand, a Position-Based Visual Servoing (PBVS) approach supposes to know a priori a model of the object, not provided in this challenge.

However, our algorithm compensates this lack with the 3D knowledge coming from the stereo camera. The approach is composed by a training phase and an iterative query phase. In the training phase, the robot is placed into a known position ${}^{w}\xi_{cd}$, where both cameras could see the same region of interest inside the object. A stereo pair template is acquired and a keypoints extraction is performed by using *FAST* algorithm [156,157] for the detection and *ORB* [158] as descriptors. A sparse triangulation is computed after a keypoints matching between left and right templates by using intrinsic and extrinsic camera parameters. In this way, a model representing the object is obtained from this set of 3D points. Finally, the object pose with respect to the camera frame ${}^{cd}\xi_{o}$ is estimated with a *Virtual Visual Servoing* algorithm [159] based on the 3D information of the object.

In the query phase, we used a single-camera approach. For each iteration, a template matching with the desired template has been performed in order to identify a valid area where extracting keypoints on the current frame (Fig. 8.8).

Then, we matched current and corresponding keypoints by using the available 3D coordinates in the trained model in order to obtain the current camera pose with respect to the object ${}^{c}\xi_{o}$. The displacement between the current and the desired position could

be easily computed as described in Eq. 8.6.

$$ {}^{c}\xi_{cd} = {}^{c}\xi_{o} * {}^{o}\xi_{cd} = {}^{c}\xi_{o} * {}^{cd}\xi_{o}^{-1} \tag{8.6} $$

After the robot movement the effective displacement could be different due to noise presence (Eq. 8.7).

$$ {}^{c}\tilde{\xi}_{cd} = {}^{c}\xi_{cd} + \Delta\xi \tag{8.7} $$

For this reason the procedure has been iterated to satisfy the condition in Eq. 8.8.

$$ \Delta\xi \leq \Delta\xi_{max} \tag{8.8} $$

where $\Delta\xi_{max}$ is related to the desired accuracy.

Experimentally, an average of 6 iterations were needed to reach a precision of $1mm$ for translation and $0.2deg$ for rotation.

The training phase allowed us to perform a PBVS task in an easy way, without any actual model of the object or a priori information. Therefore, the identification process for unknown objects becomes quite simple and immediate. In fact, the operations needed to perform the tasks have been restricted to:

- update the templates collected for the stereo camera pairs.
- recompute ${}^{cd}\xi_{o}$ through keypoints matching and triangulation.

The query phase remained unchanged.

A simple template matching in 2D has been used to find the screw and hole positions to perform the screwing task.

8.3.3 Results

In the first phase of EuRoC challenge we proposed a solution for solving a door assembly task. The uncertain position of the part needed for the assembly and the introduction of novel components make the problem non trivial. The scoring system helped us in taking into account all the aspects of the problem. The described solution has been tested with two different configurations.

The first configuration is the official setup of the benchmarking activities in Stuttgart. Unfortunately, the system was not ready for performing the entire set of tasks in the challenge during the pre-assigned temporal slot. Anyway, the available algorithms

Figure 8.9: The robot is able to insert the plastic module into the door, both in Stuttgart and laboratory facility.

worked as expected and we were able to score some points. Official results regarding the EuRoC Challenge 1 can be founded at the project website [3].

The second setup has been created in our laboratory at the University of Padova in order to independently test our methods. The use of a different hardware enhanced both reliability and portability of the whole system. We also applied the developed algorithms to a pair of door and module actually different from the ones available at the Fraunhofer IPA facility. Having more time at disposal gave us the possibility to further test and improve our system in order to achieve the objectives. Finally, we were able to correctly pick and place the module, and the framework is robust to module shifts Fig. 8.9.

[3] http://www.euroc-project.eu/

8.4 EuRoC Freestyle Phase

A key factor for the Industry 4.0 upgrade is the use of robots [160]. Nowadays, manipulators are employed for supply chains in which the same task should be accomplished several times in a repetitive manner. In the majority of the cases, human operators can understood easily how to perform the task even in complex situations, but they have not the expertise to program the robot. A useful solution could be obtained if the operator would be able to teach the robot how to perform a certain task, guiding the robot or showing himself what to do by using a robot learning by demonstration paradigm [161]. Many examples in the literature show the useful aspects of applying a robot learning by demonstration paradigm [162].

Up to now, several research groups have developed different paradigms and techniques, but only a limited number of attempts have been exploited in real industrial environments. Myers et al. [163] wanted to automatically insert a PC card into a backplane slot on the motherboard treating forces/moments as the sensed inputs and robot velocities as the control outputs. Baroglio et al. [164] believe that the robots ability to gain profit from its experiences is crucial for fully exploiting its potential. They analyzed several approaches and tested them in a classical industry-like problem: insert a peg into a hole. The task was performed while recovering from error situations, in which, for instance, the peg is stuck midway because of a wrong inclination. Neto et al. [165] presented a way to program a robot showing it what to do by using gestures and speech. The gestures are extracted from a motion sensor, namely a Wii remote controller. The Japanese company Fanuc is developing robots that use reinforcement learning to train themselves [166]. Fanucs robot learns how to pick up objects while capturing video footage of the process. The new knowledge is used to refine a deep learning model that controls robot actions. It has been proved that after about eight hours the robot reaches up to 90 percent accuracy or above, almost the same as if it was programmed by an expert. With respect to previous works, we introduce two main contributions to the stat of the art. The first is introducing in an industrial-like environment the use of a Robot Learning framework trained by means of visual information collected with no need for markers or special tools. The second is making Robot Learning and Inverse Kinematics work alongside in order to benefit from both methods.

For EuRoC challenge, ITRXCell's end user, ICPE, needed a solution for the development of an automatic electric motor coil winding and assembling. This goal has been gradually reached passing through several demonstrative phases as is shown in (Fig. 8.10).

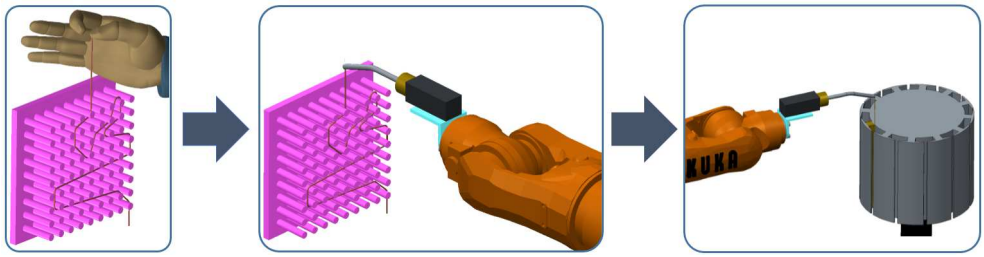


Figure 8.10: Operation sequence.

Our teams project aims to increase the competitiveness in the European electric vehicles and motors manufacturing. Automation is already applied at different levels in this field, nevertheless it is facing strong competition from Countries with low labor cost. Increasing process efficiency would strengthen a very critical sector for Europe, as it is expected to garner $22.32 billion by 2022, registering a CAGR of 3.7% during the forecast period 2016-2022 [167]. Thanks to its flexibility, the process can easily adapt to new developed motors, without the need of expensive and time-consuming changing in the layout. Particularly, the removal of auxiliary special wire guides implies a reduction of setup times of the winding machine by 50%. The robotic-based system does not require any machining of new fittings for every new production batch. This will lead to an additional reduction of setup costs by 70%, mainly in terms of effort.

8.4.1 Task and System Description

The first EuROC task-oriented phase is the Freestyle. In this phase, we focused on developing a solid learning system and a reliable human-robot interface. Indeed, the Freestyle objectives were agreed in order to put the basis of the following rounds. The same approach will be used during the Showcase to wind up the stator coils of an electric machine. The key robot action is the winding motion around a fixed point as an initial step towards the final goal to wind the coil of a real electric motor. In the Freestyle activities, the robot has to learn an arbitrary path. The framework is shown in Fig. 8.11 The team is highly skilled in learning systems and this expertise has been exploited to teach the robot how to unroll a wire following a specific path in order to pass a wire through a peg grid composing different possible routes, as shown by an operator. The trajectory is decided by a person not knowing the system to demonstrate the consistency of the approach used. The operator teaches the selected path moving a tool in a natural manner to deploy the wire through a pin table, for a relatively low number of demonstrations. The user guides the copper wire through the pole grid using

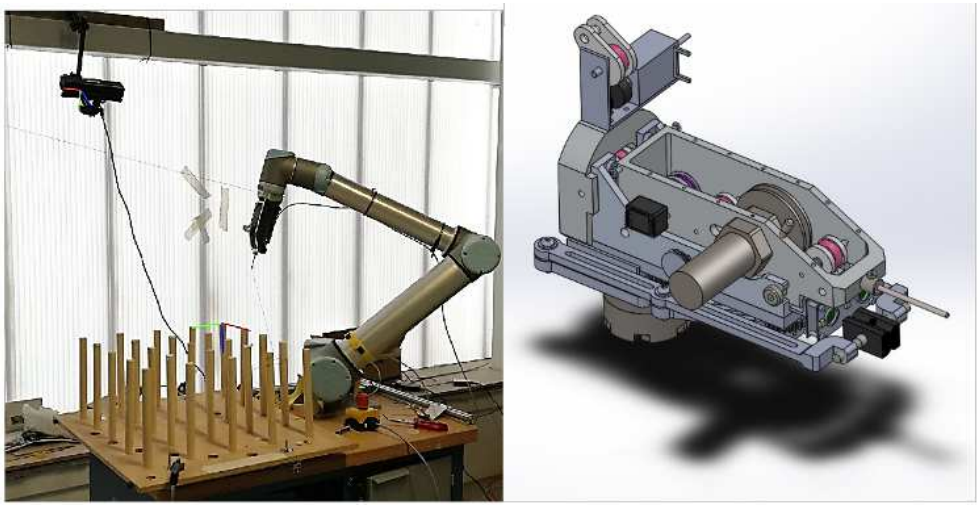


Figure 8.11: System setup.

an ad-hoc designed tool maintaining the wire constantly stretched, and preventing the possibility of knotting. In order to select a pole to pass through, the operator has to roll up the wire twice around each choice. The system records the covered trajectories by using a camera network composed by both 2D and 3D cameras. A 6 DoF robot manipulator, equipped with the same custom wire deployment system, has to replicate the motion of the operator and unroll a wire along the same path taught by them. The result is considered correct if the robot is able to replicate the pole sequnce in the exact same order selected by the operator. Both the operator and the robot starts from a fixed position. The camera network system has also the capability to monitor the workspace by detecting and tracking humans. In this phase, the robot motion stops as soon as a human being is detected in a danger zone around the robot by the cameras, as a first step of a more advanced safely controlled environment. Moreover, the tool has to maintain the tension of the wire, while allowing the user to detach the tool for demonstrating the task and to attach it back on the robot when finished. During the Freestyle round, no sensor is integrated in the end effector, since the only aim of the robot is to copy the operator motion. More advanced features will be added starting from the Showcase round to improve the system capability to work within a certain tension range.

As the robot deploys the wire, particular attention has to been kept in order to prevent the wire from getting stuck or break. To do so, it is important that the robot performs a human-like movement, choosing a smart way for moving from one pole to another, avoiding useless change of direction or turnabout. These observations have been considered by developing a learning by demonstration framework.

8.4.2 Methodology

In our solution, we made Robot Learning and Inverse Kinematics work together in order to accomplish a more general and robust solution. The basic idea is to take advantage of the human capability to find a solution by simply looking at the problems, while leaving to the robotic system the hand mathematical computation. In our context, for people is very easy to find the path to follow in order to pass through a series of poles, while it is very difficult (or even useless) for them to compute the robot joint positions to guide the robot end effector along the same trajectory without interfering with the copper wire. In order to do so, the set of useful information extracted from the trajectories performed by humans has been described by using a GMM [10], while a GMR has been used for retrieving an unified smoothed trajectory for the robot TCP. Therefore, the learned trajectory has been translated from Task Space, in the tracking system reference frame, into robot Joint Space by means of a inverse kinematic engine.

During this work, we considered mainly three aspects in order to make the robot correctly reproduce the operator actions:

- Detect the selected poles in the right order.
- Find the best entrance and exit position for the pole wrapping.
- Make the robot deploy the wire correctly.

Poles selection

Starting from the trajectory extracted from the camera network system, our goal is to detect which poles the operator selected and in which order. The information at our disposal was already transformed and projected on the 2D plane, so the input data is a sequence of (x, y) coordinates of the tool position. The algorithm used is a derived from the consensus algorithm. The solution of the consensus problem is the result of the agreement among a number of processes (or agents). The result we would like to achieve is the pole selected by a person while deploying the wire. Basically, the consensus problem requires agreement among a number of agents for a single data value as well as our poles selection algorithm seeks at which poles have been visited and on which order. Some of the processes could be unreliable since the visual system has estimated them wrongly, therefore our selection algorithm should be able to confirm the information coming from a single point the trajectory by compering it with the others. In the same way, consensus protocols must be must somehow put forth their candidate values, and agree on a single consensus value.

Algorithm1 Path detection

```
1:  ←trajectory      // sequence of (x,y) coordinates
2:  visited_poles ← 0
3:  n ←rows(trajectory)
4:  np ← 25      // num poles
5:  still points remotion
6:  for all (x,y) ∈ trajectory do
7:        clustering      // ∀(x,y) ∈ trajectory find pole ∈ [1,np] : (x,y) ∈ pole belonging area
8:  for all p ∈ [1, np] do
9:        cell_duration[p] ← ∑_{i=1}^{n} i ⇔ trajectory[i] == p
10:       count_visits[p] ← ∑_{i=1}^{n} i ⇔ trajectory[i] == p
11:       mean_time[p] ← cell_duration[p] / count_visits[p]
12:       visited_poles ← visited_poles + 1
13: count ← ∑_{i=1}^{np} count_visits[i]
14: threshold ← floor(count / visited_poles)
15: fixed_treshold ← 40
16: if threshold > fixed_threshold then
17:       threshold = fixed_threshold
18: for all p ∈ [1, np] do
19:       if count_visits[p] > threshold then
20:             ordered insertion of p in final_path based on mean_time
```

Figure 8.12: Pole selection algorithm.

In our consensus algorithm, adapted for this particular case, we starts dividing the grid in different areas belonging to the ”nearest” pole without overlapping, so that every pole is in the center of a square. The idea is to assign to each pole an afference area homogeneously distributed. After the grid division, we perform a sort of clustering operation, where each point is substituted with the relative pole area. Once we count the number of points belonging to each pole, a threshold helps in recognizing the selected poles, without mistakenly choosing poles where the tools passes often without selecting them. It is worth to notice that each pole can be visited only once in a specific trajectory. Considering the visit order helps in correctly detect the poles in the right order. A preprocessing phase is needed for remove the still periods in which the tool is motionless in a fixed point. This case could alter the outcome, since it would result like many consecutive samples in the same pole area, seeming like as that pole has been selected. The poles detection algorithm is described accurately in Fig. 8.12.

Entrance and exit position estimation

In order to avoid breaks of the wire or tangles it is important that the robot begin and end his winding motion in the correct place. The correct place depends mainly on the previously visited pole. For example, if the previously visited pole is above on the left with respect to the currently selected pole, the tool should come from left. After the implementation of poles selection, we only know which poles have been chosen,

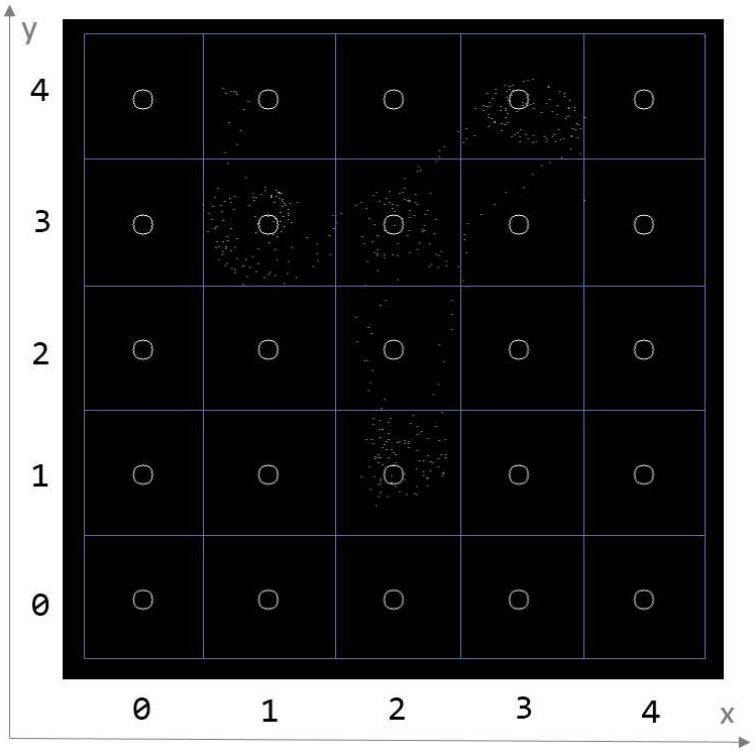


Figure 8.13: Trajectory grid.

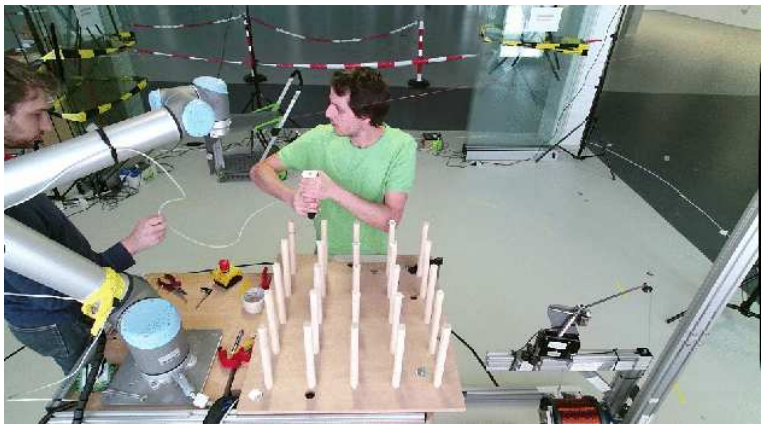


Figure 8.14: Human demonstration.

but we do not know anything about how to practically perform the winding. In order to do so, we exploit the human knowledge and expertise. Usually, a person is capable to understand which is the best way to reach a fixed point also dealing with constraints. Accordingly, we can take advantage of the human operator knowledge and overcome the planning limits.

Our goal is to compute the (x, y) coordinates of the start winding point and the end winding point. These coordinates are computed using a probabilistic framework. We use the start and the end winding point coordinates recorded from many winding tests performed by many different operators. The human expertise and knowledge give us the best way to overcome this critical issue in a probabilistic way. Furthermore, using many experiments performed by different subjects brings generality into the system, since every person could think to a different, although correct solution. The use of several executions allows the achievement of a generalized solution, which takes into account the intrinsic variability in the tasks. We obtained such results by using a Robot Learning by Demonstration paradigm, able to build a robust model of the coordinates starting from a limited number of demonstrations. Another vantage of this solution is that it could be used also by unskilled operators, since no further information or training phases are needed. The interesting coordinates are selected from the operators recorded tool trajectory. The selection has been made empirically since there is no need of great precision, in fact it could lead to overfitting.

GMM [10] Chap. 3 is used as probabilistic framework to predict the (x, y) coordinates.

In order to build the probabilistic model we introduces two fictitious poles in the system: *pole -1* and *pole* $+\infty$. The first one represents the robots starting point, while the second one represents the final goal, both outside the pole grid. The collected information are sufficient in order to allow a mapping of every possible combination of poles, profiting by the grid symmetry. Considering data collected from S subjects, each of them completed the task T times and for each task he chose P different poles. The total number of data sample is $N = S * T * (P+1)$. A single data in input at the framework ζ_j, $1 \leqslant j \leqslant N$ is described in Eq. 8.9.

$$\zeta_j = \{\alpha_w, \alpha_h, \beta_w, \beta_h, \gamma_x, \gamma_y, \lambda_x, \lambda_y\} \in \mathbb{R}^8 \tag{8.9}$$

with:

- α_w, α_h respectively width and high position of the previous pole.

- β_w, β_h respectively width and high position of the current pole.
- γ_x, γ_y respectively x and y coordinates of the exit position from the previous pole.
- λ_x, λ_y respectively x and y coordinates of the entrance position from the current pole.

The resulting probability density function is computed Eq. 8.10:

$$p\left(\zeta_j\right) = \sum_{k=1}^{K} \pi_k \mathcal{N}\left(\zeta_j; \mu_k, \Sigma_k\right) \tag{8.10}$$

with:

- π_k prior probabilities.
- $\mathcal{N}\left(\zeta_j; \mu_k, \Sigma_k\right)$ Gaussian distribution defined by μ_k and Σ_k, respectively mean vector and covariance matrix of the k-*th* distribution.

The number of Gaussian components used in the model has been estimated empirically, showing good results with k = 10. The GMR provides smooth and generalized exit and entering points for the considered poles starting from the GMM. Every exit and entering points and their covariance are estimated from the known visited poles using Eq. 8.11 and Eq. 8.12

$$\{\hat{\gamma}_x, \hat{\gamma}_y, \hat{\lambda}_x, \hat{\lambda}_y\} = E\left[\{\gamma_x, \gamma_y, \lambda_x, \lambda_y\} | \{\alpha_w, \alpha_h, \beta_w, \beta_h\}\right] = \sum_{k=1}^{K} \eta_k \{\hat{\gamma}_x, \hat{\gamma}_y, \hat{\lambda}_x, \hat{\lambda}_y\} \tag{8.11}$$

$$\hat{\Sigma}_s = Cov\left[\{\gamma_x, \gamma_y, \lambda_x, \lambda_y\} | \{\alpha_w, \alpha_h, \beta_w, \beta_h\}\right] = \sum_{k=1}^{K} \eta_k \hat{\Sigma}_{\{\hat{\gamma}_x, \hat{\gamma}_y, \hat{\lambda}_x, \hat{\lambda}_y\}, k} \tag{8.12}$$

with:

- η_k, the weight of the k*th* Gaussian component through the mixture.
- $\{\hat{\gamma}_x, \hat{\gamma}_y, \hat{\lambda}_x, \hat{\lambda}_y\}$, the conditional expectation of $\{\gamma_x, \gamma_y, \lambda_x, \lambda_y\}$ given $\{\alpha_w, \alpha_h, \beta_w, \beta_h\}$.
- $\hat{\Sigma}_{\{\hat{\gamma}_x, \hat{\gamma}_y, \hat{\lambda}_x, \hat{\lambda}_y\}, k}$, the conditional covariance of $\{\gamma_x, \gamma_y, \lambda_x, \lambda_y\}$ given $\{\alpha_w, \alpha_h, \beta_w, \beta_h\}$.

The generalized form of the motions required only weights, means and covariances of the Gaussian components calculated through the EM algorithm.

Robot movement

Once the preprocessing phase have been completed and we have obtained the selected poles, the entrance coordinates and exit coordinates, in the right order, we have the complete information needed in order to make the robot repeat the operator task by using the inverse kinematics. We use Trac-IK [168] as inverse kinematic motor. The planner is RTT connect [169] from the OMPL library [170]. The planner includes an obstacle avoidance algorithm, in order to avoid the poles. We use MoveIt [171] as interface for planning and visualization in a virtual environment. After winding around every pole, the end effector is lifted a few millimeters, in order to avoid the wire from getting stuck. With our solution, once completed the preprocessing phase everything is handled autonomously from the robot inverse kinematics.

8.4.3 Results

Since the project is structured as a challenge, we need to obtain the correct result in the shortest time in order to gain a good score. In the final test, a person shows 5 arbitrary paths previously selected from an external subject by moving the end effector in a 25 peg grid for a maximum time of 1 minute. No special marker or material has been placed on the person or on the tool. The system records the covered trajectories by using only the camera network system. An automatic tool has been developed to extract useful data from videos with almost no human intervention. A very complex and robust user interface has been implemented. *Metric I* measures the mean time needed to compute the information provided by the camera network system after the demonstration stops. The robot learns each path and passes a wire through the pegs composing the trajectory. Learning frameworks are usually based on probabilistic models built from a series of previous demonstrations called training set. An initial training set of 40 examples has been used as a basis to compute the robot trajectory. Anyway, it could not cover all possible paths, in those cases the model needs to be updated. Moreover, the operator should be able to check the validity of a novel demonstration as soon as possible. *Metric II* measures the mean time needed to update the model. Finally, the user should perform a low number of demonstrations to obtain the desired robot motion. *Metric III* measures the mean number of examples needed to learn a selected path in addition to the initial demonstration. Targets for each metric have been selected by looking at expectation of our industrial partner and taking into account the state of the art in the field.

The time needed to compute the data recorded by the camera network was in mean

56.24s. We are able to provide an updated model starting only by the initial model and the data acquired during last demonstration in 2.20s. Nevertheless, the initial model has been always sufficient in order to compute the correct path during all the 5 different paths, resulting in 0 additional demonstrations.

The system guarantees high success rate, high responsiveness and low effort for humans. In fact, even in the worst cases, we over performed the targets by obtaining 58.21s, 2.21s, and 0 additional demonstrations respectively for Metric I, II, and III. The results are summarized in Tab. 8.1.

Metric	Description	Achievement	Worst case
Metric I	Time needed to extract the demonstrated trajectory	56.24s	58.21s
Metric II	Time needed to update the robot model	2.20s	2.21s
Metric III	Number of additional demonstrations	0	0

Table 8.1: Final results for Freestyle phase.

8.4.4 Conclusions

In this project, we presents a Robot Learning framework able to acquire information from a human demonstrations by using only a camera network system. The framework has been paired to an Inverse Kinematics engine in order to make the robot deploy a wire along a grid of poles following the same pole sequence as the human operator. We measure a set of metric in order to validate our system in accordance with the requests of a real industrial partner. We uses this system as a basis for the industrial winding of coils for electrical motors to be used in the automotive market.

8.5 EuRoC Showcase Phase

As previously described, ITRXCell goal for the EuRoC challenge is the development of a robotized system for producing complete and working stators for electric motors. The interest for electric motors has increased in the last years to reduce the use of fossil fuels for environmental reasons, with the ideal goal to eliminate non-renewable energy resources in few decades. ICPE (End user) has as main goal the obtaining of innovative technologies and products, efficient and competitive, without harmful impact on the environment. During the Showcase round, UNIPD and ITR (Research team) and STAM (System integrator) collaborated to develop an automatic tool for electric motor stators winding focusing on a specific category of motors. The learning algorithm developed during the Freestyle round has been extended for winding a pole, to be mounted automatically afterwards on a motor stator. The system takes into account coil dimensions (height, width, and depth), number of turns, wire thickness and allowed tension to compute the robot trajectory. A set of quality checks has been performed both online and after the winding procedure by following ICPE instructions to validate the resulting coil in terms of electrical performances.

This work aims to reduce costs and increase flexibility with the following contributions:

- Important reduction of setup time and costs of the winding machine, thanks to the simplicity and flexibility of the proposed approach.
- Increase in the quality of the final motors, thanks to the increased amount of copper that the robot will be able to insert in each coil with respect to manual winding.
- Possibility to parallelize the winding operations, dramatically increasing production rate.
- Decreased number of defected cores, thanks to an advanced quality inspection system.
- Reduction of environmental impact of the production process, thanks to a reduction of wasted copper wire.

8.5.1 Introduction

An automatic system for coil winding has to be affordable to a wide range of users: from small-medium enterprises (SMEs), producing small batches of motors and fre-

quently changing products design, to big companies, having a market request of several thousand standard units. The low flexibility of automated winding machines [172], i.e. the time and costs required to switch from one design to another, coupled to their high cost (up to 100k Euros), force small manufacturers (especially SMEs) to employ human operators in this task. The handcrafted job is obviously much more flexible, but more expensive (because of labor cost and equipment), and for the worker it is distressing, frustrating and repetitive. Few attempts of robotic cell for coil winding have been made [172]. In this work, we aim at achieving the product flexibility required for this business sector by developing an interactive robotic cell for this task. Such a reconfigurable cell has been provided with learning capabilities. The cell is suitable for winding the coils of several kind of electric machines, starting from the information of a simple teaching interface that can be easily used by operators without specific skills in robotics. The concept of a flexible production will use a needle winding technique. The production process is divided into coils manufacturing and insertion of these on the stator. The coils are wound on frames, after which they are mounted onto the stator. For this particular application the winding process is restricted to concentrated windings. However, distributed windings or even complex winding schemes are achievable by winding individual coils. The proposed production process will have the potential to allow three dimensional shapes of the coils and complex winding schemes.

During the project, we faced the following challenges:

- Teach the robot how to properly wind the coils of stator/rotors.
- Robotize the manufacturing process of electric machines, in particular the winding of coils on stator or rotor cores.
- Detect and report non-compliances in the process of the coil winding.

In the proposed work, the selected electric motors have the following features:

- Frameless torque motors designed to be compact, high performance and cost effective.
- Allow direct coupling with the payload, eliminating parts of mechanical transmission.
- Maintenance free.
- High energy NdFeB magnets maximize torque density.

Main applications for the proposed motors are electric vehicles, machine tools, laser scanning and printing, motion simulators, rotary stage, robots, tracking systems.

8.5.2 State of the art of Electric motors

The need of an alternative power supply system for cars will be a crucial issue in the next years. Up to now, important steps forward have been made in the electrical motors. Factories like Tesla produce cars whose motors performances are comparable with traditional motors ones.

The desired goal of our end-user is the development of an automatic tool, able to create autonomously an electric motor. Up to now, there are industrial machines able to wind up coils. These machines are very expensive and they are not flexible.

The electric machines manufactures have to deal with uncertain sales volumes. As follow, the batch sizes in the manufacturing process are varying. Furthermore, continuous efforts are taken by the industries R&D departments to develop optimized electric machines with increased efficiency, increased power density, decreased manufacturing cost, etc. This also leads to currently uncertain motor designs for its manufacturing processes. In the product lifetime its design may change several times.

An automated winding machine requires stators batches of minimum 30.000 units/year. A semi-automated winding machine requires stators batches of minimum 6000 units/year. The batches are related with only one stator type. For a semi-automated machine, in order to wind a new type of stator is necessary to invest between 10.000 and 15.000 EUR in dedicated tools. The tool swapping will require between 2 - 4 hours.

Also there is a trend in the manufacturing industry to work with minimum or even zero stocks. The products will be manufactured after receiving the orders. The development of flexible production technologies that can be adapted to varying motor constructions is an existing concern as long as manufacturing uncertainties still exist. The process related to the coils manufacturing and theirs transfer/insertion into the stator are addressed.

The concept of a flexible production will use a needle winding technique. The production process is divided into coils manufacturing and insertion of these on the stator. The coils are wound on frames, after which they are mounted onto the stator. For this particular application the winding process is restricted to concentrated windings. However, distributed windings or even complex winding schemes are achievable by winding individual coils. The proposed production process will have the potential to allow three dimensional shapes of the coils and complex winding schemes.

Each of the five proposed electric motors requires 68.5 minutes for winding the stator coils by manual operation. This means a person is able to wind 7 stators during a normal working day of 8 hours. At ICPE the cost of winding one stator (from the proposed ones) is 8 EUR. A semi-automated winding machine which requires an operator

for handling and for doing some manual operations will wind the coils for one stator in 2 minutes. By using a semi-automated winding machine a person is able to wind 240 stators during 8 hours. The price of a semi-automated winding is 35.000 EUR. In order to have an economical production process, for a quantity higher than 6000 electric motors it is worth to purchase a semi-automated winding machine. By these investments a person will be able to wind 37.000 (includes the machine maintenance time) stator coils over one year. The same number of the stator core can be made manually by 24 workers. By using a fully automated winding machine with a purchasing cost starting from 150.000 EUR, a minimum quantity of 30.000 electric motors/year is required.

The system takes into account the coils dimensions (height, width, and depth), the number of turns in the coil, the wire thickness and allowed tension. These characteristics improve the system capabilities to compute the trajectory to be covered by the robot tool. In fact, the considered features are used to plan the path for winding a coil never seen before by the system. Of course, it will still be possible to refine the computed trajectory by teaching a better route through human demonstration. Novel demonstrations can be acquired by the learning system to iteratively improve its internal model and increase the performances of the whole winding procedure. The parameters of pole dimensions, number of turns in the coil, wire thickness and desired tension are provided as input to the robot by the operator, without the need of specific sensors to identify them. Based on the given information, the system chooses the proper coil from the coils hub and the robot tool gripper picks and places it on the adjustable winding stage. Later, the tool clamps its wire to the winding stage and starts winding the coil. A set of basic quality inspection protocols, based on turns count, wire tension and wire round distribution unity have been introduced, in order to guarantee a high standard of the winding process. A tension sensor has been integrated into the robot end effector in order to control the wire tension (Fig. 3). The output of the sensor will be used to close the loop in the controller, adjusting the joint trajectories to match the desired output. This feature allows the robot to keep the wire tension as much as possible within the prescribed range, in order to reach optimal performances of the winded coil. Finally, the robot gripper picks the wound coil and places it on the empty stator slot.

8.5.3 Achieved results

Task division

The desired task has been carried out thanks to the collaboration between all the partners. They put together their different abilities and efforts facing different aspects:

- UNIPD as Academic part of the Research Team, has developed the robot learning framework, the visual detection algorithms, the robot motion system, and the main software interface.
- ITR as Industrial part of the Research Team, has provided his expertise in quality inspection by using an online visual control of the winding process.
- STAM as System Integrator, has implemented the needed hardware providing the necessary bridge between scientific knowledge and actual end user needs.
- ICPE as end user has projected a novel concept for electric motors suitable for electrical vehicles in which the coils could be wound separately and subsequently assembled in the stator in order to facilitate the hardware and software development in the strict challenge time schedule.

8.5.4 Data

The manufacturer worked on redesign the electric motors in order to allow the winding process by mean of a robotized device: stator core electromagnetic design, stator core mechanical design, winding design, and coil frame mechanical design in order to allow an interlocking function with the stator core. For an accurate performance prediction of the motors redesign structures, Finite Element (FE) analysis is used. Five different types of electric motors were designed and optimized in order to allow a flexible production of the windings. The stator consists of a laminated steel core in whose slots is located a three phase star connected winding. The rotor consists of a magnetic steel ring on which there are placed high energy permanent magnets. Applications for the proposed motors are electric vehicles, machine tools, laser scanning and printing, motion simulators, rotary stage, robots, tracking systems.

The specifications of the proposed outer rotor frameless motors are reported in **Table 8.2**, while **Figure 8.15** shows some of the real components derived from the manufacturer designs.

8.5.5 System

The system takes into account the coils dimensions (height, width, and depth), the number of turns in the coil, the wire thickness and allowed tension. The considered features are used to plan the path for winding a coil. The planning is based on a learning framework previously developed in [11] [173]. These characteristics improve the previous system capabilities [174] [175] to compute the trajectory to be covered by

Table 8.2: Specifications of the motors produced by the manufacturer during the project.

Params	Unit	M1	M2	M3	M4	M5
Ext. diam.	mm	128	178	178	252	252
Inner diam.	mm	80	120	120	160	160
Act. length	mm	30	20	30	30	50
Rated power	W	1600	2400	3100	3000	4500
Conn. torque	N m	13	24	30	36	54
Peak torque	N m	30	53	69	82	124
Rated speed	rpm	1200	1000	1000	800	800
Noload speed	rpm	1500	1350	1200	900	850
Inertia	Kg m^2	0.09	0.03	0.032	0.055	0.06
Weight	Kg	3.7	5.65	6.65	8.5	10
Phase conn.		Y	Y	Y	Y	Y
N. of poles		14	20	20	20	20

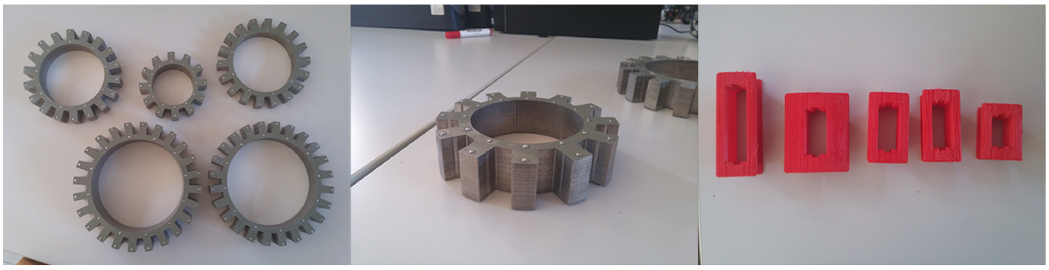


Figure 8.15: Stators and coils composing the motors used in the project.

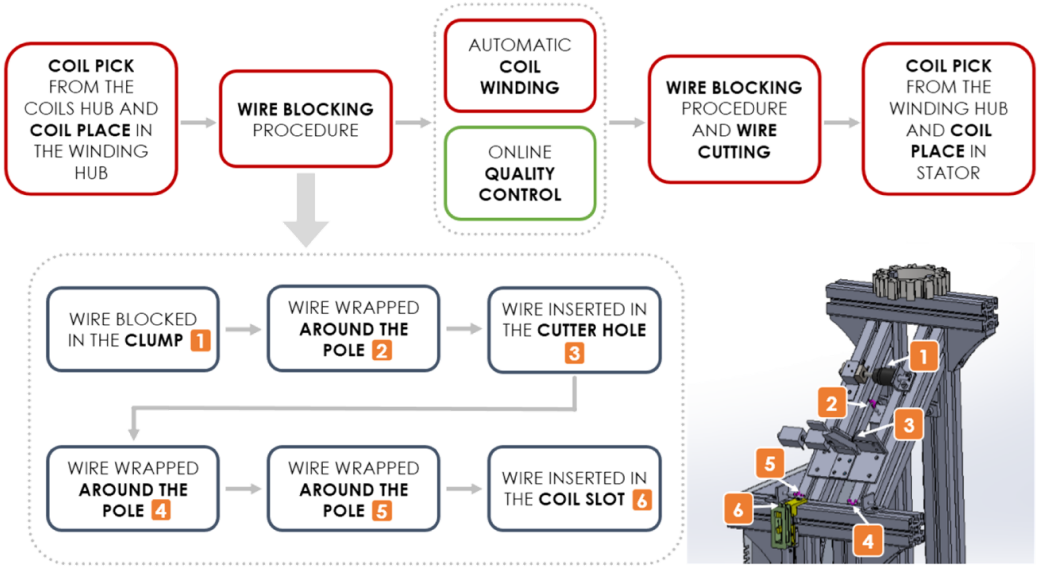


Figure 8.16: Sequence of automated winding procedure.

the robot tool. The learning system has been trained with several examples generated by using an initial set of human demonstrations, with the idea to improve the internal model in an iterative manner and increase the performances of the whole winding procedure. Of course, it is still possible to refine the computed trajectory by teaching the robot a better route through human demonstration. The objective is to enable the system to wind up a coil to be mounted on a stator never seen before. The parameters of pole dimensions, number of turns in the coil, wire thickness and desired tension are provided as input to the robot by the operator, without the need of specific sensors to identify them. Based on the given information, the system chooses the proper coil from a coil hub and the robot tool gripper picks and places it on the adjustable winding stage. Later, the tool clamps its wire to the winding stage and starts winding the coil. A set of basic quality inspection protocols, based on turns count, wire tension and wire round distribution unity have been introduced, in order to guarantee a high standard of the winding process. A tension sensor has been integrated into the robot end effector

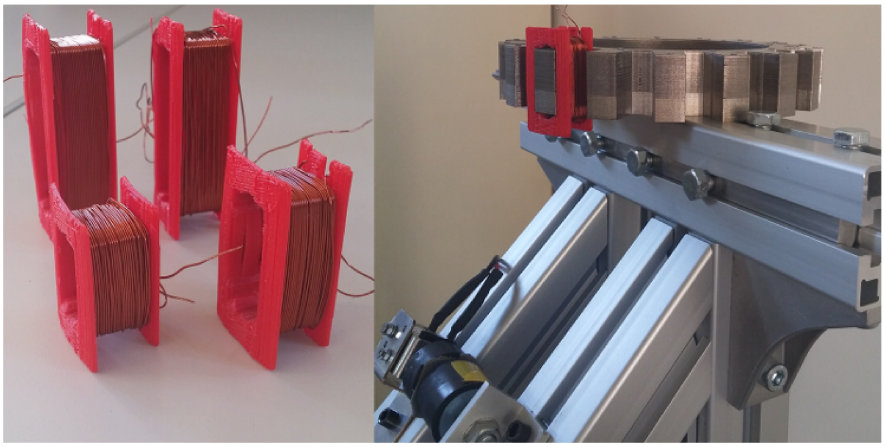


Figure 8.17: A set of wound coils ready for testing and a coil mounted on the stator.

in order to control the wire tension. The output of the sensor has been used to close the loop in the controller, adjusting the joint trajectories to match the desired output. This feature allows the robot to keep the wire tension as much as possible within the prescribed range, in order to avoid picks in the tension and reach optimal performances of the winded coil. Finally, the robot gripper picks the wound coil and places it on the empty stator slot. The process is summarized in Fig. 8.16.

The robot working process has to be able to adapt to different sizes of poles and input parameters to control the winding process. The system does not need human intervention in wire handling, online quality control, pick and place. In fact, the robotic platform has been provided with automating wire clamping and cutting, sensors for determining wire tension and wire rounds position, and a custom gripper for coil pick and place from the hub to the winding stage and from winding stage to the stator.

The whole system is completely integrated in the ROS framework [151], with the idea to easily modify and update the framework with different robots or sensors.

8.5.6 Results

We tested the complete system in a new robotic cell set up at Fraunhofer IPA in Stuttgart. This field-test was meant to prove the robot capability to work with different sizes and characteristics of the coils, and the correctness of both pick and place tasks. Tests concentrated on the winding and pick and place tasks Fig. 8.17.

A monitor shows wire tension during winding operations, in order to identify in real time picks that could break the wire. The framework has been tested with several different coils, the robot performed the required tasks after an operator manually inserted the parameters for the considered motor. The performance have been tested in

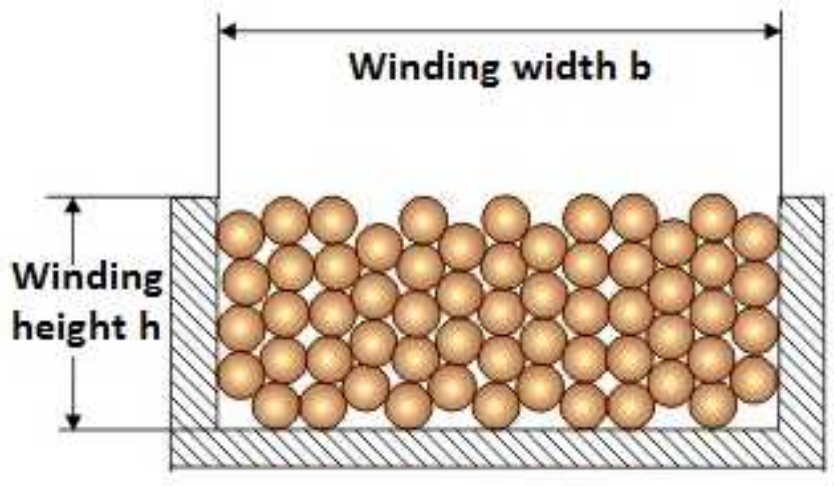


Figure 8.18: Cross section of a coil slot.

terms of productivity, repeatability, reduced manufacturing costs, flexibility, and setup time. At the end of each winding process, the coil has been compared with standards coming from actual industrial manufacturer by checking copper fill factor, inductances, resistances, and conductance at high voltage. In particular, the copper fill factor is the ratio of the copper conductors area over the total slot area.

For a section of coil (Fig. 8.18), the copper fill factor is computed as explained in Eq. 8.13.

$$copperFieldFactor = \frac{nS}{bh} \tag{8.13}$$

where:

- n is the number of copper turns (conductors).
- S is the part of cross section composed by copper conductor.
- b is the base of the cross section of a coil slot.
- h is the height of the cross section of a coil slot.

During the tests in Stuttgart, we were able to reach a Copper Fill Factor of 0.5. For electric motors used in standard applications, the copper fill factor is usually around 0.2. Moreover we achieved a valuable advantage in terms of productivity and repeatability. At the current state the production system in the manufacturer facility is able

Table 8.3: Results obtained during the testing phase at Fraunhofer IPA.

Metric	Manufacturer	Our system
Number of stator wound every 8 hours	7	15
Correct wound coils	90%	50%
Mean Copper fill factor	0.2	0.5
High voltage	Pass	Pass
Resistance	Pass	Pass
Repeatability	20%	100%

to wind 7 stators in 8 hours. In ours case the robotic arm provides 15 completed stators every 8 hours. Finally, the repeatability has also increased in a very significant manner due to the robotized approach. On the other hand, we faced a major problem with an increased number of faulty products. The results can be certainly improved with a more accurate tuning of the overall system, but some parts of the framework should be revised in order to avoid failures. A simple example regards the material used for the coils. It was too fragile and sometimes it broke while winding the copper wire. The breaking problem could be avoided by 3D printing the coils perpendicularly with respect to the winding direction. The geometry of the piece would have helped in making it more robust. Another possibility could have been to completely change the plastic material, for example by using nylon instead of polylactic acid. Anyway, it is worth to notice that the copper material (the most expensive one) used for the spare parts can be recycled in the very same process, and it had not been wasted.

Another very important aspect of our system is the high flexibility provided with respect to industrial winding machines available in the market. The capability to switch between different types of motors with minimum cost for additional tooling is essential. Commercial winding machines usually do provide very limited flexibility with expensive additional tools requested to wind different stators types. Moreover, the time needed for switching from one tool to another is quite long taking from some hours to a day. With our system, a complete change of the entire production from a motor type to a different one require more or less 15 minutes, reducing drastically the minimum number of pieces for a sustainable production, and opening the market to small-medium enterprises with a low margin of investment.

8.5.7 Conclusions

In this work, we introduced an innovative automatic technology able to increase considerably the competitiveness of European enterprises operating in the electric motors manufacturing field. The system is meant to introduce a flexible approach for winding motor coils of different types, sizes, and power with a reduced and limited human intervention in the process. Nevertheless, if necessary, the human operator can still be part of the loop in order to improve the performance of the system. The proposed technology has the potential to reduce costs, time, risks for electrical machine manufacturers in the near future and put the basis for a different way of producing high performance electrical components with an improved copper fill factor.

Chapter 9

Conclusions

The thesis proposed a set of general subject-independent frameworks to model human motion in order to actuate online robotic devices, with both industrial and rehabilitation purposes. In particular, we developed subject-independent frameworks with three different objectives:

- use physiological signals to control wearable devices;
- exploit physiological signals in industrial contexts;
- apply Robot Learning in Industry 4.0 dynamic environments.

Physiological signals, like Electromyography (EMG) signals, are non-stationary information collected from human beings. Using such signals to generalize among several repetitions of a certain task is challenging due to the peculiar characteristics of physiological data. The usual approach when working with this kind of signals focuses on subject-specific solutions. In this thesis, we proposed a subject-independent framework to control robotics prosthesis or exoskeletons by using physiological signals. A robust preprocessing phase has been developed in order to clean and smooth the signals from variability due to noise or outliars. Wavelet Transform (WT) have been applied to the raw signals to perform an analysis in both time and frequency. In conjunction to this first transformation, the preprocessing phase included signals smoothing and normalization. An additional preprocessing phase is possible when working offline. Dynamic Time Warping (DTW) computed a time stretching and distortion in order to face changes in the movement velocity. By removing artifacts, we have been able to highlight the peculiar common characteristics of the specific motion along several different trials. The refined signals have been used to train offline a probabilistic model, namely a Gaussian Mixture Model (GMM), able to represent the signal as a weighted

sum of Gaussian components. A regression technique, namely Gaussian Mixture Regression (GMR), has been used to continuously estimate the joint bending angles to control a robotic device by exploiting only physiological signals. Instead, we exlpoited a classification technique, namely Gaussian Mixture Classification (GMC), to choose the performed movement among a set of possible choices.

At first, the feasibility of the subject-independent solution has been tested by considering EMG signals from three subjects performing a very simple task, i.e. kicking a ball from a sitting position. The framework reached a mean accuracy of 85%, with a maximum accuracy of 93%. Tests have been expanded by considering a much larger dataset (i.e. 40 healthy persons), and more complex movements. The model was able to estimate the motion of both single ($\rho_{\alpha,\hat{\alpha}} = 0.8224$) and multiple ($\rho_{\alpha,\hat{\alpha}} = 0.8067$) joints for different movements. The performance the previous framework involving exclusively EMG signals have been compared with a framework considering additional information from Inertial Measurement Unit (IMU). Tests proved that the introduction of IMU data helped in improving significantly the generalization capabilities of the framework and consequently in obtaining a better subject-independent model, increasing the accuracy up to 10%.

Finally, a complete low cost framework has been proposed, building a 3D printed hand, recording EMG signals with a low cost device, and developing a subject-independent model from these data. The system was able to reach a mean accuracy of 76.8% showing that expensive devices or long training sessions from a specific subject are not necessary to control a prosthesis.

Moreover, we demonstrated that a priori information from a hierarchic taxonomy of hand grasps is able to provide a guideline able to speed-up and improve classification results. We also developed for the first time a taxonomy including quantitative information considering data by many different subjects.

EMG signals have been used also to predict the human movement in an industrial context. We presented a general framework for human-machine interface in industrial applications. The proposed framework improved the robot capabilities to work alongside humans or cooperate with them. It exploited two classifiers running in parallel, a HMM for motion detection and a GMM for direction classification, triggering the state of a FSM. The system has been implemented under ROS for an easier applicability to several robotic devices and applications. The main advantage of this approach is the flexibility of speed-accuracy trade-off obtained thanks to the novel dynamic stopping criteria introduced in this work.

Many industrial solutions have been deeply investigated in this thesis with the goal of solving the challenges proposed in the EuRoC competition. The goal of EuRoC is to boost the collaboration between universities and industries, in order to improve the industrial production with innovative solutions. EuRoC was organized in three steps of increasing complexity. We developed innovative Machine Learning solutions at each step, by exploiting the Robot Learning by Demonstration paradigm oriented to subject-independence, generality and robustness of the proposed frameworks.

During the first phase, we solved a relevant industrial task (i.e. mounting a door module into the car door) by dividing the main assignment in easier subtasks and by exploiting the human expertise in the most complex situations. In the second phase, different subjects taught the robot how to perform a task. The system learn how to wind up a wire through a peg table via visual human demonstrations. The third phase was an extension of the second one. The robot had to learn how to wind copper wire around 3D printed electric motor coils and finally mount the winded coils in the motor stator. The system worked with different types of coils and stators, in a robust and flexible way. We developed a flexible framework able to adapt to different situations and boost the production procedure in electric motors. With our solution, it is possible to switch from one motor model to another in a few minutes, without changing technology or reprogramming the robot. The proposed system can keep the costs low, while improving the production and the operators quality of work, by substituting them in alienating and repetitive tasks.

In this thesis, we pursued one of the most important objectives in robotics: to improve the people quality of life. The application of our framework could alleviate operators from heavy tasks, or help injured people to regain their lost functionality. The development of a subject-independent approach aimed at making these technologies available for everyone at low-cost and without long training sessions.

Bibliography

[1] R. N. Khushaba, S. Kodagoda, M. Takruri, and G. Dissanayake, "Toward improved control of prosthetic fingers using surface electromyogram (emg) signals," *Expert Systems with Applications*, vol. 39, no. 12, pp. 10 731–10 738, 2012.

[2] F. Orabona, C. Castellini, B. Caputo, A. E. Fiorilla, and G. Sandini, "Model adaptation with least-squares svm for adaptive hand prosthetics," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009, pp. 2897–2903.

[3] P. Konrad, "The abc of EMG," *A practical introduction to kinesiological electromyography*, no. April, pp. 1–60, 2005.

[4] K. Kiguchi and Y. Hayashi, "An EMG-Based Control for an Upper-Limb Power-Assist Exoskeleton Robot." *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 42, no. 4, pp. 1064–1071, 2012.

[5] S. Calinon and A. Billard, "Active Teaching in Robot Programming by Demonstration," *Robot & Human Interactive Communication*, pp. 702–707, 2007.

[6] T. Matsubara and J. Morimoto, "Bilinear modeling of emg signals to extract user-independent features for multiuser myoelectric interface," *Biomedical Engineering, IEEE Transactions on*, vol. 60, no. 8, pp. 2205–2213, 2013.

[7] C. J. De Luca, "The use of surface electromyography in biomechanics," *Journal of applied biomechanics*, vol. 13, no. 2, pp. 135–163, 1997.

[8] A. Hartwell, V. Kadirkamanathan, and S. Anderson, "Person-specific gesture set selection for optimised movement classification from emg signals," in *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*. IEEE, 2016, pp. 880–883.

[9] L. Peternel, N. Tsagarakis, and A. Ajoudani, "Towards multi-modal intention interfaces for human-robot co-manipulation," in *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*. IEEE, 2016, pp. 2663–2669.

[10] S. Calinon, F. Guenter, and A. Billard, "On learning, representing, and generalizing a task in a humanoid robot," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 37, no. 2, pp. 286–298, 2007.

[11] S. Michieletto, N. Chessa, and E. Menegatti, "Learning how to approach industrial robot tasks from natural demonstrations," in *Advanced Robotics and its Social Impacts (ARSO), 2013 IEEE Workshop on*. IEEE, 2013, pp. 255–260.

[12] H. Lasi, P. Fettke, H.-G. Kemper, T. Feld, and M. Hoffmann, "Industry 4.0," *Business & Information Systems Engineering*, vol. 6, no. 4, pp. 239–242, 2014.

[13] B. Siciliano, F. Caccavale, E. Zwicker, M. Achtelik, N. Mansard, C. Borst, M. Achtelik, N. O. Jepsen, R. Awad, and R. Bischoff, "Euroc-the challenge initiative for european robotics," in *ISR/Robotik 2014; 41st International Symposium on Robotics; Proceedings of*. VDE, 2014, pp. 1–7.

[14] S. Pizzolato, L. Tagliapietra, M. Cognolato, M. Reggiani, H. Müller, and M. Atzori, "Comparison of six electromyography acquisition setups on hand movement classification tasks," *PloS one*, vol. 12, no. 10, p. e0186132, 2017.

[15] B. N. Taylor and C. E. Kuyatt, *Guidelines for evaluating and expressing the uncertainty of NIST measurement results*. Citeseer, 1994.

[16] Z. Ju, G. Ouyang, M. Wilamowska-Korsak, and H. Liu, "Surface emg based hand manipulation identification via nonlinear feature extraction and classification," *IEEE Sensors Journal*, vol. 13, no. 9, pp. 3302–3311, 2013.

[17] N. Bu, M. Okamoto, and T. Tsuji, "A hybrid motion classification approach for emg-based human–robot interfaces using bayesian and neural networks," *IEEE Transactions on Robotics*, vol. 25, no. 3, pp. 502–511, 2009.

[18] A. Krasoulis, S. Vijayakumar, and K. Nazarpour, "Evaluation of regression methods for the continuous decoding of finger movement from surface emg and accelerometry," in *Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on*. IEEE, 2015, pp. 631–634.

[19] D. G. Lloyd and T. F. Besier, "An emg-driven musculoskeletal model to estimate muscle forces and knee joint moments in vivo," *Journal of biomechanics*, vol. 36, no. 6, pp. 765–776, 2003.

[20] P. Gerus, M. Sartori, T. F. Besier, B. J. Fregly, S. L. Delp, S. A. Banks, M. G. Pandy, D. D. D'Lima, and D. G. Lloyd, "Subject-specific knee joint geometry improves predictions of medial tibiofemoral contact forces," *Journal of biomechanics*, vol. 46, no. 16, pp. 2778–2786, 2013.

[21] R. Valentini, S. Michieletto, F. Spolaor, Z. Sawacha, and E. Pagello, "Processing of semg signals for online motion of a single robot joint through gmm modelization," in *2015 IEEE International Conference on Rehabilitation Robotics (ICORR)*, Aug 2015, pp. 943–949.

[22] C. Castellini, A. E. Fiorilla, and G. Sandini, "Multi-subject/daily-life activity emg-based control of mechanical hands," *Journal of neuroengineering and rehabilitation*, vol. 6, no. 1, p. 41, 2009.

[23] R. Khushaba, "Correlation analysis of electromyogram signals for multiuser myoelectric interfaces," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 22, no. 4, pp. 745–755, July 2014.

[24] C. W. Antuvan, S.-C. Yen, and L. Masia, "Simultaneous classification of hand and wrist motions using myoelectric interface: Beyond subject specificity," in *Biomedical Robotics and Biomechatronics (BioRob), 2016 6th IEEE International Conference on*. IEEE, 2016, pp. 1129–1134.

[25] D. Yang, W. Yang, Q. Huang, and H. Liu, "Classification of multiple finger motions during dynamic upper limb movements," *IEEE journal of biomedical and health informatics*, vol. 21, no. 1, pp. 134–141, 2017.

[26] A. E. Gibson, M. R. Ison, and P. Artemiadis, "User-independent hand motion classification with electromyography," in *ASME 2013 Dynamic Systems and Control Conference*. American Society of Mechanical Engineers, 2013, pp. V002T26A002–V002T26A002.

[27] T. Tommasi, F. Orabona, C. Castellini, and B. Caputo, "Improving control of dexterous hand prostheses using adaptive learning," *Robotics, IEEE Transactions on*, vol. 29, no. 1, pp. 207–219, 2013.

[28] M. Atzori, A. Gijsberts, S. Heynen, A.-G. M. Hager, O. Deriaz, P. Van Der Smagt, C. Castellini, B. Caputo, and H. Müller, "Building the ninapro database: A resource for the biorobotics community," in *Biomedical Robotics and Biomechatronics (BioRob), 2012 4th IEEE RAS & EMBS International Conference on*. IEEE, 2012, pp. 1258–1265.

[29] M. Atzori, A. Gijsberts, C. Castellini, B. Caputo, A.-G. M. Hager, S. Elsig, G. Giatsidis, F. Bassetto, and H. Müller, "Electromyography data for non-invasive naturally-controlled robotic hand prostheses," *Scientific data*, vol. 1, 2014.

[30] M. Vitor-Costa, L. Pereira, R. Oliveira, R. Pedro, T. Camata, T. Abrao, M. Brunetto, and L. Altimari, "Fourier and wavelet spectral analysis of emg signals in maximal constant load dynamic exercise," in *Conference Proceedings (IEEE Engineering in Medicine and Biology Society Conf)*, vol. 1, 2010, pp. 4622–4625.

[31] S. Lee and G. Saridis, "The control of a prosthetic arm by EMG pattern recognition," *Automatic Control, IEEE Transactions on*, vol. 29, no. 4, pp. 290–302, Apr 1984.

[32] C. Loconsole, S. Dettori, A. Frisoli, C. A. Avizzano, and M. Bergamasco, "An EMG-based approach for on-line predicted torque control in robotic-assisted rehabilitation," in *Haptics Symposium (HAPTICS), 2014 IEEE*. IEEE, 2014, pp. 181–186.

[33] T. Lalitharatne, Y. Hayashi, K. Teramoto, and K. Kiguchi, "A study on effects of muscle fatigue on EMG-based control for human upper-limb power-assist," in *Information and Automation for Sustainability (ICIAfS), 2012 IEEE 6th International Conference on*, Sept 2012, pp. 124–128.

[34] M. Sartori, M. Reggiani, E. Pagello, and D. Lloyd, "Modeling the Human Knee for Assistive Technologies," *Biomedical Engineering, IEEE Transactions on*, vol. 59, no. 9, pp. 2642–2649, Sept 2012.

[35] A. Ismail and S. Asfour, "Continuous wavelet transform application to EMG signals during human gait," in *Signals, Systems amp; Computers, 1998. Conference Record of the Thirty-Second Asilomar Conference on*, vol. 1, Nov 1998, pp. 325–329 vol.1.

[36] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *Information Theory, IEEE Transactions on*, vol. 36, no. 5, pp. 961–1005, Sep 1990.

[37] F. Laterza and G. Olmo, "Analysis of EMG signals by means of the matched wavelet transform," *Electronics Letters*, vol. 33, no. 5, pp. 357–359, Feb 1997.

[38] P. Guglielminotti and R. Merletti, "Effect of electrode location on surface myoelectric signal variables: a simulation study," in *9th Int. Congress of ISEK*, vol. 188, 1992.

[39] U. Sahin and F. Sahin, "Pattern recognition with surface emg signal based wavelet transformation," in *Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on*. IEEE, 2012, pp. 295–300.

[40] A. Phinyomark, C. Limsakul, and P. Phukpattaranont, "Application of Wavelet Analysis in EMG Feature Extraction for Pattern Classification," *Measurement Science Review*, vol. 11, no. 2, pp. 45–52, 2011.

[41] R. H. Chowdhury, M. B. I. Reaz, M. A. B. M. Ali, A. A. A. Bakar, K. Chellappan, and T. G. Chang, "Surface Electromyography Signal Processing and Classification Techniques," *Sensors*, vol. 13, no. 9, pp. 12 431–12 466, 2013.

[42] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," *Expert Systems with Applications*, vol. 32, no. 4, pp. 1084–1093, 2007.

[43] ——, "Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using wavelet feature extraction," *Computers in biology and medicine*, vol. 37, no. 2, pp. 227–244, 2007.

[44] L. Chun-Lin, "A tutorial of the wavelet transform," *NTUEE, Taiwan*, 2010.

[45] Y. Sheng, "Wavelet transform," *The transforms and applications handbook*, pp. 747–827, 1996.

[46] G. Tomasi, F. van den Berg, and C. Andersson, "Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data," *Journal of Chemometrics*, vol. 18, no. 5, pp. 231–241, 2004.

[47] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 1–38, 1977.

[48] G. Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, pp. 461–464, 1978.

[49] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *2nd International Symposium on Information Theory*, 1973, pp. 267–281.

[50] A. Barron, J. Rissanen, and B. Yu, "The minimum description length principle in coding and modeling," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2743–2760, 1998.

[51] C. S. Wallace and D. L. Dowe, "Minimum message length and kolmogorov complexity," *The Computer Journal*, vol. 42, no. 4, pp. 270–283, 1999.

[52] S. Calinon and A. Billard, "Incremental learning of gestures by imitation in a humanoid robot," in *Proceedings of the ACM/IEEE international conference on Human-robot interaction*. ACM, 2007, pp. 255–262.

[53] A. Keil, T. Elbert, and E. Taub, "Relation of accelerometer and emg recordings for the measurement of upper extremity movement." *Journal of Psychophysiology*, vol. 13, no. 2, p. 77, 1999.

[54] A. Gijsberts and B. Caputo, "Exploiting accelerometers to improve movement classification for prosthetics," in *Rehabilitation Robotics (ICORR), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1–5.

[55] L. Liu, X. Chen, Z. Lu, S. Cao, D. Wu, and X. Zhang, "Development of an emg-acc-based upper limb rehabilitation training system," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 3, pp. 244–253, 2017.

[56] R. N. Khushaba, A. Al-Timemy, S. Kodagoda, and K. Nazarpour, "Combined influence of forearm orientation and muscular contraction on emg pattern recognition," *Expert Systems with Applications*, vol. 61, pp. 154–161, 2016.

[57] K. Demet, N. Martinet, F. Guillemin, J. Paysant, and J.-M. Andre, "Health related quality of life and related factors in 539 persons with amputation of upper

and lower limb," *Disability and Rehabilitation*, vol. 25, no. 9, pp. 480–486, 2003.

[58] J. Pons, E. Rocon, R. Ceres, D. Reynaerts, B. Saro, S. Levin, and W. Van Moorleghem, "The manus-hand dextrous robotics upper limb prosthesis: mechanical and manipulation aspects," *Autonomous Robots*, vol. 16, no. 2, pp. 143–163, 2004.

[59] G. McGimpsey and T. C. Bradford, "Limb prosthetics services and devices," *Bioengineering Institute Center for Neuroprosthetics Worcester Polytechnic Institution*, 2008.

[60] N. Unwin, "Epidemiology of lower extremity amputation in centres in europe, north america and east asia," *BJS*, vol. 87, no. 3, pp. 328–337, 2000.

[61] I. J. Krajbich, "Lower-limb deficiencies and amputations in children," *JAAOS-Journal of the American Academy of Orthopaedic Surgeons*, vol. 6, no. 6, pp. 358–367, 1998.

[62] R. R. Ma, L. U. Odhner, and A. M. Dollar, "A modular, open-source 3d printed underactuated hand," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*. IEEE, 2013, pp. 2737–2743.

[63] J. Yang, E. P. Pitarch, K. Abdel-Malek, A. Patrick, and L. Lindkvist, "A multi-fingered hand prosthesis," *Mechanism and Machine Theory*, vol. 39, no. 6, pp. 555–581, 2004.

[64] U. Scarcia, G. Berselli, G. Palli, and C. Melchiorri, "Modeling, design, and experimental evaluation of rotational elastic joints for underactuated robotic fingers," in *Humanoid Robotics (Humanoids), 2017 IEEE-RAS 17th International Conference on*. IEEE, 2017, pp. 353–358.

[65] A. Beneteau, G. Di Caterina, L. Petropoulakis, and J. Soraghan, "Low-cost wireless surface emg sensor using the msp430 microcontroller," in *Education and Research Conference (EDERC), 2014 6th European Embedded Design in*. IEEE, 2014, pp. 264–268.

[66] Y. Fang, N. Hettiarachchi, D. Zhou, and H. Liu, "Multi-modal sensing techniques for interfacing hand prostheses: a review," *IEEE Sensors Journal*, vol. 15, no. 11, pp. 6065–6076, 2015.

[67] T. Feix, J. Romero, H.-B. Schmiedmayer, A. M. Dollar, and D. Kragic, "The grasp taxonomy of human grasp types," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 1, pp. 66–77, 2016.

[68] M. R. Cutkosky, "On grasp choice, grasp models, and the design of hands for manufacturing tasks," *IEEE Transactions on robotics and automation*, vol. 5, no. 3, pp. 269–279, 1989.

[69] M. Atzori and H. Müller, "Control Capabilities of Myoelectric Robotic Prostheses by Hand Amputees : A Scientific Research and Market Overview," *Frontiers in systems neuroscience*, pp. 1–13, 2015.

[70] D. Farina, N. Jiang, H. Rehbaum, A. Holobar, B. Graimann, H. Dietl, and O. C. Aszmann, "The extraction of neural information from the surface EMG for the control of upper-limb prostheses: Emerging avenues and challenges," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 4, pp. 797–809, 2014.

[71] C. Castellini, P. Artemiadis, M. Wininger, A. Ajoudani, M. Alimusaj, A. Bicchi, B. Caputo, W. Craelius, S. Dosen, K. Englehart, *et al.*, "Proceedings of the first workshop on peripheral machine interfaces: Going beyond traditional surface electromyography," *Frontiers in neurorobotics*, vol. 8, p. 22, 2014.

[72] M. Santello, M. Flanders, and J. F. Soechting, "Postural hand synergies for tool use," *The Journal of Neuroscience*, vol. 18, no. 23, pp. 10 105–10 115, 1998.

[73] C. Häger-Ross and M. H. Schieber, "Quantifying the independence of human finger movements: Comparisons of digits, hands, and movement frequencies," *Journal of Neuroscience*, vol. 20, no. 22, pp. 8542–8550, 2000. [Online]. Available: http://www.jneurosci.org/content/20/22/8542

[74] C. Eppner, R. Deimel, J. Álvarez-Ruiz, M. Maertens, and O. Brock, "Exploitation of environmental constraints in human and robotic grasping," *The International Journal of Robotics Research*, vol. 34, no. 7, pp. 1021–1038, 2015.

[75] F. Stival, S. Michieletto, and E. Pagello, "Subject Independent EMG Analysis by using Low-Cost hardware," in *Systems, Man, and Cybernetics (SMC), 2018 IEEE International Conference on*. IEEE, 2018.

[76] V. Patel, P. Thukral, M. K. Burns, I. Florescu, R. Chandramouli, and R. Vinjamuri, "Hand grasping synergies as biometrics," *Frontiers in bioengineering and biotechnology*, vol. 5, p. 26, 2017.

[77] S. A. Overduin, A. d'Avella, J. Roh, and E. Bizzi, "Modulation of muscle synergy recruitment in primate grasping," *Journal of Neuroscience*, vol. 28, no. 4, pp. 880–892, 2008.

[78] R. Prevete, F. Donnarumma, A. d'Avella, and G. Pezzulo, "Evidence for sparse synergies in grasping actions," *Scientific reports*, vol. 8, no. 1, p. 616, 2018.

[79] J. R. Napier, "The prehensile movements of the human hand," *Bone & Joint Journal*, vol. 38, no. 4, pp. 902–913, 1956.

[80] J. M. F. Landsmeer, "Power Grip and Precision Handling," *Annals of the Rheumatic Diseases*, vol. 21, no. 2, pp. 164–170, 1962. [Online]. Available: http://ard.bmj.com/cgi/doi/10.1136/ard.21.2.164

[81] N. Kamakura, M. Matsuo, H. Ishii, F. Mitsuboshi, and Y. Miura, "Patterns of static prehension in normal hands," *American Journal of Occupational Therapy*, vol. 34, no. 7, pp. 437–445, 1980. [Online]. Available: +http://dx.doi.org/10.5014/ajot.34.7.437

[82] S. K. Skerik, M. W. Weiss, and A. E. Flatt, "Functional evaluation of congenital hand anomalies." *The American journal of occupational therapy : official publication of the American Occupational Therapy Association*, vol. 25, no. 2, pp. 98–104, mar 1971. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/5551518

[83] M. R. Cutkosky, "On grasp choice, grasp models, and the design of hands for manufacturing tasks," *IEEE Transactions on Robotics and Automation*, vol. 5, no. 3, pp. 269–279, Jun 1989.

[84] K. Wolf, A. Naumann, M. Rohs, and J. Müller, "A taxonomy of microinteractions: Defining microgestures based on ergonomic and scenario-dependent requirements," in *IFIP Conference on Human-Computer Interaction*. Springer, 2011, pp. 559–575.

[85] I. M. Bullock, R. R. Ma, and A. M. Dollar, "A hand-centric classification of human and robot dexterous manipulation," *IEEE transactions on Haptics*, vol. 6, no. 2, pp. 129–144, 2013.

[86] M. F. Cohen, *An introduction to logic and scientific method*. Read Books Ltd, 2013.

[87] G. D. Kessler, L. F. Hodges, and N. Walker, "Evaluation of the cyberglove as a whole-hand input device," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 2, no. 4, pp. 263–283, 1995.

[88] S.-W. Lee and X. Zhang, "Development and evaluation of an optimization-based model for power-grip posture prediction," *Journal of biomechanics*, vol. 38, no. 8, pp. 1591–1597, 2005.

[89] P. Cerveri, E. De Momi, N. Lopomo, G. Baud-Bovy, R. Barros, and G. Ferrigno, "Finger kinematic modeling and real-time hand motion estimation," *Annals of biomedical engineering*, vol. 35, no. 11, pp. 1989–2002, 2007.

[90] I. Carpinella, P. Mazzoleni, M. Rabuffetti, R. Thorsen, and M. Ferrarin, "Experimental protocol for the kinematic analysis of the hand: definition and repeatability," *Gait & Posture*, vol. 23, no. 4, pp. 445–454, 2006.

[91] Y.-H. Lee and C.-Y. Tsai, "Taiwan sign language (tsl) recognition based on 3d data and neural networks," *Expert systems with applications*, vol. 36, no. 2, pp. 1123–1128, 2009.

[92] D. Farina, R. Merletti, and R. M. Enoka, "The extraction of neural strategies from the surface EMG," *J Appl Physiol*, vol. 96, pp. 1486–1495, 2004.

[93] D. Farina, N. Jiang, H. Rehbaum, A. Holobar, B. Graimann, H. Dietl, and O. C. Aszmann, "The extraction of neural information from the surface EMG for the control of upper-limb prostheses: Emerging avenues and challenges," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 4, pp. 797–809, 2014.

[94] C. J. De Luca, "The Use of Surface Electromyography in Biomechanics," *Journal of Applied Biomechanics*, vol. 13, pp. 135–163, 1997.

[95] M. Atzori, A. Gijsberts, C. Castellini, B. Caputo, A.-G. M. Hager, E. Simone, G. Giatsidis, F. Bassetto, and H. Müller, "Clinical parameter effect on the capability to control myoelectric robotic prosthetic hands," *Journal of Rehabilitation Research and Development*, vol. 53, no. 3, pp. 345–358, 2016.

[96] V. Gregori, A. Gijsberts, and B. Caputo, "Adaptive learning to speed-up control of prosthetic hands: A few things everybody should know," *arXiv preprint arXiv:1702.08283*, 2017.

[97] H. Tsuji, H. Ichinobe, K. Ito, and M. Nagamachi, "Discrimination of forearm motions from emg signals by error back propagation typed neural network using entropy," *IEEE Transactions, Society of Instrument and Control Engineers*, vol. 29, no. 10, pp. 1213–1220, 1993.

[98] O. Fukuda, T. Tsuji, M. Kaneko, and A. Otsuka, "A human-assisting manipulator teleoperated by EMG signals and arm motions," *IEEE Transactions on Robotics and Automation*, vol. 19, no. 2, pp. 210–222, April 2003.

[99] E. Scheme and K. Englehart, "Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use," *The Journal of Rehabilitation Research and Development*, vol. 48, no. 6, p. 643, 2011. [Online]. Available: http://www.rehab.research.va.gov/jour/11/486/pdf/scheme486.pdf

[100] L. J. Hargrove, K. Englehart, and B. Hudgins, "A comparison of surface and intramuscular myoelectric signal classification," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 5, pp. 847–853, 2007.

[101] S. J. Edwards, D. J. Buckland, and J. McCoy-Powlen, *Developmental & functional hand grasps*. Slack, 2002.

[102] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 7, pp. 848–854, 2003.

[103] B. Hudgins, P. Parker, and R. N. Scott, "A New Strategy for Multifunction Myoelectric Control," *IEEE Transactions on Biomedical Engineering*, vol. 40, no. 1, pp. 82 – 94, 1993.

[104] A. Gijsberts, M. Atzori, C. Castellini, H. Müller, and B. Caputo, "Movement Error Rate for Evaluation of Machine Learning Methods for sEMG-Based Hand Movement Classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 4, pp. 735–744, 2014.

[105] B. Mulgrew, P. Grant, and J. Thompson, *Digital Signal Processing: Concepts and Applications*. Crinan Street London N1 9XW: Macmillan Education UK, 1999.

[106] A. D. C. Chan and G. C. Green, "Myoelectric control development toolbox," in *Proceedings of 30th Conference of the Canadian Medical & Biological Engineering Society*, vol. 1, no. September, 2007, pp. M0100–1. [Online]. Available: http://www.sce.carleton.ca/faculty/chan/matlab/myoelectriccontroldevelopmenttoolbox.pdf

[107] A. Gijsberts, M. Atzori, C. Castellini, H. Müller, and B. Caputo, "Movement Error Rate for Evaluation of Machine Learning Methods for sEMG-Based Hand Movement Classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 4, pp. 735–744, 2014.

[108] I. Kuzborskij, A. Gijsberts, and B. Caputo, "On the Challenge of Classifying 52 Hand Movements from Surface Electromyography," in *Engineering in Medicine and Biology Society (EMBC), 34th Annual International Conference of the IEEE*, 2012, pp. 4931 – 4937.

[109] M.-F. Lucas, A. Gaufriau, S. Pascual, C. Doncarli, and D. Farina, "Multi-channel surface {EMG} classification using support vector machines and signal-based wavelet optimization," *Biomedical Signal Processing and Control*, vol. 3, no. 2, pp. 169 – 174, 2008, surface ElectromyographySurface Electromyography. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1746809407000791

[110] M. A. Oskoei and H. Hu, "Myoelectric control systems—a survey," *Biomedical Signal Processing and Control*, vol. 2, no. 4, pp. 275 – 294, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1746809407000547

[111] M. Zecca, S. Micera, M. C. Carrozza, and P. Dario, "Control of Multifunctional Prosthetic Hands by Processing the Electromyographic Signal," *Critical Reviews in Biomedical Engineering*, vol. 30, no. 4-6, pp. 459–485, 2002. [Online]. Available: http://www.dl.begellhouse.com/journals/4b27cbfc562e21b8,5b51f41866ab7f77,5d2cbc051cfa16a4.html

[112] M. Zardoshti, B. C. Wheeler, K. Badie, and R. Hashemi, "Evaluation of emg features for movement control of prostheses," in *Proceedings of the 15th Annual*

International Conference of the IEEE Engineering in Medicine and Biology Societ, Oct 1993, pp. 1141–1142.

[113] W. Krzanowski, *Principles of multivariate analysis*. OUP Oxford, 2000.

[114] P. C. Mahalanobis, "On the generalised distance in statistics," vol. 2, no. 1, pp. 49–55, 1936.

[115] M. Pawlik and N. Augsten, "Tree edit distance: Robust and memory-efficient," *Information Systems*, 2016.

[116] M. Pawlik and N. J. Augsten, "Efficient computation of the tree edit distance," *ACM Trans. Database Syst.*, vol. 40, no. 1, pp. 3:1–3:40, mar 2015.

[117] F. Stival, S. Michieletto, E. Pagello, H. Müller, and M. Atzori, "Quantitative hierarchical representation and comparison of hand grasps from electromyography and kinematic data," in *in Workshop Proceedings of the 15th International Conference on Autonomous Systems IAS-15, Workshop on Learning Applications for Intelligent Autonomous Robots (LAIAR-2018)*. ISBN: 978-3-00-059946-0, 2018.

[118] C. Whidden, N. Zeh, and R. G. Beiko, "Supertrees based on the subtree prune-and-regraft distance," *Systematic biology*, p. syu023, 2014.

[119] P. Eccarius, R. Bour, and R. A. Scheidt, "Dataglove measurement of joint angles in sign language handshapes," *Sign Language & Linguistics*, vol. 15, no. 1, pp. 39–72, 2012. [Online]. Available: http://www.jbe-platform.com/content/journals/10.1075/sll.15.1.03ecc

[120] V. Gracia-Ibáñez, M. Vergara, J. H. Buffi, W. M. Murray, and J. L. Sancho-Bru, "Across-subject calibration of an instrumented glove to measure hand movement for clinical purposes," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 20, no. 6, pp. 587–597, apr 2017. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/10255842.2016.1265950

[121] F. Stival, M. Moro, and E. Pagello, "A first approach to a taxonomy-based classification framework for hand grasps," in *in Workshop Proceedings of the 15th International Conference on Autonomous Systems IAS-15, Workshop on Learning Applications for Intelligent Autonomous Robots (LAIAR-2018)*. ISBN: 978-3-00-059946-0, 2018.

[122] N. Dimitrova and G. Dimitrov, "Interpretation of emg changes with fatigue: facts, pitfalls, and fallacies," *Journal of Electromyography and Kinesiology*, vol. 13, no. 1, pp. 13–36, 2003.

[123] A. A. Abdullah, A. Subasi, and S. M. Qaisar, "Surface emg signal classification by using wpd and ensemble tree classifiers," in *CMBEBIH 2017*. Springer, 2017, pp. 475–481.

[124] F. H. Chan, Y.-S. Yang, F. Lam, Y.-T. Zhang, and P. A. Parker, "Fuzzy emg classification for prosthesis control," *IEEE transactions on rehabilitation engineering*, vol. 8, no. 3, pp. 305–311, 2000.

[125] X. Tang, Y. Liu, C. Lv, and D. Sun, "Hand motion classification using a multi-channel surface electromyography sensor," *Sensors*, vol. 12, no. 2, pp. 1130–1147, 2012.

[126] F. Stival, S. Michieletto, and E. Pagello, "Online subject-independent modeling of semg signals for the motion of a single robot joint," in *Biomedical Robotics and Biomechatronics (BioRob), 2016 6th IEEE International Conference on*. IEEE, 2016, pp. 1110–1116.

[127] ——, "Subject-independent modeling of semg signals for the motion of a single robot joint," in *Workshop of Robotics: Science and Systems*, 2015.

[128] S. Krishnan, R. Fox, I. Stoica, and K. Goldberg, "Ddco: Discovery of deep continuous options for robot learning from demonstrations," in *Conference on Robot Learning*, 2017, pp. 418–437.

[129] F. Stival, S. Michieletto, and E. Pagello, "How to deploy a wire with a robotic platform: Learning from human visual demonstrations," *Procedia Manufacturing*, vol. 11, pp. 224 – 232, 2017, 27th International Conference on Flexible Automation and Intelligent Manufacturing, FAIM2017, 27-30 June 2017, Modena, Italy. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2351978917304389

[130] M. Laskey, C. Chuck, J. Lee, J. Mahler, S. Krishnan, K. Jamieson, A. Dragan, and K. Goldberg, "Comparing human-centric and robot-centric sampling for robot deep learning from demonstrations," in *Robotics and Automation (ICRA), 2017 IEEE International Conference on*. IEEE, 2017, pp. 358–365.

[131] L. Peternel, W. Kim, J. Babič, and A. Ajoudani, "Towards ergonomie control of human-robot co-manipulation and handover," in *Humanoid Robotics (Humanoids), 2017 IEEE-RAS 17th International Conference on*. IEEE, 2017, pp. 55–60.

[132] A. Pfister, A. M. West, S. Bronner, and J. A. Noah, "Comparative abilities of microsoft kinect and vicon 3d motion capture for gait analysis," *Journal of medical engineering & technology*, vol. 38, no. 5, pp. 274–280, 2014.

[133] N. T. Fitter and K. J. Kuchenbecker, "Using imu data to demonstrate hand-clapping games to a robot," in *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*. IEEE, 2016, pp. 851–856.

[134] C. Yang, J. Chen, and F. Chen, "Neural learning enhanced teleoperation control of baxter robot using imu based motion capture," in *Automation and Computing (ICAC), 2016 22nd International Conference on*. IEEE, 2016, pp. 389–394.

[135] V. Venek, W. Kremser, and C. Schneider, "Towards an imu evaluation framework for human body tracking." *Studies in health technology and informatics*, vol. 248, pp. 156–163, 2018.

[136] F. Stival, S. Michieletto, and E. Pagello, "Toward a better robotic hand prosthesis control: using emg and imu features for a subject independent multi joint regression mode," in *Biomedical Robotics and Biomechatronics (BioRob), 2018 7th IEEE International Conference on*. IEEE, 2018.

[137] J. Vogel, C. Castellini, and P. van der Smagt, "Emg-based teleoperation and manipulation with the dlr lwr-iii," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sept 2011, pp. 672–678.

[138] L. Peternel, N. Tsagarakis, D. Caldwell, and A. Ajoudani, "Adaptation of robot physical behaviour to human fatigue in human-robot co-manipulation," in *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, Nov 2016, pp. 489–494.

[139] S. Tortora, S. Michieletto, and E. Menegatti, "Synergy-based gaussian mixture model to anticipate reaching direction identification for robotic applications," in *Workshop Proceedings of the 15th International Conference on Autonomous Systems IAS-15, Workshop on Learning Applications for Intelligent Autonomous Robots (LAIAR-2018) on*, 2018.

[140] K. P. F.R.S., "Liii. on lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, no. 11, pp. 559–572, 1901. [Online]. Available: https://doi.org/10.1080/14786440109462720

[141] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, p. 788, 1999.

[142] A. d'Avella, P. Saltiel, and E. Bizzi, "Combinations of muscle synergies in the construction of a natural motor behavior," *Nature neuroscience*, vol. 6, no. 3, p. 300, 2003.

[143] S. Y. Kung, *Kernel methods and machine learning*. Cambridge University Press, 2014.

[144] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains," *The annals of mathematical statistics*, vol. 41, no. 1, pp. 164–171, 1970.

[145] A. Lenhardt, M. Kaper, and H. J. Ritter, "An adaptive p300-based online brain–computer interface," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 16, no. 2, pp. 121–130, 2008.

[146] D. Christiansen, "Spectral lines: Announcing the amazing micro-mouse maze contest," *IEEE Spectrum*, vol. 14, no. 5, pp. 27–27, May 1977.

[147] H. Kitano, M. Asada, Y. Kuniyoshi, I. Noda, E. Osawa, and H. Matsubara, "Robocup: A challenge problem for ai," *AI magazine*, vol. 18, no. 1, p. 73, 1997.

[148] K. Iagnemma and M. Buehler, "Editorial for journal of field robotics-special issue on the darpa grand challenge," *Journal of Field Robotics*, vol. 23, no. 9, pp. 655–656, 2006.

[149] S. Michieletto, E. Tosello, F. Romanelli, V. Ferrara, and E. Menegatti, "Ros-i interface for comau robots," in *International Conference on Simulation, Modeling, and Programming for Autonomous Robots*. Springer, 2014, pp. 243–254.

[150] M. Carraro, M. Munaro, and E. Menegatti, "Cost-efficient rgb-d smart camera for people detection and tracking," *Journal of Electronic Imaging*, vol. 25, no. 4, p. 041007, 2016.

[151] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "Ros: an open-source robot operating system," in *ICRA workshop on open source software*, vol. 3, no. 3.2. Kobe, Japan, 2009, p. 5.

[152] H. Chen and Y. Liu, "Robotic assembly automation using robust compliant control," *Robotics and Computer-Integrated Manufacturing*, vol. 29, no. 2, pp. 293–300, 2013.

[153] M. Goyal, S. Netessine, and T. Randall, "Deployment of manufacturing flexibility: An empirical analysis of the north american automotive industry," *Available at SSRN 2077659*, 2012.

[154] P. Tavares and A. Sousa, "Flexible pick and place architecture using ros framework," in *Information Systems and Technologies (CISTI), 2015 10th Iberian Conference on*, June 2015, pp. 1–6.

[155] A. Vakanski, F. Janabi-Sharifi, I. Mantegh, and A. Irish, "Trajectory learning based on conditional random fields for robot programming by demonstration," in *Proceedings of the IASTED International Conference on Robotics and Applications*, 2010.

[156] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking." in *IEEE International Conference on Computer Vision*, vol. 2, October 2005, pp. 1508–1511. [Online]. Available: http://www.coxphysics.com/work/rosten_2005_tracking.pdf

[157] ——, "Machine learning for high-speed corner detection," in *European Conference on Computer Vision*, vol. 1, May 2006, pp. 430–443. [Online]. Available: http://www.coxphysics.com/work/rosten_2006_machine.pdf

[158] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "Orb: An efficient alternative to sift or surf," in *Proceedings of the 2011 International Conference on Computer Vision*, ser. ICCV '11. Washington, DC, USA: IEEE Computer Society, 2011, pp. 2564–2571. [Online]. Available: http://dx.doi.org/10.1109/ICCV.2011.6126544

[159] E. Marchand and F. Chaumette, "Virtual visual servoing: a framework for real-time augmented reality," in *EUROGRAPHICS 02 Conf. Proceeding of Computer Graphics Forum, Saarebrucken, Germany*, vol. 21, no. 3, 2002, pp. 289–298.

[160] M. Hägele, K. Nilsson, and J. N. Pires, "Industrial robotics," in *Springer handbook of robotics*. Springer, 2008, pp. 963–986.

[161] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, "Robot programming by demonstration," in *Springer handbook of robotics*. Springer, 2008, pp. 1371–1394.

[162] R. Dillmann, "Teaching and learning of robot tasks via observation of human performance," *Robotics and Autonomous Systems*, vol. 47, no. 2-3, pp. 109–116, 2004.

[163] D. R. Myers, M. J. Pritchard, and M. D. Brown, "Automated programming of an industrial robot through teach-by showing," in *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*, vol. 4. IEEE, 2001, pp. 4078–4083.

[164] C. Baroglio, A. Giordana, R. Piola, M. Kaiser, and M. Nuttin, "Learning controllers for industrial robots," *Machine learning*, vol. 23, no. 2-3, pp. 221–249, 1996.

[165] P. Neto, J. N. Pires, and A. P. Moreira, "Highlevel programming and control for industrial robotics: using a handheld accelerometerbased input device for gesture and posture recognition," *Industrial Robot: An International Journal*, vol. 37, no. 2, pp. 137–147, 2010. [Online]. Available: https://doi.org/10.1108/01439911011018911

[166] "This factory robot learns a new job overnight," https://www.technologyreview.com/s/601045/this-factory-robot-learns-a-new-job-overnight/, accessed: 2016-03-15.

[167] "Europe electric motor market by output power (integral horsepower (ihp), fractional horsepower (fhp) output), by application (industrial machinery, motor vehicles, heating ventilating and cooling (hvac) equipment, aerospace & transportation, household appliances, commercial applications) - opportunity analysis and industry forecast, 2014 2022," https://www.alliedmarketresearch.com/europe-electric-motor-market, accessed: 2016-03-15.

[168] P. Beeson and B. Ames, "Trac-ik: An open-source library for improved solving of generic inverse kinematics," in *Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on*. IEEE, 2015, pp. 928–935.

[169] J. J. Kuffner and S. M. LaValle, "Rrt-connect: An efficient approach to single-query path planning," in *Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on*, vol. 2. IEEE, 2000, pp. 995–1001.

[170] I. A. Sucan, M. Moll, and L. E. Kavraki, "The open motion planning library," *IEEE Robotics & Automation Magazine*, vol. 19, no. 4, pp. 72–82, 2012.

[171] S. Chitta, I. Sucan, and S. Cousins, "Moveit![ros topics]," *IEEE Robotics & Automation Magazine*, vol. 19, no. 1, pp. 18–19, 2012.

[172] K. Feldmann, A. Dobroschke, and S. Junker, "Advanced processes and systems for the automated assembly of magnetic components for electrical machines," in *1st CIRP-International Seminar on Assembly Systems, Stuttgart*, 2006, pp. 83–88.

[173] E. Tosello, S. Michieletto, A. Bisson, E. Pagello, and E. Menegatti, "A learning from demonstration framework for manipulation tasks," in *ISR/Robotik 2014; 41st International Symposium on Robotics; Proceedings of*. VDE, 2014, pp. 1–7.

[174] F. Stival, S. Michieletto, and E. Pagello, "How to deploy a wire with a robotic platform: Learning from human visual demonstrations," *Procedia Manufacturing*, vol. 11, pp. 224–232, 2017.

[175] F. Castelli, S. Michieletto, S. Ghidoni, and E. Pagello, "A machine learning-based visual servoing approach for fast robot control in industrial setting," *International Journal of Advanced Robotic Systems*, vol. 14, no. 6, p. 1729881417738884, 2017.